

На правах рукописи

ХОРОШКО ВАРВАРА АНДРЕЕВНА

**ЛОКАЛИЗАЦИЯ ГЕНОВ В МОРФОЛОГИЧЕСКИХ
СТРУКТУРАХ ПОЛИТЕННЫХ ХРОМОСОМ
DROSOPHILA MELANOGASTER**

03.01.07 – молекулярная генетика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата биологических наук

Новосибирск 2022

Работа выполнена в лаборатории молекулярной цитогенетики Федерального государственного бюджетного учреждения науки Института молекулярной и клеточной биологии Сибирского отделения Российской академии наук, г. Новосибирск

Научный руководитель: **Жимулёв Игорь Фёдорович**, доктор биологических наук, академик РАН, заведующий лабораторией молекулярной цитогенетики, научный руководитель Федерального государственного бюджетного учреждения науки Института молекулярной и клеточной биологии Сибирского отделения Российской академии наук, г. Новосибирск

Оппоненты: **Шидловский Юлий Валерьевич**, доктор биологических наук, профессор РАН, ведущий научный сотрудник, заведующий лабораторией регуляции экспрессии генов в развитии Федерального государственного бюджетного учреждения науки Института биологии гена Российской академии наук, г. Москва

Фёдорова Светлана Александровна, кандидат биологических наук, ведущий научный сотрудник, заведующая Сектором генетики клеточного цикла Федерального исследовательского центра Института цитологии и генетики Сибирского отделения Российской академии наук, г. Новосибирск

Ведущее учреждение: Федеральное государственное бюджетное учреждение Институт молекулярной генетики Национального исследовательского центра «Курчатовский институт», г. Москва

Защита состоится: 08.06.2022 г. в 10:00 часов на заседании диссертационного совета Д 003.074.01 по защите диссертаций на соискание ученой степени кандидата наук при ФГБУН Институт молекулярной и клеточной биологии СО РАН (630090, Новосибирск, пр. академика Лаврентьева, 8/2, тел. (383)373-02-49, e-mail: ovant@mcb.nsc.ru).

С диссертацией можно ознакомиться в библиотеке и на сайте Института молекулярной и клеточной биологии СО РАН https://www.mcb.nsc.ru/diss_council/autoref

Автореферат разослан «___» _____ 2022 г.

Ученый секретарь
диссертационного совета,
кандидат биологических наук



О.В. Антоненко

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность исследования. Биологические процессы, протекающие в эукариотических клетках, во многом зависят от уровня экспрессии генов. Особый интерес представляют интерфазные хромосомы, в которых происходит активная экспрессия генетического материала, а их интенсивность функционирования максимальна. Именно в этот период хромосомы диплоидных клеток максимально деконденсированы, однако не являются доступными для изучения с помощью световой микроскопии в силу своего небольшого размера. Регуляция генетической активности интерфазных хромосом в значительной степени определяется особенностями структурной организации хроматина, поэтому для изучения этой взаимосвязи необходимо сопоставить цитологическую и молекулярную карты хромосом.

Политенные хромосомы дрозофилы широко используются в качестве модели интерфазных хромосом, поскольку имеют гигантский размер и в интерфазном ядре с помощью светового микроскопа их можно визуализировать по всей длине. Такие хромосомы образуются в тканях слюнных желез в процессе цикла последовательных репликаций хроматид без их расхождения. Благодаря тому, что все дочерние хроматиды остаются связанными, на всем протяжении каждой из хромосом четко визуализируется рисунок из чередующихся темных и светлых полос; он постоянен и зависит от уровня конденсации хроматина вдоль хромосом и образован темными, плотно упакованными дисками и деконденсированными светлыми промежутками между ними – междисками. До недавнего времени точная локализация этих цитологических структур на физической карте ДНК была невозможна.

Благодаря проектам по секвенированию генома дрозофилы и изучению хроматина культур делящихся клеток modENCODE [Adams et al., 2000; modENCODE Consortium et al., 2010], получено большое количество полногеномных данных, систематизация которых привела к созданию компьютерных моделей, разделяющих хроматин дрозофилы на различные состояния. Существует несколько моделей, описывающих состояния хроматина в зависимости от локализации белков, модификаций гистонов и общей чувствительности к ДНКазе I, однако подход к локализации междисков одновременно на цитологическом и молекулярно-генетическом уровне так и не был найден. Для осуществления этой задачи в нашей лаборатории был разработан

алгоритм, позволяющий присваивать молекулярные координаты районам, соответствующим междискам на политенных хромосомах в соответствии с локализацией белковых маркеров [Zhimulev et al., 2014]. Модель позволила проводить масштабные сопоставления междисков и дисков молекулярной карте, что позволило оценить количество черных дисков на хромосоме 2R и впервые составить точную карту четвертой хромосомы *D. melanogaster* [Демидова и др., 2016; Колесникова, Жимулев, 2016].

Связь структурных элементов, составляющих интерфазные хромосомы, с генами и исследование принципов организации хромосом являются интересными задачами для изучения с помощью методов современной молекулярной генетики. Точная карта сайтов связывания белковых факторов с генетическими структурами, локализация их в определенных доменах хроматина и морфологических структурах хромосом имеет большое значение для расшифровки механизмов регуляции генов. Для изучения этих процессов необходимо локализовать гены и их элементы (промоторы, экзоны, интроны) в морфологических структурах интерфазных политенных хромосом, сопоставить расположение генов разных типов (развития и домашнего хозяйства) с распределением различных состояний хроматина на карте интерфазных хромосом. Данная работа посвящена исследованию организации генов и генома в составе интерфазных хромосом.

Цель исследования состояла в выявлении закономерностей генетической организации структур политенных хромосом – дисков и междисков, а также в точной локализации их границ на молекулярной карте. Для выполнения этой цели были поставлены следующие задачи:

1. С помощью модели четырех состояний хроматина локализовать на физической карте ДНК выборки из значительного количества междисков и черных дисков.
2. В локализованных структурах хромосом изучить распределение различных состояний хроматина, генетический и белковый состав.
3. С помощью FISH исследовать в политенных хромосомах цитологическую локализацию ряда генов, имеющих в своем составе протяженные интроны, провести анализ сигналов зондов, сопоставить молекулярную локализацию этих генов с цитологической, описать различия в устройстве и локализации генов в геноме.

Научная новизна и практическая значимость. В настоящей работе впервые исследована организация морфологических структур

политенных хромосом слюнных желез дрозофилы: междисков, черных и серых дисков и активных участков (пуфов). На физической карте ДНК с большой точностью локализованы 32 междиска, изучена их генетическая организация. Показано, что в междисках расположены 5'-концы генов домашнего хозяйства, описаны принципы расположения и ориентации таких генов. Определена генетическая организация серых и черных дисков. Серые диски содержат кодирующие части генов с повсеместной активностью, черные диски полигенны, содержат протяженные тканеспецифические гены и большие межгенные промежутки. На физической карте ДНК с большой точностью локализованы границы 62 индивидуальных черных дисков, состоящих из плотно компактизованного, поздно реплицирующегося хроматина. Установлено своеобразие их генетического, нуклеосомного и белкового содержания; изучена молекулярно-генетическая структура таких дисков. На краях черных дисков впервые обнаружены участки хроматина с промежуточными уровнями компактизации и времени завершения репликации. Изучена выборка из 15 генов развития, имеющих размер от 37 до 393 Кб; впервые показано, что на политенных хромосомах из материала интронов этих генов формируются как индивидуальные диски, так и их серии.

