

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ НАУЧНОЕ
УЧРЕЖДЕНИЕ «ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
ИНСТИТУТ ЦИТОЛОГИИ И ГЕНЕТИКИ СИБИРСКОГО ОТДЕЛЕНИЯ
РОССИЙСКОЙ АКАДЕМИИ НАУК»

На правах рукописи

ЛУКЪЯНЧИКОВА ВАРВАРА АЛЕКСЕЕВНА

**ОСОБЕННОСТИ ТРЕХМЕРНОЙ ОРГАНИЗАЦИИ ХРОМАТИНА
У ПРЕДСТАВИТЕЛЕЙ КОМАРОВ РОДА
*ANOPHELES***

1.5.22 - клеточная биология (биологические науки)

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата биологических наук

Научный руководитель:
к.б.н. Фишман Вениамин Семенович

Новосибирск - 2022

ОГЛАВЛЕНИЕ

ОГЛАВЛЕНИЕ	2
СПИСОК СОКРАЩЕНИЙ И ТЕРМИНОВ	6
ВВЕДЕНИЕ.....	10
Актуальность работы.....	10
Научная новизна.....	11
Теоретическая и практическая значимость исследования.	12
Методы диссертационной работы	12
Основные положения, выносимые на защиту	13
Апробация результатов и публикации	13
Вклад автора	15
Структура и объем работы	15
Благодарности	15
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ.....	17
1.1 Первые исследования организации хроматина и структуры хромосом.....	17
1.2 Исследования биофизических свойств хроматинового волокна	19
1.2.1 Типы хроматина.....	20
1.2.2 Описание структуры хроматина в терминах физики биополимеров	22
1.3 Новые подходы в исследовании 3D-структуры генома.....	23
1.3.1 Группа 3C-методов, вариации подхода	23
1.3.2 Особенности Hi-C.....	26
1.3.3 Применение технологии Hi-C для улучшения геномных сборок	28
1.4 Уровни пространственной организации хроматина в ядре, ключевые структурные единицы	29
1.4.1 Активные и неактивные хроматиновые компартменты.....	30
1.4.2 Взаимодействия хроматина с ядерной ламиной. Ламины. ЛАДы	31
1.4.3 Доменные структуры, связанные с локализацией участков хроматина вблизи ядрышка. Ядрышко-ассоциированные хроматиновые домены (ЯАДы)	33
1.4.4 Домены ранней и поздней репликации	33
1.4.5 ТАДы или топологически ассоциированные домены - фундаментальная структурно- функциональная единица генома живых организмов, их основные характеристики	34
1.4.6 Хроматиновые петли	37
1.4.7 Участие длинных некодирующих РНК в организации хроматина	39

1.5 Молекулярные механизмы, обеспечивающие 3D-организацию генома.....	40
1.5.1 Механизм фазовой сепарации (phase separation)	41
1.5.2 Механизм loop extrusion (выпетливание хроматина)	43
1.5.3 Пространственная организация митотических хромосом	46
1.6 Особенности организация хроматина у насекомых.....	47
1.6.1 Основные белки и белковые комплексы, участвующие в поддержании архитектуры ядра у насекомых.....	48
1.6.2 Доменные структуры – ТАДы или компартментные домены?	49
1.6.3 Хроматиновые петли насекомых	52
1.6.4 Домены в цис- и транс-взаимодействиях. Явление трансекции в регуляции пространственных взаимодействий.	53
1.6.5 Пересечение дискового/междискового паттерна политенных хромосом со структурой ТАДов/интер-ТАДов у <i>Drosophila melanogaster</i>	54
1.7 Анофелесы или малярийные комары – уникальный объект исследования пространственной организации генома	55
1.7.1 Классификация и филогения	55
1.7.2 Распространение	57
1.7.3 Цикл развития	60
1.7.4 Накопление и аккумуляция хромосомных перестроек - двигатель эволюционного процесса.....	61
1.7.5 Способность переносить малярию.....	62
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ.....	66
2.1 Разведение и поддержание комариных колоний различных видов <i>Anopheles</i> и работа с клеточной линией MSQ43	66
2.1.1 Линии комаров, использованные в эксперименте, условия их содержания, кормления и разведения	66
2.1.2 Получение эмбрионального материала для подготовки Hi-C-библиотек	66
2.1.3 Культивирование клеток линии MSQ43.....	67
2.2 Молекулярные методы	69
2.2.1 Протокол приготовления Hi-C-библиотек из эмбрионов комаров	69
2.2.2 Очистка фрагментов ДНК на магнитных шариках (Size selection, Ampure beads).....	71
2.2.3 Проверка качества Hi-C-библиотек	72
2.2.4 Полимеразная цепная реакция с градиентом температур и без	74
2.2.5. Гель-электрофорез в агарозном геле.....	74

2.2.6 Иммуноокрашивание на антитела.....	75
2.2.7 Выделение белков из тканей и проведение вестерн блот – анализа	76
2.2.8 Иммунопреципитация хроматина (ChIP)	77
2.2.9 Выделение РНК из тканей (очистка от рибосомальной фракции) для анализа транскриптома	79
2.2.10 Подготовка ДНКовых и РНКовых библиотек для проведения секвенирования нового поколения на платформе Illumina	79
2.2.11 Выделение геномной ДНК из тканей.....	79
2.2.12 Полимеразная цепная реакция (ПЦР) с Q5-полимеразой	80
2.2.13 Подготовка специфических флуоресцентных проб с помощью избирательного мечения участков ДНК методом ник-трансляции	80
2.2.14 Мечение участков ДНК методом случайных праймеров.....	81
2.2.15 Высаждение и подготовка флуоресцентных проб к гибридизации	82
2.3 FISH-эксперименты и кариотипирование.....	83
2.3.1 Диссекция и фиксация отдельных органов насекомых для проведения 3D-FISH	83
2.3.2 Подготовка образцов к гибридизации	83
2.3.3 Протокол 3D-FISH (флуоресцентная иммуногибридизация).....	83
2.3.4 Приготовление давленных препаратов политенных хромосом, выделенных из яичников и слюнных желез.....	84
2.3.5 Протокол 2D-FISH (флуоресцентная гибридизация)	85
2.3.6 Кариотипирование клеточной линии MSQ43	85
2.4 Сборка геномов	86
ГЛАВА 3. РЕЗУЛЬТАТЫ.....	88
3.1 Оптимизация протокола Hi-C для эмбрионов комаров рода <i>Anopheles</i> . Оценка качества и характеристика полученных библиотек	88
3.2 Улучшение существующих сборок геномов для пяти видов комаров рода <i>Anopheles</i>	91
3.3 Детекция ранее известных и новых хромосомных перестроек с помощью данных Hi-C. Полиморфные и фиксированные инверсии у представителей родов <i>Anopheles</i> и <i>Aedes</i>	95
3.3.1 Характеристика инверсий у <i>Anopheles</i>	96
3.3.2 Полиморфные и фиксированные инверсии у различных представителей рода <i>Aedes</i> . Инверсионный полиморфизм как фактор формирования пищевых и экологических предпочтений	100
3.4 Основные консервативные принципы и особенности 3D-организации генома комаров рода <i>Anopheles</i> , выявленные методом Hi-C	103

3.4.1 Rabl-конфигурация хроматина в интерфазном ядре у комаров рода <i>Anopheles</i> . Визуализация с помощью FISH.....	103
3.4.2 Характеристика выделенных структур: компартменты.....	105
3.4.3 Характеристика выделенных структур: домены и хроматиновые петли	106
3.4.4 Свойства и особенности эухроматина и гетерохроматина на тепловой карте пространственных контактов	108
3.5 Интенсивные дальние взаимодействия хроматина, специфичные для комаров <i>Anopheles</i> . Их визуализация на картах пространственных контактов и подтверждение методом FISH.....	113
3.6 Сравнение выявленных принципов укладки хроматина у малярийных комаров с другими таксонами 120	
ГЛАВА 4. ОБСУЖДЕНИЕ.....	123
4.1 Улучшение существующих сборок геномов для пяти видов комаров рода <i>Anopheles</i>	123
4.2 Технология Hi-C для выявления полиморфных и фиксированных инверсий у представителей рода <i>Anopheles</i> и <i>Aedes</i>	123
4.3 Особенности пространственных контактов хроматина у малярийных комаров.....	125
4.4 Уникальные хроматиновые петли	127
4.5 Заключение	131
ВЫВОДЫ.....	133
СПИСОК ЛИТЕРАТУРЫ.....	135
ПРИЛОЖЕНИЕ 1	156
ПРИЛОЖЕНИЕ 2	158
ПРИЛОЖЕНИЕ 3	160
ПРИЛОЖЕНИЕ 4	161
ПРИЛОЖЕНИЕ 5	163
ПРИЛОЖЕНИЕ 6	165
ПРИЛОЖЕНИЕ 7	167

СПИСОК СОКРАЩЕНИЙ И ТЕРМИНОВ

“Греческие острова” (от англ. Greek islands) – геномные локусы межгенных энхансеров, специфично взаимодействующих с транскрипционными факторами LHX2 и EBF, осуществляющие регуляторные взаимодействия в обонятельных нейронах [1]

2D-FISH – флуоресцентная *in situ* гибридизация на плоских препаратах (в двух измерениях)

3D-FISH – флуоресцентная *in situ* гибридизация в пространстве (в трех измерениях)

3D-SIM-микроскопия (от англ. three-dimensional Structured Illumination Microscopy) – микроскопия структурированного освещения

BG3 – клеточная культура, полученная из тканей центральной нервной системы личинок третьего возраста *Drosophila melanogaster*

BUSCO (от англ. Benchmarking Universal Single-Copy Orthologs) – показатель представленности известных для данного таксона консервативных генов; характеристика новых геномных сборок, выражается в процентном эквиваленте.

Hi-C (от англ. capture Hi-C) – полногеномный метод захвата конформации хромосом с высоким разрешением с обогащением

ChIP-seq (от англ. Chromatin Immunoprecipitation) – объединение технологий иммунопреципитации хроматина и секвенирования нового поколения (NGS)

CND (от англ. Chromatin NanoDomains) - хроматиновые нанодомены

CpG-островки – участки генома, линейная последовательность которых состоит из множественных повторяющихся пар цитозин/гуанин (CpG). Как правило, метилированы по цитозину (C); обогащены вблизи районов промоторов

CTCF – высоко консервативный транскрипционный фактор, распознающий и связывающий мотив CCCTC

C-технологии – группа методов, в основе которых лежит принцип лигирования сближенных в пространстве молекул, впервые открыт Деккером [2]

DamID – молекулярный метод, позволяющий идентифицировать участки связывания белков интереса с ДНК/хроматином; в результате действия искусственно созданного химерного белка, состоящего из белка интереса и аденин метилтрансферазы, участки хроматина, с которыми связывается целевой белок, получают метильную метку

GO (от англ. Gene Ontology) – платформа, посвященная созданию унифицированной терминологии для аннотации генов и генных продуктов всех биологических видов [3]

HR-метод (от англ. High-resolution Chromatin Conformation Capture method) – метод захвата конформации хромосом с высоким разрешением, предложен в работе [4]

HCNE (от англ. **H**ighly **C**onservative **N**on-coding **E**lements) – высококонсервативные некодирующие элементы

Hi-C (от англ. **H**igh-throughput **C**hromosome **C**onformation **C**apture) – полногеномный метод захвата конформации хромосом с высоким разрешением, предложен в работах [5,6]

Kc167 – клеточная культура, полученная из эмбрионов на 13-15 стадиях развития

MAPQ (от англ. **M**apping **Q**uality) – одна из важных характеристик данных после NGS-секвенирования, позволяющая отфильтровать прочтения (риды) хорошего качества

NGS-секвенирование (от англ. **N**ext **G**eneration **S**equencing) – секвенирование нового поколения

PacBio-секвенирование (от англ. **P**acific **B**iosciences) или **SMRT** (от англ. **S**ingle **M**olecule **R**eal **T**ime) секвенирование – альтернативная NGS технология секвенирования, позволяющая генерировать длинные прочтения

PBMCs (от англ. **P**eripheral **B**lood **M**ononuclear **C**ells) – мононуклеары периферической крови

PBS (от англ. **P**hosphate **B**uffered **S**aline) – фосфатно-солевой буфер

PRC1, PRC2 (от англ. **P**olycomb **R**epressive **C**omplex) – комплексы белков группы поликомб

Rabl-конфигурация или конфигурация по Раблю – тип организации интерфазных хромосом, особенностью которого является объединение/кластеризация центромерных областей хромосом на одном полюсе ядра, а теломерных – на другом [7,8]

RPM – (от англ. **R**evolutions **P**er **M**inute) – единица измерения, альтернативная G ($G = 1.12 \times \text{Radius (mm}^2) \times (\text{RPM}/1000)^2$), характеризует количество вращений ротора в центрифуге за минуту

single-nucleus Hi-C или **snHi-C** – полногеномный метод захвата конформации хромосом с высоким разрешением в одном ядре

SNP (от англ. **S**ingle **N**ucleotide **P**olymorphisms) – однонуклеотидные полиморфизмы

SNV (от англ. **S**ingle **N**ucleotide **V**ariations) – однонуклеотидные замены

SSC (от англ. **S**aline-**S**odium **C**itrate) – цитратно-солевой буфер

Анкеры хроматиновых петель – участки генома, вовлеченные в пространственные взаимодействия, детектируемые на Hi-C –карте в виде дальних хроматиновых петель.

ДДТ (от англ. **D**DT) – трихлорметилди(п-хлорфенил)метан) представляет инсектицид, применяемый против комаров, вредителей хлопка, соевых бобов, арахиса; по официальной рекомендации ВОЗ применялся в целях профилактики малярии; на сегодняшний день запрещен к производству и применению Стокгольмской конвенцией, которую на 2021 год подписали 184 страны

ДНКаза - дезоксирибонуклеаза - нуклеаза, катализирующая деградацию ДНК

ДТТ (от англ. DTT) – дитиотреитол используют для «снятия защиты» с тиолированной ДНК, так как концевой атом тиолированной ДНК в растворе имеет склонность к формированию димеров, что значительно снижает эффективность последующих реакций соединения

Кб – 1 тысяча пар нуклеотидов или килобаза

Комплекс малярийных комаров *Anopheles gambiae* – комплекс, включающий в себя несколько видов-сублингов, в следующем составе: *An. coluzzii*, *An. gambiae*, *An. arabiensis*, *An. quadriannulatus*, *An. melas*, *An. merus*, *An. bwambae* [9]

ЛАДы – ламин-ассоциированные домены

Мб – 1 миллион пар нуклеотидов или мегабаза

мЭСК – мышинные эмбриональные стволовые клетки

Показатель N50 – самый короткий контиг при суммировании длин минимального числа контигов для получения суммы, большей или равной половине длины сборки генома (~центр масс распределения длин контигов)

ПЦР – полимеразная цепная реакция

Разрешение HiC-карт – показатель количества нуклеотидов, приходящихся на один маленький квадрат на карте пространственных контактов (бин). Стандартно используются размеры 1 Кб, 5 Кб, 10 Кб, 25 Кб на бин. Данная величина зависит от рестрикционного фермента и от глубины секвенирования. Чем больше глубина секвенирования, тем больше прочтений попадают на один бин, тем более мелкие структуры хроматина можно обнаружить на картах пространственных контактов.

Рид (от англ. read) – одна прочитанная в результате NGS-секвенирования молекула. Прочтение может быть как с одного (одиночный), так и с двух концов (парный). Далее по тексту термин “рид” будет использоваться в значении парный рид.

РНКаза – рибонуклеаза - нуклеаза, катализирующая деградацию РНК

рРНК – рибосомальная РНК

ТАДы (от англ. **T**opologically **A**ssociated **D**omains) – топологически ассоциированные домены

эЛАДы – эухроматиновые ламин-B1-ассоциированные домены

ЭМ-переход – эпителиально-мезенхимальный переход

ЯАДы – ядрышко-ассоциированные домены

Модификации гистонов:

H2AK119Ubiq – метка моноубиквитинилирования, устанавливаемая комплексом PRC1; наблюдается в репрессированных районах хроматина

H3K27ac – метка активного хроматина, характерна для областей активной транскрипции

H3K27me3 – метка триметилирования, устанавливаемая комплексом PRC2; наблюдается в репрессированных районах хроматина

H3K36me3 – метка, ассоциированная с активной транскрипцией; наблюдается в телах генов

H3K4me1 – метка, ассоциированная с активными/ключевыми энхансерами

H3K79me2 – гистоновая модификация, распознаваемая белком HP1

H3K9me2 – метка, ассоциированная с гетерохроматином ядерной периферии

H3K9me3 – метка, ассоциированная с гетерохроматиновыми областями

H4K20me3 – метка, ассоциированная с гетерохроматиновыми областями

ВВЕДЕНИЕ

Актуальность работы. На сегодняшний день комары представляют серьезную угрозу для человека в связи с высоким патогенным потенциалом, активным переносом трансмиссивных болезней и возрастающей миграционной способностью. Род малярийных комаров *Anopheles* включает более 400 видов, которые широко распространены по всем континентам планеты за исключением Антарктиды, а около 60 представителей рода способны переносить паразитов – возбудителей малярии (*Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae*, *Plasmodium knowlesi*), а также другие инфекционные и вирусные заболевания (гельминтозы, филяриатозы, вирус лихорадки о'ньонг-ньонг, вирус Западного Нила, вирусы восточного и западного лошадиного энцефалита и некоторые другие). Среди перечисленных патогенных последствий малярия является наиболее опасным заболеванием, которое по последним оценкам ВОЗ насчитывает ~241 миллион случаев в год в 85 странах мира и уносит жизни ~0.627 миллиона человек ежегодно, а процент детской смертности в возрасте до 5 лет в 2020 году составил 77% [10]. Потенциал к переносу вирусов и других патогенов животных и человека, в том числе различных подтипов малярийного плазмодия, в значительной степени варьирует между представителями рода [11,12]. Причины таких различий до сих пор являются объектом активных дискуссий и считаются не до конца выясненными. Согласно последним научным работам именно геномная и генная пластичность среди представителей рода *Anopheles* позволяет им так быстро и успешно адаптироваться к новым условиям, средам обитания, источникам питательных веществ [12]. Становится понятно, что только полномасштабные исследования геномов, полученные для многих представителей малярийных комаров, позволят выявить ключевые механизмы, ответственные за приспособляемость к условиям среды, а также предполагаемую коэволюцию с патогенами, которых они переносят [12].

Геном эукариотических клеток представляет сложную многоуровневую систему, которая осуществляет тонкие биологические процессы на протяжении всего клеточного цикла. Поскольку для правильной реализации таких процессов необходимо сохранить доступность хроматиновой нити для множественных транскрипционных факторов и других белков, именно то, каким образом геном организован в пространстве клеточного ядра, его архитектура, оказывает огромное влияние на активность генов и функции клеточных белков. Было показано, что хроматин в интерфазном ядре иерархично уложен в нескольких уровнях, каждый из которых образуется и функционирует по уникальному механизму. В последние годы демонстрируется универсальность основных принципов укладки

хроматина, а также вовлеченность пространственной организации генома во многие клеточные процессы, а именно: в активность транскрипции, репликации, репарации повреждений ДНК, регуляции генной экспрессии и эволюционный процесс.

Основной **целью** представленной **диссертационной работы** является характеристика пространственной организации хроматина у пяти видов комаров рода *Anopheles* (*Anopheles coluzzii*, *Anopheles merus*, *Anopheles stephensi*, *Anopheles atroparvus* и *Anopheles albimanus*).

Для достижения поставленной цели были сформулированы следующие **задачи**:

1. оптимизировать протокол технологии Hi-C (метод захвата конформации хромосом) для его применения на эмбрионах комаров рода *Anopheles*, подготовить и провести секвенирование Hi-C-библиотек для 15–18-часового эмбрионального материала каждого из пяти видов комаров *Anopheles* в двух репликах;
2. на основании полученных данных по частотам пространственных контактов провести сборку геномов пяти видов анофелесов до хромосомного уровня;
3. построить тепловые карты Hi-C и выполнить детальный анализ ключевых структур пространственной организации геномов: топологически ассоциированных доменов (ТАДов), хроматиновых петель и компартментов;
4. с использованием полученных данных о пространственных контактах описать хромосомные aberrации, характерные для представителей комаров рода *Anopheles*; расширить использование данного подхода для комаров рода *Aedes*, в том числе для поиска полиморфных и фиксированных инверсий в популяциях;
5. подтвердить основные выявленные закономерности 3D-организации хроматина при помощи независимых методов 2D-FISH и 3D-FISH;
6. сравнить выявленные закономерности организации хроматина у малярийных комаров с данными для других организмов.

Научная новизна. Hi-C-технология, основной метод диссертационной работы, позволяющий получать информацию о частоте пространственных контактов геномных локусов, является новейшим перспективным подходом, надежность и точность которого демонстрируются публикациями в ведущих научных изданиях. В представленной работе метод Hi-C был впервые применен для улучшения геномных сборок представителей рода *Anopheles*, в результате чего геномы трех видов малярийных комаров - *An. coluzzii*, *An. merus*, *An. stephensi*, были впервые собраны до хромосомного уровня. Комбинация

нескольких подходов, основанных на различных молекулярных механизмах (FISH, Hi-C, PacBio-секвенирование, сравнение с результатами цитогенетического и физического картирования), позволила получить максимально полную на текущий момент информацию о 3D-структуре генома малярийных комаров, а также избежать неточностей в выводах, связанных с техническим исполнением методов. Впервые было проведено сравнение и показана схожесть принципов укладки хроматина у представителей родов *Anopheles* и *Drosophila*: экспоненциальная зависимость частоты контактов хроматина от расстояния, соответствие хроматиновых доменов компартментам, наличие конфигурации хромосом по Раблю, а также наличие хроматиновых петель, образуемых белками группы поликомб. В работе было впервые показано, что компартментализация эу- и гетерохроматина в совокупности с организацией по Раблю хромосомных территорий определяют паттерн дистальных контактов ДНК в геномах комаров рода *Anopheles*. Впервые с помощью метода Hi-C был проведен успешный поиск инверсионных полиморфизмов в колониях *Aedes aegypti*, в результате чего были описаны 25 новых хромосомных инверсий у представителей данного вида. Среди пяти видов малярийных комаров было выявлено четыре хромосомных инверсии, две из которых были описаны впервые.

Теоретическая и практическая значимость исследования. На сегодняшний день данные о трехмерной организации генома ограничены лишь небольшим числом модельных видов. Исследование этого вопроса в таксоне комаров рода *Anopheles* является не только важным шагом на пути к пониманию глобальных закономерностей в эволюции 3D-укладки хроматина у живых организмов, но также позволит использовать данные принципы для решения практических задач, связанных с генетическими системами контроля, адаптивным поведением и распространением комаров рода *Anopheles* по земному шару. Кроме того, в работе была впервые показана возможность использования технологии Hi-C для поиска инверсионных полиморфизмов в популяциях комаров, в том числе для тех видов, где применение стандартных цитологических подходов затруднено.

