

# ОТЧЕТ О ПРОДЕЛАННОЙ РАБОТЕ С ИСПОЛЬЗОВАНИЕМ ОБОРУДОВАНИЯ ИВЦ НГУ

## 1. Аннотация

Гены локуса - *Pdgfra*, *Kit* и *Kdr* являются важными регуляторами развития, а нарушения их активности у человека связано с развитием ряда онкологических заболеваний. С помощью системы CRISPR/Cas9 были созданы несколько линий мышей, несущих различные делеции в локусе. Анализ последствий этих модификаций - на организменном уровне (анализ фенотипа), на уровне изменений пространственных контактов хроматина (capture Hi-C), а также на уровне активности генов (RNA-seq), позволил продемонстрировать тканеспецифичную роль 3D структуры генома в регуляции транскрипции генов локуса *Kit*.

## 2. Тема работы

Исследование роли пространственной организации генома в регуляции активности генов на примере локуса *Pdgfra/Kit/Kdr*, с использованием системы редактирования генома CRISPR/Cas9

## 3. Состав коллектива

Баттулин Нариман Рашитович, к.б.н., ИЦиГ СО РАН, заведующий лабораторией генетики развития, НГУ, преподаватель. Руководитель

Рыжкова Анастасия Сергеевна, м.н.с., ИЦиГ СО РАН. Исполнитель

## 4. Информация о гранте

РФФИ № 18-29-07022. Исследование роли пространственной организации генома в регуляции активности генов на примере локуса *Pdgfra/Kit/Kdr*, с использованием системы редактирования генома CRISPR/Cas9. Руководитель – Баттулин Нариман Рашитович. 2018-2022

## 5. Научное содержание работы

### 5.1 Постановка задачи

Известно, что существует множество уровней регуляции активности генов, таких как изменения метилирования ДНК, ковалентных модификаций гистонов и негистоновых белков хроматина. С развитием технологий высокопроизводительного секвенирования для изучения стал доступен новый уровень регуляции – изменение пространственной организации генома. Полногеномное картирование хроматиновых контактов методом Hi-C показало, что в трехмерном пространстве геном разделен на инсулированные от соседних участков фрагменты, которые были названы топологически ассоциированными доменами (ТАД) [1]. Считается, что взаимодействие энхансеров с генами происходит в пределах одного домена, т.е. ТАДы определяют, какие взаимодействия между промоторами и энхансерами могут реализоваться в организме, а какие запрещены [2]. Однако, результаты работ по экспериментальному нарушению границ ТАДов противоречивы, в ряде исследований удаление границ ТАДов не приводило к значимым изменениям активности генов. Кроме того, на сегодняшний день информация об эффекте делеций границ ТАДов имеется лишь для ограниченного числа локусов. Таким образом,

проект был направлен на оценку роли инсуляторных элементов, ограничивающих пространственные контакты между генами развития *Pdgfra*, *Kit* и *Kdr*, в поддержании их паттерна экспрессии *in vivo*. Каждый из этих генов располагается в своем собственном ТАДе, имеет свою клеточную специфичность. Для того, чтобы выяснить какую роль в регуляции активности данного локуса играет 3D организация хроматина мы решили нарушить структуру ТАДов путем удаления сайтов связывания CTCF в границах ТАДов. Для анализа последствий манипуляций я границами ТАДов мы использовали три типа клеток с контрастной активностью упомянутых генов – фибробласты, тучные клетки и меланоциты. Первичные культуры клеток были получены от мышей, несущих соответствующие делеции в локусе.