Положения, выносимые на защиту. Впервые предложен подход, совмещающий использование биоинформатических методов и FISH, позволяющий картировать структуры интерфазных политенных хромосом (междиски, серые и черные диски) на физической карте ДНК и выявлять их генетическую организацию и хроматиновый состав. Выводы картирования получены на основании исследования больших выборок из 32 междисков и 62 черных дисков; впервые с большой точностью на молекулярной карте локализованы их границы. С помощью модели четырех состояний хроматина установлена корреляция активного и открытого для транскрипции материала междисков состоянию aquamarine, рыхлых серых дисков – lazurite, а транскрипционно неактивному материалу черных дисков – ruby. Показано, что черные диски по краям во всех случаях ограничены фрагментами malachite, имеющими промежуточные уровни компактизации хроматина и времени репликации ДНК. Генам с повсеместной активностью впервые установлено соответствие двум морфологическим структурам хромосом – междиску и серому диску; промоторы таких генов локализуются в междиске. Показано, что «гибридные» диски, возникающие в результате слияния компактных

черных и рыхлых серых дисков, на краях содержат сайты хроматина lazurite, в которых располагается активно транскрибируемый ген или гены. Длинные гены, в составе которых преобладают протяженные интроны, имеют сложную организацию на уровне хромосомных структур и занимают серию из нескольких дисков и междисков. Впервые показано, что в таких генах формирование дисков происходит из материала интронов.

Степень достоверности и апробация результатов. По материалам диссертации опубликовано 6 статей, все в журналах списка ВАК. Результаты работы были представлены в виде устных и стендовых докладов на международных конференциях: Annual International Conference on Biology в 2015 (Греция), «Chromosome» в 2015 и 2018 (Россия), «6th Annual World Congress of Molecular & Cell Biology» в 2016 (Китай), «Drosophila Heterochromatin» в 2017 (Италия), «Chromosomes and Mitosis» в 2016; а также на отечественных конференциях: «Лекция для молодых ученых» в 2017, «Дрозофила в генетике и медицине» в 2017.

Вклад автора. Основные результаты получены автором самостоятельно. Работы по флуоресцентной *in situ* гибридизации проведены совместно с Похолковой Г. В. (ИМКБ СО РАН). Автор осуществлял подбор праймеров и амплификацию, очистку и мечение зондов для последующей гибридизации и микроскопии, выполненных Похолковой Г. В. Результаты *in situ* гибридизации обработаны автором самостоятельно с помощью программ Adobe. В исследованиях использована модель четырех состояний хроматина, созданная Гончаровым Ф. П. с помощью алгоритма Hidden Markov Model (Гончаров Ф. П., Болдырева Л. В., Демакова О. В., Жимулев И. Ф. Программа “НММ4Vand” зарегистрирована в Государственном реестре программ для ЭВМ 01.09.2014, свидетельство №2014618888). Обработка полногеномных данных осуществлена автором самостоятельно при помощи программ R, Galaxy, Microsoft Excel и совместно с Левицким В. Г. (ИЦиГ СО РАН) при использовании Microsoft Excel и C++. Статьи написаны в соавторстве с Жимулевым И. Ф., Беязевой Е. С., Зыковой Т. Ю. С Макуниным И. В. получены результаты по исследованию времени репликации хроматина malachite. Совместно с Зыковой Т. Ю. и Левицким В. Г. выполнена работа по взаимной локализации состояний хроматина и корреляционным отношениям длин генов и интронов.

Структура и объем диссертации. Диссертация состоит из введения, обзора литературы, описания материалов и методов, результатов и обсуждения, а также выводов и списка цитируемой литературы, включающего 165 ссылок. Работа изложена на 128 страницах текста, содержит 6 таблиц и 53 рисунка.

Благодарности автора. Работа выполнена в лаборатории молекулярной цитогенетики Института молекулярной и клеточной биологии СО РАН. Автор выражает глубокую благодарность руководителю академику РАН, д.б.н., проф. Игорю Фёдоровичу Жимулёву за помощь на всех этапах выполнения диссертации, д.б.н. Елене Сергеевне Беляевой, к.б.н. Татьяне Юрьевне Зыковой за помощь в постановке задач исследований и в работе; к.б.н. Лидии Валерьевне Болдыревой за уверенное руководство процессом подготовки дипломной работы и освоения молекулярных и цитогенетических методов; друзьям за их искренний интерес к работе и помощь в освоении методик. Автор особенно благодарен коллегам д.б.н. Галине Витальевне Похолковой за помощь по части цитологических исследований на препаратах политенных хромосом и к.б.н. Виктору Георгиевичу Левицкому за помощь в обработке данных с помощью биоинформатических методов.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Обзор литературы содержит сведения об открытии политенных хромосом, их организации, явлении политении, разделении хроматина на эу- и гетерохроматин, постоянстве дискового рисунка. Рассмотрены гипотезы о генетической организации дисков и междисков, описаны различные модели состояний хроматина и отдельно представлена модель 4НММ, домены которой исследовались в данной работе.

Материалы и методы исследования.

Из слюнных желез личинок третьего возраста линии *w*, *SUUR*-изготавливали препараты давленных политенных хромосом. Для подготовки зондов для флуоресцентной ДНК/ДНК гибридизации *in situ* на политенных хромосомах использовали методы выделения и очистки геномной ДНК, получения ДНК-зондов с помощью ПЦР, их очистки и мечения флуорохромами, горизонтальный электрофорез в агарозном геле. После проведения гибридизации *in situ* препараты анализировали при помощи световой и эпифлуоресцентной микроскопии, обработку полученных изображений проводили при помощи программ Adobe. Анализ полногеномных данных по распределению белков,

регуляторных элементов генома и состояний хроматина различных моделей проводили с помощью программ R, Galaxy, Microsoft Excel, C++.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

1. Организация доменов открытого хроматина.

Междиски содержат декомпактизованный, открытый для репликации хроматин, генетический состав доменов которого до недавнего времени оставался загадкой. Для исследования мы выбрали 19 междисков, определяемых под электронным микроскопом как постоянно присутствующие промежутки между дисками на всех срезах серий, полученных В. Ф. Семешиним, а так же 13 междисков, идентифицированных с помощью встроек Р-элементов [Ватолина и др., 2011]. Упростив структуру генов на молекулярной карте и представив самые длинные транскрипты в виде стрелок, мы проанализировали принцип их расположения в доменах открытого хроматина, расположенных в различных районах политенных хромосом. (для примера рассмотрены девять междисков в районе 9F13...10B3, Рис. 1).