Методы диссертационной работы. В ходе данного исследования были использованы различные генно-инженерные и молекулярные научные методы, а именно: полимеразная цепная реакция (ПЦР) с градиентом температур и без, гель-электрофорез в агарозном/полиакриламидном геле, выделение из тканей ДНК, РНК и белков. Освоен и оптимизирован протокол приготовления Hi-C-библиотек и библиотек после иммунопреципитации хроматина (ChIP-seq библиотеки), выполнена подготовка ДНКовых и РНКовых проб для проведения секвенирования нового поколения (NGS) на платформе Illumina; получены навыки проведения очистки фрагментов ДНК на магнитных шариках, а

также - навыки работы с агарозными и стрептавидиновыми шариками. В процессе выполнения экспериментов по иммунопреципитации хроматина получен навык работы с белками (вестерн блот, иммуноокрашивание). В рамках выполнения экспериментов по флуоресцентной гибридизации *in situ* (FISH) освоена техника работы с бактериальными культурами (BACs), дизайн и подготовка специфических флуоресцентных проб с помощью избирательного мечения участков ДНК методом ник-трансляции, методом случайных праймеров, а также ПЦР с добавлением в реакцию меченных нуклеотидов. Приобретены навыки приготовления давленных препаратов политепных хромосом, выделенных из яичников и слюнных желез, а также освоена методика диссекции и фиксации отдельных органов насекомых для проведения 3D-FISH-экспериментов и иммуноокрашивания. Приобретены навыки работы с комариными колониями различных видов, разведение и поддержание колоний, получение необходимого количества комариных яиц. Кроме того, была освоена техника получения и обработки изображений со светового, флуоресцентного и конфокального микроскопов; получены навыки работы со вспомогательными компьютерными программами (BLAST, Oligocalculator, SnapGene, PrimerPlus, Ensemble, Juice box/Juice box assembly tools, геномный браузер IGV, ZEN, Adobe Photoshop, Inkscape и некоторые другие).

Основные положения, выносимые на защиту:

1. Границы точек разрыва полиморфных инверсий у комаров рода *Aedes*, впервые установленные на основе карт трехмерных контактов хроматина, хорошо согласуются с полученными ранее цитогенетическими данными.
2. Распределение хроматиновых доменов и компартментов в клеточных ядрах комаров рода *Anopheles* соответствует профилю генной экспрессии и распределению эпигенетических модификаций гистонов, что является особенностью представителей отряда *Diptera*.
3. 3D-организация генома комаров рода *Anopheles* характеризуется наличием уникальных, консервативных в пределах рода, хроматиновых петель, образование которых не может быть объяснено активностью генов, расположенных в данных локусах, и связыванием оснований петель с белками группы Polysomb.

Апробация результатов и публикации. Научные результаты, изложенные в данной работе, были представлены на нескольких крупных международных конференциях в виде стендовых и устных докладов. А именно:

1. **Varvara Lukyanchikova**, Miroslav Nuriddinov, Polina Belokopytova, Igor V. Sharakhov, Veniamin Fishman. “Insights into the 3D-genome organization in malaria mosquitoes”, BGRS-2022, Новосибирск, Россия, 4-8 июля, 2022;
2. **Varvara Lukyanchikova**, Miroslav Nuriddinov, Veniamin Fishman, Polina Belokopytova, Alexander Tyapkin, Nariman Battulin, Igor Sharakhov “Anopheles mosquitoes revealed new principles of 3D genome organization in insects”, MCCMB-2021, Москва, Россия, 30 июля – 2 августа, 2021;
3. **Varvara Lukyanchikova**, Igor V. Sharakhov, Veniamin Fishman, Miroslav Nuriddinov, Nariman Battulin, Oleg L. Serov "Revealing the principles of 3D genome organization and improving the accuracy of genome assemblies of malaria mosquitoes using Hi-C approach", EMBO Workshop "The Genome in Three Dimensions", Килини, Греция, 20-24 мая, 2019;
4. **Varvara Lukyanchikova**, Veniamin Fishman, Miroslav Nuriddinov, Nariman Battulin, Oleg L Serov, Igor V Sharakhov "Improving the accuracy of genome assemblies of Anopheles species and revealing the principles of 3D genome organization in dipteran insects using the Hi-C approach", 2018 ESA, ESC, and ESBC Joint Annual Meeting, Ванкувер, Британская Колумбия, Канада, 11-14 ноября, 2018;
5. **Varvara Lukyanchikova**, Veniamin Fishman, Miroslav Nuriddinov, Nariman Battulin, Oleg L Serov, Igor V Sharakhov “The Hi-C approach improved genome assemblies of Anopheles species and revealed principles of 3D genome organization in dipteran insects”, CSH Meeting: Nuclear Organization & Function, Колд Спринг Харбор, Нью-Йорк, США, 1-5 мая, 2018.

По теме диссертации было опубликовано 3 работы. Основные результаты были изложены в рецензируемом журнале *Nature Communications*.

1. **Lukyanchikova V**, Nuriddinov M, Belokopytova P, Taskina A, Liang J, Reijnders MJMF, Ruzzante L, Feron R, Waterhouse RM, Wu Y, Mao C, Tu Z, Sharakhov IV, Fishman V. 2022. *Anopheles* mosquitoes reveal new principles of 3D genome organization in insects. *Nat Commun* 13:1960., <https://doi.org/10.1101/2020.05.26.114017>;
2. Zamyatin A, Avdeyev P, Liang J, Sharma A, Chen C, **Lukyanchikova V**, Alexeev N, Tu Z, Alekseyev MA, Sharakhov IV. Chromosome-level genome assemblies of the malaria vectors *Anopheles coluzzii* and *Anopheles arabiensis*. *GigaScience*. 2021 Mar; 10(3): giab017; <https://doi.org/10.1093/gigascience/giab017>;
3. Compton A, Liang J, Chen C, **Lukyanchikova V**, Qi Y, Potters M, Settlage R, Miller D, Deschamps S, Mao C, Llaca V, Sharakhov IV, Tu Z. The beginning of the end: a

chromosomal assembly of the New World malaria mosquito ends with a novel telomere. *G3: Genes, Genomes, Genetics*. 2020 Oct 1; 10(10): 3811-9; <https://doi.org/10.1534/g3.120.401654>.

Вклад автора. Автор диссертационной работы самостоятельно провела большинство молекулярных и клеточных экспериментов (оптимизация методов и проведение экспериментов Hi-C, ChIP-seq, RNA-seq, FISH, иммуноокрашивание антителами, поддержание комариных колоний и культивирование клеточной линии MSQ43), а также подробный анализ карт пространственных контактов комаров родов *Anopheles* и *Aedes*, треков распределения гистоновых меток и данных по активной транскрипции. Масштабная биоинформатическая обработка полученных данных была выполнена м.н.с. М.А.Нуриддиновым (ИЦиГ СО РАН, Новосибирск), отдельные технические задачи – А.К.Таскиной (ИЦиГ СО РАН, Новосибирск) и В.С.Фишманом (ИЦиГ СО РАН, Новосибирск). Анализ данных транскриптомов и эпигенетических профилей был выполнен м.н.с. П.С.Белокопытовой (НГУ, Новосибирск). Анализ BUSCO был проведен сотрудниками лаборатории Роберта Уотерхауса (Лозанна, Швейцария). Hi-C-библиотеки для комаров рода *Aedes* были подготовлены сотрудником лаборатории И.В.Шарахова Джиангтао Лиангом (Virginia Tech, США), карты контактов построены н.с., к.б.н. И.И.Брусенцовым (ИЦиГ СО РАН, Новосибирск).

Структура и объем работы. Диссертация состоит из введения, четырех глав, выводов, списка литературы и семи приложений. Работа изложена на 167 страницах, проиллюстрирована 33 рисунками и содержит 9 таблиц.

Благодарности. Автор диссертационной работы выражает искреннюю благодарность следующим лицам: д.б.н., профессору И.В.Шарахову (Virginia Tech, Блексбург, США) за то, что был идейным вдохновителем работы, оказал неоценимую помощь в теории проведения экспериментов, а также предоставил технические возможности и оборудование своей лаборатории; в.н.с, к.б.н. В.С.Фишману (ИЦиГ СО РАН, Новосибирск), так как он был чутким наставником, руководил научным процессом, идеологически направлял движение; зав.лабораторией, к.б.н. Н.Р.Баттулину (ИЦиГ СО РАН, Новосибирск) за первое знакомство с методом Hi-C, за руководство и поддержку на первых этапах научного пути, за интерес к окружающему миру и за то, что личным примером вдохновляет на продолжение исследовательской деятельности; м.н.с. М.А.Нуриддинову (ИЦиГ СО РАН, Новосибирск) - за продуктивное научное взаимодействие, проведение обширной биоинформатической части исследования и

помощь в освоении азов программирования; м.н.с. П.С.Белокопытовой (НГУ, Новосибирск) – за анализ данных транскриптомов и эпигенетических профилей, а также за критический подход к делу; к.б.н. И.И.Брусенцову (ИЦиГ СО РАН, Новосибирск) – за дискуссии, построение многочисленных Hi-C-карт для комаров рода *Aedes*; к.б.н., профессору М.В.Шараховой (Virginia Tech, Блексбург, США) за ценные советы и рекомендации по части инверсионных полиморфизмов у *Aedes*; д.б.н., профессору П.М.Бородину – за дружескую поддержку и возвращение научного потенциала; всем коллегам и сотрудникам отдела молекулярных механизмов онтогенеза во главе с д.б.н., профессором О.Л.Серовым (ИЦиГ СО РАН, Новосибирск), а также сотрудникам лаборатории И.В. Шарахова в Политехническом Университете Вирджинии (Virginia Tech, Блексбург, США), которые своим исследовательским оптимизмом поддерживали автора на его пути.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1 Первые исследования организации хроматина и структуры хромосом

Геном высших эукариот организован в пространстве ядра не случайным образом. Основные принципы укладки хроматина были открыты в первой половине прошлого века благодаря базовым исследованиям с использованием светового микроскопа, а также классическим биохимическим и молекулярным подходам. Первые работы с применением светового микроскопа выявили существование районов активного и неактивного хроматина, а также занимаемое ими определенное положение внутри ядерного пространства (обзорно представлено в [13]).

Дальнейшие исследования в этой области привели к открытию хромосомных территорий, околософерических образований отдельных хромосом, диаметр которых меньше диаметра клеточного ядра. Хромосомные территории можно наблюдать на стадии интерфазы в клетках животных (млекопитающих, курицы), растений (арабидопсиса), грибов [14–17]. Как оказалось впоследствии, белковый комплекс конденсина II, компактизуя хромосомы вдоль собственной оси, играет определяющую роль в формировании хромосомных территорий, предупреждая неконтролируемое перемешивание генетического материала [18]. В то же время для клеток дрожжей, злаков, саламандр, некоторых круглых червей, двукрылых насекомых характерно наличие несколько другой системы организации интерфазных хромосом – Rab1-конфигурация или конфигурация по Раблю [7,8,19]. Особенностью данной структуры является объединение/кластеризация центромерных областей хромосом на одном полюсе ядра, а теломерных – на другом. Существуют и промежуточные варианты, когда центромеры остаются сцеплены вместе, а теломеры расходятся по периферии ядра, либо теломеры прикреплены к одному участку ядерной оболочки, а центромеры расположены по периферии (конформация «bouquet»). Действительно, такое расположение хромосом характерно для некоторых полиплоидных злаков [20,21]. Существует гипотеза, согласно которой Rab1-конфигурация хромосом в интерфазе представляет остаточный после митоза вариант организации ядра, когда соединенные на полюсах центромеры не успевают разойтись [21]. Забегая вперед отметим, что новейшие данные свидетельствуют о том, что хромосомные территории и Rab1-конфигурация хромосом являются биологическими антагонистами или крайними состояниями одного процесса и контролируются на протяжении клеточного цикла субъединицами комплекса конденсина II [22].

Двукрылые, в частности род *Drosophila*, являются уникальным объектом, поскольку в их организме можно наблюдать различные типы организации интерфазного хроматина. Так, для диплоидных клеток раннего эмбриона, развивающихся элементов нервной системы у личинок, а также политенных клеток слюнных желез характерна Rabl-конфигурация [23–25], в то время как для сперматоцитов G2-фазы первого мейоза характерно наличие классических хромосомных территорий [26]. Наиболее любопытная картина наблюдается в питающих клетках яичников, содержащих политенные хромосомы, где организация хроматина непостоянна и изменяется в процессе развития фолликула, претерпевая суммарно три различных состояния [27]. На протяжении первых четырех эндоциклов хромосомы плотно конденсированы, достаточно узкие и короткие, проходят полную репликацию ДНК и наблюдаются в виде пяти отдельных хромосомных плеч. Затем, в конце пятого эндоцикла репликационный процесс прекращается раньше полного завершения синтеза ДНК, клетка проходит M-подобную фазу клеточного цикла, во время которой образуются 32 пары сестринских хроматид, хромосомы постепенно проходят деконденсацию и приобретают структуру, называемую “5-blob stage”, когда их плечи соединены узкими перемычками центромерного гетерохроматина. Данное образование очень напоминает хромосомные территории, наблюдаемые в клетках млекопитающих. На протяжении последующих эндоциклов репликация в клетке проходит не полностью, хромосомы деконденсированы, перемешаны между собой, занимают весь объем клеточного ядра и представляют 32 субполитенные структуры. Таким образом, за три эндоцикла оогенеза питающие клетки яичников переходят из стадии конденсированных политенных хромосом в стадию 32-плоидной политении [27].

Дальнейшие исследования в области организации интерфазного хроматина и хромосомных территорий привели к открытию взаимосвязи между плотностью генов на хромосоме и ее положением в пространстве ядра [28–31]. В частности, была замечена предпочтительная локализация транскрипционно активных районов в центре ядра, в то время как бедные генами участки или районы со сниженной транскрипцией располагались по периферии. Проявление данной закономерности мы можем наблюдать у насекомых, птиц, млекопитающих, что свидетельствует о строгой консервативности принципа. Исключение в данном правиле составляют клетки с высокоспециализированной функцией, например, палочки сетчатки у ночных млекопитающих, ядра которых выполняют роль собирающей свет линзы, обеспечивая возможность ночного зрения [32]. В этом случае на терминальном этапе нейральной дифференцировки происходит постепенная реорганизация классической укладки хроматина и возникает инвертированная организация ядра, где

неактивные участки собраны в центре ядра в виде 1-2 хромоцентров, а активные локусы размещены по периферии [32–34].

Помимо избирательного расположения хроматина относительно периферии и центра ядра, были обнаружены постоянные и тканеспецифичные районы контактов с ядерной ламиной (ЛАДы) [35–37] и ядрышком (ЯАДы) [38–40]. Также была продемонстрирована предпочтительная локализация некодирующих повторенных последовательностей, не содержащих генов, в центральной части хромосомной территории, что позволяет суммарно занимать меньший объем по сравнению с активными генными локусами [41].

Значительным событием стало возникновение такого метода как флуоресцентная *in situ* гибридизация – FISH [42], с помощью которой стали возможны наблюдения за пространственной локализацией индивидуальных генов и генных кластеров, была показана зависимость 3D-положения геномного локуса от его транскрипционной активности [43–45]. Но несмотря на успехи использования микроскопа, серьёзным лимитирующим фактором данного подхода в течение нескольких десятилетий являлось ограниченное количество участков, которые удавалось исследовать в одном эксперименте, а также довольно низкое разрешение метода. Тем не менее, последние достижения в области микроскопии со сверхразрешением [46–48], а также возникшая возможность прижизненных наблюдений за индивидуальными клетками, появление 3D-SIM-микроскопии или микроскопии структурированного освещения (three-dimensional structured illumination microscopy) [49,50] оставляет за данным подходом многообещающие перспективы в изучении внутриядерных процессов и хроматиновых взаимодействий [48].

1.2 Исследования биофизических свойств хроматинового волокна

Молекула ДНК в клеточном ядре функционирует не в виде самостоятельной структуры, а в тесной взаимосвязи с различными ядерными белками, которые определяют её локальное положение в пространстве ядра, доступность для контактов с другими белками и элементами генома, а также осуществляют вспомогательные процессы при транскрипции и репликации генетического материала. По этой причине в рамках пространственной организации выделяют комплексную структуру, называемую хроматиновой нитью или хроматином, который включает в себя молекулу ДНК и различные вспомогательные белки, и белковые конгломераты, находящиеся с ней в тесной взаимосвязи.

1.2.1 Типы хроматина

Исторически сложилось, что весь хроматин клетки принято делить на 2 основных типа: эухроматин - открытый, обогащенный активными метками и активно транскрибируемыми генами, и гетерохроматин - закрытый хроматин с низкой плотностью генов и отсутствием их активности. Более поздние исследования демонстрируют существование подтипов гетерохроматина и эухроматина, что связано с избирательным взаимодействием со специфическими структурными белками и наличием характерных гистоновых меток на достаточно протяженных участках генома. Так, два известных типа гетерохроматина - PcG и HP1 - выполняют принципиально различные функции в ядре. Первый тип PcG обогащен белками комплекса Polycomb и гистоновой модификацией H3K27me₃, формирует протяженные гетерохроматиновые блоки и участвует в регуляции генов развития и дифференцировки [51]. Тогда как HP1-гетерохроматин связан с белком HP1 и некоторыми ассоциированными с ним белками, содержит гистоновые метки H3K9me₃ и H3K9me₃, преимущественно замечен в прицентромерных областях, и, помимо репрессированных регионов, часто содержит транскрипционно активные гены [52]. Прямые сравнения данных о сайтах посадки указанных белков не показали какого-либо пересечения, что говорит о двух независимых типах гетерохроматина [53].

Полногеномное картирование сайтов связывания 53 структурных белков с молекулой ДНК, а также данные по четырем ключевым гистоновым модификациям (H3K4me₂, H3K9me₂, H3K27me₃ и H3K79me₃) у плодовой мушки *Drosophila* позволили подразделить хроматин на несколько типов [54]. Примечательно, что разделение данных типов хроматина было основано не только на взаимодействии со специфическими белками, но и на биохимических свойствах хроматина, транскрипционной активности, гистоновых метках, времени репликации. В результате в эксперименте на клеточной линии Kc167 были выделены пять типов хроматина, два из которых представляют эухроматин, а три – гетерохроматин. Два типа гетерохроматина, голубой и зеленый, соответствуют описанным ранее PcG- и HP1-обогащенным гетерохроматиновым регионам, соответственно. Черный хроматин, который покрывает большую часть генома ~48%, состоит из протяженных участков, относительно обедненный генами, транскрипционной активностью и активными метками H3K4me₂ и H3K79me₃, связан преимущественно с гистонами H1, белками D1, IAL и SUUR, а также SU(HW), LAM и EFF, функции которых способствуют выполнению роли глобального репрессирующего механизма. В случае встройки трансгенов в область черного хроматина наблюдается преимущественное их замолкание, что говорит о его глобальной роли в качестве транскрипционного репрессора. При этом отмечается, что фрагменты

черного хроматина могут приобретать свойства других типов в различных тканях и при дифференцировке. Для желтого и красного типов хроматина, которые реплицируются в ранней S-фазе и представляют эухроматин, характерны высокий уровень матричных РНК, РНК-полимеразы и меток активного хроматина, H3K4me2 и H3K79me3, обогащение по белкам HDAC (активная транскрипция), DF31 (деконденсация хроматина), ASH2 (субъединица H3K4-метилтрансферазы). Различия между красным и желтым хроматином, вероятно, связаны с функцией и спецификой генов, которые они содержат, а также характерными белками. Так, согласно данным GO-анализа (от англ. **Gene Ontology** - генная онтология), красный хроматин обогащен генами с более специфичной функцией, тогда как желтый хроматин - генами, функционирующими в большинстве тканей организма. Авторы предполагают, что такое различие функций может подкрепляться также двумя различными регуляторными механизмами. В подтверждение данной гипотезы в участках красного хроматина отмечено двукратное обогащение по H3K9me3 (высококонсервативные некодирующие элементы) относительно желтого хроматина, которые широко известны за свою регуляторную роль.

Стоит также отметить существование схожей номенклатуры хроматина у дрозофилы, выполненной группой И. Ф. Жимулева, где за основу были взяты распределение и состав белков в районах междисков политенных хромосом и выделены четыре типа хроматина, а именно: *аквамарин* (aquamarine), *лазурит* (lazurite), *малахит* (malachite) и *руби* (ruby) [55,56]. Характерной особенностью *аквамарина* является обогащение белками ремоделинга хроматина, фактором CHRIZ/CHROMATOR и транскрипционными факторами, участвующими в инициации транскрипции; *лазурит* обогащен белками – участниками элонгации транскрипции, и преобладает в экзонах и интронах генов; *малахит* обогащен структурными белками, такими как CP190, SU(HW) и mod 2.2; хроматин типа *руби*, в свою очередь, содержит ламин, SUUR и белки группы Поликомб (Polycomb-mediated repression), а также перекрывается с ранее определенным в Filion et al. 2010 черным хроматином.

Подобная классификация была выполнена и для генома *Arabidopsis thaliana*, что позволило выявить девять типов хроматина, основываясь на специфических гистоновых метках [57,58]. Примечательно, что выявленные типы хроматина отлично совпадают/коррелируют с известными ранее пятью функциональными элементами генома: сайтами старта транскрипции (TSS) и промоторными областями, областями тел генов, регуляторными межгенными областями, хроматином, связанным с комплексом Polycomb, репрессированным гетерохроматином [57,59]. Интересно, что оставшаяся часть генома,

которая не попала ни в одну из выделенных групп, представляет межгенные участки с промежуточными характеристиками хроматина.

1.2.2 Описание структуры хроматина в терминах физики биополимеров

Самый первый уровень, на котором проявляются сильные ДНК-белковые взаимодействия, это уровень нуклеосом, где хроматиновая нить представлена в виде длинного гибкого полимера и, вероятно, подчиняется тем же физическим законам и фундаментальным принципам, что и классическая полимерная молекула. Поскольку пространственная укладка и биологическая динамика хроматинового веретена представляет огромный научный интерес для исследователей, предпринимались неоднократные и очень успешные попытки построения 3D-моделей и описания данной структуры в терминах физики биополимеров [60–63]. При отсутствии сторонних ограничивающих воздействий существуют три основных физических параметра, определяющих форму полимера и средний занимаемый им объем: персистентная длина (количественная характеристика гибкости полимера), плотность массы (характеристика степени компактизации полимера) и одно из трех описанных далее состояний или типов полимера [64,65]. Первый тип - это *random coil* (или *random walk*), при котором вероятность двум локусам встретиться и сформировать хроматиновую петлю очень быстро уменьшается с увеличением расстояния между ними (при изменении линейного расстояния между двумя участками полимера данного типа в 4 раза расстояние между ними в пространстве изменяется вдвое). Второй тип, когда составляющие полимера обладают свойством стерически взаимодействовать друг с другом (отталкиваться или притягиваться), называется *swollen globule*. Объем, занимаемый полимером данного типа, в силу меньшей подвижности последнего будет увеличиваться по сравнению с *random coil*. И, наконец, состояние полимера, которое, как оказалось, максимально точно описывает пространственную укладку молекулы хроматина – *fractal globule* [5,66]. Этот тип занимает минимальный объем в пространстве, не образует узлов и запутанных участков, сохраняет достаточно высокую свободу перемещений (вероятность встречи двух фрагментов при увеличении дистанции между ними падает медленнее, чем для *random walk* или *swollen coil*), но при этом не является равновесным и при достаточном времени ожидания обычно переходит в *random walk* или *swollen coil*.

Хотя биофизические модели пространственной укладки полимеров предписывают высокую подвижность молекул хроматина внутри ядра, в реальности на объем, занимаемый хроматином, оказывают воздействие несколько факторов: ядерная оболочка,

необходимость существования отдельных хромосомных территорий, а также чрезвычайно низкая скорость свободного перемешивания длинных полимеров (Rosa and Everaers 2008), что позволяет значительно сократить возможные пространственные конформации и предположить существование неких общих схем организации хроматинового веретена [67].

Основным принципиальным отличием между 3D-структурой белка и пространственной укладкой молекулы хроматина является невозможность последнего сформировать две одинаково уложенные молекулы: каждая является уникальной по своей пространственной структуре. Именно по этой причине в исследованиях на клеточных популяциях, где каждая отдельная единица имеет свою хроматиновую структуру, для описания укладки хромосомной нити внутри ядра необходимо использовать статистический подход. Эксперименты с использованием метода FISH подтвердили существование большого межклеточного разнообразия в расположении в пространстве ядра одних и тех же локусов даже среди клеток одного типа [47,68].