## 5.2 Современное состояние проблемы

Одним из центральных вопросов в 3D геномике остается определение биологической функции ТАДов. В их границах обнаруживаются конвекгентные мотивы связывания белка CTCF. Высокая эволюционная консервативность районов, определяющих формирование границ ТАДов свидетельствует об их важности. Согласно основной гипотезе, ТАДы ограничивают цис-действие энхансеров, таким образом препятствуя эктопической активации промоторов. В пользу этой гипотезы свидетельствуют несколько работ, в которых было показано, что нарушение структуры ТАДов приводит к изменению активности соседних генов, что, в свою очередь, приводит к фенотипическим проявлениям [3]. Однако в ряде работ ставится под сомнение универсальность инсуляторной функции всех ТАДов. Так, Desprang с соавторами исследовали роль ТАДов в регуляции генов *Sox9* и *Kcnj2* [4]. Оба гена задействованы в развитии конечностей и лежат в смежных ТАДах. Авторы показали, что делеция границы между ТАДами не приводит к значительным нарушениям в активности генов и никак не сказывается на фенотипе животных. При этом для полного слияния ТАДов понадобилось удалить не только границу, но и часть внутри ТАДовых сайтов посадки CTCF. И даже после полного слияния ТАДов эспрессия генов в них не менялась. Наиболее сильные изменения экспрессии наблюдались лишь при инверсиях крупных фрагментов, включающих границу ТАДов. По-видимому, такие инверсии сильно меняют репертуар формируемых хроматиновых петель. Кроме того, авторы подчеркивают, что доминирующее на сегодняшний день представление о неразборчивости большинства энхансеров (т.е. способности формировать эктопические контакты с неродными промоторами) тоже может быть неверно.

1. Dixon, J.R.; Selvaraj, S.; Yue, F.; Kim, A.; Li, Y.; Shen, Y.; Hu, M.; Liu, J.S.; Ren, B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 2012, 485, 376–380, doi:10.1038/nature11082.
2. Rowley, M.J.; Corces, V.G. Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* 2018, doi:10.1038/s41576-018-0060-8.
3. Lupiáñez, D.G.; Kraft, K.; Heinrich, V.; Krawitz, P.; Brancati, F.; Klopocki, E.; Horn, D.; Kayserili, H.; Opitz, J.M.; Laxova, R.; et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 2015, 161, 1012–1025, doi:10.1016/j.cell.2015.04.004.

4. Despagne, A.; Schöpflin, R.; Franke, M.; Ali, S.; Jerković, I.; Paliou, C.; Chan, W.L.; Timmermann, B.; Wittler, L.; Vingron, M.; et al. Functional dissection of the Sox9–Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. *Nat. Genet.* 2019, doi:10.1038/s41588-019-0466-z.

### 5.3. Подробное описание работы, включая используемые алгоритмы

Основными инструментами исследования в настоящей работе являлись методы 3D геномики – capture Hi-C и ChIP-seq. Анализ экспериментальных данных осуществлялся с использованием ресурсов кластера. Для обработки данных capture Hi-C использовался алгоритм Juicer. Прочтения были картированы на версию генома мыши mm10, матрицы контактов были построены из прочтений с показателем MAPQ $\geq$ 30. Нормализация осуществлялась при помощи метода VC\_SQRT.

Консервативность локуса Kit была оценена при помощи инструмента C-InterSecture [5], разработанного в нашей лаборатории. Для межвидового сравнения были сгенерированы карты областей синтении на основе попарных выравниваний различных геномов. Сначала геномы были выровнены при помощи LastZ. Затем данные выравнивания были преобразованы в net-файлы посредством Kent Utils [<https://github.com/ucscGenomeBrowser/kent>]. Из этих файлов были построены карты синтении при помощи C-InterSecture.

В ходе анализа данных ChIP-seq сырые парные прочтения были предварительно обработаны при помощи инструмента Cutadapt для исключения последовательностей адаптеров [6]. Треки геномного покрытия (bigWig файлы) были сгенерированы с использованием алгоритма Aquas с параметрами «TF» или «histone» [[https://github.com/kundaje/lab/chipseq\\_pipeline](https://github.com/kundaje/lab/chipseq_pipeline)].

5. Nuriddinov, M.; Fishman, V. C-InterSecture-a computational tool for interspecies comparison of genome architecture. *Bioinformatics* 2019, doi:10.1093/bioinformatics/btz415.

6. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011, doi:10.14806/ej.17.1.200.