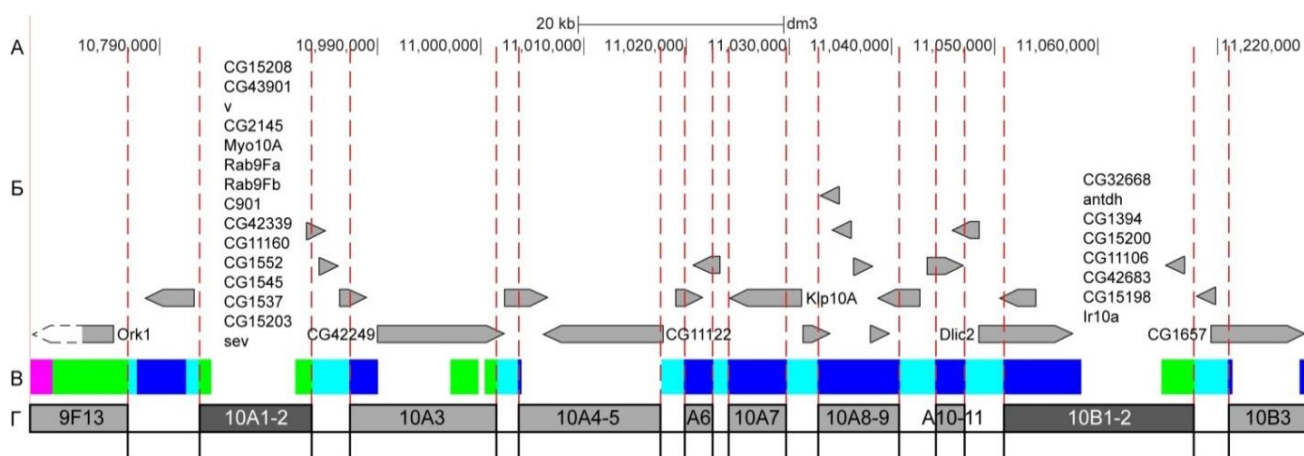


Рисунок 1. Сопоставление рисунка дисков на молекулярной карте с расположением генов в районе 9F13...10B3 [Zhimulev et al., 2014]. (А) – масштаб (Кб) и геномные координаты (п. н.); (Б) – локализация генов по версии FlyBase 5.50; (Б) – модель четырех состояний хроматина 4HMM [Zhimulev et al., 2014]; (Г) – схематическое представление междисков, серых и черных дисков в районе.

В исследовании мы сопоставили цитологические данные о расположении генов в районе 9F13...10B3 с физической картой, представленной в модели четырех состояний хроматина [Zhimulev et al., 2014]. По сравнению с черными дисками интеркалярного гетерохроматина (ИГХ), в которых локализуются, в основном длинные,

тканеспецифические поздно реплицирующиеся гены, в серых дисках лежат короткие и активные гены, в основном домашнего хозяйства. В представлении модели 4НММ на молекулярной карте между дисками 10А1-2 и 10В1-2 можно увидеть семь фрагментов aquamarine (Рис. 1), равных числу междисков. Этим фрагментам соответствуют 5'-регуляторные области активных генов, а совсем короткие гены находятся в них целиком. Получается, что фрагментам lazurite будут соответствовать соседствующие с ними серые диски, образованные структурными частями этих генов. На молекулярной карте чередующиеся фрагменты aquamarine и lazurite идеально отражают картину чередования дисков и междисков на цитологической карте.

1.1 Общий принцип организации междисков

Мы впервые прокартировали границы 32 междисков с точностью до 200 п. н. (разрешение модели 4НММ). Каждому изученному междиску (за исключением 21D1-2/D3 и 21D3/D4) соответствует единственный фрагмент хроматина состояния aquamarine. В 31 междиске из изученных 32 (в междиске 61С7/С8 лежит транскрибируемая РНК, но отсутствуют белок-кодирующие гены) локализуются промоторы активных генов или их 5'-концы. По типу расположения эти гены могут лежать «голова к голове» ($\leftarrow \rightarrow$), «голова к хвосту» ($\leftarrow \leftarrow$ или $\rightarrow \rightarrow$) или находиться в междиске целиком по причине своего небольшого размера. Если активный ген достаточно протяженен и его 3'-конец лежит за пределами фрагмента aquamarine, то структурная часть такого гена образует фрагменты lazurite, на цитологической карте они будут представлены серыми дисками.

2. Организация дисков интеркалярного гетерохроматина

По данным модели 4НММ, транскрипционно неактивного хроматина в составе генома дрозофилы определяется около 65%, из которых ruby (черные диски) составляет 42% или 43906 Кб (Рис. 2А). Суммарная длина выбранных нами 62 районов составляет 14587 ± 24.8 Кб (примерно 12.3% генома). Практически все исследованные черные диски из выборки содержат только транскрипционно неактивные состояния хроматина ruby (87.1%) и malachite (11.0%), что составляет практически 98.1% суммарной длины 62 районов (Рис. 2Б).

Как было показано, aquamarine модели 4НММ соответствует междискам, lazurite – серым дискам, ruby – транскрипционно неактивному хроматину. Malachite оказался интересен для изучения, поскольку этому состоянию хроматина не соответствует ни одна из

изученных структур хромосом. Краевой malachite соединяет диски с прилегающими к ним междисками, являясь переходной зоной между репрессированным ruby и активными, постоянно экспрессирующимися районами междисков. При подсчете числа точек перехода между состояниями хроматина оказалось, что malachite всегда ограничивает ruby при соседстве с aquamarine и lazurite, независимо от того, относится ли район к ИГХ; на геномном уровне не было обнаружено ни одного контакта ruby с aquamarine или lazurite – между ними всегда располагается фрагмент хроматина malachite (Рис 3). Функциональное значение этой закономерности в локализации malachite скорее всего заключается в обеспечении перехода между доменами хроматина, несущими различные функции.

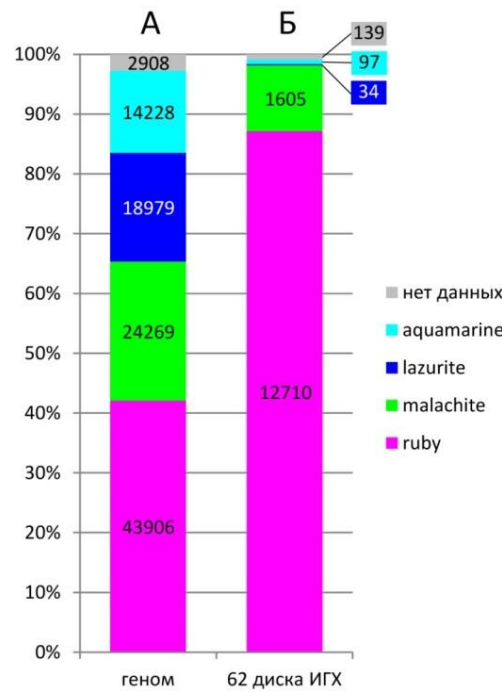
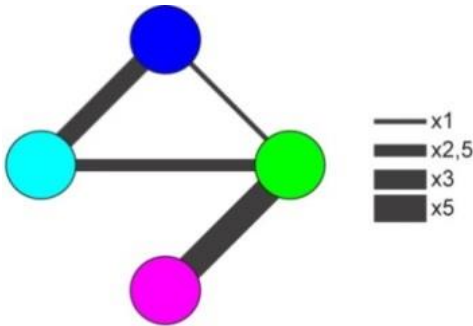


Рисунок 2. Хроматиновый состав генома (А) в сравнении с 62 исследованными дисками ИГХ (Б). Размер последовательностей указан в Кб внутри соответствующих состояниям хроматина блоков.



тип хроматинового контакта	количество
aquamarine-lazurite	5763
lazurite-malachite	2013
malachite-aquamarine	5227
ruby-malachite	9631
ruby-aquamarine	0
ruby-lazurite	0

Рисунок 3. Число границ при переходах из одного состояния хроматина модели 4НММ в другой [Khoroshko et al., 2016]. Толщина линии соответствует частоте контактов, их точное число указано в таблице.