1.3 Новые подходы в исследовании 3D-структуры генома

1.3.1 Группа 3C-методов, вариации подхода

Интенсивное развитие молекулярных и биоинформатических подходов, накопление данных о неслучайных контактах участков генома внутри ядра, описание биофизических свойств нити хроматина – факторы, которые напрямую и косвенно привели к появлению технологии, называемой *методом захвата конформации хроматина* (chromosome conformation capture или 3C) [2,69]. Используя набор биохимических и молекулярных реакций - ковалентная сшивка/фиксация ДНК-белковых взаимодействий, последовательные реакции фрагментации и межмолекулярного лигирования сближенных фрагментов генома, - 3C-метод позволяет оценить частоту пространственных взаимодействий двух определенных участков хроматина. Имея данные частот таких контактов с покрытием для всего генома, появляется возможность идентифицировать положение участков хроматина относительно друг друга в пространстве ядра. На сегодняшний день доступны различные модификации данного метода и в зависимости от целей и приоритетов исследования протоколы на основе метода захвата конформации хроматина предоставляют информацию о различных контактах. Когда известны два участка генома, например, промотор гена и его предполагаемая регуляторная последовательность, и необходимо оценить, насколько часто они взаимодействуют в пространстве ядра в определенном типе клеток (частота взаимодействий одного геномного

локуса с другим) используется базовая вариация метода - 3C [2]. В случае, если стоит задача выявить контакты (регуляторные последовательности) целевого локуса с регионами из списка кандидатов (контакты одного локуса с несколькими участками генома) подойдет 4C-подход [70]. Когда в исследовании участвуют одновременно несколько целевых последовательностей и несколько их предполагаемых регуляторов, то изучают хроматиновые взаимодействия “многих со многими” методом 5C [71]. Если же возникает задача охарактеризовать 3D-структуру всего генома, необходимо идентифицировать пространственные контакты “всех со всеми”.

Объединение 3C-протокола с технологией массового параллельного секвенирования, дополненное мечением химерных молекул, образуемых из сближенных в пространстве участков генома, получило название Hi-C-метода [72,73] и открыло для исследователей возможность изучать частоты взаимодействий всех участков хроматина в полногеномном масштабе. Среднее разрешение метода составляет несколько тысяч пар нуклеотидов, что в десятки раз превосходит возможности микроскопии. Использование комбинации различных эндонуклеаз рестрикции позволяет достичь разрешения метода в ~500 п.н. (данный параметр может варьировать в зависимости от обогащения конкретного генома сайтами узнавания эндонуклеазы рестрикции), что открывает беспрецедентные возможности для исследования пространственных взаимодействий хроматина с высокой точностью ([74,75], коммерчески доступные наборы реактивов Arima Hi-C kit+ и Arima High Covarage Hi-C kit). В настоящее время известны несколько эндонуклеаз рестрикции, которые активно используются в Hi-C экспериментах. К ним относятся: HindIII, BglII, SacI, BamHI, EcoRI, AclI, DpnII, MboI [76].

Поскольку использование эндонуклеаз рестрикции напрямую связано с распределением их сайтов узнавания по геному, покрытие прочтениями после секвенирования получается неравномерным. В большинстве работ разрешение метода с использованием рестриктаз оказывается достаточным для получения качественных данных, но в задачах, где, например, требуется идентифицировать точечные взаимодействия на близких расстояниях, необходима высокая глубина прочтений и исключительная точность разрешения, поэтому было предложено использовать менее зависимый от нуклеотидного состава фермент – дезоксирибонуклеазу I (ДНКазу I). Данный подход получил название ДНКазного Hi-C или Echo-C [74,77,78] и активно используется в прикладных научных задачах, как, например, изучение экзома, детекция хромосомных перестроек, поиск SNV (single nucleotide variations). Еще одной альтернативой эндонуклеазам рестрикции, узнающим 4 или 6 нуклеотидов в цепи ДНК, может выступать

микрোকковая нуклеаза - MNase. Модификация метода носит название Micro-C и поднимает разрешение карт пространственных контактов на уровень нуклеосом [79,80].

Технология сHi-C (capture Hi-C) позволяет обогатить популяцию полученных Hi-C данных за счет специфической гибридизации с целевой библиотекой [81–83]. В частности, матрицей может служить РНК-библиотека (RNA bait) активно транскрибируемых генов, или меченые олигонуклеотиды, специфичные для исследуемого района генома. В результате данный подход позволяет получить информацию о пространственных взаимодействиях исследуемого района (или нескольких локусов) с более высоким разрешением и покрытием по сравнению с обычными Hi-C библиотеками.

В том же году, что и Hi-C, была впервые описана технология ChIA-PET (Chromatin interaction analysis by paired-end tag sequencing), представляющая молекулярный гибрид между ChIP (иммунопреципитацией хроматина) и Hi-C-методом и позволяющая выявлять пространственные взаимодействия участков хроматина, опосредованные определенным белком [84–86]. В процессе обработки клеточного лизата параформальдегидом происходит ковалентная сшивка комплексов ДНК с белками, после чего они избирательно преципитируются на магнитных шариках, связанных со специфическими антителами. Далее происходит обработка ультразвуком, после чего на преципитированные фрагменты хроматина пришивают линкерные метки, затем происходит лигирование в сильно разбавленных условиях, рестрикция по сайтам внутри линкеров для получения линейных молекул, подготовка фрагментов к секвенированию и прочтение полученных химерных молекул. Данная технология активно используется при поиске регуляторных последовательностей, а также изучении энхансер-промоторных взаимодействий [87–89].

Идеологически сходные с ChIA-PET методы – HiChIP и PLAC-seq– позволяют использовать в 100 раз меньшее количество клеток без потери в разрешении карт контактов. Данное улучшение достигается благодаря сохранению интактности ядер на всех этапах вплоть до межмолекулярного лигирования, за которым следуют соникация и иммунопреципитация хроматина. Изменение порядка действий по сравнению с ChIA-PET позволило заметно улучшить эффективность метода и соотношение уровня сигнала к шуму, а также сделать подход воспроизводимым для различных организмов и клеточных типов [90,91].

Помимо изучения пространственной организации в популяции клеток, где нивелируются все индивидуальные различия, C-технологии позволяют работать и с единичными клетками. Так, был разработан подход – snHi-C (single cell или single nucleus

Hi-C), оптимизированный для работы с хроматином одной клетки [92,93]. Были получены тепловые карты контактов для отдельных T_H1-клеток, мышинных ооцитов, гаплоидных и диплоидных мышинных эмбриональных стволовых клеток (мЭСК), мононуклеаров периферической крови, лимфобластоидной клеточной линии GM12878 и других [92–95]. Оказалось, что данные по клеточной популяции представляют усреднение пространственных взаимодействий, тогда как в индивидуальных клетках становится возможным наблюдать динамичность геномных образований. Ключевые структуры 3D-организации генома - компартменты, ЛАДы (ламин-ассоциированные домены), энхансер-промоторные взаимодействия - сохраняются в каждой клетке, при этом межхромосомные контакты, внутридоменные взаимодействия и хроматиновые петли имеют определенную долю вариативности [96]. Стоит также отметить модификацию snHi-C метода – Sci-Hi-C, которая позволяет без физического выделения отдельных ядер изучать разнообразие пространственных контактов в единичных клетках путем комбинаторного мечения хроматина популяции клеток введением индивидуальных баркодов [75,97]. Особенностью DipC-подхода, еще одной вариации snHi-C, является возможность разделять пространственные контакты отдельных гомологичных хромосом [94], что может быть существенно при изучении молекулярных механизмов импринтинга или трансекции.

1.3.2 Особенности Hi-C

Требуемое количество стартового материала для Hi-C эксперимента в среднем составляет более пяти миллионов клеток, но может быть существенно меньше, что зависит от типа и свойств исследуемых клеток. Существующие наборы готовых реактивов (Arima genomics, 10XGenomics, Dovetail Genomics и другие), позволяющие провести эксперимент Hi-C от стадии фиксированных образцов до подготовки библиотек к секвенированию, оптимизированы для использования 1-5 мкг ДНК фиксированного материала.

Протокол приготовления Hi-C-библиотек включает несколько основных этапов. Образец ткани, клеточная культура или же целый организм фиксируются в 1-4% параформальдегиде с образованием ДНК-белковых (ДНК-ДНК, белок-белок) сшивков с одновременной или последующей гомогенизацией до одноклеточной суспензии. Затем следует этап лизиса клеточных мембран с сохранением структуры клеточных ядер и обработка ядерной фракции выбранной эндонуклеазой рестрикции. Мечение образующихся липких концов осуществляется биотинилированным А- или С-нуклеотидом с помощью фрагмента Клёнова. Смесь маркированных и немаркированных фрагментов хроматина обрабатывается T4-лигазой в сильно разбавленных условиях, в результате чего

ДНК-фрагменты, сшитые с одним белком/белковым комплексом, оказываются связаны в одну молекулу. Далее, происходит ультразвуковая обработка (соникация) полученных химерных молекул с целью получения фрагментов в диапазоне 300-500 п.н., необходимых для последующего NGS-секвенирования. Затем выполняется обогащение целевой библиотеки за стрептовидиновые магнитные шарики, которые связывают молекулы встроенного биотина с высокой аффинностью, что позволяет включить в анализ именно химерные молекулы участков ДНК, контактирующих в пространстве ядра (Рисунок 1).

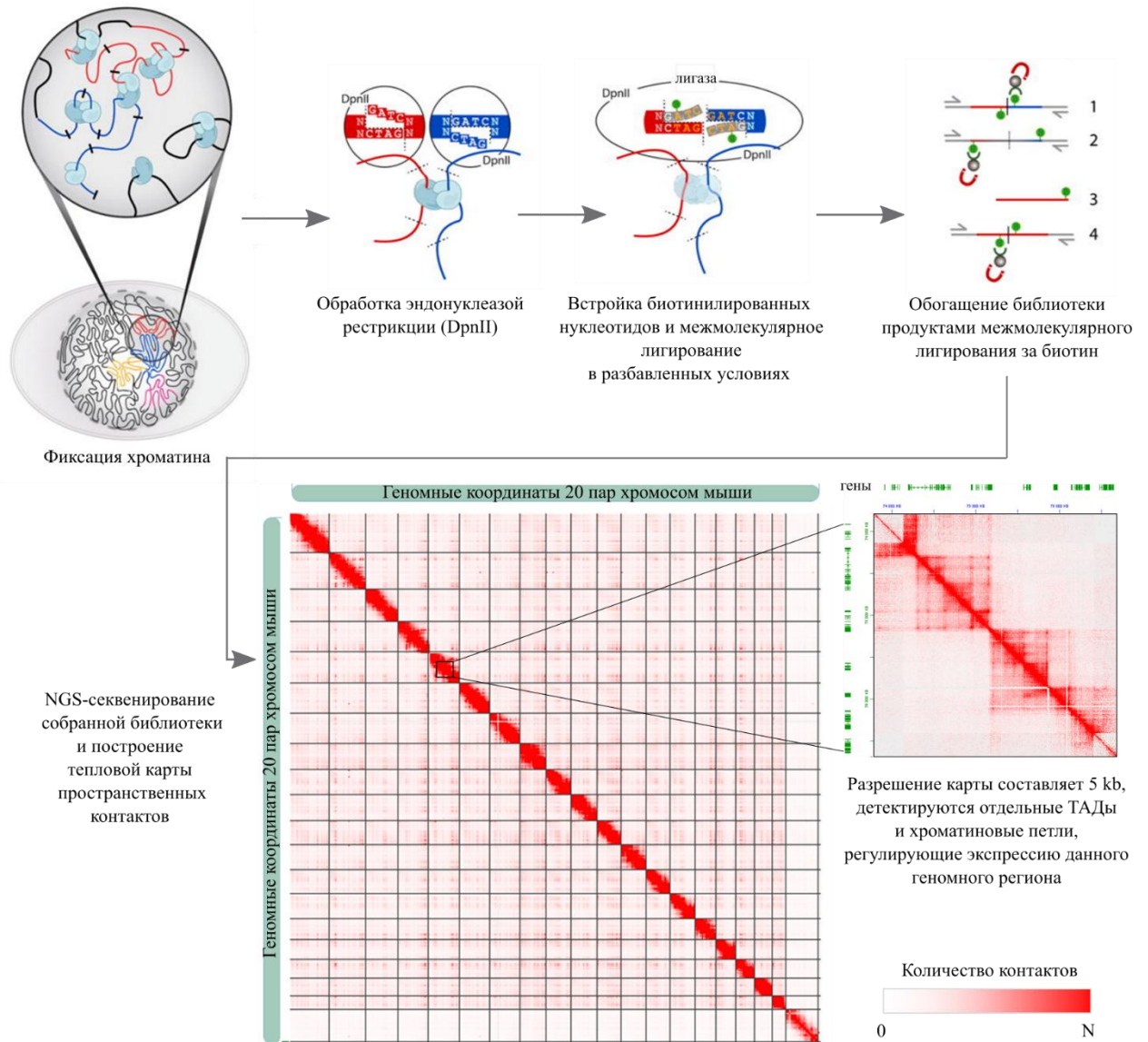


Рисунок 1. Схема протокола Hi-C на примере мышинного генома. Зелеными кружками отмечены биотинилированные нуклеотиды. Частично адаптирован из [98].

Дальнейшая подготовка Hi-C библиотек к полногеномному секвенированию осуществляется с использованием специализированных наборов, адаптированных для секвенирующей платформы Illumina (KAPA HyperPrep Kit, NEBNext Ultra II DNA Library

Prep Kit, TruSeq DNA Library Prep kit, Nextera DNA Flex Library Prep kit и другие). После получения данных полногеномного секвенирования, прочтения (или риды) выравниваются на референсный геном и затем обрабатываются в специализированном программном обеспечении, например, Hi-C Pro [99], Juicer [100] и других. В результате алгоритм формирует матрицы пространственных контактов, которые визуальны представлены в виде тепловой карты, где по обеим осям отложены геномные координаты, а насыщенность цвета каждого квадрата (или бина) карты прямо пропорциональна частоте взаимодействий фрагментов генома, на пересечении которых находится квадрат (Рисунок 1).

1.3.3 Применение технологии Hi-C для улучшения геномных сборок

Одним из побочных, но чрезвычайно важных результатов применения метода Hi-C стала возможность создавать новые и качественно улучшать существующие геномные сборки живых организмов, которые состоят из многочисленных отдельных контигов и содержат множественные повторенные последовательности [101,102]. Используя данные о частотах пространственных взаимодействий фрагментов генома, исследователи получают точную информацию о порядке расположения отдельных контигов и ридов. Особенно впечатляющие результаты дает объединение технологий, продуцирующих длинные прочтения (Oxford Nanopore, PacBio sequencing, 10X genomics), и Hi-C-метода. За последние 2-3 года по описанной стратегии были собраны геномы представителей разных царств, которые ранее состояли из тысячи отдельных разрозненных контигов [103–112].

На момент начала диссертационной работы хромосомные сборки геномов существовали для двух из пяти видов анофелесов, выбранных в эксперимент (*An. albimanus* и *An. atroparvus*). Ранее геномные сборки комаров получали преимущественно секвенированием на платформе Illumina, где максимальный размер прочтения составляет 2x300 п.н. [113], а в отдельных экспериментах может достигать 750 п.н. [114]. И, хотя вероятность возникновения ошибки в единичном прочтении достаточно мала, 0,1% на нуклеотид [115], общее качество итоговой сборки получалось довольно низким в силу сложности объединения отдельных контигов в более крупные скаффолды. С появлением новых технологий, таких как ONT (Oxford Nanopore Technology), PacBio (Pacific Biosciences), 10XGenomics, позволяющих прочитывать молекулы ДНК длиной до нескольких мегабаз [116], прочтение геномов вышло на новый уровень. Именно объединение нескольких технологий позволяет обойти ограничения каждой, в результате чего количество искусственных вставок NNN снижается, поднимается континуальность

отдельных скаффолдов, аккуратность и достоверность сборок существенно возрастают [117].

1.4 Уровни пространственной организации хроматина в ядре, ключевые структурные единицы

Основным результатом развития молекулярных подходов и C-технологий стало открытие и укрепление представлений о специфических формированиях нити хроматина - хромосомных территориях, А/В-компартаментах, ТАДах (топологически ассоциированных доменах), ЛАДах (ламино-ассоциированных доменах), ЯАДах (ядрышко-ассоциированных доменах), хроматиновых петлях - как о важных и значимых структурных единицах в пространстве клеточного ядра. Таким образом, утверждается, что геном в пространстве интерфазного ядра иерархично организован и состоит из несколько уровней, каждый из которых образуется и функционирует по уникальному механизму (Рисунок 2). В данном разделе будет дано современное представление о каждой из упомянутых выше структур.

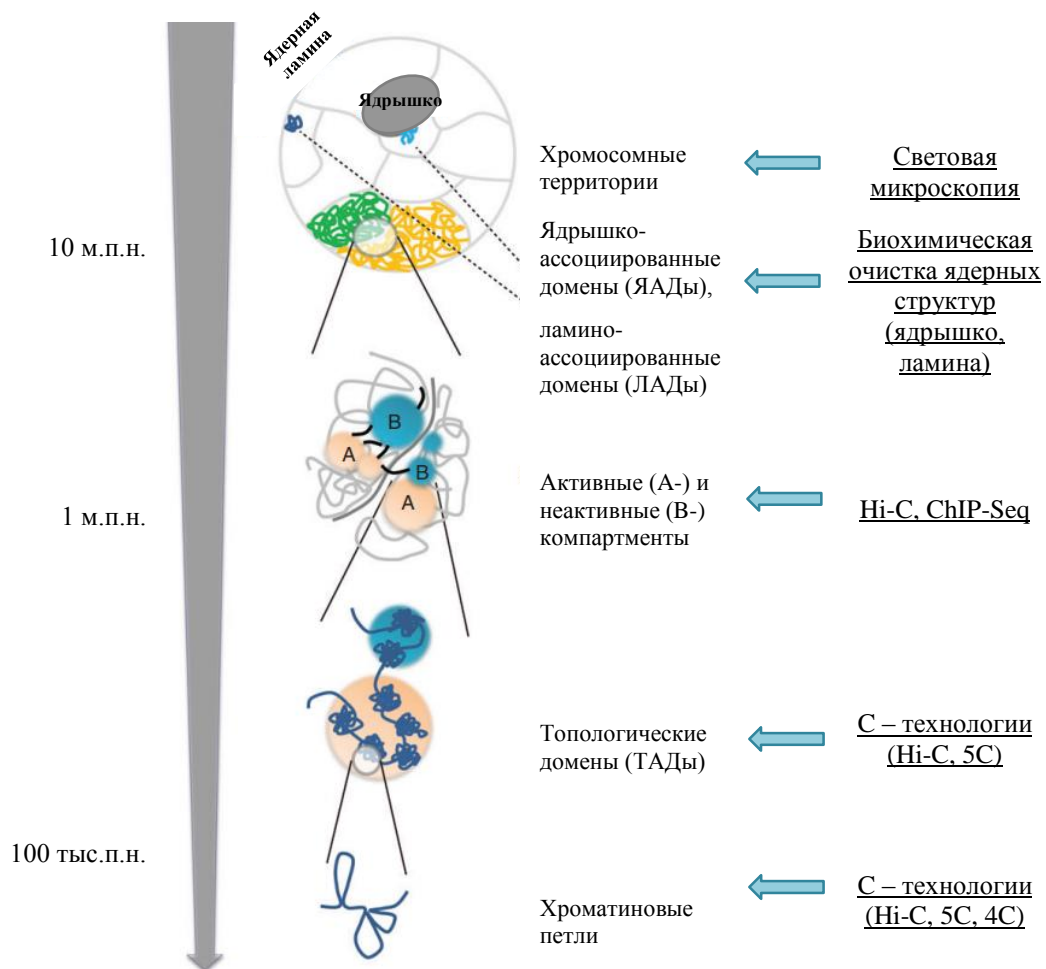


Рисунок 2. Иерархическая лестница пространственной организации хроматина в интерфазном ядре. Адаптирован из [123].

1.4.1 Активные и неактивные хроматиновые компартменты

Анализ тепловых карт пространственных взаимодействий, полученных с помощью Hi-C, позволил выявить три основных уровня организации интерфазного хроматина – компартменты, ТАДы и петли. В данном разделе речь пойдет о наиболее крупных структурах хроматина, ядерных компартментах. Они представлены достаточно протяженными блоками, представляющими районы активного и неактивного хроматина, которые могут изменяться в процессе дифференцировки [118], и которые принято называть А- и В-компартментами [5,119]. Важной особенностью в процессе дифференцировки считается усиление взаимодействий между В-компартментами и ослабление взаимодействий между А-компартментами [120]. Типичная визуализация данных структур на картах контактов – это так называемый паттерн “шахматной доски” или “клетчатый паттерн”, где контактирующие в пространстве ядра фрагменты активного или репрессированного хроматина формируют “шахматную клетку”. А-компартменты представляют участки генома, для которых характерно обогащение генами и очагами активной транскрипции, преобладание активных гистоновых меток H3K36me3, H3K79me2, H3K27ac, H3K4me1, ранняя репликация, а также полное отсутствие перекрывания с ЛАДами и ЯАДами. Активные компартменты можно подразделить на А1 и А2, где для А2 характерен более низкий G/C-состав, преобладание длинных генов и чуть более долгое время репликации, продолжающееся до середины S-фазы клеточного цикла. В-компартменты, представляющие условно неактивный хроматин, подразделяются на четыре субкомпартмента – В1, В2, В3 и В4. Так, для субкомпартмента В1 была показана позитивная корреляция с H3K27me3 и негативная с H3K36me3. Субкомпартмент В1 представляет факультативный гетерохроматин, репликация которого приходится на середину S-фазы. Подтипы В2 и В3 реплицируются в самом конце S-фазы и не содержат специфических гистоновых меток, характерных для других субкомпартментов. Более половины фрагментов В2 приходится на прицентромерный гетерохроматин, а также пересекается с ЛАДами и ЯАДами. В3 попадает в районы ЛАДов, но совершенно отсутствует в ЯАДах. Четвертый тип субкомпартментов был открыт несколько позднее остальных, занимает всего 11 Мб и расположен на 19 хромосоме человеческого генома. Так, для него типично содержание гистоновых меток как активных H3K36me3, так и неактивных, H3K9me3 и H4K20me3, ассоциированных с гетерохроматиновыми областями. Данный субкомпартмент уникален тем, что в нем сосредоточено около половины генов суперсемейства KRAB-ZNF по геному [119]. Белковые единицы, входящие в состав

семейства KRAB-ZNF, являются специфичными для тетрапод и представляют самый большой класс ZNF-белков.

1.4.2 Взаимодействия хроматина с ядерной ламиной. Ламины. ЛАДы

Как известно, ядерная оболочка, состоящая из внутренней и внешней мембран и подстилающей их фибриллярной сети многочисленных белков и белковых комплексов, активно участвует в организации хроматина на протяжении всего клеточного цикла. Одними из важнейших компонентов фибриллярной сети или ядерной ламины являются ламины, белковые молекулы, способные связывать ДНК, гистоны, а также белки, взаимодействующие с гистонами, и оказывать косвенное влияние на регуляцию генов. Недавние обзоры свидетельствуют о том, что ламины способствуют организации и закреплению больших блоков гетерохроматина возле ядерной ламины [121,122]. Микрофотографии, сделанные с помощью электронного микроскопа, также свидетельствуют о преимущественном накоплении блоков гетерохроматина вдоль ядерной ламины, которая может обеспечить необходимое инертное/неактивное окружение [123].