### 5.4. Полученные результаты

На первом этапе алгоритм C-InterSecture был использован для сравнения пространственной организации хроматина в локусе Kit у человека (hg38), мыши (mm10), кролика (oryCun2),

собаки (canFam3) и курицы (galGal5) с использованием полученных в данной работе (для мыши), а также опубликованных данных Hi-C. Результаты сравнения показали высокую консервативность архитектуры хроматина в локусе, организованном в три смежных ТАДа, каждый содержащий гены *Pdgfra*, *Kit* и *Kdr*, соответственно (Рис.1).

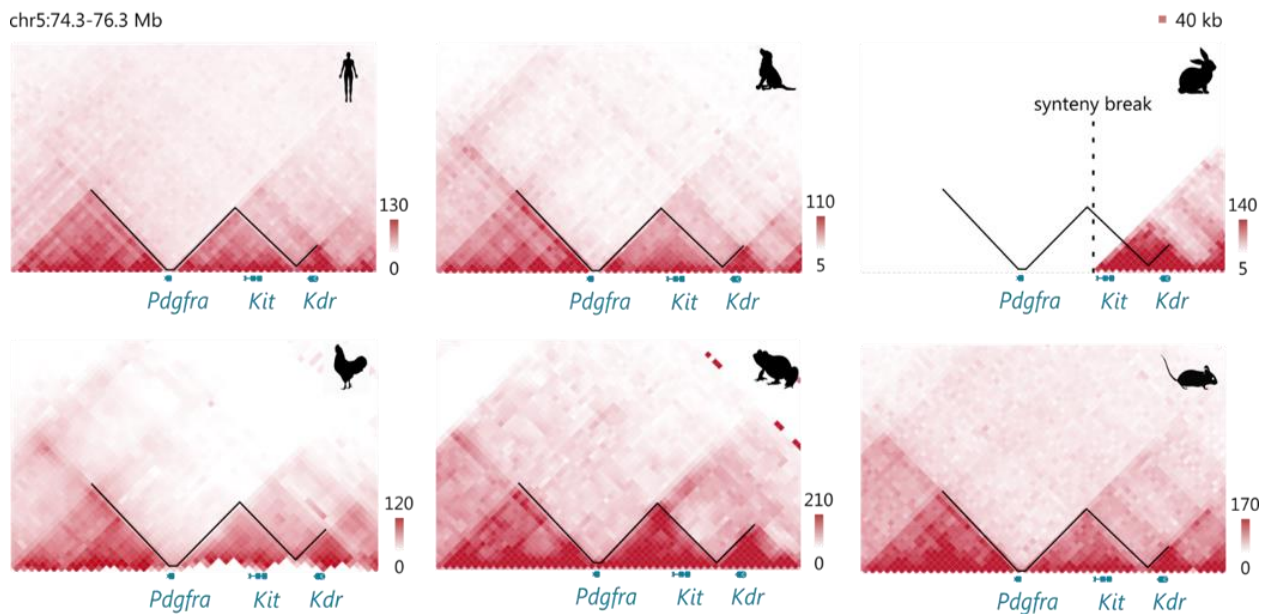


Рис. 1. Hi-C контакты представителей различных видов были перекартированы на геном мыши. Алгоритм C-InterSecture был использован для нейтрализации разницы в геномном расстоянии между синтенными локусами у разных видов.

Проведенный в проекте анализ пространственной организации исследуемого локуса методом capture Hi-C показал, что удаление границы действительно во всех случаях приводило к слиянию соседних ТАДов. Однако надо подчеркнуть, что выраженность объединения районов, относящихся к соседним ТАДам различалась в разных клеточных типах. Наиболее показательно, это можно продемонстрировать на примере удаления границы между ТАДами *Kit* и *Kdr*. Удаление этой границы приводит к исчезновению инсуляции между ТАДами, что хорошо видно при сравнении с соответствующим клеточным типом дикого типа. Самое интересное, что одна и та же мутация в разных клеточных типах (тучных клетках и меланоцитах) приводит к разным пространственным конфигурациям хроматина (Рис.2). Мы считаем, что такие разительные отличия в пространственной конфигурации локуса можно объяснить разницей в активности генов и эпигенетическим ландшафте локуса в разных типах клеток. Так хотя ген *Kit* активен в обоих типах клеток, уровень его активности сильно отличается. В тучных клетках *Kit* входит в пятерку самых активных генов в геноме (топ-5). В меланоцитах *Kit* тоже активно экспрессируется, но ближе к сотому месту (топ-100). Чрезвычайная активность *Kit* в тучных

клетках по-видимому объясняет инсуляцию пространственных контактов в непосредственно в районе гена.