Мы рассчитали время репликации для aquamarine, lazurite, malachite и ruby. Для каждого состояния хроматина в геноме было определено число проб, соответствующих стадиям репликации от

ранней до поздней (для анализа были использованы данные по репликации в культуре клеток Кс, конвертированные в BED-формат [Schwaiger et al., 2009]). Исследование показало, что активные aquamarine и lazurite в культуре клеток Кс реплицируются рано, репрессированный ruby – поздно, как и ожидалось. Malachite в геноме

имеет промежуточное время репликации – середину S фазы.

2.1 Структура границ дисков интеркалярного гетерохроматина

Мы исследовали края 62 дисков ИГХ и по принципу организации их границ разделили на два класса, представители которых – 13B3-4 и 10B1-2 – рассмотрены на Рис. 4. В случае диска 13B3-4 (Рис. 4, А-Д) его граница проходит между фрагментом aquamarine междиска 13B1-2/13B3-4 и фрагментом malachite, ограничивающим блок неактивного материала диска. В этом междиске локализуются белки, свойственные открытому хроматину (CHRIZ, RNA pol II, BEAF). Во втором случае граница диска 10B1-2 проходит между фрагментом aquamarine междиска 10A10-11/10B1-2 и фрагментом lazurite (Рис. 4, Е-К). Так как белки, свойственные открытому хроматину, локализуются только в междиске, мы можем отнести фрагмент lazurite в состав диска. Все остальные диски из исследуемой выборки

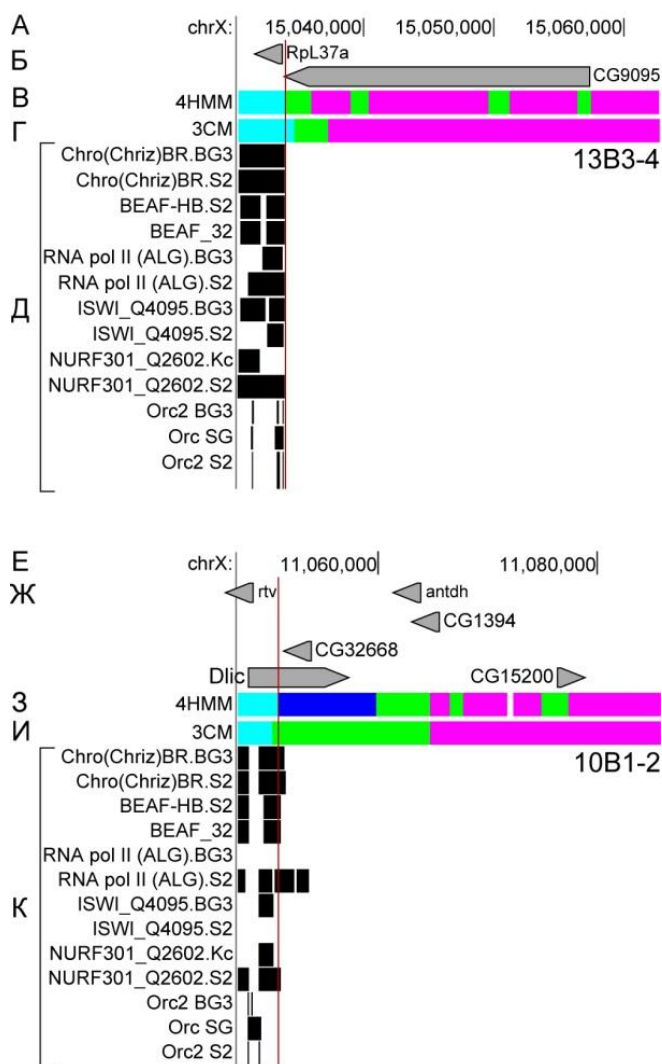


Рисунок 4. Локализация структурных доменов в дисках 13B3-4 (А-Д) и 10B1-2 (Е-К) [Хорошко и др., 2018]. (А, Е) – координаты на молекулярной карте; (Б, Ж) – локализация генов, версия FlyBase5.50; (В, З) – граница дисков согласно модели 4HMM [Zhimulev et al., 2014]; (Г, И) – модель 3CM трех уровней конденсации хроматина [Milon et al., 2014]; (Д, К) – белки открытого хроматина. Граница диска и междиска обозначена красной линией.

были рассмотрены сходным образом. На Рис. 5 представлена схема организации краев исследованных дисков ИГХ из нашей выборки (Рис. 5, тип I и II) в сравнении с серыми дисками (Рис. 5, серый диск).

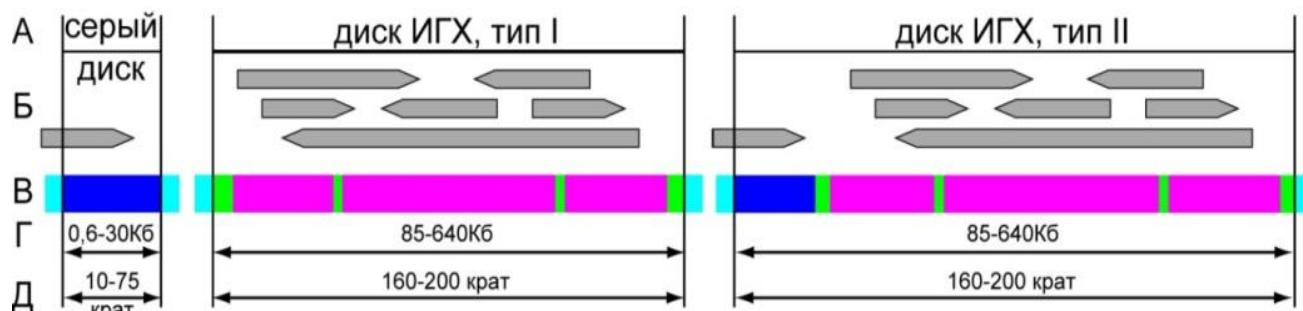


Рисунок 5. Схема организации серых дисков и дисков ИГХ на молекулярной карте [Хорошко и др., 2018]. (А) – диск и тип границы; (Б) – локализация генов; (В) – диски согласно модели 4НММ [Zhimulev et al., 2014]; (Г) – пределы варьирования размеров дисков.

Оба типа дисков ИГХ, изображенных на Рис. 5 (тип I и II), на цитологической карте выглядят одинаково крупными и черными. Различающую их особенность можно увидеть лишь на молекулярной карте с помощью биоинформатических методов. В диске ИГХ I типа на краях, кроме фрагментов malachite, так же располагаются фрагменты aquamarine, соответствующие междисковым участкам хромосом (Рис. 5, тип I). На краях дисков ИГХ II типа могут располагаться фрагменты lazurite, плотно прилегающие к блокам ruby и malachite, составляющим материал диска. В этом случае под микроскопом оба диска (типа I и II) будут выглядеть одинаково, но существенно различаться в генетическом смысле. Диск I типа содержит только поздно реплицирующиеся тканеспецифические гены, в то время как диск II типа содержит такой же генетический материал, как и диск I типа и, кроме этого, прилегающую к нему структурную часть гена домашнего хозяйства в соединенном с ним сером диске.