Состав белков ядерной периферии непостоянен и изменяется в течение клеточного цикла и процесса дифференцировки, создавая возможность для образования ЛАДов – хроматиновых структур в размерном диапазоне от 0.1 до нескольких млн.п.н., активно взаимодействующих с ламиной, содержащих в большинстве своем неактивные гены, обогащенных репрессивными модификациями гистонов, покрывающих ~40% генома, с плотностью генов, вдвое меньшей по сравнению с другими участками генома [35–37,124]. ЛАДы были выявлены независимо несколькими методами, а именно: технологией DamID, где белок Dam сливали с компонентами ядерной ламины [35,125,126]; иммунопреципитацией хроматина за антитела против ламинных белков; а также подтверждены при помощи FISH. Отмечается, что у млекопитающих границы ЛАДов (а также участки в пределах нескольких тыс.п.н. от них) обычно четкие и отмечены особыми характеристиками: обогащены сайтами посадки белка CTCF, содержат активные промоторы или CpG-островки [127]. Данная особенность может быть связана с необходимостью блокировки распространения ЛАДов по хроматину [128]. Существенным различием ЛАДов млекопитающих и насекомых является несовпадение паттернов белков HP1 и ламина В у дрозофилы [35,54], в то время как у млекопитающих наблюдается обогащение по H3K9me2 (одной из гистоновых модификаций, распознаваемых HP1) в районах, принадлежащих ЛАДам [36,129]. При этом метка H3K27me3 ассоциированной с комплексом Polycomb репрессии хроматина обнаруживается в ЛАДах и млекопитающих, и

насекомых, что позволяет предположить существование не охарактеризованного типа гетерохроматина, локализованного в районах ЛАДов. Кроме того, для геномов человека и дрожофилы был отмечен тот факт, что некоторые хромосомы имеют относительно низкое содержание ЛАДов (хромосома 19 у человека), что, вероятно, связано с их более центральным положением относительно периферии [36,37]. Как выяснилось позднее, ЛАДы консервативны между видами, но в то же время представляют довольно динамичный класс элементов, который меняется в зависимости от стадии развития клетки. Подвижность структуры ЛАДов в зависимости от фазы циркадного ритма была показана в работе [130], где авторы выявили более 200 доменов, за 30 часов меняющих свой размер и состав, а также разобщенность событий активности генов и пространственных взаимодействий с ламинной в случае циркадного цикла.

На основании профилей ЛАДов, построенных для многих клеточных типов, их принято разделять на два класса – конститутивные и факультативные [131]. В работе по изучению гепатотоксичности циклоспорина А (CsA), проведенной на клетках гепатокарциномы HepG2, авторы выделяют А-ЛАДы, В-ЛАДы, А/В-ЛАДы, которые преимущественно взаимодействуют с ламинной А, В или обоими соответственно [132]. В другой работе авторам удалось выделить так называемые эЛАДы (эухроматиновые ламин-В1-ассоциированные домены), которые обеспечивают активную экспрессию эухроматиновых генов во время эпителиально-мезенхимального перехода. Они построили полногеномный профиль расположения ламин В1, и оказалось, что он хорошо коррелирует с областями открытого эухроматина и активно транскрибируемых генов. Снижение экспрессии ламин В1 приводило к масштабному изменению транскрипционного профиля и направления эпителиально-мезенхимального (ЭМ) перехода, что подтверждает необходимость правильной 3D-организации для многих клеточных событий, в том числе для осуществления ЭМ-перехода [133]. Сопоставление данных с Hi-C-картами показало пересечение районов, обогащенных по ламину В1, с сильными границами ТАДов.

Удивительно, что в некоторых клеточных типах можно наблюдать инвертированную архитектуру ядра, когда периферический гетерохроматин занимает центральное положение, а эухроматин с активно транскрибируемыми генами уходит на периферию в виде тонкого ободка. В работах Соловей [32,33] удалось показать, что инвертированный тип ядра является следствием отсутствия экспрессии ламин В и А/С. Возникла гипотеза, позднее подтвержденная компьютерным моделированием, что для компартиментализации эу- и гетерохроматина в инвертированном ядре достаточно только гетерохроматиновых взаимодействий, добавление же гетерохроматиновых контактов с ламинной является

необходимым условием построения архитектуры обычного неинвертированного типа [134]. Таким образом, взаимодействие хроматина с ламиной представляет чрезвычайно важный класс контактов, который приводит к формированию ядра по стандартному типу. Нарушение взаимодействий гетерохроматина с ламиной приводит к серьезным порокам развития, обычно несовместимым с жизнью, вследствие необратимых нарушений генной экспрессии как результат изменений пространственной организации ядра [135–137].

1.4.3 Доменные структуры, связанные с локализацией участков хроматина вблизи ядрышка. Ядрышко-ассоциированные хроматиновые домены (ЯАДы)

Другой тип хроматиновых доменов был описан группами Лангст и Ламонд, когда им впервые удалось выделить и очистить ядрышко и связанные с ним участки хроматина. В результате эксперимента была показана неслучайность районов связывания, характерную для всех хромосом, и выявлены хроматиновые домены, избирательно связывающиеся с ядрышковой структурой [38,39]. Как известно, ядрышко является крупнейшей ядерной структурой, расположение которой определяется рибосомальным локусом повторенных последовательностей и рибосомальными некодирующими молекулами РНК. Данная структура необходима для обеспечения масштабной транскрипции рРНК и рибосомального биосинтеза [40]. Благодаря успеху авторов стало известно, что помимо основной функции, вокруг ядрышка происходит скопление гетерохроматиновых локусов (в том числе районов центромер), которые формируют хроматиновые домены, названные ЯАДами или ядрышко-ассоциированные хроматиновые домены. Данные структуры обогащены репрессивными гистоновыми метками (H3K9me3 и другие), также им характерна низкая плотность генов и сниженная транскрипционная активность. Нетрудно заметить, что оба типа хроматиновых доменов, и ЛАДы, и ЯАДы, обладают схожими характеристиками в отношении замолкания транскрипции. Кроме того, было показано частичное перекрывание участков, связанных с ЛАДами и ЯАДами, что может свидетельствовать о способности репрессированного хроматина контактировать с обоими структурами в равной степени [38,39].

1.4.4 Домены ранней и поздней репликации

Давно известно, что в течение S-фазы время репликации ДНК для различных участков генома может варьировать в достаточно широком диапазоне. Так, еще в конце прошлого века в цитологических исследованиях было замечено, что хроматин, расположенный у периферии ядра, начинает реплицироваться преимущественно в средней и поздней S-фазе, тогда как центрально расположенный эухроматин в ранней и средней

[138]. Недавние эксперименты показали, что пространственную организацию генома можно представить в виде ранних и поздних репликативных доменов, причем подобное разделение подкрепляется различиями по плотности генов, активностью транскрипции, нуклеотидным составом [139,140]. Репликативные домены были выявлены у человека, мыши, дрозофилы [139–142], что говорит об их сохранении в процессе эволюции организмов. Оказалось, что для геномов мыши и человека, претерпевших многочисленные хромосомные перестройки относительно друг друга, принадлежность участков хроматина к репликативным доменам сохраняется с высокой консервативностью. При этом важно подчеркнуть, консервативность репликативных доменов имеет независимое от структуры ДНК (плотность генов, GC состав) происхождение [143]. Сравнение доменов поздней репликации и ЛАДов позволило выявить достаточно хорошее перекрывание, но не полную идентичность друг другу [35,37].

1.4.5 ТАДы или топологически ассоциированные домены - фундаментальная структурно-функциональная единица генома живых организмов, их основные характеристики

Накопление данных, полученных методом Hi-C, позволило выявить обособленные структуры в пространстве ядра, приоритетно взаимодействующие внутри себя, имеющие у млекопитающих характерный размер ~0,1-1 м.п.н. и названные топологически ассоциированными доменами или ТАДами [144,145]. Оказалось, что ключевые архитектурные белки и белковые структуры, такие как CTCF, когезиновый (SMC1, SMC3 и RAD21) и конденсиновый (CAP-H2, CAP-D3) комплексы, релиз-фактор WAPL и его партнер PRDM5, белок-загрузчик NIPBL, субъединицы TFIIIC и кадгерина, некоторые другие клеточные белки играют заметную роль в формировании структурной и функциональной целостности ТАДов, что было отмечено в многочисленных исследованиях [146]. Было показано, что большинство граничных районов ТАДов позвоночных в значительной степени обогащены сайтами посадки белка CTCF и субъединицами когезинового комплекса SMC1 и SMC3, что позволяет предполагать участие этих факторов в формировании ТАДов и субТАДов [119]. В недавних исследованиях по иммунопреципитации хроматина (ChIP-seq), предоставляющих информацию о сайтах взаимодействия различных белков с геномной ДНК, было выяснено, что именно однонаправленные сайты посадки CTCF обеспечивают «заякоривание» хроматиновой петли в 90% случаев [147]. Предполагается, что ориентация сайта посадки CTCF принципиально важна [119,147–149]. В случае делеции сайта посадки CTCF ожидаемо исчезал сигнал белка на ChIP-треке и обеспечиваемый им пик (хроматиновая петля) на карте Hi-C. В случае же восстановления сайта посадки CTCF, но в инвертированном

варианте, ChiP-сигнал восстанавливался, а пик не возвращался [148]. Нокаут по белку CTCF приводит к сокращению числа внутридоменных взаимодействий и увеличению контактов между локусами разных доменов [150,151]. Тем не менее, глобальная пространственная организация и дальние взаимодействия по геному в этом случае сохраняются. Данный факт, а также то обстоятельство, что большая часть существующих CTCF-связывающих сайтов расположена вне границ ТАДов (~75%) [144], свидетельствует о наличии дополнительных факторов, вовлеченных в поддержание 3D-структуры хроматиновых доменов.

Как оказалось, ТАДы являются высоко консервативными структурными единицами генома среди представителей класса млекопитающих [144,149], а также среди различных клеточных типов [144,152], как в процессе развития организма, так и на протяжении эволюции. Существование топологической организации было показано и для геномов двукрылых, *Drosophila melanogaster* [153,154] и *Aedes Aegypti* [101], характерный размер ТАДов в этих случаях уменьшается прямо пропорционально размеру геномов и составляет ~100-200 тыс.п.н. и ~700 тыс.п.н. соответственно. При этом плотность генов на один ТАД сохраняется на том же уровне, что у млекопитающих.

Ранее предполагалось, что уровень ТАДов является чрезвычайно древним функциональным образованием и встречается у бактерий, растений и животных. Hi-C исследования, проведенные на растениях - арабидопсисе, рисе, ячмене, кукурузе, сорго, просо, томатах, хлопке - подтверждают существование 3D-организации ядра у данного царства. Согласно последним литературным обзорам [155], у растений можно наблюдать всю иерархическую лестницу пространственной организации ядра, начиная от хромосомных территорий и заканчивая ТАДами и петлями. Но на уровне ТАДов у них проявляется довольно сильное разнообразие. Так, у риса по результатам Hi-C было выявлено несколько тысяч ТАДов, тогда как у арабидопсиса отсутствовали четкие ТАДы и наблюдались лишь элементы, подобные граничным участкам у животных (TAD-boundry-like regions). Подобный феномен может быть следствием высокой вариативности пространственной организации хроматина у растений [155]. Тем не менее, в сравнительном анализе было показано, что для растений характерно разделение генома на доменоподобные структуры четырех типов на основании эпигенетических меток: метилирование ДНК, открытый хроматин, гистоновая модификация H3K27me3, и промежуточный тип хроматина, не обогащенный ни одной из представленных меток. Оказалось, выделенные домены отлично совпадают с локальными компартментами, что также наблюдается в геноме насекомых [154]. Важно также, что, в отличие от животных, в

сравнительном анализе не было выявлено сохранения консервативности доменов между разными растительными видами [155].

Поскольку Hi-C-эксперименты – это исходно усредненные данные по популяции клеток, визуализирующие лишь парные взаимодействия, закономерно возник вопрос – насколько точно структуры, показанные благодаря технологии Hi-C, отражают действительную картину событий, происходящих в клетке. Прижизненная визуализация отдельных клеточных ядер убедительно показала, что фрагменты генома, входящие в один ТАД, сближены в пространстве ядра и контактируют заметно чаще по сравнению с локусами, принадлежащими разным ТАДам [46,156]. Интересно, что расположение доменов имеет достоверную корреляцию с состоянием хроматина и временными рамками репликации в конкретном участке ДНК [36,142,157–159], что позволяет предполагать прямое или косвенное влияние гистоновых модификаций и репликативной системы на формирование или/и поддержание ТАДов. В более поздних исследованиях были выделены структуры внутри ТАДов, имеющие средний размер порядка 200 т.п.н., называемые субдоменами, или хроматиновыми петлями [119,160], характерные для определенных стадий развития и этапов клеточной дифференцировки, обеспечивающие локальную организацию хроматина на уровне регуляции конкретного гена.

В работе Хаархиус и коллег авторы предполагают, что ТАДы отражают поликлональную популяцию хроматиновых петель, создавая физическую границу для динамичного протягивания петли через когезиновое кольцо между соседними сайтами CTCF [163]. В поддержку динамичности структуры доменов также свидетельствуют исследования единичных клеток различной спецификации у *Drosophila* с помощью микроскопии сверхразрешения/суперразрешения и single-cell Hi-C. А именно, было показано, что физические контакты между основаниями ТАДов, наблюдаемые на Hi-C картах в виде пиков или петель, в реальности встречаются чрезвычайно редко, независимо от размера домена, эпигенетического профиля участка или клеточного типа. Тогда как дальние взаимодействия, также являясь стохастическими, могут быть более частыми, благодаря схожему эпигенетическому статусу участка. Авторы пришли к выводу, что, хотя внутриТАДовые взаимодействия достаточно стохастичны и редки, эпигенетические метки по геному направляют и способствуют объединению топологических доменов в хроматиновые структуры более высокого порядка [164]. Таким образом, геном чрезвычайно пластичен и подвижен, и его пространственная организация основана на достаточно редких временных контактах хроматина, вероятность которых может модулироваться на нескольких уровнях.

Эволюция ТАДов неотделима от эволюции белков, ответственных за поддержание данных геномных структур. Исследования сайтов связывания различных транскрипционных факторов убедительно показали, что изменений только в нуклеотидной последовательности недостаточно, чтобы объяснить всю эволюционную динамику [149,165–168]. Вполне вероятно, что сайты посадки транскрипционных факторов и организация хроматина высокого порядка эволюционируют вместе, таким образом направляя ключевые изменения в регуляции генов. Интересная работа была проведена, где авторы попытались выявить дивергировавшие и консервативные черты пространственной организации для относительно отдаленных геномов четырех представителей класса млекопитающих [149]. Авторы показали, что консервативность пространственной организации присутствует в синтенных областях и коррелирует с консервативностью сайтов связывания CTCF. Кроме того, было отмечено, что сила межТАДовой границы находится в прямой зависимости от степени консервативности сайта связывания. При изменении ориентации сайта CTCF с высокой вероятностью происходит смена паттерна экспрессии гена, расположенного в том же домене, что может свидетельствовать в пользу косвенного участия архитектурных белков в регуляции генной экспрессии. Авторы наблюдали также, что в случае, когда в процессе эволюции возникают хромосомные перестройки, они преимущественно проходят по границам доменов, оставляя интактными сами доменные структуры. Таким образом, они делают вывод, что домены служат структурными единицами в эволюции, поддерживая целостность и селективно избегая нарушений при хромосомных перестройках [149].

В одной из недавних работ были выявлены убедительные свидетельства пересечений между отдельными консервативными некодирующими участками ДНК (CNEs – conserved non-coding elements) и целыми кластерами таких участков (GRBs – genomic regulatory blocks) со структурами ТАДов и их границ в экспериментах на человеческих клетках и плодовой мушки *Drosophila*. Поскольку консервативные некодирующие районы являются действительно очень древними структурами, это позволило авторам спекулировать о возможном возникновении ТАДов или же их предшественников сотни миллионов лет назад, как минимум у общего предка позвоночных и артропод [169].

1.4.6 Хроматиновые петли

Большинство хроматиновых петель на карте пространственных контактов у млекопитающих представлены точечными взаимодействиями в вершинах ТАДов, либо внутри них, обогащены метками активного хроматина и опосредуют ткане-специфичные

энхансер-промоторные взаимодействия [119,170,171]. Известны примеры, когда энхансеры располагаются на расстоянии нескольких млн.п.н. от целевых генов, но при этом вовлечены в жизненно важные регуляторные процессы [172]. Предполагается, что формирование данного типа петель может происходить по когезин-зависимому пути [119,170] или по механизму “CTCF enhancer-docking” [151,173], по которому, в частности, супер-энхансеры могут активировать онкогены в раковых клетках. Ауксин-индуцибельное снижение количества CTCF и когезина приводит к тотальному исчезновению данного типа взаимодействий за небольшим исключением некоторых контактов [174]. Кроме того, было отмечено возникновение новых точечных промотор-энхансерных контактов на коротких расстояниях, вероятно связанное с нарушением инсуляции в результате деплеции архитектурных белков.

Помимо CTCF, другой транскрипционный фактор - YY1, впервые описанный у млекопитающих в 1991 году [175–177], может участвовать в формировании регуляторных петель хроматина [178–180]. Кроме привлечения белков и белковых комплексов для активации или репрессии ДНК – в роли транскрипционного фактора, YY1 выполняет и структурную функцию, связывая гипометилированные регуляторные и промоторные области генов с образованием гомодимера. В процессе нейральной дифференцировки происходит массовое исчезновение белка CTCF совместно с утратой регуляторных взаимодействий генов, поддерживающих плюрипотентность. В свою очередь новые регуляторные связи обеспечиваются белком YY1, который совместно с фактором CTCF способствует образованию хроматиновых петель для регуляции специфичных для нейральных клеток энхансер-промоторных взаимодействий [178].

Примечателен вариант интенсивных цис- и транс-взаимодействий, наблюдаемых в генных кластерах ольфакторных рецепторов в постмитотических клетках обонятельных нейронов, наблюдаемых как на картах пространственных контактов, так в экспериментах по гибридизации *in situ* – FISH [1]. Известно, что в предшественниках обонятельных нейронов (базальных клетках) обнаруженные регуляторные взаимодействия практически отсутствуют, что свидетельствует об их участии в направленной спецификации клеточной функции.

Необычным примером дальних хроматиновых взаимодействий являются петли, обнаруженные в гемоцитобластах (гемопоэтических стволовых клетках) в обширных гипометилированных областях – “долинах” ДНК-метилирования [181]. Долинами ДНК-метилирования называют протяженные участки (3,5 – 25 Кб) довольно древнего происхождения с характерно низким паттерном метилирования по цитозину,

обогащенными высококонсервативными генами и регуляторами развития, а также метками H3K4me3 или H3K27me3 [182–184]. Внутри “долин” было обнаружено несколько сотен специфических хроматиновых взаимодействий, имеющих наивысший среди “долин” профиль H3K27me3, формирующих сети дальних контактов как внутри одной, так и между разными хромосомами. Основания петель могут быть разделены расстоянием в несколько млн.п.н, а наблюдаемые взаимодействия независимы от CTCF и когезина, но зависимы от уровня метилирования региона. Предполагается, что найденные образования хроматина, встречающиеся преимущественно в стволовых клетках, выполняют регуляторную функцию и вовлечены в поддержание мультипотентности [181].

1.4.7 Участие длинных некодирующих РНК в организации хроматина

Любопытный феномен был описан в работе Падекен и Хьюн [185] о том, что длинные некодирующие РНК способны действовать в качестве триггеров и вовлекаться в процесс перемещения неактивного хроматина к периферии ядрышка, подтверждая предположения о глобальном участии некодирующих РНК в организации хроматина.

Наиболее хорошо исследованный пример – это запуск инактивации X – хромосомы, который начинается с физического взаимодействия длинной некодирующей РНК Xist с целевой последовательностью на X-хромосоме, вследствие чего происходит активация и масштабный приток белковых комплексов ремоделинга хроматина, ответственных за репрессирование участка [186,187]. Подробное полногеномное исследование взаимодействий Xist выявило физические контакты с более чем 80 белками, большинство из которых являются субъединицами эпигенетических регуляторов, как например (SPEN)/SMRT (spen family transcriptional repressor) и SHARP (HDAC-Associated Repressor Protein). Таким образом, длинная некодирующая РНК Xist служит в качестве физического триггера организации нового хроматинового домена в ядре – репрессированного ядерного компартмента инактивированной X-хромосомы [123].

Другой интересный пример, когда длинные некодирующие РНК опосредуют образование функционального хроматинового домена – это РНК Firre (Functional intergenic repeating RNA element) [188,189]. Исходно обнаруженная как необходимый компонент при дифференцировке в адипоциты, Firre может формировать физические взаимодействия с несколькими районами вблизи генов, расположенными на разных хромосомах, благодаря привлечению hnRNPU, белкового компонента ядерного матрикса, образуя фокус локальной экспрессии в ядре или РНК-овый фокус. Вовлеченные гены обычно объединены временем транскрипции и сближены в пространстве ядра с локусом транскрипции Firre. Интересно,

что Firre локализуется в ядре на расстоянии 5 млн.п.н. от места своей транскрипции, где и образует РНК-овый фокус. Было обнаружено, что для образования взаимодействий с hnRNP и локализации транскриптов в правильном месте ядра необходима повторенная последовательность в 156 пн. Примечательно, что в экспериментах по удалению участка Firre из генома происходила диссоциация транс-хромосомного хроматинового комплекса, сформированного Firre, тогда как возобновление экспрессии путем ретровирусной интеграции приводило к восстановлению структуры РНК-ового фокуса [189]. Таким образом, Firre действует как платформа для организации транс-хромосомных взаимодействий. Данная некодирующая РНК расположена на X-хромосоме, способна избежать X-инактивации и экспрессироваться как до, так и после нее. После X-инактивации Firre опосредует закрепление X-хромосомы на периферии ядрышка, тем самым обеспечивая поддержание ее репрессированного статуса [190]. Уникальна и пространственная организация локуса Firre, который расположен на границе ТАДов и обогащен 15 сайтами посадки CTCF. В случае направленного удаления этих сайтов топологическая организация остается неизменной, что говорит о вовлечении дополнительных механизмов, поддерживающих эволюционно консервативное образование [191].

Биологические примеры, приведенные выше, иллюстрируют как цис-, так и транс-хромосомные взаимодействия, опосредованные длинными некодирующими молекулами РНК, которые приводят к формированию функциональных хроматиновых структур внутри ядра. Выполняя роль своеобразной приманки, некодирующие РНК направленно связываются с целевым фрагментом генома, привлекая в этот участок необходимый белковый аппарат, который, в свою очередь, запускает процессы ремоделинга, релокации или других структурных модификаций хроматина.

1.5 Молекулярные механизмы, обеспечивающие 3D-организацию генома

За последние несколько лет произошел глобальный скачок в понимании молекулярных механизмов, обеспечивающих поддержание пространственной структуры ядра у млекопитающих. Принятая научным сообществом гипотеза утверждает, что в интерфазном ядре действуют два главных независимых механизма, на которых держится и функционирует специфическая пространственная организация генома [170,192]. Первый – это когезин-независимый механизм фазовой сепарации, благодаря которому в ядре формируются активные и неактивные компартменты, тогда как второй – механизм, зависимый от присутствия когезинового комплекса белков, в процессе которого

происходит образование ГАДов и хроматиновых петель, что обеспечивает спецификацию функции и правильное взаимодействие промоторов генов и их энхансеров. В разделе ниже будут рассмотрены оба молекулярных механизма.