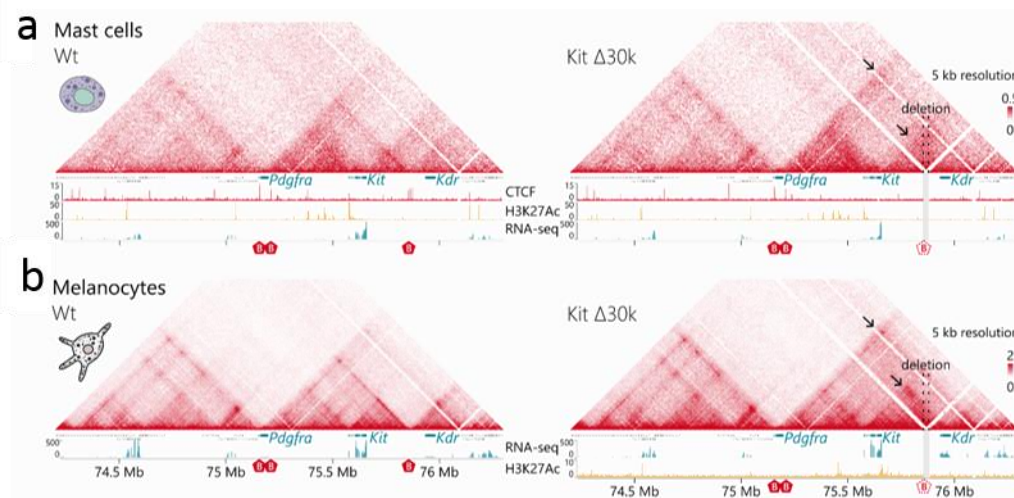


Рис. 2. Трехмерная организация генома локуса *Kit* в тучных клетках и меланоцитах. Тепловые карты сHi-C, сигналы RNA-seq и ChIP-Seq в локусе *Kit* в тучных клетках дикого типа (слева) и *Kit* Δ30k (справа) (a) и меланоцитах (b); эктопические взаимодействия между ТАДами отмечены стрелками;

## 6. Эффект от использования кластера в достижении целей работы

Обработка данных высокопроизводительного секвенирования требует больших объемов 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

0. Evelyn Kabirova, Anastasiya Ryzhkova, Varvara Lukyanchikova, Anna Khabarova, Alexey Korablev, Tatyana Shnaider, Miroslav Nuriddinov, Polina Belokopytova, Galina Kontsevaya, Irina Serova, Nariman Battulin TAD border deletion at the *Kit* locus causes tissue-specific ectopic activation of a neighboring gene. bioRxiv 2022.12.29.522177; doi: <https://doi.org/10.1101/2022.12.29.522177> (проходит рецензирование в Nature Communications)
1. Kabirova E., Lukyanchikova V., Ryzhkova A., Khabarova A. 3D organization of the genome tissue-specifically regulates activity of the development genes *Pdgfra*, *Kit*, *Kdr* in mouse. Международная конференция: 29th International Student Congress Of (bio)Medical Sciences, 8-10 июня 2022, Гронинген, Нидерланды

2. Kabirova E., Lukyanchikova V., Ryzhkova A., Khabarova A., Battulin N. Influence of TADs boundary disruption on 3D genome organisation and genes expression at the mouse *Pdgfra/Kit/Kdr* locus. Международная конференция: BGRS/SB-2022, 4-8 июля 2022, Новосибирск, Россия