2.2 Общий принцип организации черных дисков интеркалярного гетерохроматина

Мы впервые прокартировали границы с точностью до 200 п. н. (разрешение модели 4НММ) для 62 черных дисков ИГХ, состоящих в основном из транскрипционно неактивного хроматина (доля ruby 87.1%, доля malachite 11.0%). Оказалось, что домены ruby по краям всегда ограничены фрагментами malachite, являющимися переходной зоной между неактивным материалом диска и открытым для транскрипции материалом междиска, что также подтверждается данными по экспрессии для различных состояний хроматина модели

4НММ – malachite имеет промежуточное время репликации. Тела генов домашнего хозяйства образуют фрагменты хроматина состояния lazurite, поэтому, если такой ген лежит в соседнем с диском ИГХ междиске и направлен в него, то на краю диска формируется фрагмент lazurite, входящий в состав диска. Таким образом, черные диски ИГХ состоят из протяженных блоков ruby, перемежающихся фрагментами malachite и иногда содержащие небольшие домены lazurite по краям.

3. Интроны «длинных» генов

При анализе распределения длин 13832 генов и их интронов, приведенных в списке FlyBase для генома дрозофилы (Release 5.57) очевидно, что с ростом длины гена растет и суммарная длина интронов,

а самые длинные гены почти полностью состоят из материала интронов (Рис. 6А).

Для соотнесения данных о локализации генов в разных состояниях хроматина модели 4НММ мы разбили все гены на четыре класса по длине наибольшего транскрипта: длины до 50 Кб, 50-100 Кб, 100-150 Кб и свыше 150 Кб. Численность классов составила 13547, 214, 55 и 16 генов. Оказалось, что с ростом длины гена увеличивается покрытие этих генов ruby хроматином от 30% (для длин генов до 50 Кб) до 55%, 57% и 66% (для длин генов в диапазонах 50-100 Кб, 100-150 Кб и свыше 150 Кб) (Рис. 6Б). В то же время покрытие malachite хроматином меняется незначительно (22%, 26%, 22% и 22%). Активные состояния хроматина, соответствующие междискам и рыхлым дискам (промоторам и кодирующим фрагментам активных генов) суммарно составляют 45% (17% и 28% для aquamarine и lazurite), тогда как для генов с большей длиной их доля не превышает 3% (Рис. 6Б). Следовательно, увеличение длины генов происходит за счет увеличения длины интронов, при этом

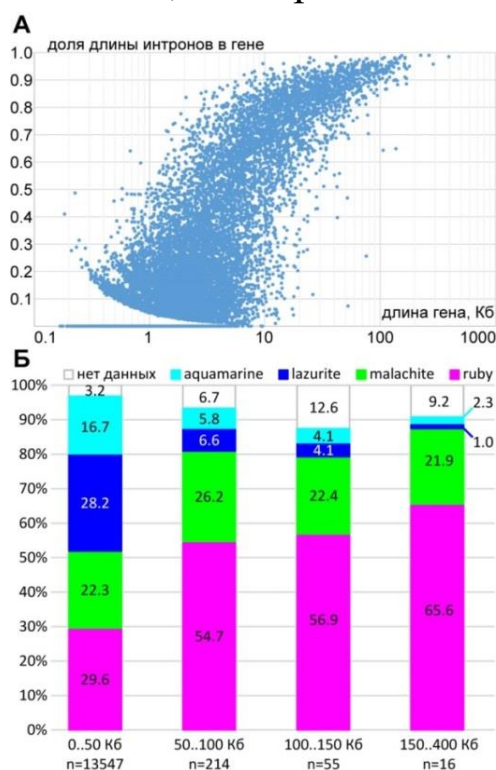


Рисунок 6. Распределение интронов и различных состояний хроматина в соответствии с моделью 4НММ [Khoroshko et al., 2020]. (А) – доля длины интронов в гене по отношению к его длине в Кб; (Б) – распределение четырех состояний хроматина в генах с длиной <2, 2-50, 50-100 и >100 Кб.

растет доля покрытия репрессированного *rubu* хроматина и сокращаются доли активного *aquamarine* и транскрибируемого *lazarite*.

Мы изучили выборку генов, имеющих размер от 100 Кб и выше, их в геноме дрозофилы оказалось 70, при этом доля интронной ДНК у 31 гена оказалась выше 95%, у 52 генов выше 93%, у 63 генов выше 90%, у 2 генов 64% и 60% соответственно. Из 70 генов для исследования мы выбрали три, имеющих большую длину и высокий процент содержания интронов, а также дополнительно – группу наиболее длинных генов, имеющих длину менее 100 Кб (12 генов), в том числе гены, участвующие в регулировании экдизонового каскада. В этих 15 генах максимальную длину составляют *malachite* и *rubu*, занимаемые интронами, а их суммарная длина составляет 41.9% и 28.2% соответственно. С помощью FISH на препаратах политенных хромосом мы изучили локализацию выбранных генов (для примера рассмотрены три из них).

CG43867 Длина гена составляет 119.7 Кб, 94.2% которой занимают интроны. Зонды для *in situ* гибридизации были выбраны из 3'- и 5'-концов гена *CG43867-RA* (красная и зеленая стрелки на Рис. 7Г, Д). Анализ показал, что зонды, разделенные 116.0 Кб, определяются как два отдельных сигнала в междисках 1C4-5/D1-2 и 1D1-2/D3, по краям рыхлого серого диска 1D1-2 (Рис. 7Г-3). В соответствии с картой К. Бриджеса [Bridges, 1938], зонды ограничивают два междиска и один

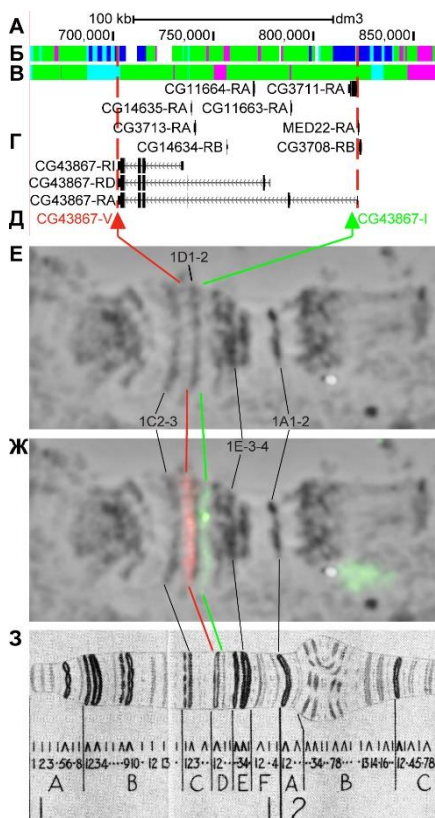


Рисунок 7. Локализация гена CG43867 в районе 1CD (ген ограничен красной пунктирной линией) [Khoroshko et al., 2020]. (А) – масштаб (Кб) и геномные координаты (п. н.); (Б) – модель четырех состояний хроматина 4НММ [Zhimulev et al., 2014]; (В) – модель трех уровней конденсации хроматина 3СМ [Milon et al., 2014]; (Г) – локализация генов; (Д) – локализация зондов; (Е) – фазово-контрастная микрофотография района; (Ж) – *in situ* сигналы зондов; (З) – хромосомная карта К. Бриджеса [Bridges, 1938].

серый диск, состоящий из интронного материала длиной около 112 Кб.