1.5.1 Механизм фазовой сепарации (phase separation)

Фундаментальный вопрос, который возникает относительно механики процесса функционирования клетки, каким образом пространственно разобщить множество молекулярных реакций, создав необходимое биохимическое окружение для каждой из них? Идеи разделения и объединения макромолекул в виде коацерватов были предложены еще на заре развития биологии в 1930е годы Александром Опариным. Вопрос, каким же образом происходит подобная сепарация молекул, навел исследователей на мысль о физико-химическом описании клетки и представлении цитоплазмы в виде водной субстанции, плотно заполненной жидкими коллоидными частицами. Именно эти частицы, объединяясь в группы и привлекая все новые компоненты со схожими свойствами, запускают процесс фазовой сепарации, создавая центры протекания биохимических реакций. Так, кажется понятным, каким образом мембранные клеточные органеллы выполняют эту функцию, но совершенно неясно, каким образом немембранные элементы, находящиеся в одной водной среде, не перемешиваются с помощью обычной диффузии и формируют функционально и пространственно отграниченные юниты. Недавние исследования предлагают рассматривать клеточные компартменты в виде жидких капель, элементы которых взаимодействуют с себе подобными с высокой аффинностью. Именно различие в силе таких взаимодействий и обуславливает возникновение фазовой сепарации. Проведя физико-химический анализ существующих клеточных компартментов (ядрышко, тельца Кахаля, сплайсосомы, стрессорные гранулы и другие) было показано, что практически все из них обладают свойствами жидкости (смешиваемость, текучесть, вязкость). В случае разделения мембранных органелл и водной цитоплазмы, причиной возникновения фазовой сепарации является гидрофобный эффект. Но если речь идет о немембранных белковых структурах, то здесь уже играют роль стереоспецифические взаимодействия гидрофильных белковых доменов. Так, в случае большого скопления и плотной упаковки белковых молекул может возникать состояние кристаллической структуры, в котором молекулярные элементы уже не способны перестроиться. Но жидкое состояние также остается возможным в том случае, если при высокой концентрации белка упаковка не настолько плотная, а между участниками могут возникать характерные взаимодействия [по обзору 193]. В *in vitro* экспериментах было показано, что слабых мультивалентных взаимодействий сигнальных белков может быть достаточно для запуска формирования и объединения

жидких капель, в которых концентрация данных белков в 100 раз выше, чем в остальном растворе [194].

На сегодняшний день предполагается, что хроматин в клеточном ядре подразделяется на компартменты и субкомпартменты согласно своим молекулярным свойствам (эпигенетические модификации) и биологической функции (функциональные ядерные образования, такие как ядрышко) именно по механизму фазовой сепарации, в самоорганизующейся манере и без затрат дополнительной энергии. Современные биофизические теории стараются объяснить, как водорастворимые белковые комплексы связываются и организуют транскрипционно активные и неактивные хроматиновые локусы в водной ядерной фракции. Ключевая идея фазовой сепарации состоит в разделении участков генома благодаря специфическим биофизическим характеристикам и свойствам, где подобное взаимодействует с подобным с высокой аффинностью. Так, были предложены два молекулярных механизма данного процесса. Первый – фазовая сепарация по типу полимер-полимер, где происходит образование упорядоченной глобулы в результате формирования перемычек между отдельными нуклеосомами участков хроматина со схожими свойствами, индуцируемое структурными белками. Такие взаимодействия могут преобразовывать нуклеосомную цепь и связанные с ней транскрипционные факторы в упорядоченную глобулярную структуру, образующую ядерный субкомпартмент. При этом стоит отметить, что образование функциональных связей между белками не является обязательным условием. В основе второго предполагаемого механизма лежит представление о ядерном субкомпартменте и нуклеоплазме как о жидкостях, формирующих разные фазы. Стабилизация данных фаз происходит за счет мультивалентных взаимодействий растворимых белковых молекул, связанных с хроматином. Интересно, что подобного рода структуры могут существовать в ядре даже в отсутствие хроматинового кора. Оба описанных механизма подразумевают существование ядерных субкомпартментов в виде динамичных сферических образований, которые могут сливаться и разделяться, а также находятся в постоянных биохимических взаимодействиях с окружающей нуклеоплазмой [195].

За последние пару лет были опубликованы работы, в которых утверждается, что именно гетерохроматин является триггером процесса фазовой сепарации [134]. Интересным остается вопрос, как в условиях постоянно меняющихся биохимических состояний и динамично функционирующего ядра удастся поддерживать разделение участков генома. В работе Ларсон и коллег утверждается, что фосфорилированные молекулы белка HP1 α (структурный белок, преимущественно ассоциированный с

областями «молчащего» гетерохроматина), а также связанные с ДНК молекулы способны формировать отдельные «капли», которые привлекают и аккумулируют другие HP1 α -молекулы, образуя ядерные субкомпарменты. Авторы показали, что как фосфорилированные, так и нефосфорилированные формы HP1 α ускоряют компактизацию цепей ДНК, даже обладающих различными физико-химическими характеристиками. Таким образом, они предположили изоляцию и формирование HP1 α -субкомпарментов в качестве основного молекулярного механизма, по которому происходит замолкание генов в гетерохроматиновых локусах [196]. В другом исследовании авторы в экспериментах *in vitro* подтвердили жидкостные свойства молекул белка HP1 α , наблюдая их чувствительность к разрушению слабых гидрофобных связей и уменьшение диффузных процессов между разными фазами, а также показали разделение гетерохроматиновых локусов по механизму жидкостно-жидкостной фазовой сепарации с образованием отдельных компарментов внутри ядра [197].

В работе Фалк и коллег авторы, развивая предложенные идеи, утверждают, что гетерохроматин является ключевым элементом процесса фазовой сепарации [134]. На основании построенной ими полимерной модели хромосом и результатов компьютерного моделирования они убедительно показывают, что для фазовой сепарации необходимо и достаточно только гетерохроматиновых взаимодействий, чтобы обеспечить компарментализацию эу- и гетерохроматина как в обычном, так и инвертированном ядрах. Согласно их модели, состояние инвертированного ядра является нативным, тогда как обычная организация с гетерохроматином на периферии и эухроматином внутри – результатом добавления контактов гетерохроматиновых локусов с ламинной. Данная гипотеза действительно кажется логичной и простой, но требует экспериментального подтверждения.

1.5.2 Механизм *loop extrusion* (выпетливание хроматина)

Идеи о существовании молекулярного механизма, по которому два отдаленных элемента на хромосоме оказываются сближены в результате протягивания хроматиновой петли через белковый комплекс возникли довольно давно [171,198–200]. Наконец, было установлено, что белок CTCF связывается с молекулой ДНК за счет специфических сайтов узнавания и, находясь в конвергентной ориентации, может физически блокировать скольжение комплекса когезина по нити хроматина, создавая тем самым границы ТАДов [147]. На сегодняшний день утверждается, что у позвоночных именно когезиновый комплекс, состоящий из субъединиц SMC1, SMC3, и RAD21 (или SCC1), формирует

белковое кольцо, к которому крепятся обе молекулы ДНК [201–204], и интенсивно протягивает нить хроматина через петлю до тех пор, пока не встречает два конвергентно направленных сайта CTCF, которые формируют физический барьер и маркируют границы хроматиновых доменов и петель [148,171,200]. NIPBL-опосредованная загрузка когезина на ДНКовую молекулу и его высвобождение, обеспечиваемое релиз-фактором WAPL, определяет динамичность ТАДовых структур [163,192,205]. Тот факт, что протягивание петель – динамический процесс, а границы хроматиновых доменов во всех клетках маркированы одинаково, хорошо согласуется с данными микроскопии для единичных клеток, где элементы одного ТАДа не всегда расположены физически близко в пространстве ядра [47,156], а также с данными single cell Hi-C, где структура ТАДов является отражением наиболее распространённого варианта контактов для клеточной популяции [92,93]. Стоит отметить, что, хотя механизм протягивания хроматина через когезиновое кольцо функционирует именно для конвергентных сайтов CTCF в основании петель, не любой сайт CTCF способен служить границей ТАДа согласно “правилу конвергентности CTCF” [147–149], поскольку встречается много одиночных, в том числе конвергентно направленных сайтов, расположенных внутри ТАДов, не участвующих в формировании доменов [144,148,206].

Эксперименты, где на клеточных культурах позвоночных получали нокауты по белкам WAPL и NIPBL, служат ярким подтверждением механизма “loop extrusion”. Подавление активности белка WAPL, который отвечает за снятие когезина с молекулы ДНК путем открывания когезинового кольца [207–209], приводило к увеличению общего количества и времени нахождения когезина на ДНК, удлинению хроматиновых петель и снижению их общего числа, радикальному ослаблению компартментных взаимодействий, а также к глобальной компактизации хроматина по всему геному [205,207,210]. Результаты микроскопии сверхразрешения свидетельствуют об увеличении пространственных контактов между доменами при элиминации WAPL [211]. Индуцибельный нокаут по гену *Nipbl* приводил к масштабному исчезновению ТАДов по геному, более тонкому и выраженному разделению на активные и неактивные компартменты с сохранением общего паттерна у мутантных клеток, а активность компартмента хорошо коррелировала с эпигенетическим состоянием хроматина. Интересное наблюдение было сделано относительно транскрипционных изменений вследствие NIPBL-нокаута: для генов, расположенных в более крупных ТАДах, а также окруженных протяженными межгенными участками, наблюдалась тенденция к снижению экспрессии. Примечательно, что гистоновые метки активных промоторов сменились согласно изменениям в

транскрипционном статусе, тогда как метки активных энхансеров для большинства локусов остались интактными. Кроме того, возникла эктопическая транскрипция некодирующих молекул РНК, преимущественно расположенных в экзогенных областях, которая в норме не детектируется, а также участились случаи двунаправленной транскрипции. Авторы предполагают, что в нокаутных по NIPBL клетках, в связи со снижением дальних пространственных взаимодействий в отсутствие ТАДов, активные энхансеры, сохранившие соответствующие гистоновые метки, не всегда могут добраться до целевых промоторов и начинают активировать гены, расположенные физически ближе [192].

Сходный эффект наблюдается для клеток со снижением активности элементов когезинового комплекса (RAD21, SMC1, SMC3), когда элиминируются ТАДы, но сохраняются компартменты и гистоновые модификации. Любопытно, что восстановление экспрессии белков приводит к восстановлению первоначальной структуры [170]. В некоторых случаях меняется экспрессия генов, в особенности генов, расположенных ближе к граничным участкам ТАДов [170,211], а суперэнхансеры, утрачивая дальние контакты, начинают взаимодействовать с близлежащими промоторами генов, опосредуя их активацию [170,205]. Полученные факты свидетельствуют о центральной роли комплекса когезина в локальной компактизации хромосом путем постоянного образования динамичных хроматиновых петель, формирующихся между граничными элементами и визуализирующимся на Hi-C-карте в виде отдельных ТАДов и петель.

Наконец, было продемонстрировано прямое доказательство существования механизма “loop extrusion” в *in vitro* экспериментах на культуре дрожжей [212]. С помощью микроскопии высокого разрешения авторы во временной динамике визуализировали конденсиновый комплекс, протаскивающий молекулу ДНК и формирующий хроматиновую петлю. Авторы показали, что данный процесс происходит по одностороннему механизму, когда одна цепь ДНК фиксирована, а вторая протягивается через кольцо АТФ-зависимым мотором/мотором конденсинового комплекса.

Биоинформационные симуляции механизма “loop extrusion” очень хорошо коррелируют с наблюдаемыми сигналами ChIP-seq и распределением белков-участников в живой клетке. Так, при одностороннем протягивании хроматина наблюдается повышенная загрузка когезина на одного из оснований ТАДа по сравнению с другим и на карте пространственных взаимодействий наблюдается характерная полоса. В реальности, в случае таких архитектурных полос или страйпов, в основании домена лежит видимый пик субъединиц когезина и повышенный пик белка NIPBL [213].

Интересная существующая гипотеза предполагает влияние транскрипционного статуса района и присутствия РНК-полимеразы на хроматине на возможность его связывания с молекулой когезина, таким образом участвуя в формировании петель [214]. Серия элегантных экспериментов, выполненных Бусслингер и коллегами убедительно показала, что транскрипционная активность участка генома может определять сайты связывания когезина с хроматином в отсутствие CTCF и WAPL [215]. Что служит серьезным аргументом в поддержку молекулярной модели, где РНК-полимераза является движущей силой процесса выпетливания хроматина, опосредованного когезиновым комплексом.

Стоит, однако, отметить, что вопрос о том, каким именно образом происходит протягивание нити хроматина, нагруженной белковыми молекулами, через когезиновое кольцо фиксированного размера (35 нм), остается открытым. Данные о скольжении комплекса когезина по хроматину неоднозначны: с одной стороны, показано, что рибосомы, РНК-полимераза с транскриптами, нуклеосомы, реплисомы – могут свободно проходить сквозь хроматиновое кольцо [202,203]; с другой стороны, есть серия работ, где были описаны физические препятствия в виде отдельных белковых молекул, блокирующие или замедляющие скольжение когезинового комплекса [215–218]. В исследовании Прадан и коллег, опубликованном на платформе BioRxiv, авторы наблюдают за образованием хроматиновой петли в условиях *in vitro* и фиксируют возможность прохождения белковых частиц крупных размеров (~200 нм) через когезиновый комплекс, что невозможно по общепринятому топологичному механизму, когда нить хроматина протягивается сквозь когезиновое кольцо [204]. Они исключают вариант размыкания кольца перед препятствием, и утверждают, что выпетливание хроматина с крупными белковыми молекулами действительно происходит, но по альтернативному, не топологичному, механизму.

1.5.3 Пространственная организация митотических хромосом

В свете открытия молекулярных механизмов, опосредующих пространственную организацию генома на стадии интерфазы, любопытно посмотреть, что же происходит со структурами хроматина в процессе митоза. Согласно последним исследованиям, предполагается, что выпетливание хроматина, опосредованного когезиновым комплексом, является универсальным механизмом, участвующим в формировании хроматиновых доменов на стадии интерфазы и обеспечивающим компактизацию и сегрегацию хромосом в метафазе [173,200,219,220]. Исследователи изучили пространственные взаимодействия в динамике конденсации хроматина до митотических хромосом на клетках HeLa S3.

Согласно картам пространственных контактов, в митозе происходит полная утрата ТАДов и компартментов, на стадии прометафазы наблюдается возникновение второй диагонали, отражающей глобальное усиление взаимодействий на расстоянии ~ 3 млн.п.н., при этом в момент формирования максимально компактизованных метафазных хромосом вторая диагональ смещается на расстояние 10 млн.п.н.. Данное наблюдение прекрасно коррелирует с наблюдениями микроскопии, теоретическими расчетами и результатами биофизического моделирования [63,219,221]. Также формирование второй диагонали и практически полное отсутствие ТАДов и хроматиновых петель, напоминающее митотические хромосомы, характерно для ядер эритроцитов позвоночных на терминальной стадии дифференцировки, когда происходит глобальная реорганизация хроматина, вероятно, в связи со спецификацией функции и необходимостью более компактной укладки хроматина [222].

1.6 Особенности организация хроматина у насекомых

Подтверждение универсальности и консервативности основных принципов пространственной организации хроматина в ядре стало значительным научным событием последних лет [144,154]. Сравнение пространственной укладки хроматина у насекомых и млекопитающих позволяет говорить об общности фундаментальных принципов, что включает наличие хромосомных территорий, активных и неактивных компартментов, доменов (или ТАДов) и хроматиновых петель. Основные представления об организации хроматина у насекомых получены на основе данных по *Drosophila* [154,223,224]. ТАДы, опосредованные белком CTCF, являются ключевым компонентом в системе пространственной организации у позвоночных, тогда как у беспозвоночных и у растений участие CTCF и его гомологов в образовании подобных хроматиновых структур является предметом активных дискуссий [155,224]. Большинство известных принципов пространственной организации было подтверждено и описано на клетках и тканях мыши и человека, тогда как у насекомых и других живых организмов был показан феномен, но не выяснены механизмы. Предполагается, что домены, наблюдаемые у данных организмов, отражают транскрипционное состояние отдельных локусов генома (Рисунок 3). Граничные участки формируются между активными и неактивными районами, что выражается в характерных гистоновых модификациях и связанных с такими районами комплексами белков. Наблюдаемая картина разделения замечательно совпадает с той, что выявляется при элиминации белков CTCF и RAD21 у млекопитающих, а именно - наблюдается исчезновение ТАДов с сохранением структур компартментов [150,170,173], а выполненные

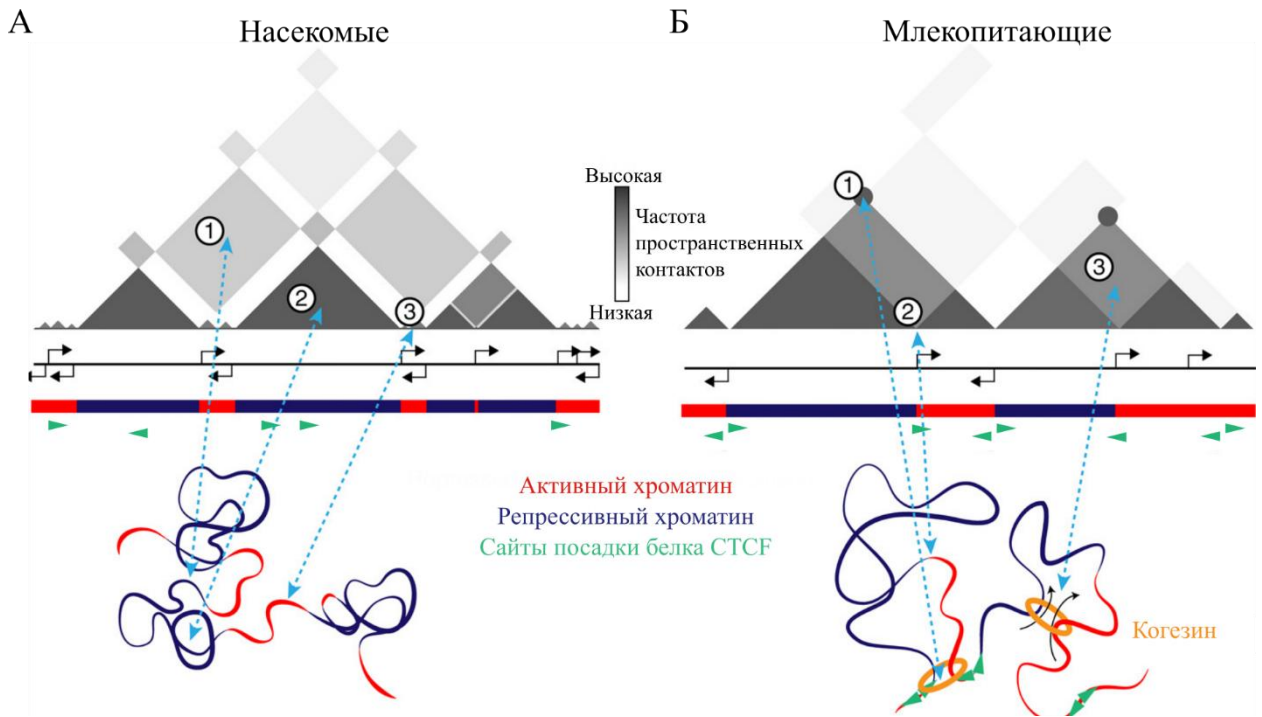


Рисунок 3. Сравнение пространственной организации у насекомых (панель А) и млекопитающих (панель Б). Иллюстрация адаптирована из [225].

на клетках *Drosophila* Hi-C-эксперименты с высоким покрытием подтверждают точное разделение доменов между активными и неактивными компартментами [224]. Компьютерная симуляция сепарации хроматина, основанная на данных о транскрипционной активности локуса, достоверно воспроизводит наблюдаемую *in vivo* картину (по Rowley et al. 2019). В данном разделе будут подробно рассмотрены характерные особенности организации хроматина у насекомых в сравнении с млекопитающими.

1.6.1 Основные белки и белковые комплексы, участвующие в поддержании архитектуры ядра у насекомых

У представителей насекомых, в частности у *Drosophila melanogaster*, белки, участвующие в поддержании архитектуры ядра, можно подразделить на две группы: находящиеся в составе первой способны напрямую взаимодействовать с ДНК при помощи ДНК-связывающих доменов, а представители второй группы, связываясь с белками из первой, формируют комплексные структуры, участвующие в образовании целевых взаимодействий. Таким образом, к первой группе относятся: dCTCF (дрозофилиный ортолог CTCF), Su(Hw) (Suppressor of Hairy-wing), BEAF-32 (Boundary Element Associated Factor-32), GAF (GAGA factor), Zelda, DREF, TFIIC, Z4, Elba, ZIPIC, Zw5 (Zest white 5), Ibf1 и Ibf2 (Insulator binding factors 1 и 2), тогда как вспомогательными являются белки CP190, Mod(mdg4), RAD21 (компонент когезинового комплекса), CAP-H2 (компонент

конденсинового комплекса), Fs(1) h-L, L3mbt и Chromator [226,227]. Несколько лет назад были описаны еще два белка – ZIPIC и Pita, локализация которых часто перекрывается хотя бы с одним из CP190, BEAF-32 или dCTCF, но совершенно не пересекается с Su(Hw) [228]. Данное наблюдение только подкрепляет гипотезу, выдвинутую в работе Ван Боргл, что в зависимости от напарника может меняться выполняемая функция белка, а также о кооперативности инсуляторных факторов для выполнения различных функций [229]. Кроме того, было показано, что сайты посадки ZIPIC и Pita преимущественно совпадают с промоторными областями.

Ключевые белки, обеспечивающие доменную организацию у млекопитающих (CTCF, SMC1, SMC3, RAD21, CAP-H2, CAP-D3 и некоторые другие) или их гомологи (функциональные аналоги) обнаружены и полноценно функционируют у насекомых. В пределах границ доменов были отмечены такие транскрипционные факторы, как BEAF-32, Su(Hw), GAF, Zw5, Ibf1 и Ibf2, которые притягивают вспомогательные белки, например, CP190 и Mod(mdg4), что делает их вероятными кандидатами на выполнение функции CTCF и когезинового комплекса у насекомых [226,227,230,231]. В одной из работ предполагается, что комбинации BEAF-32/CP190 и BEAF-32/Chromator могут являться функциональными аналогами комплекса CTCF/когезин на границах доменов у млекопитающих [231]. В то же время, биоинформационный анализ генов-ортологов и их белковых доменов показал, что, например, BEAF-32 специфичен именно для дрозофилид [232].

Несмотря на то, что в пионерских исследованиях пространственной организации генома была показана консервативность архитектуры хроматина у всех эукариот, а также позднее подтвержденная интенсивная кластеризация архитектурных белков на границах доменов у *Drosophila melanogaster* [153,154,233], участие CP190, Beaf-32, dCTCF и Chromator в формировании ТАДов в клетках насекомых, как и существование самих ТАДов, до сих пор представляет открытый вопрос.

1.6.2 Доменные структуры – ТАДы или компартментные домены?

Интенсивное развитие методов захвата конформации хромосом и получение более детальных карт пространственных контактов привело к выявлению множества субдоменов внутри регионов активного хроматина, которые ранее выступали в анализе как отдельные ТАДы или междоменные участки [158,226,231]. На сегодняшний день существует несколько точек зрения касательно вопроса о доменных структурах в геноме насекомых. В работе Роули и соавторов утверждается, что домены, наблюдаемые у дрозофилы, являются компартментными доменами и при высоком разрешении Hi-C-карт полностью совпадают с

А/В-компартаментами [224]. Визуально на Hi-C-картах с высоким разрешением субдомены млекопитающих выглядят как треугольники с ярко выраженными вершинами. У *Drosophila* доменные структуры представлены сходным образом, но при увеличении разрешения Hi-C-карты наблюдаемые пики оказываются небольшими компартаментами, демонстрируя взаимодействия не только внутри своего домена, но и с подобными структурами по геному (Рисунок 3). Примечательно, что отличием ТАДов насекомых от ТАДов млекопитающих является отсутствие угловых пиков усиления взаимодействий в участке заякоривания ТАДа, которое наблюдается у последних [119,231,234]. Сходные с петлями млекопитающих образования были обнаружены у дрозофилы (~500 петель), но, во-первых, они не совпадали с границами доменов, как правило, попадая внутрь доменной структуры, а, во-вторых, их границы характеризовались обогащением по сайтам посадки различных архитектурных белков, а не преимущественно CTCF [226].

В другом исследовании авторы придерживаются прямо противоположной точки зрения, согласно которой ТАДы являются консервативными элементами 3D-организации хроматина среди всех эукариот [231]. Проведя *in situ* Hi-C и получив 353 млн чистых прочтений и максимальное разрешение в 200 пн (уровень нуклеосом), им удалось выделить более 4000 ТАДов, что почти в 7 раз превышает количество, описанное в ранних работах [154,158]. Авторы утверждают, что крупные суперТАДы, которые удается визуализировать на 20 Кб-разрешении карт контактов, разбиваются на более мелкие ТАДы при 1 Кб-разрешении, но межТАДовые районы остаются без изменений. Распределение по геному ключевых архитектурных белков и 15 гистоновых модификаций позволило авторам разделить ТАДы на 4 основных типа, преимущественно связанное с делением хроматина на активный и неактивный [231]. В других работах на *Drosophila* было показано, что ТАДы соответствуют разным эпигенетическим состояниям хроматина. Так, были описаны “нулевые” домены, которые не ассоциированы с какими-либо эпигенетическими характеристиками; активные домены - районы активного хроматина; поликомбные домены, представленные локусами, обогащенными по эпигенетическим меткам белков группы поликомб; HP1-домены - участки классического и прицентромерного гетерохроматина [154,231].

Наличие сайтов посадки белка dCTCF в границах доменов у *Drosophila melanogaster* позволяет предположить, что механизм выпетливания хроматина, результатом которого у млекопитающих является образование ТАДов, может существовать и у дрозофилид. Однако, РНК-интерференция против dCTCF у *Drosophila* не привела к полному исчезновению CAP-D3-субъединицы конденсинового комплекса с мест посадки на ДНК,

что может свидетельствовать о дополнительных механизмах или вспомогательных участниках данного взаимодействия [233]. Более детальное изучение принципов пространственной укладки хроматина на нескольких клеточных линиях *Drosophila melanogaster* (S2, Kc167, DmBG3-c2 и OSC) показало, что границы доменов в первую очередь определяются активной транскрипцией. Авторы обнаружили, что границы доменов характеризуются высоким уровнем транскрипции, обогащенностью генами домашнего хозяйства и эпигенетическими метками активного хроматина. При этом выраженного обогащения белками dCTCF и Su(Hw) в границах доменов не наблюдалось. Ещё одним аргументом в пользу того, что белок dCTCF не играет существенной роли в процессах формирования доменов у насекомых, является тот факт, что биоинформатические предсказания границ доменов на основании распределения dCTCF не позволяют достоверно определять положение границ [158]. В то же время информации о расположении меток активного хроматина и/или активности экспрессии является достаточным. В пользу данного утверждения свидетельствует тот факт, что регионы с низким содержанием эпигенетической метки репрессированного хроматина H3K27me3 оказываются существенно обогащены инсуляторными белками Beaf-32 и Chromator. В свою очередь, белки dCTCF и Su(Hw) не имеют какого-либо предпочтения к связыванию с хроматином в активном или неактивном состояниях [230]. Кроме того, анализ взаимной ориентации соседних сайтов посадки dCTCF в геноме *Drosophila melanogaster* не выявил конвергентную закономерность формирования ТАДов, наблюдаемую у млекопитающих [148,149]. Что может говорить об иной природе образования доменов у насекомых, обусловленной не только взаимодействиями белка CTCF.

Исследования укладки хроматина у дрозофилид в нейрональной клеточной линии BG3 показало довольно существенное расхождение в структуре ТАДов по сравнению с эмбриональной линией Kc167. По мнению авторов, это связано с серьезными изменениями транскрипционного статуса в результате дифференцировки, а также специфическими энхансер-промоторными взаимодействиями, уникальными для каждой из линий. Авторы утверждают, что двунаправленная транскрипция, а также целевые взаимодействия энхансеров с промоторами могут играть принципиальную роль в формировании ТАДов у насекомых [235].

В ходе недавних экспериментов с полным нокаутом по dCTCF удалось выяснить, что этот белок имеет критическое значение для регуляции экспрессии определенной группы генов, вовлеченной в нейрогенез, и организации хроматина в соответствующих локусах, но не выполняет определяющую роль в формировании доменов по всему геному.

На глобальном уровне пространственной организации потеря dCTCF не оказала влияние на 84% всех детектированных доменных границ [236]. Следует отметить, что в данном конкретном случае только связывания dCTCF оказалось недостаточно для обеспечения правильной экспрессии генов или формирования пространственной организации: свои регуляторную и архитектурную функции dCTCF выполняет за счёт привлечения белка CP190 и связывания с ним.

Можно заключить, что пространственная организации хроматина у насекомых на примере *Drosophila* обеспечивается преимущественно делением хроматина на активный и неактивный в соответствии с эпигенетическим статусом геномного региона. Наблюдаемое обогащение доменной границы архитектурными белками задает силу данной границы, где более нагруженные белками области создают более мощную инсуляцию домена, а менее нагруженные - ответственны за ген-специфичную регуляцию [233]. На более тонком уровне разграничение на домены задаётся связыванием CP190 и Chromator с белками, способными узнавать специфические последовательности ДНК, такими как Beaf-32 и dCTCF. Таким образом, домены у насекомых, компартментные или топологические, представляют динамичные в развитии организма структуры и отражают функциональную взаимосвязь между энхансерами, инсуляторными белками и транскрипционной активностью конкретного участка [235].

1.6.3 Хроматиновые петли насекомых

Согласно последним работам, выполненным на насекомых, хроматиновые петли неизвестной природы были отмечены у дрозофилы [158,237]. В частности, Hi-C высокого разрешения позволило выявить несколько сотен подобных петель в дрозофилином геноме, не связанным по локализации с белком CTCF или его аналогами [214]. У дрозофилы было показано, что участки заякоривания (анкоры) специфических хроматиновых петель (point-to-point дальние взаимодействия) обогащены белками комплекса Polycomb, Rad21, а также энхансерами, участвующими в развитии, тогда как белок CTCF, опосредующий образование петель у млекопитающих, не наблюдается в этих локусах [224,226,237,238]. Природа образования таких петель и поддерживающий их молекулярный механизм пока остается невыясненными.

Примечательно недавнее эволюционное исследование, показавшее отсутствие нескольких элементов комплекса конденсина II у некоторых групп насекомых, в частности CAP-G2 у представителей *Diptera*, в состав которого входит и род комаров *Anopheles* [239]. Вместе с отсутствием CAP-G2 у них наблюдается синапс гомологичных хромосом в

соматических клетках на протяжении клеточного цикла, что является уникальной чертой данной группы насекомых. Интересный вопрос заключается в том, за счет чего происходит образование хроматиновых петель у тех видов живых организмов (в том числе насекомых), которым характерно отсутствие элементов конденсинового комплекса, и возможен ли в этом случае механизм выпетливания хроматина, опосредованный когезином.

Еще один тип дальних взаимодействий был показан в эмбриональном развитии дрозофилы - дальние хроматиновые петли временного характера, которые возникали в промежутке 13-14 цикла делений и исчезали сразу после, стадиоспецифично контролируя генную экспрессию [238]. Взаимодействующие хроматиновые локусы оказались обогащены эпигенетическими метками активной транскрипции, а также транскрипционным фактором Zelda и РНК-полимеразой II. Некоторые типы дальних петель, более ранние, оказались чувствительны к элиминации Zelda, тогда как более поздние - нет, из чего авторы предположили существование двух типов хроматиновых взаимодействий, зависимых и независимых от транскрипционного фактора Zelda. Из другой работы известно, что на ранних стадиях развития Zelda формирует особые микрокомпарменты в ядре или компартментные хабы для пространственной инсуляции целевых генов [240].

1.6.4 Домены в цис- и транс-взаимодействиях. Явление трансвекции в регуляции пространственных взаимодействий.

Процесс трансвекции, обеспечивающий регуляторные взаимодействия участков гомологичных хромосом, является уникальным геномным феноменом и был впервые описан для регуляции/контроля генов развития *bithorax* [241], *white* [242], *decapentaplegic* [243], *yellow* [244] у насекомых. Регуляторные взаимодействия подобного рода становятся физически возможны благодаря плотному синапсису гомологичных хромосом в интерфазных ядрах соматических клеток, характерному для двукрылых. В то же время у млекопитающих, хромосомы которых на стадии интерфазы организованы в хромосомные территории и разобщены в пространстве ядра, процесс соматического хромосомного спаривания также был описан, например, при некоторых патологических состояниях [245]. В клетках других организмов феномен трансвекции тоже встречается, но значительно реже и требует предварительного спаривания гомологичных хромосом [246,247]. У млекопитающих было описано всего несколько случаев трансвекции [248], зато известны примеры регуляторных взаимодействий, сопряженных с образованием хроматиновых

контактов между элементами разных хромосом [249–252], в том числе импринтированных локусов. Зачастую экспрессия таких генов является моноаллельной [252].

Предполагается, что трансвекция может участвовать и в пространственной организации хроматиновых структур, необходимой, в частности, для совместной регуляции генов и межгенной комплементации [253,254]. Так, была предложена модель, по которой инсуляторные элементы, расположенные на гомологичных хромосомах, инициируют образование транс-топологического хроматинового домена, схожего по структуре с доменами в цис-положении, что способствует контактам общего энхансера со своими промоторами [254–257]. Предполагается, что формирование домена может направляться взаимодействием комплекса PolII, расположенном на промоторных областях, с транскрипционными факторами, сидящими на энхансерах [258]. В результате наблюдаемых пространственных взаимодействий в домен привлекается транскрипционный белковый комплекс и обеспечивается совместная транскрипция. Стоит подчеркнуть, что трансвекция носит не тотальный, а клеточноспецифичный характер, где тонкая настройка осуществляется цис-регуляторными последовательностями [259].

Оказалось, что архитектурный белковый комплекс конденсин II является антагонистом хромосомного синапсиса/спаривания, и повышенная экспрессия CapH2, одной из субъединиц комплекса, приводит к дизагрегации политенных хромосом и ингибированию трансвекции [260]. Современная гипотеза утверждает, что исходное/нативное состояние хроматина – это как раз спаривание гомологичных хромосом, а регуляция подвижного баланса между синапсисом и асинапсисом хромосом осуществляется на уровнях разного масштаба и обеспечивает необходимое пространственное окружение [253]. Утрата субъединиц конденсинового комплекса в некоторых таксонах напрямую коррелирует с переходом от хромосомных территорий к Rab1-конфигурации, вероятно, вследствие снятия блока на хромосомный синапсис [22].

*1.6.5 Пересечение дискового/междискового паттерна политенных хромосом со структурой ТАДов/интер-ТАДов у *Drosophila melanogaster**

Политенные хромосомы представляют удобную биологическую модель для изучения интерфазного ядра. На этой стадии становятся визуально детектируемы многие геномные структуры - области активной транскрипции, представленные в виде объемных пуфов, инверсионные петли, участки интеркалярного и прицентромерного гетерохроматина, а также диски и междиски. Последние две структуры отражают степень конденсации хроматина, которая варьирует от 1:3 - 1:15 для междисков до 1:200 для

плотных гетерохроматиновых дисков [55]. Оказалось, дисковый рисунок идентичен в разных тканях, а также стабилен и уникален среди особей одного вида. Данный факт способствовал возникновению гипотезы о том, что политенные хромосомы отражают универсальную пространственную организацию хроматина у двукрылых в увеличенном масштабе. Способа проверить данную гипотезу до последнего времени не существовало по причине отсутствия точного соответствия между геномной картой и картой морфологических структур политенных хромосом. Результаты Hi-C-эксперимента, выполненного на политенных хромосомах слюнных желез [234,237], свидетельствуют в пользу данной гипотезы, демонстрируя пересечение 95% блоков интеркалллярного гетерохроматина с границами ТАДов у *Drosophila melanogaster*. Сравнительные исследования доменной организации между интерфазными диплоидными и политенными ядрами у *Drosophila melanogaster* убедительно показали полное совпадение ТАДового рисунка [234]. Кроме того, архитектурные белки, предположительно участвующие в образовании границ ТАДов (перечислить), были обнаружены в междисковых районах политенных хромосом. Гибридизация флуоресцентных проб на границы нескольких ТАДов также подтвердила их соответствие диск/междисковому рисунку политенных хромосом [234,237,261].

1.7 Анофелесы или малярийные комары – уникальный объект исследования пространственной организации генома

1.7.1 Классификация и филогения

Комары рода *Anopheles* представляют достаточно уникальный биологический объект исследования, поскольку виды, входящие в состав рода, при крайне высоком фенотипическом сходстве эволюционно разделены очень большим расстоянием. Впервые род малярийных комаров *Anopheles* был описан в 1818 году Иоганном Вильгельмом Майгеном, увлеченным немецким исследователем двукрылых насекомых (Рисунок 4, А). Первичная классификация, достаточно сложная, была основана на характерном жилковании крыльев и структуре усиков у самцов и позволяла разделить разных представителей *Diptera* между собой, в том числе комаров рода *Anopheles* (Рисунок 4, Б). Позднее, по поручению Лондонского Королевского общества в связи с распространением малярии в регионах Британских колоний, Фредериком Теобальдом была проведена полноценная работа по описанию малярийных комаров в период с 1899 по 1910 годы. Классификация была основана на фенотипических отличиях чешуек на груди и брюшке, разделяла всех известных малярийных комаров на 18 родов. Позднее 4 из них

переквалифицировали в подроды рода *Anopheles* (*Cellia*, *Kerteszia*, *Nyssorhynchus* и *Stethomyia*), а оставшиеся 14 оказались синонимичны между собой и принадлежали одному из подродов - *Anopheles*, *Cellia* или *Nyssorhynchus*. Нахождение новых представителей рода убедительно свидетельствовало о несовершенстве предложенной Теобальдом системы. Современная фенотипическая внутривидовая классификация малярийных комаров основана на количестве и положении щетинок на гонококситах самцовых гениталий, была введена сэром Ричардом Кристоферсом в 1915 году и на сегодняшний день включает 7 подродов - *Anopheles*, *Cellia*, *Kerteszia*, *Lophopodomyia*, *Nyssorhynchus*, *Stethomyia* и *Baimaia*, последний из которых был включен в классификацию Харбахом в 2005 году (согласно обзору [262]). Появление цитогенетических и молекулярных подходов позволило успешно классифицировать существующие виды на основании характерного рисунка/паттерна политенных хромосом [263], а также геномного анализа молекулярных маркеров, таких как митохондриальные гены цитохромоксидаз I и II (COI, COII) и рибосомальные локусы (18S, 28S, ITS1, ITS2) [264–266].

Предполагается, что малярийные комары эволюционно отделились в Юрском периоде (145 – 200 млн лет назад). Молекулярные оценки свидетельствуют, что раскол между двумя подсемействами *Culicinae* и *Anophelinae* произошел 197,5 миллионов лет назад, в раннюю юрскую эпоху, но серьезная диверсификация случилась только в меловом периоде [267]. Наиболее древние ископаемые находки представителей рода *Anopheles* были обнаружены в бирманском янтаре из Мьянмы и датируются ранним сеноманским этапом позднего мелового периода: *Priscoculex burmanicus*, 110-100 млн лет назад [268] и *Burmaculex antiquus*, 100-90 млн лет назад [269]. Ископаемый вид *Paleoculicis minutus* был найден в канадском янтаре из Альберты и отнесен к кампанскому этапу позднего мелового периода (76-79 млн лет назад [270]). Более современные образцы древних анофелесов были обнаружены в Доминиканском янтаре и предположительно относятся к середине третичного периода (15-45 млн лет назад, *Anopheles dominicanus* [271]), а также на территории Германии - представители позднего олигоцена (23-28 млн лет назад, *Anopheles rottensis*, [272]) (Рисунок 4, В). Тем не менее, так как сохранность образцов ископаемых насекомых не всегда приемлема для визуальной идентификации, филогенетические оценки расхождения видов внутри рода основываются на методе молекулярных часов, который фиксирует скорость накопления нуклеотидных замен в белок-кодирующих районах [273].

Виды, выбранные в диссертационное исследование, разделены эволюционным расстоянием в 0.5-100 млн лет (Рисунок 4, Г). Так, наиболее близкими являются *Anopheles coluzzii* и *Anopheles merus* с дистанцией в ~0.5 млн лет. Оба вида относятся к комплексу

Anopheles gambiae, могут скрещиваться с образованием частично фертильного потомства [274,275]. Далее, в 35-40 млн лет от них на филогенетическом дереве расположен *Anopheles stephensi*, в 75-80 млн лет *Anopheles atroparvus*, и, наконец, разошедшийся с комплексом *Gambiae* около 100 млн лет назад наиболее отдаленный вид – *Anopheles albimanus*.

1.7.2 Распространение

География представителей рода *Anopheles* охватывает все материки планеты, за исключением Антарктиды (Рисунок 4, Д). История распространения и заселения данным родом доступных мест обитания до сих пор является объектом дискуссий, но основные пути миграции были связаны с началом активного перемещения человека с целью торговли или захвата новых территорий, а на сегодняшний день интенсивно развиваются в результате климатических изменений. Неравномерность заселения земного шара связана с сезонностью и минусовыми температурами, во время которых комары вынуждены диапаузировать, а далеко не все представители семейства *Culicidae* благоприятно переносят долгие заморозки. Диапауза оказалась одним из стратегически выгодных адаптационных механизмов, когда на время неблагоприятных сезонов происходит тотальная приостановка активности организма [276]. В частности, наблюдается снижение интенсивности метаболических процессов, остановка морфогенеза и репродуктивной функции, повышение жизнеспособности в экстремальных условиях. Примечательно, что комары в зависимости от конкретного вида входят в диапаузу на разных стадиях жизненного цикла (личинка, куколка, имаго), что также задает адаптационную вариативность и конкурентноспособность в сходных условиях обитания [277]. Глобальное потепление климата способствует сокращению длительности холодных сезонов, что позволяет многим комарам осваивать и закрепляться на ранее непригодных для существования территориях, расширяя свой ареал обитания. Представленная диссертационная работа посвящена изучению закономерностей пространственной организации генома пяти представителей рода *Anopheles*, принадлежащих трем различным под родам: *Cellia* (*An. coluzzii*, *An. merus*, *An. stephensi*), *Anopheles* (*An. atroparvus*) и *Nyssorhynchus* (*An. albimanus*) (Рисунок 4, Г) [278]. В данном разделе будет дана краткая характеристика мест обитания и пищевых привычек указанных видов (более подробно описаны в таблице Приложения 1).

An. albimanus является жителем прибрежных территорий Нового света, преимущественно распространен в странах Карибского бассейна, на берегах Центральной и Южной Америки и занимает лидирующие позиции по переносу малярии в Латинской Америке [279,280]. На территории Колумбии обитает на обоих побережьях, где

Тихоокеанская сторона представлена тропическими лесами с высокой влажностью и большим количеством осадков, а берег Карибского бассейна - лесной зоной с сухим климатом и более высокими температурами [281]. Обычно комар питается на домашних животных, но вариативность экологических ареалов, доступность источников питания и обширная география распространения могут смещать предпочтения данного представителя *Anopheles* в сторону человека [279,282,283].

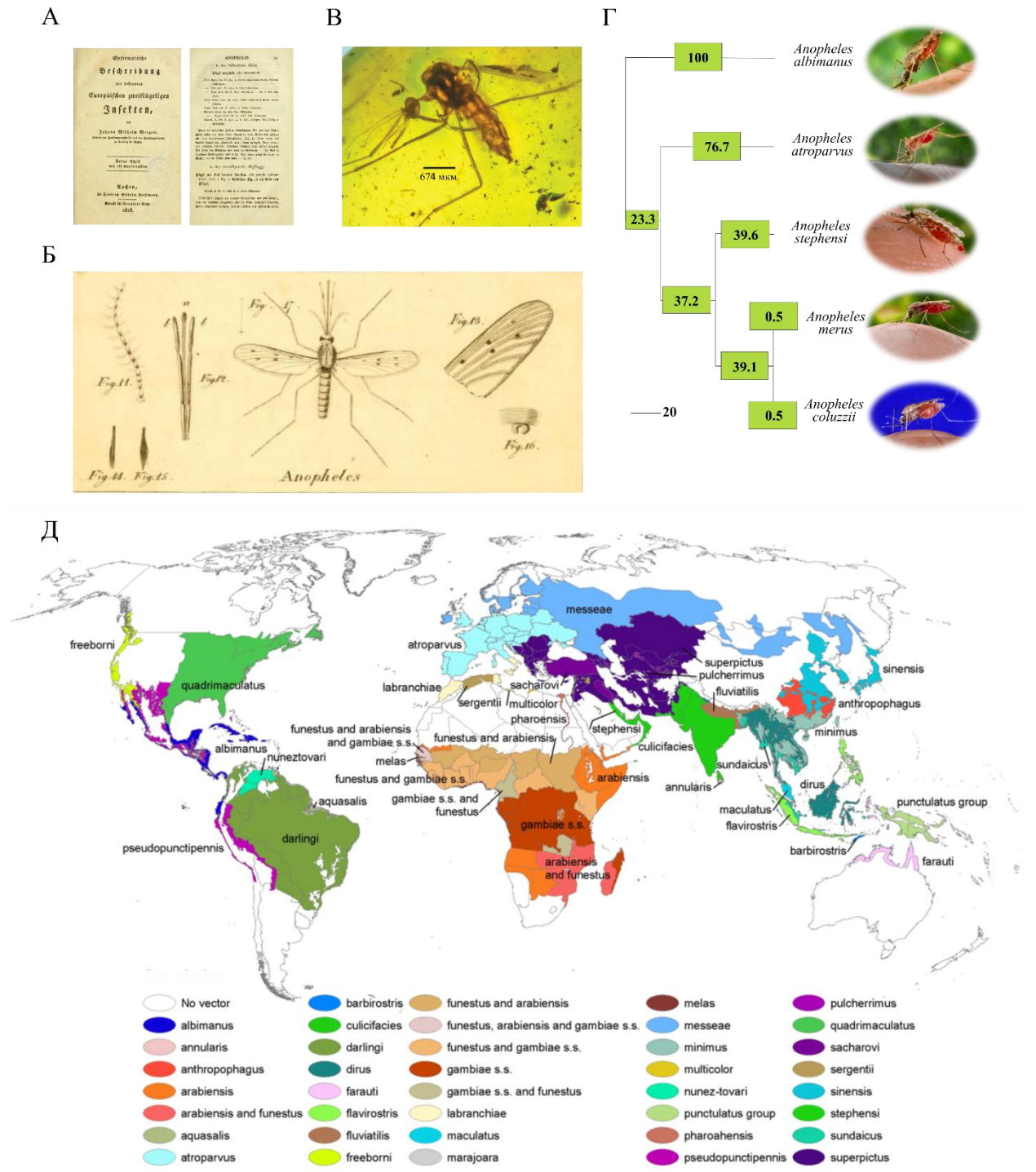


Рисунок 4. История исследований, филогения и распространение малярийных комаров. А. и Б. Фрагменты из книги Systematische Beschreibung der bekannten europäischen zweiflügeligen Insekten, изданной Иоганном

Майгеном в 1818 году, где впервые приводится описание комаров рода *Anopheles*; В. Предок современных малярийных комаров - *Priscoculex burmanicus* [267]; Г. Филогения видов, выбранных в диссертационное исследование (фотографии адаптированы с [284], благодарность James Gathany); Д. Карта-схема распространения малярийных комаров по планете. Цветом отмечен 41 вид рода *Anopheles*, подтвержденный переносчик малярии (адаптировано из [278]).

An. atroparvus представляет биологический интерес в связи с сохраняющимся высоким малярийным потенциалом и распространенностью данного вида на территории Европы и Европейской части России. В XX веке данный комар был успешным переносчиком малярии в Европе [285], но даже после того, как малярия была искоренена на территории Европейских стран, *An. atroparvus* остается под пристальным наблюдением ученых, особенно в свете изменения климата на планете. Массовая миграция населения из стран ближнего и дальнего зарубежья в сочетании с глобальным потеплением создает серьезные условия для возникновения вспышек малярии и других заболеваний в северных регионах Евразии, включая Центральную Азию, Кавказ и такие регионы России как Ростовская область, Краснодарский край, Чеченская Республика и Республика Дагестан.