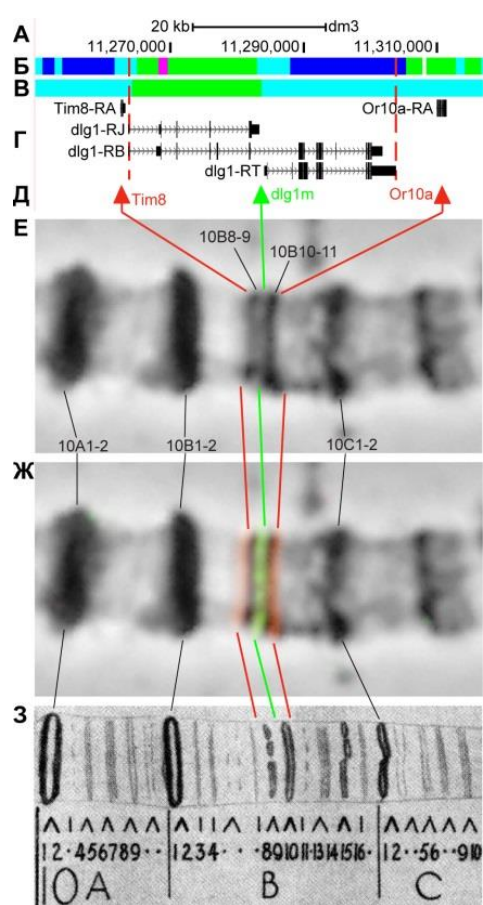


Рисунок 8. Локализация гена *dlg1* в районе 10В (ген ограничен красной пунктирной линией) [Khoroshko et al., 2020]. Подпись к рисунку аналогична таковой у Рис. 7.

и 3С7/С8. Два зонда (*kirre-IV*, *kirre-V*), разделенные 37.3 Кб, определяются как два отдельных сигнала в центральной и проксимальной частях диска 3С5-6 (Рис. 9Г-И). Хромосомные районы между зондов представлены: тонким черным диском 3С1 (*kirre-I* / *kirre-II*), большими черными дисками 3С2-3 (*kirre-II* / *kirre-III*) и 3С5-6 (*kirre-III* / *kirre-IV*) (Рис. 9Г-И). Таким образом, в пределах интронов гена *kirre* расположены три черных диска и три междиска района 3С1-6.

dlg1 Длина гена составляет 40.1 Кб, 81.9% которой занимают интроны. Зонды для *in situ* гибридизации были выбраны из последовательностей генов *Tim8*, *Or10a* (*Odorant receptor 10a*) и 5'-конца гена *dlg1-RT* (красные и зеленая стрелки на Рис. 8Г, Д). Анализ показал, что зонды, разделенные 20.5 Кб и 26.1 Кб, определяются как три отдельных сигнала в междисках 10В7/В8-9, 10В8-9/10-11 и 10В10-11/В12 (Рис. 8Г-З). В соответствии с картой К. Бриджеса [Bridges, 1938], зонды *Tim8* и *Or10a* ограничивают три междиска и два серых диска.

kirre Длина гена составляет 393.7 Кб, 98.4% которой занимают интроны. Зонды для *in situ* гибридизации были выбраны из фрагментов хроматина aquamarine, соответствующих 5'-концу *kirre-RF* (*kirre-I*), 5'-области *CG32795-RD* (*kirre-II*), 3'-концу *CG3603-RB* (*kirre-III*), 5'-области *kirre-RD* (*kirre-IV*) и 3'-концу *kirre-RF* (*kirre-V*) (красные и зеленые стрелки на Рис. 9Г, Д). Анализ показал, что зонды (*kirre-I*, *kirre-III*, и *kirre-V*) определяются как три отдельных сигнала в междисках 3В3-4/С1-2, 3С3-4/С5-6

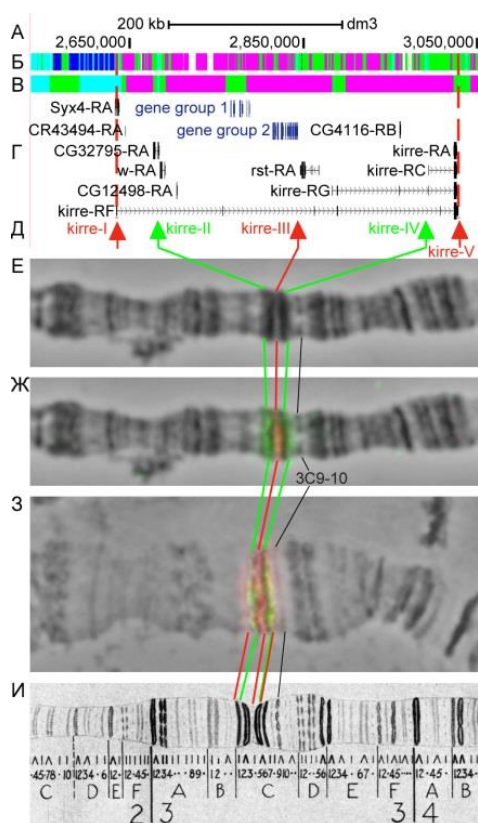


Рисунок 9. Локализация гена *kirre* в районе 3C (ген ограничен красной пунктирной линией) [Khoroshko et al., 2020]. Подписи (А-Д) аналогичны таковым у Рис. 7; (Е) – фазово-контрастная микрофотография района; (Ж) – *in situ* сигналы зондов *kirre-II*, *kirre-III* и *kirre-IV* на потянутой хромосоме; (З) – *in situ* сигналы зондов *kirre-I*, *kirre-II*, *kirre-III*, *kirre-IV* и *kirre-V*; (И) – хромосомная карта К. Бриджеса [Bridges, 1938].

В наиболее частом случае диски, образованные из материала интронов, расположены между промотором на 5'-конце и межгенной областью (междиском) на 3'-конце гена. Междиск такого типа генетически и молекулярно неоднороден, поскольку он содержит интронный материал (*ruby* и *malachite*) и экзоны гена (*lazurite*). Внутри некодирующего интрона преобладающего длинного «основного» гена расположены один или несколько разнонаправленных генов; в некоторых случаях таких генов нет. К этому типу относятся диски, образованные из интронов в генах *CG3777* (1A8–1B1-2), *CG43867* (1D1-2), *CG42666* (2B9-10), *trol* (3A1-3), *sgg* (3A9-10), *Nrg* (7F2). -4), *EcR* (42A14), *Eip74EF* (диски в районе 74EF) и *Eip75B* (диски в районе 75B) (Рис. 10 И).

br, *Hr46* и *Eip78C* частично аналогичны типу, описанному выше; их 5'-концы, как гены развития, не лежат в *aquamarine*. Однако в теле такого гена много материала, и он образует плотный серый сгусток, не являющийся диском, но сигналы зондов из начала и конца гена легко различимы. Материал гена выглядит серым, рыхлым и полуконденсированным, и это не обычный диск, поскольку он генетически гетерогенен и содержит как кодирующий, так и интронный материал.

В некоторых случаях диски формируются из интронного материала, как показано в генах *dlg1* и *dnc*. Промотор гена *dlg1* и

альтернативный промотор (локализованный в середине фрагмента ДНК, занятого геном) расположены в aquamarine, на цитологической карте, совпадающей с междисками. Между дистальным и средним промоторами расположен интрон размером ~20 Кб, состоящий из malachite с одним включением ruby. Он образует диск 10B8-9, который содержит только интронную ДНК гена *dlg1*. Второй диск 10B10-11 образован из материала, расположенного между средним промотором и 3'-концом гена; он состоит из интронов и частей гена, кодирующих белок. Таким образом, *dlg1* занимает пять хромосомных структур политенной хромосомы X, то есть два диска и три междиска (Рис. 10 II).