An. coluzzii – один из представителей малярийных комаров комплекса *An. gambiae*, группы из 9 видов-двойников, населяющих территории Тропической Африки (Африки к югу от Сахары). *An. coluzzii* был выделен в отдельный вид совсем недавно, а до этого считался М-молекулярной формой *An. gambiae*. На сегодняшний день вид имеет обширный ареал обитания, преобладая на западе и северо-западе континента, и является одним из важнейших переносчиков малярии в мире (по материалам [286]). В лабораторных условиях вид представлен в виде двух колоний - МОПТИ и Ngoussou. Колония *An. coluzzii* Ngoussou была основана в 2006 году профессором Ланзаро из кладки диких самок *An. coluzzii*, пойманных в Камеруне [287], тогда как колония *An. coluzzii* МОПТИ берет начало от самок *An. coluzzii*, отловленных в Мали.

Вид *An. merus* распространен на восточном и юго-восточном побережьях Африки, а также на Мадагаскаре и предпочитает водоемы повышенной солености (до 50%). Изначально не входивший в группу основных переносчиков малярии, *An. merus* за последние годы прочно закрепился в лидирующих позициях по количеству случаев переноса малярийного плазмодия в Восточной и Юго-Восточной Африке [288,289].

An. stephensi встречается преимущественно в Южной и Средней Азии, и является ключевым переносчиком малярии для таких стран как Индия, Китай, Афганистан, Бахрейн, Бангладеш, Египет, Иран, Ирак, Оман, Пакистан, Саудовская Аравия, Тайланд [290] (Рисунок 5, А). Диплоидная линия комаров MSQ 43 была впервые получена в 1969 году

группой И.Шнайдер из личинок первого возраста малярийного комара *An. stephensi* [291] (Рисунок 5, А-Д). На протяжении нескольких десятилетий культура использовалась в научных исследованиях и на сегодняшний день доступна в Ресурсном центре по исследованию малярии (MR4, www.beiresources.org).

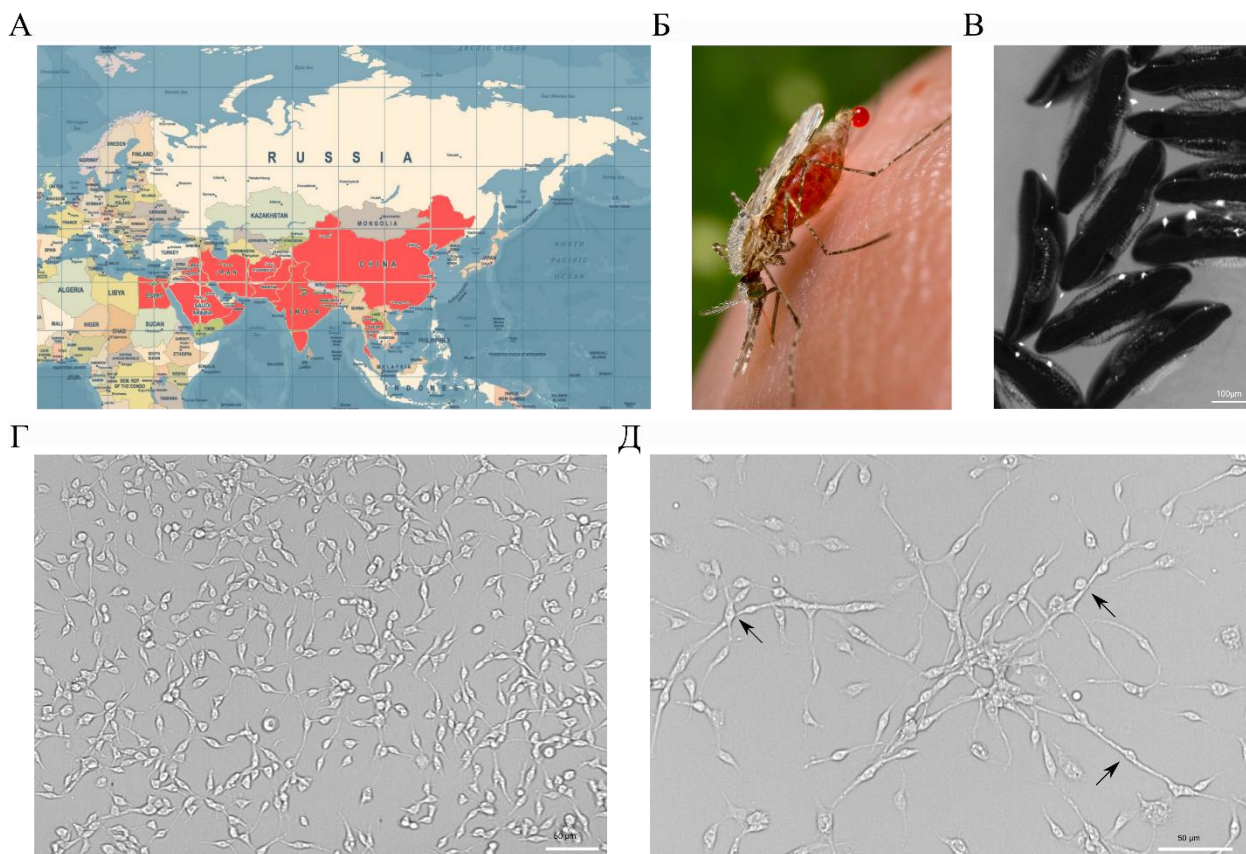


Рисунок 5. Вид *An. stephensi*, ключевой переносчик малярии в странах Южной и Средней Азии. А. Географическое распространение вида *An. stephensi*, красным выделены регионы, где *An. stephensi* был замечен; Б. Фотография взрослой самки в процессе кормления (фото взято с [284]); В. Изображение яиц *An. stephensi* через 15-20 часов после откладки; Г. Клетки линии MSQ43 в культуре; Д. Формирование межклеточных контактов, стрелками отмечены веретеновидные образования.

1.7.3 Цикл развития

Жизненный цикл малярийных комаров, который длится в среднем 3-4 недели, включает 4 выраженных стадии развития: яйцо, личинка, куколка и взрослая особь или имаго. Спустя 48-72 часа после кормления самка малярийных комаров откладывает 50-200 яиц на поверхность воды, либо на хорошо увлажненный субстрат (в случае лабораторных колоний – влажная фильтровальная бумага). Комариные яйца покрыты твердыми защитными оболочками – экзо- и эндохорионом, и имеют поплавки, позволяющие удерживаться на поверхности воды. Затем, через 1-3 суток происходит выход личинок 1го возраста (в холодных регионах процесс может занимать 2-3 недели), которые в течение 6-

11 дней претерпевают еще 3 дополнительные стадии развития прежде, чем превратиться в куколку. Скорость развития зависит от конкретного вида и температуры воды. Личинки малярийных комаров развиваются в водной среде, отфильтровывают мелкие организмы и нитчатые водоросли, периодически всплывая к поверхности воды для поддержания дыхания. Данная стадия завершается линькой и вылуплением взрослого комара или имаго. После вылета из куколки имаго необходимо 3-7 дней, чтобы достичь половозрелого состояния, после чего возможно спаривание и оплодотворение. После копуляции сперма сохраняется в семяприемнике самки и её бывает достаточно для 2-3-х кладок. Питание взрослых самцов и самок резко различается: самцы питаются исключительно нектаром растений в природе или 10% раствором сахарозы в лабораторных условиях, тогда как самкам для развития яиц дополнительно требуется богатая белками пища. Для активации работы яичников, развития ооцитов и перехода в стадию вителлогенеза необходимо питание кровью, поэтому после спаривания самки малярийных комаров приступают к поиску источника питания. Поступление высокобелковых веществ стимулирует продукцию гормонов – ювенильного и 20-гидроксиэкдизона, которые, в свою очередь, активируют работу яичников, происходит развитие ооцитов и откладка яиц [292]. Короткая длительность жизненного цикла (~3–4 недели для представителей рода *Anopheles*), возможность получения и ведения лабораторных колоний комаров, а также относительная неприхотливость данного рода насекомых к условиям содержания делают представителей рода *Anopheles* действительно удобным научным объектом, в том числе для изучения эволюционных процессов пространственной организации генома.

1.7.4 Накопление и аккумуляция хромосомных перестроек - двигатель эволюционного процесса

Экологический успех вида зависит от его умения приспосабливаться и преодолевать препятствия биотического и абиотического характера [293]. Впервые идея о влиянии хромосомных инверсий на адаптационные способности вида была сформулирована Добжанским после его наблюдений за частотами встречаемости инверсионных перестроек у *Drosophila pseudoobscura* в различных природных биотопах [294,295]. На сегодняшний день известно, что хромосомные перестройки являются ключевыми драйверами эволюционного процесса у комаров [296,297], способны приводить к изменению физиологических свойств, жизненных характеристик, адаптационных способностей их носителей, а также репродуктивной изоляции и, предположительно, к видообразованию [263,298]. В недавних работах было показано, что закрепление определенного набора генов внутри полиморфных инверсий у эволюционно далеких видов комаров может

поддерживать адаптацию к схожим условиям обитания [293,299]. На примере комаров рода *Anopheles* было замечено, что некоторые из его представителей обитают на небольшом фиксированном ареале, тогда как другие виды занимают более крупные пространства, активно и успешно осваивают новые территории, демонстрируя высокую адаптационную способность к изменению влажности воздуха [293,300]. Наиболее известный и хорошо изученный пример – это инверсия 2La и вариации 2R-инверсии (2Rb, -c, -d, и -u) в комплексе *An. gambiae*. Частоты встречаемости данных инверсий в популяциях возрастают с увеличением степени засушливости обитаемых территорий [301], а результаты исследований свидетельствуют в пользу того, что гены, расположенные внутри инвертируемого региона и, соответственно, подверженные запрету рекомбинации, вовлечены в поддержание водного баланса [302] и устойчивости к переносу малярийного плазмодия [303]. Утверждается, что именно высокий инверсионный полиморфизм определил эпидемиологически важные маршруты распространения малярийных комаров [304].

Последние работы по изучению эволюции хромосом между видами данного рода свидетельствуют о наличии многочисленных геномных перестроек, произошедших как внутри одного хромосомного элемента (понятие “хромосомный элемент” было введено в работе [305]), так и между ними [297]. Кроме того, согласно недавним исследованиям, некоторые участки генома обладают способностью быстрее накапливать и фиксировать хромосомные перестройки, чем другие [306]. А именно, в работе, проведенной на нескольких видах рода *Anopheles* было показано, что в процессе эволюции абберации хромосом наиболее быстро закрепляются именно на половой X-хромосоме [297]. Точно такая же закономерность, но в меньшем масштабе, наблюдалась и для плодовых мушек *Drosophila* [307]. Кроме того, пространственные контакты между элементами генома способствуют возникновению и увеличивают частоту хромосомных перестроек между сближенными контактирующими участками [308].

1.7.5 Способность переносить малярию

Известно, что комары рода *Anopheles* или малярийные комары, являются переносчиками многих патогенов, вирусных и инфекционных заболеваний, в том числе малярийного плазмодия. В 1880 году Альфонс Лаверан, офицер французской армии в Алжире, обнаружил и описал малярийных паразитов в крови заболевших. К концу XIX века стала определяться роль комаров в переносе малярии и различных филяриатозов и интерес к роду *Anopheles* стремительно вырос. В 1898 году Лондонское Королевское общество

избрало специальную комиссию и поручило ей заняться полноценным исследованием малярии [262]. В том же году был опубликован основополагающий научный труд Рональда Росса, который, изучив цикл развития *Plasmodium relictum* – паразита, ответственного за малярию у птиц, заключил, что заражение происходит путем впрыскивания малярийного плазмодия в кровь хозяина из слюнных желез комара. Затем следует размножение паразита в кровяном русле (позднее были выделены несколько стадий – шизонты и мерозоиты из печени, трофозоиты, шизонты и мерозоиты из эритроцитов), завершающаяся образованием мужских и женских гамет (микрогаметоциты и макрогаметоциты), которые, попадая в кишечник комара при кормлении кровью инфицированной особи, формируют споры (оокинеты и ооцисты), которая, в свою очередь, делясь на палочковидные структуры (позднее – спорозоиты), достигает слюнных желез комара и цикл замыкается. Росс предполагал, что у человека малярия развивается по такому же сценарию, но доказать экспериментально не имел возможности. Его итальянские коллеги, Бигнами и Грасси, имея доступ к пациентам с малярией вблизи Рима и на Сицилии, проследили все стадии развития плазмодия, и, наконец, показали, что именно самки комаров рода *Anopheles* передают малярийного плазмодия человеку при укусах (согласно обзорам [309] и [262]). Краткая схема передачи малярийного плазмодия проиллюстрирована на Рисунке 6.

На сегодняшний день считается, что шесть видов плазмодия могут отвечать за развитие заболевания (*Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale curtisi*, *Plasmodium ovale wallikeri*, *Plasmodium malariae*, *Plasmodium knowlesi*), тогда как 2 из них - *P. falciparum* и *P. vivax* - являются наиболее распространенными и опасными патогенами [310,311]. В регионах Центральной, Южной и Юго-Восточной Азии преобладает *P. vivax*, тогда как в странах Африки более распространен вид *P. falciparum*. Возникновение мутаций в β -глобиновом локусе обусловило развитие и закрепление серповидно-клеточной анемии в человеческой популяции как способ защиты от малярийного плазмодия [312]. Дело в том, что точечные мутации в гене *HBB* могут влиять на уровень экспрессии гемоглобина (β -талассемия), либо существенно изменять биофизические свойства белка (серповидно-клеточная анемия) и, как следствие, свойства клеток эритроидного ряда, которые становятся не пригодными для размножения плазмодия. Наиболее распространенным мутантным вариантом считается гемоглобин HbS с заменой аминокислоты глутамин на валин в 6м положении ($\alpha 2\beta^S 2$). Гомозиготы, как и гетерозиготы по мутациям в гене *HBB*, обладают повышенной врожденной устойчивостью к заражению малярией [313]. Существование данного заболевания является прямым следствием противостояния защитных систем организма человека малярийному плазмодию в тех регионах, где процент

заболевающих ежегодно особенно высок, и по этой причине малярийные комары считаются одними из немногих эукариот, оказавших влияние на эволюцию человека [262].

Потенциал к переносу вирусов и других патогенов животных и человека, в том числе различных подтипов малярийного плазмодия, в значительной степени варьирует между представителями рода *Anopheles* [11,12] и зависит от многих биологических факторов, в том числе от плотности популяции, процента успешного кормления, продолжительности жизни самок [314]. Молекулярные причины таких различий до сих пор являются объектом активных дискуссий и считаются не ясными. Ряд геномных локусов предположительно связан со способностью комаров быть переносчиками малярии [303,315–320]. Так, было показано, что ген *APL1*, присутствующий в единичной копии у *An. stephensi* и в трех копиях у представителей комплекса *Gambiae*, ответственен за выживаемость при инфицировании и устойчивость к малярийному плазмодию [321]. Другой белок *TEP1*, ключевой компонент врожденной системы иммунного ответа, связываясь с зараженными плазмодием клетками эпителия средней кишки, уничтожает их, тем самым снижая патогенную нагрузку [322]. Примечательно, что совсем недавно была показана плеотропная природа данного белка, когда по тому же молекулярному механизму он участвует в элиминации неликвидных сперматозоидов у самцов, повышая фертильность [323]. Оказалось, что вид *An. coluzzii* является преимущественным носителем варианта белка *TEP1*R1*, обеспечивающего устойчивость комара к малярийному плазмодию, и видовая структура популяций комплекса *Anopheles gambiae* играет существенную роль в динамике распространения малярийного плазмодия на протяжении летнего сезона [300].

Недавние исследования хроматина в эпителиальных клетках средней кишки у инфицированных и неинфицированных комаров показали существенные различия профилей экспрессии и эпигенетических меток [324,325]. При изучении пространственных контактов генома малярийного плазмодия наблюдалась кластеризация в пространстве ядра особой группы генов, ответственных за проникновение и закрепление плазмодия в организме хозяина, предположительно для совместной регуляции их экспрессии и возникновения дополнительной вариативности в результате рекомбинации [326–328]. На стадии спорозоитов, перед внедрением в организм человека, была описана глобальная репрессия транскрипции и трансляции, что выражалось в утрате центромерных взаимодействий по сравнению с другими стадиями, возникновении дополнительных дальних пространственных контактов, а также массовом отсутствии белкового продукта для уже синтезированных транскриптов [327,328]. Накопление геномных данных свидетельствует о глобальных транскрипционных и пространственных изменениях

хроматина на протяжении клеточного цикла развития плазмодия, вероятно подверженных эпигенетической регуляции [325]. Давно известно, что механизмы работы некоторых ферментативных систем паразита могут подстраиваться и адаптироваться под механизмы хозяина. Можно предполагать, что и адаптационные механизмы паразита и хозяина также могут быть схожи.

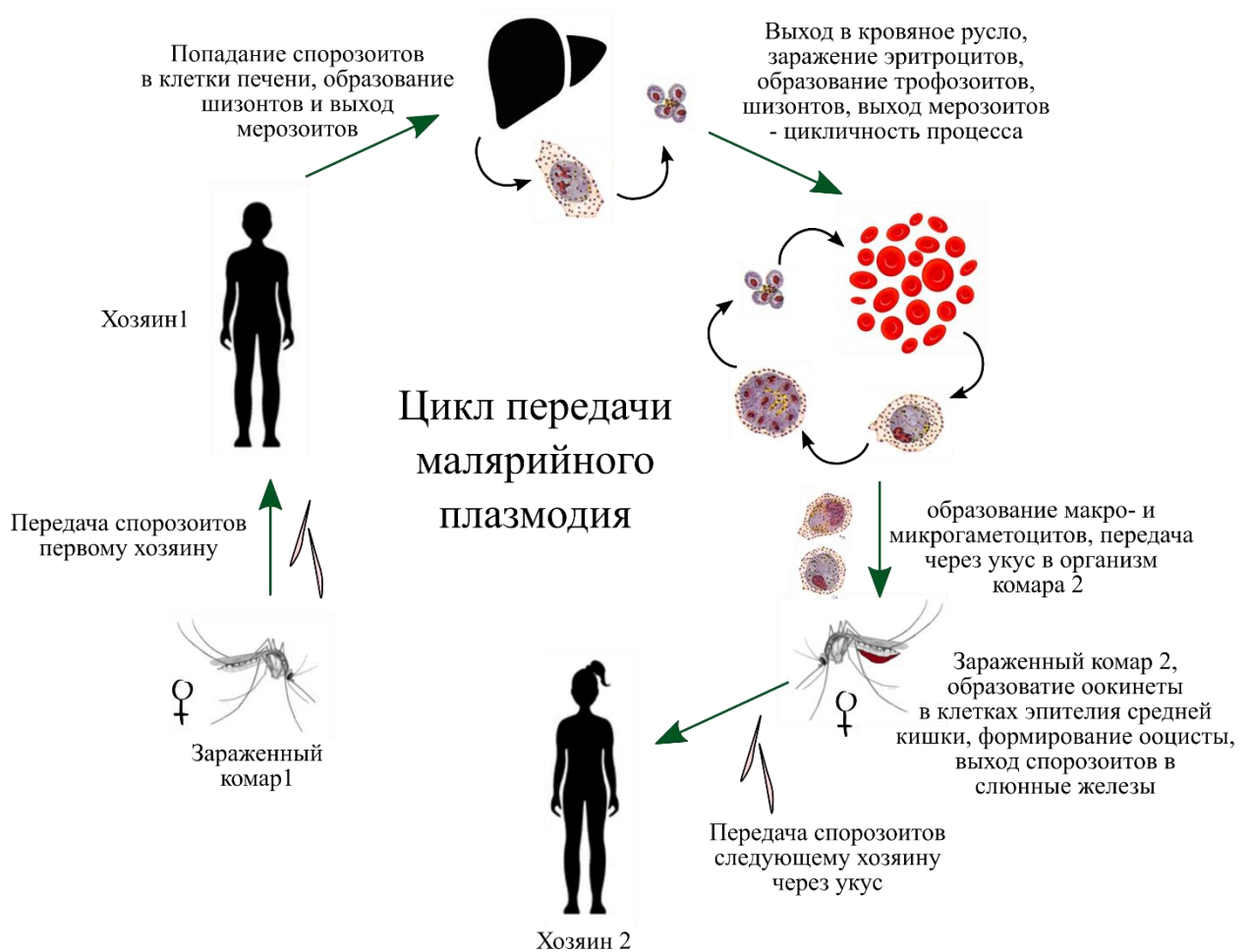


Рисунок 6. Цикл передачи малярийного плазмодия: зараженная самка *Anopheles*, кусая человека, передает ему спорозоитов малярийного плазмодия. Они, попадая в клетки печени, начинают активно размножаться с образованием мерозоитов, которые достигая кровяного русла, закрепляются в эритроцитах и циклично размножаются с образованием трофозоитов, шизонтов и мерозоитов. Кроме того, из трофозоитов происходит образование мужских и женских гаметоцитов, которые при укусе попадают в организм комара, где, достигая эпителия средней кишки, образуют зиготу. Она, в свою очередь, растет, формируя оокинету и затем ооцисту, из которой выходят новые спорозоиты. При укусе следующего хозяина спорозоиты попадают в организм человека и цикл повторяется.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ.

2.1 Разведение и поддержание комариных колоний различных видов *Anopheles* и работа с клеточной линией MSQ43

2.1.1 Линии комаров, использованные в эксперименте, условия их содержания, кормления и разведения

Для диссертационной работы были использованы лабораторные колонии комаров рода *Anopheles*, полученные из специализированного Ресурсного центра по исследованию малярии, Атланта, США (MR4, www.beiresources.org) - линия MOPTI для вида *An. coluzzii* (MRA-763), линия MAF для *An. merus* (MRA-1156), линия EBRO для *An. atroparvus* (MRA-493), линия STECLA для *An. albimanus* (MRA-126), а также колония индийского подвида комара *An. stephensi* – Indian strain, являющаяся собственностью Вирджинского Политехнического Университета (Virginia Tech).

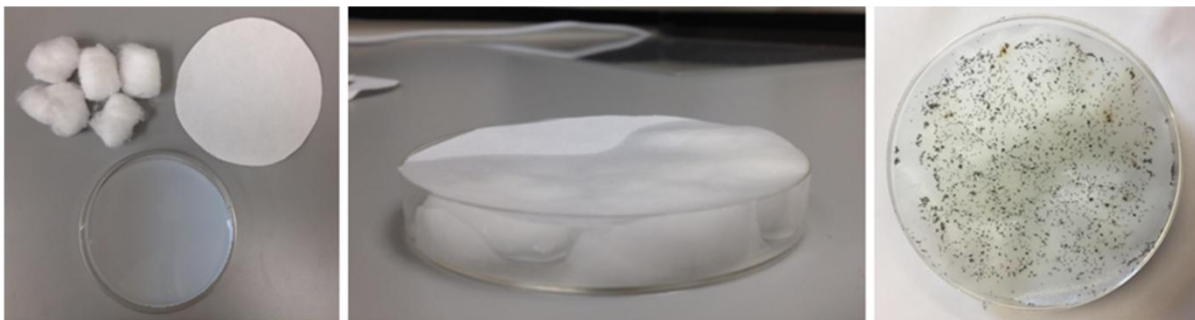
Все из указанных выше колоний содержались в инсектарии на базе Института Естественных Наук им. Фралина в Вирджинском Политехническом Университете, Блексбург, США. Комариные эмбрионы вылупляли в пресной воде, начиная прикорм в виде сухого рыбьего корма на 1-2 день после появления личинок первого возраста. Далее, личинки культивировали в специальных контейнерах с пресной водой в течение 10-12 дней (от личинок 1 до 4 возраста), периодически добавляя прикорм и контролируя плотность особей, до момента появления куколок. Последних отлавливали и переносили в отдельные небольшие емкости с водой, которые помещали непосредственно в клетку для взрослых особей. Вылетевшие взрослые особи содержались в инкубаторах при температуре 27°C, относительной влажности 75% и 12-часовыми циклами освещения день/ночь.