Особый интерес представляет ген *kirre*. В политенных хромосомах этот ген имеет еще более сложную генетическую организацию. Это один из самых длинных генов в геноме дрозофилы, так как он занимает 393.7 Кб от промотора (зонд *kirre*-I) до 3'-конца (зонд *kirre*-V), а 98.4% его длины приходится на интроны. Почти весь материал гена представлен компактными, транскрипционно неактивными и поздно реплицирующимися дисками; на уровне хроматина этот материал также очень компактен. В пределах интронов *kirre* расположены 23 других гена, некоторые из них являются генами домашнего хозяйства умеренного уровня активности, остальные принадлежат к генам

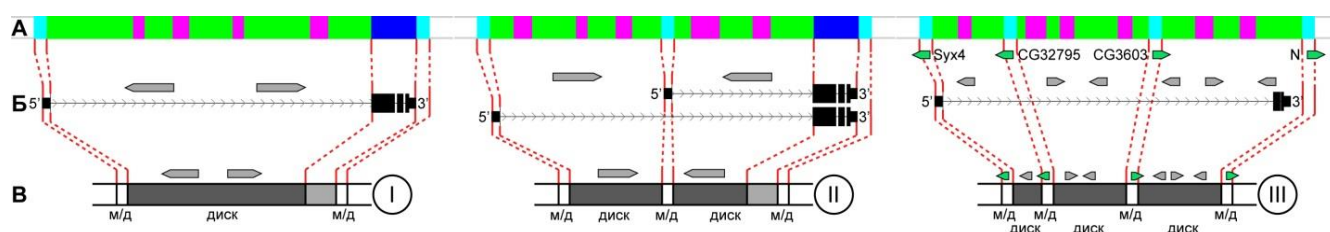


Рисунок 10. Схема организации дисков политенных хромосом. «Основной» ген содержит экзоны и интронный материал; серые стрелки обозначают более короткие гены, расположенные в интронах «основного» гена. II. Схема образования двух дисков и трех междисков из материала гена, состоящего из двух интронов и двух альтернативных промоторов «основного» гена. III. Схема образования четырех дисков и четырех междисков из генного материала с альтернативными промоторами и трех интронов, состоящих из поздно реплицирующегося хроматина; серые стрелки обозначают более короткие гены, расположенные в интронах «основного» гена [Khoroshko et al., 2020]. (A) – модель 4HMM [Zhimulev et al., 2014]; (Б) – изученные гены; (В) – структура дисков и междисков (м/д) политенных хромосом, включающая длинный «основной» ген, а так же короткие гены развития (обозначены серым) и домашнего хозяйства (обозначены зеленым цветом).

развития (Рис. 10 III).

В данной работе было проведено исследование организации генов и генома в составе интерфазных хромосом. С помощью модели четырех состояний хроматина 4НММ [Zhimulev et al., 2014; Zyкова et al., 2018] изучены морфологические структуры политенных хромосом: междиски, черные и серые диски и активные участки (пуфы). С большой точностью на физической карте ДНК мы локализовали 32 междиска, восемь серых дисков и 62 черных диска, установили их хроматиновый состав. Междискам соответствуют фрагменты хроматина состояния aquamarine, в которых локализуются промоторы активных генов или их 5'-концы. Если промоторов генов в междиске находится несколько, то по типу расположения они могут лежать «голова к голове» ($\leftarrow \rightarrow$), «голова к хвосту» ($\leftarrow \leftarrow$ или $\rightarrow \rightarrow$) или находиться в междиске целиком по причине своего небольшого размера. В случае, когда активный ген достаточно протяженен и его 3'-конец лежит за пределами фрагмента aquamarine, то структурная часть такого гена образует фрагменты lazurite и на цитологической карте они будут представлены серыми дисками.

Черные диски полигенны и состоят из транскрипционно неактивного хроматина ruby с небольшими включениями malachite. Диск, сформированный протяженным доменом ruby, на краях всегда отграничен от активного хроматина (aquamarine, lazurite) двумя короткими фрагментами malachite, являющимися переходной зоной между неактивным материалом диска и открытым для транскрипции материалом междиска. Они имеют промежуточные уровни компактизации хроматина и времени репликации ДНК; malachite так же входит в состав диска, разбивая домен ruby на фрагменты.

В некоторых случаях на краях черных дисков встречаются участки хроматина состояния lazurite, связанные с наличием структурных частей активных генов, ориентированных в диск; такие гены характеризуются как повсеместно активные, а их промоторы локализованы в прилежащих междисках. Края дисков декомпактизуются из-за локализации в них частей транскрипционно активных генов, они разрыхляют плотно конденсированный материал, формируя из него более декомпактный серый диск, вплотную примыкающий к черному диску. Таким образом, не все диски, при микроскопии выглядящие одинаково черными, одинаковы по своей природе, они различны по генетическому и хроматиновому составу, а

различающую их особенность можно увидеть лишь на молекулярной карте с помощью биоинформатических методов.

Гены в составе черных дисков по своей протяженности зачастую сопоставимы с ними по размерам, характеризуются протяженными интронами, мало изучены и имеют неясную локализацию в структурах хромосом; такие гены мы изучили более подробно. При анализе распределения длин генов и их интронов, приведенных в списке FlyBase для генома дрозофилы, мы показали, что с ростом длины гена растет и суммарная длина интронов, а в самых длинных генах размер кодирующих экзонов пренебрежимо мал по сравнению с длиной интронов в их составе. С помощью модели четырех состояний хроматина 4НММ [Zhimulev et al., 2014] изучена выборка из 15 генов, в которую входят три гена, имеющих длину от 100 Кб и высокий процент содержания интронов, а также дополнительно 12 наиболее длинных генов, имеющих длину менее 100 Кб, в том числе гены, участвующие в регулировании экдизонового каскада. Локализация выборки генов на политенных хромосомах методом FISH показала, что из материала интронов этих генов формируются как индивидуальные диски, так и их серии. Сопоставление расположения сигналов зондов показало, что по принципу локализации генов в структурах хромосом их можно условно разделить на группы, в которых из материала интронов формируется один или серия черных дисков.

Важным результатом этого исследования является открытие нового принципа расположения генов в политенных хромосомах. В прошлом диски политенных хромосом считались участками локализации единственного гена или промоторной части и тела гена [Bauer, Beermann, 1952; Beermann, 1952; Mechelke, 1953; Breuer, Pavan, 1955; Beermann, 1956; Lifschytz, 1971]. Настоящее исследование показывает, что генный материал в зависимости от его хромосомного состава может быть организован двумя способами: кодирующие части генов расположены в серых дисках (lazarite), а интронный материал (ruby и malachite) образуют черные диски, которые могут содержать как экзоны, так и только интроны или интронный материал. Промоторные области активных генов с повсеместной активностью образуют междиски (aquamarine), в некоторых случаях разбивая материал генов, в составе которых преобладают интроны, на серии дисков.

В целом полученные результаты расширяют существующие представления о генетическом составе и организации хроматина морфологических структур интерфазных политенных хромосом.

Исследования, проведенные в данной работе, способствуют более полному пониманию принципов организации и локализации генов в геноме дрозофилы, что имеет большое значение для расшифровки механизмов регуляции генов. Предложенный подход, совмещающий использование биоинформатических методов и FISH, может быть применен для дальнейшего картирования и построения более точных и полных карт структур политенных хромосом: междисков, серых и черных дисков на физической карте ДНК, для выявления их генетической организации и хроматинового состава.

ВЫВОДЫ

В настоящей работе предложен подход, совмещающий использование биоинформатических методов и FISH, позволяющий картировать структуры интерфазных политенных хромосом (междиски, серые и черные диски) на физической карте ДНК и выявлять их генетическую организацию и хроматиновый состав. Выводы картирования получены на основании исследования больших выборок из 32 междисков и 62 черных дисков, в том числе:

1. С помощью модели четырех состояний хроматина на физической карте ДНК с большой точностью прокартированы границы 32 междисков и 62 черных дисков; определены состояния хроматина, входящие в состав этих морфологических структур.
2. Междиски соответствуют доменам aquamarine, в них локализуются белки, свойственные открытому хроматину, и промоторы транскрипционно активных генов. В зависимости от длины генов, междиск может содержать такие гены целиком или только их промоторы; в таком случае 3'-концы генов локализуются за пределами междиска.
3. Серые диски сформированы кодирующими частями транскрипционно активных генов и соответствуют доменам lazurite. Таким образом, на цитологической карте структура междиск / серый диск указывает на локализацию генов с повсеместной активностью.
4. Черные диски полигенны, им соответствуют два состояния хроматина – ruby и malachite. Протяженный домен транскрипционно неактивного ruby, формирующий диск, на краях всегда ограничен от активного хроматина (aquamarine, lazurite) двумя короткими участками malachite, имеющими промежуточные уровни

компактизации хроматина и времени репликации ДНК; malachite так же входит в состав диска, разбивая домен ruby на фрагменты.

5. Края черных дисков могут декомпактизоваться из-за локализации в них частей транскрипционно активных генов. Промоторы таких генов расположены в соседнем междиске, а гены направлены в диск. Кодирующие части генов, соответствующие фрагментам lazurite, разрыхляют плотно конденсированный материал, формируя из него более декомпактный серый диск, вплотную примыкающий к черному диску.
6. В составе 15 изученных генов развития, размером от 31 до 393 Кб, преобладают интроны. Локализация этих генов с помощью FISH на цитологической карте показала, что интроны формируют диски, а гены локализуются в протяженных участках хромосом, состоящих из серий дисков и междисков.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Приведенные в диссертации результаты опубликованы в следующих работах:

1. Zhimulev I. F., Zyкова T. Yu., Goncharov F. P., **Khoroshko V. A.**, Demakova O. V., Semeshin V. F., Pokholkova G. V., Boldyreva L. V., Demidova D. S., Babenko V. N., Demakov S. A., Belyaeva E. S. Genetic organization of interphase chromosome bands and interbands in *Drosophila melanogaster*. PLoS One 9(7): e101631, 2014.
2. Жимулев И. Ф., Болдырева Л. В., Демакова О. В., Похолкова Г. В., **Хорошко В. А.**, Зыкова Т. Ю., Лавров С. А., Беляева Е. С. Формирование дисков политенных хромосом из интронов генов у дрозофилы. Доклады Академии Наук 466(5): 607–610, 2016
3. **Khoroshko V. A.**, Levitsky V. G., Zyкова T. Yu., Antonenko O. V., Belyaeva E. S., Zhimulev I. F. Chromatin heterogeneity and distribution of regulatory elements in the late-replicating intercalary heterochromatin domains of *Drosophila melanogaster* chromosomes. PLoS ONE 11(6): e0157147, 2016
4. **Хорошко В. А.**, Зыкова Т. Ю., Попова, О. О., Жимулев И. Ф. Структура границ дисков интеркалярного гетерохроматина в политенных хромосомах *Drosophila melanogaster*. Доклады Академии Наук 479(5), 589–592, 2018 DOI: 10.7868/S0869565218110257
5. **Хорошко В. А.**, Похолкова Г. В., Зыкова Т. Ю., Осадчий И. С., Жимулев И. Ф. Локализация гена *dunce* в большой серии

близлежащих структур политенной хромосомы *Drosophila melanogaster*. Доклады Академии Наук 484(4), 55-58, 2019 DOI: 10.31857/S0869-56524844507-510

6. **Khoroshko V. A.**, Pokholkova G. V., Levitsky V. G., Zyкова T. Yu., Antonenko O. A., Belyaeva E. S., Zhimulev I. F. Genes containing long introns occupy series of bands and interbands in *Drosophila melanogaster* polytene chromosomes. *Genes (Basel)* 11(4):417, 2020 doi: 10.3390/genes11040417

Приведенные в диссертации результаты были представлены:

1. В устном докладе «Molecular and genetic structure of polytene chromosome banding pattern in *Drosophila melanogaster*», Zyкова (Vatolina) T. Yu., Sidorenko (Demidova) D. V., Levitsky V. G., Khoroshko V. A., Belyaeva E. S., Zhimulev I. F. Annual International Conference on Biology, 22-25 June 2015. Athens, Greece
2. В устном докладе «Genetic organization of *Drosophila* interphase chromosomes», Zhimulev I. F., Zyкова (Vatolina) T. Yu., Belyaeva E. S., Goncharov F. P., Semeshin V. F., Demakova O. V., Khoroshko V. A., Boldyreva L. V., Demidova D. V. International conference «Chromosome», 2015. Novosibirsk, Russia
3. В устном докладе «Structural and genetic domains of interphase chromosomes in *Drosophila melanogaster*», Zyкова T. Yu., Levitsky V. G., Sidorenko (Demidova) D. V., Sidorenko I. A., Khoroshko V. A., Zhimulev I. F. 6th Annual World Congress of Molecular & Cell Biology, April 25-28, 2016. Dalian, China
4. В устном докладе «Chromosomal organization of *Drosophila* genome», Zyкова T. Yu., Levitsky V. G., Sidorenko (Demidova) D. V., Khoroshko V. A., Zhimulev I. F. International mini-conference «Chromosomes and Mitosis». November 25, 2016, Novosibirsk, Russia.
5. В устном докладе «Gene location in intercalary heterochromatin bands in polytene chromosomes of *Drosophila melanogaster*» Khoroshko V. A., Pokholkova G. V., Levitsky V. G., Zyкова T. Yu., Antonenko O. V., Zhimulev I. F. 13th International Conference on *Drosophila* Heterochromatin. 4-10 June 2017. Cagliari, Italy
6. В устном докладе «Диски интеркалярного гетерохроматина в политенных хромосомах *Drosophila melanogaster*» Хорошко В. А., Зыкова Т. Ю., Похолкова Г. В., Левицкий В. Г., Антоненко О. В., Жимулев И. Ф. Всероссийская конференция «Дрозофила в генетике и медицине». 4-6 октября 2017. Гатчина, Россия

7. В устном докладе «Политенные хромосомы – портрет функциональной организации генома *Drosophila melanogaster*», Жимулев И. Ф., Зыкова Т. Ю., Хорошко В. А. Лекция для молодых ученых. 5 октября 2017. Новосибирск, Россия
8. В стендовом сообщении «Gene location in intercalary heterochromatin bands of *Drosophila melanogaster* polytene chromosomes», Khoroshko V. A., Pokholkova G. V., Zyкова T. Yu., Zhimulev I. F. International conference «Chromosome», 2018. Novosibirsk, Russia