2.1.2 Получение эмбрионального материала для подготовки Hi-C-библиотек

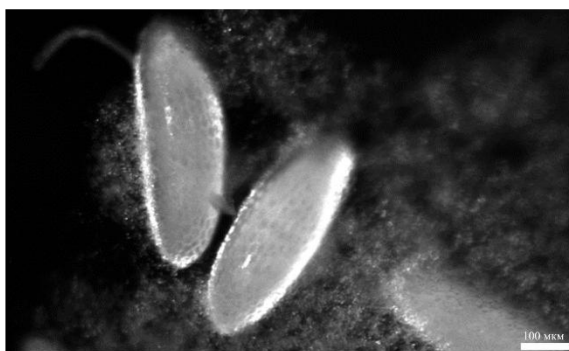
С использованием специального устройства, имитирующего поверхность кожи животного, 5-7-дневных взрослых особей кормили овечьей кровью, нагретой до 37°C. Во время кормления контролировали, чтобы брюшки самок большей части колонии были наполнены кровью. Через 48-72 часа после кормления (время определялось индивидуально для каждого вида в указанном диапазоне) в клетку помещали чашку для сбора яиц (Рисунок 7, А). Через 15-18 часов чашку вынимали из клетки, подсчитывали количество отложенных яиц и проводили фиксацию эмбрионального материала для дальнейших экспериментов. Примеры эмбрионов рода *Anopheles* в сравнении с эмбрионами *Drosophila melanogaster*

показаны на Рисунке 7, Б-Д. Для получения максимального количества яиц сбор эмбрионов на эксперимент проводили после 2го и 3го кормления.

А



Б



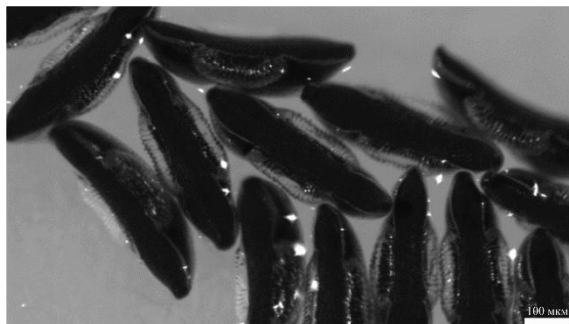
Drosophila melanogaster

В



Anopheles albimanus

Г



Anopheles stephensi

Д



Anopheles atroparvus

Рисунок 7. Получение эмбрионального материала для подготовки Hi-C-библиотек. А) Пример чашки для яиц, используемой в работе (на правом изображении показана чашка, содержащая ~5000 эмбрионов *An. coluzzii* Mopti); Б-Д) Эмбрионы *Dr. melanogaster*, *An. albimanus*, *An. stephensi*, *An. atroparvus* перед фиксацией материала.

2.1.3 Культивирование клеток линии MSQ43

Ампула с клетками линии MSQ43 была получена из Ресурсного центра по исследованию малярии, Атланта, США (MR4, www.beiresources.org). Далее, клетки были разморожены на плато T25 и культивировались на среде MEM E5 (Таблица 1) в течение нескольких пассажей при плотности 60-80%.

Таблица 1. Состав среды MEM-E5 для культивирования клеточной линии MSQ43.

КОМПОНЕНТ СРЕДЫ	КОНЕЧНАЯ КОНЦЕНТРАЦИЯ	ОБЪЕМ НА 100 МЛ MEM E5-СРЕДЫ:
среда Игла с солями Эрла (MEM 10x, Sigma Aldrich, 0275)	1x	10 мл
L-глутамин (200 mM, Thermo Fisher Scientific, 35050038)	2 mM	2 мл
10% D-глюкоза (Sigma Aldrich, G7021)	0.1%	1 мл
набор витаминов для среды Игла (100x, Thermo Fisher Scientific, 11120052)	1x	1 мл
набор незаменимых аминокислот (100x, Sigma Aldrich, M7145)	1 mM	2 мл
смесь антибиотиков пенициллина и стрептомицина (100x, Pen/Strep, Sigma Aldrich, P4333)	1x	1 мл
инактивированная эмбриональная бычья сыворотка (FBS, Sigma Aldrich, F9665)	5%	5 мл
NaHCO ₃	2.2 мг/мл	10 мл
mQ H ₂ O (автоклавированная)	-	68 мл

Линия MSQ43, первоначально полученная из личинок первого возраста малярийного комара *An. stephensi* (Рисунок 8, А), была использована в качестве модельного объекта для изучения эпигенетических меток и структур дальних хроматиновых петель. Она представляет собой адгезионную монослойную культуру (Рис.8, Б), с оптимальной плотностью рассадки 60-80%, предпочтительным культивированием на среде MEM E5 при температуре 28°C и 5%-м содержании CO₂. Клетки линии MSQ43 - округлые, зачастую имеющие 2-3 отростка по полюсам. Диаметр срединной части клетки составляет 7-15 мкм, тогда как длина между концами отростков может составлять 20-50 мкм. Отдельные клетки тяготеют к образованию межклеточных контактов, формируя веретеновидные структуры (пример показан стрелками на Рисунке 5, Д и Рисунке 8, Б). На основании анализа 4 x 100 метафазных пластинок нами было установлено, что линия MSQ43 содержит ~93-97% диплоидных клеток, тогда как оставшиеся 3-7% приходятся на три- и тетраплоидные клетки (Рисунок 8, В - Е). В единичных случаях мы наблюдали хромосомные перестройки (Рисунок 8, Д). Примечательно, что в первой работе по получению MSQ43 авторы указывают на наличие метафазных пластинок с 7-8 хромосомами в 5% случаев [291]. Мы, в свою очередь, таких пластинок не наблюдали.

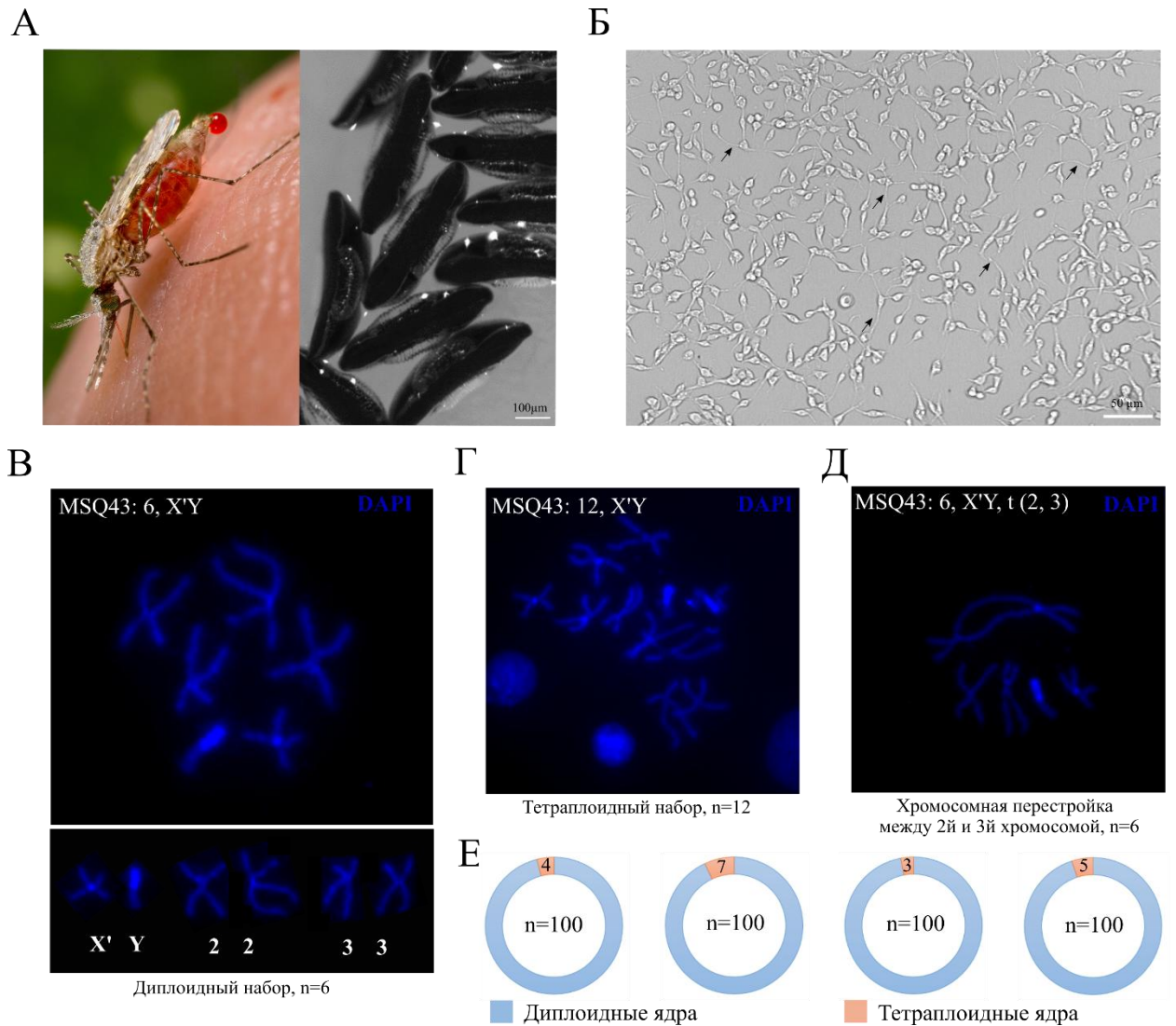


Рисунок 8. Клеточная линия MSQ43, полученная из личинок первого возраста малярийного комара *An. stephensi* (A). На панели Б. обозначены межклеточные контакты в виде веретеновидных структур. В. Диплоидный набор хромосом. Г. Тетраплоидный набор хромосом. Д. Хромосомная пластинка, содержащая перестройку между 2 и 3 хромосомами. Е. Анализ метафазных пластинок линии MSQ43.

2.2 Молекулярные методы

2.2.1 Протокол приготовления Hi-C-библиотек из эмбрионов комаров

Оригинальный протокол *in situ* Hi-C [4,119] был оптимизирован нашей исследовательской группой для 15–18 часовых комариных эмбрионов [329] и успешно применен на нескольких видах комаров родов *Anopheles* (*An. coluzzii*, *An. merus*, *An. stephensi*, *An. albimanus* и *An. atroparvus*), *Aedes* (две лабораторных линии: *Ae. aegypti* IB12 и Uganda) и *Culex* (*Culex pipiens pipiens* и *Culex pipiens molestus*). Возраст комариных эмбрионов был выбран исходя из оптимального соотношения количества клеток развивающегося эмбриона, периода начала дифференцировки клеточных тканей, а также

особенностей защитных свойств внешних мембран эмбриона (снижение восприимчивости оболочки эмбриона к обработке хлорной известью в процессе фиксации [330]).

~1000-2000 эмбрионов на стадии 15-18 часов были собраны, посчитаны и обработаны 3%-м водным раствором гипохлорита натрия в течение 10 минут. Затем, эмбрионы трижды промывали в автоклавированной чистой воде, переносили в гомогенизатор Даунса, обрабатывали фиксирующим лизис-буфером (15 mM HEPES, 60 mM KCl, 15 mM NaCl, 4 mM MgCl₂, 0.1% тритон X-100, 0.5 mM DTT, смесь протеазных ингибиторов, 3% параформальдегид), в котором гомогенизировали и инкубировали на льду в течение 10 минут. Затем, для удаления крупных фрагментов тканей и получения одноклеточной суспензии раствор пропускали через последовательную серию фильтров с размером пор 100, 60 и 10 мкм соответственно и инкубировали еще 15 минут. Далее, для инактивации параформальдегида, добавляли двойной объем буфера с глицином (15 mM HEPES, 60 mM KCl, 15 mM NaCl, 4 mM MgCl₂, 0.1% тритон X-100, 0.5 mM DTT, смесь протеазных ингибиторов, 1.25 M глицин), держали реакцию на льду в течение 5 минут. Затем осаждали клетки центрифугированием на 3000 rpm в течение 15 минут в предварительно охлажденной до 4°C центрифуге. Промывали клеточный осадок дважды в отмывочном буфере (15 mM HEPES, 60 mM KCl, 15 mM NaCl, 4 mM MgCl₂, 0.1% тритон X-100, 0.5 mM DTT, смесь протеазных ингибиторов), переносили в пробирки на 1.5 мл, промывали в 1.25X буфере для рестрикции (Neb 3.1) и замораживали осадок в жидком азоте. Хранили на -80°C до момента начала Hi-C эксперимента.

Далее, клеточный осадок ресуспендировали в 330 мкл 1.25X буфера для рестрикции (Neb 3.1), забирали 30 мкл на контроль целостности хроматина (K₀), добавляли додецилсульфат натрия (SDS) до финальной концентрации 0.3% SDS и инкубировали 1 час при 37°C с постоянным перемешиванием. Данный шаг позволяет удалить белки, несвязанные с ДНК, и сделать хроматин более доступным/открытым для действия ферментов (эндонуклеаза рестрикции и лигаза). Добавляли тритон X-100 до финальной концентрации 1.8-2% и оставляли реакцию еще на 1 час на 37°C с интенсивным перемешиванием. После добавляли 150-200U эндонуклеазы рестрикции MboI и оставляли на ночь при 37°C с постоянным эффективным перемешиванием. На следующий день инактивировали фермент нагреванием до 62°C в течение 25 минут и забирали 0.1 часть реакции на контроль порезки хроматина (K₁). Затем проводили маркирование рестрикционных концов путем введения меченных биотином нуклеотидов. Для этого готовили смесь обычных нуклеотидов без дАТФ с финальной концентрацией каждого 0.2 mM, добавляли дАТФ-биотин в концентрации 0.3 mM и 40U фрагмента Кленова.

Полученную смесь (50 мкл) добавляли к основной реакции и оставляли на 90 минут инкубироваться при 37°C с постоянным перемешиванием. Далее готовили лигазную смесь (1% тритона X-100, 0.1 нг/мкл БСА, 1X T4-лигазный буфер, 1 mM АТФ, 2000U T4-лигазы), которую затем добавляли к основной реакции и оставляли при комнатной температуре с постоянным перемешиванием. Через 2 часа добавляли новую порцию АТФ и лигазы и оставляли на ночь при комнатной температуре с постоянным перемешиванием. На следующий день проводили удаление РНК-молекул инкубацией с РНКазой при 37°C в течение 1 часа, последующую деградацию белков и снятие сшивок/кросслинков добавлением протеиназы и инкубацией на 68°C в течение 14 часов. Данный этап проводили как для основной реакции, так и для контролей (K₁ и K₀). На следующий день проводили выделение ДНК фенол-хлороформным методом с последующим высаждением в спирте. Осадок растворяли в 135 мкл воды, 5 из которых использовали для проведения тестов качества полученной библиотеки. Оставшиеся 130 мкл обрабатывали ультразвуком на приборе COVARIS M220 для получения фрагментов длиной 300-500 п.н.. Затем проводили отбор фрагментов необходимого размера очисткой реакционной смеси на магнитных шариках AMPure XP. Затем, проводили обогащение Hi-C-библиотек биотинилированными молекулами, которые в результате аффинного связывания биотина со стрептавидином преципитируются на магнитных стрептавидиновых шариках (Dynabeads, 10mg/ml, Thermo Fisher Scientific) согласно рекомендуемому протоколу производителя. Далее, проводили ПЦР-амплификацию собранных обогащенных библиотек, прикрепленных к магнитным шарикам, с использованием набора олигонуклеотидных праймеров NEBNext® Multiplex Oligos for Illumina® и набора для приготовления библиотек NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® согласно рекомендациям производителя. После амплификации шарики отбирали на магнитном штативе, а полученную итоговую библиотеку очищали от олигонуклеотидных праймер-димеров. Финальную концентрацию оценивали на спектрофотометре Qubit, готовые Hi-C-библиотеки отправляли на NGS-секвенирование на платформе Illumina. Краткая схема протокола проиллюстрирована на рисунке 9.

2.2.2 Очистка фрагментов ДНК на магнитных шариках (*Size selection, Ampure beads*)

После обработки ультразвуком мы проводили обогащение библиотек фрагментами длиной 300-500 п.н. с помощью дополнительной очистки реакционной смеси на магнитных шариках AMPure beads XP (Beckman, USA). Магнитные частицы направленно связывают молекулы ДНК определенного размера в зависимости от соотношения рабочего буфера к объему основной реакции. Согласно рекомендациям производителя, мы проводили очистку

в два этапа: на первом этапе добавляли 0.55X шариков от объема реакции, в результате чего шарики связывали молекулы крупнее 500 п.н. и затем концентрировались на магнитном штативе. Затем, переносили жидкую фракцию без магнитных частиц в новую пробирку, куда добавляли 0.35X шариков от исходного объема, что приводило к связыванию молекул длиннее 200 п.н.. Далее закрепляли шарики на магнитном штативе, промывали смесь в 80% этаноле, подсушивали шарики на воздухе для удаления остатков этанола и растворяли в буфере для элюции. В результате молекулы ДНК короче 500 п.н., но длиннее 200 п.н. оказывались в растворе.

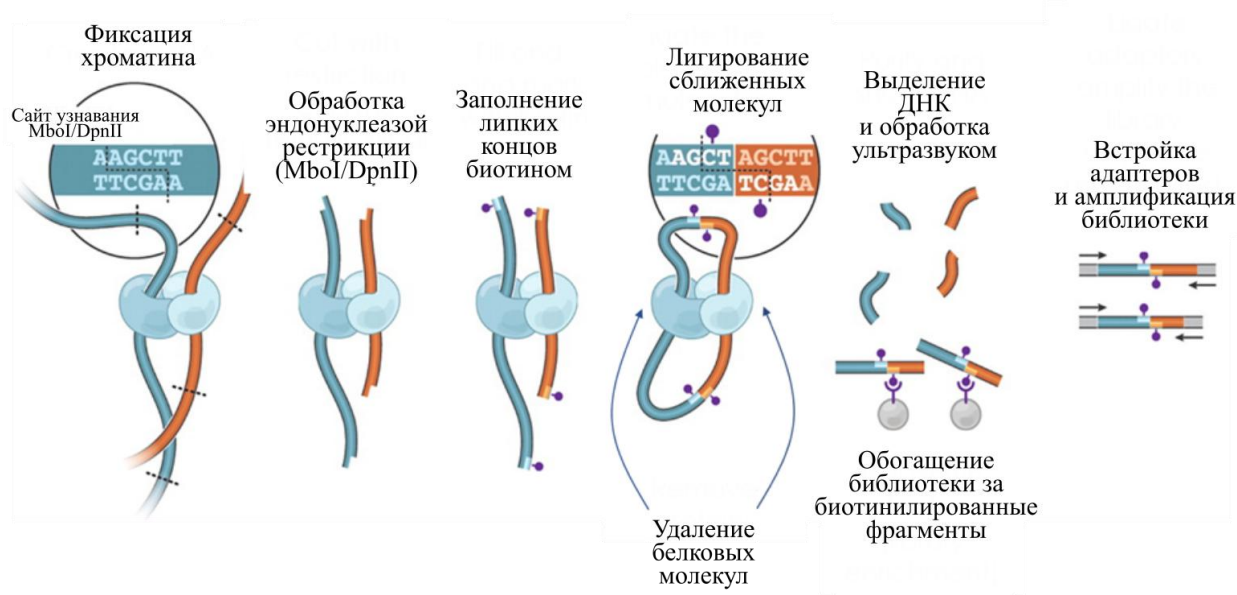


Рисунок 9. Классическая схема Hi-C-эксперимента. Адаптирована из работы Lieberman-Aiden et al., 2009.

2.2.3 Проверка качества Hi-C-библиотек

Используя модифицированный протокол, мы подготовили 10 библиотек, по 2 реплики на каждый вид. Конечная концентрация библиотеки соответствовала рекомендуемым требованиям для NGS-секвенирования, с использованием минимально возможного числа циклов ПЦР при амплификации. Предварительную проверку качества полученных библиотек мы проводили реакцией ПЦР со специфических 3С-праймеров, продуктом которых является гибридный фрагмент, образованный в случае успешного прохождения реакций рестрикции и лигирования (Рисунок 10, А - Г). Мы подобрали по 5 пар праймеров на каждый исследуемый геном, провели тестовую амплификацию ПЦР с градиентом, выявили оптимальную температуру, при которой наблюдали бэнд характерного размера (Рисунок 11, Г).

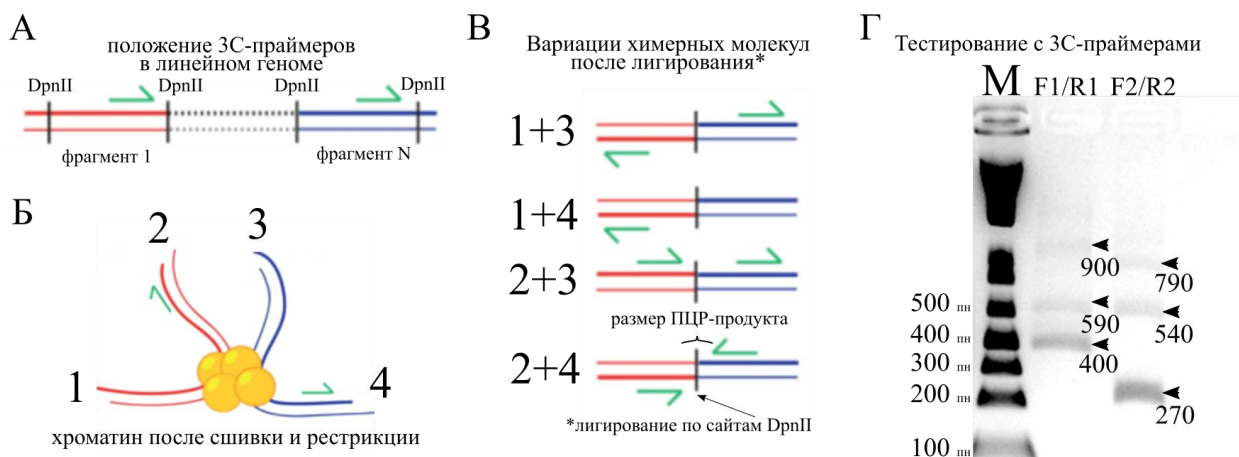


Рисунок 10. Схема, иллюстрирующая подход с 3С-праймерами. Расположения рестриционных фрагментов А) в линейном геноме и Б) после фиксации материала и рестрикции; В) вариации образующихся молекул, где целевые продукты Hi-C могут быть детектированы по наличию бэнда/сигнала характерного размера; Г) электрофореграмма с двумя разными парами праймеров, целевые продукты имеют размер 400 и 270 п.н. соответственно, более крупные фрагменты являются следствием неполной рестрикции.

Кроме того, после этапа выделения ДНК оценивали целостность исходного хроматина, а также эффективность последующих молекулярных реакций. Для этого отобранные контроли K_1 и K_0 и итоговую библиотеку наносили на 1%-й агарозный гель и разгоняли фрагменты в гель-электрофорезной камере. В случае успешного эксперимента наблюдалась следующая картина: исходный образец ДНК (гДНК или K_0) визуализируется в виде единого высокомолекулярного фрагмента, лежащего выше верхнего бэнда в маркерной лунке; после обработки рестриктазой (после MboI) происходит расщепление на градиент более мелких фрагментов (в случае клеточной линии основной пик спускается к ~1-3 Кб); а успешная обработка T4-лигазой приводит к восстановлению исходного размера молекул ДНК. Пример электрофореграмм Hi-C эксперимента представлен на рисунке 11. При сравнении панелей А и Б на рисунке 11 заметно, что в случае клеточной линии все реакции проходят более эффективно. Тем не менее, результаты NGS-секвенирования библиотек, полученных как из эмбрионального материала, так и из клеточной линии, свидетельствуют о хорошем качестве Hi-C-эксперимента.

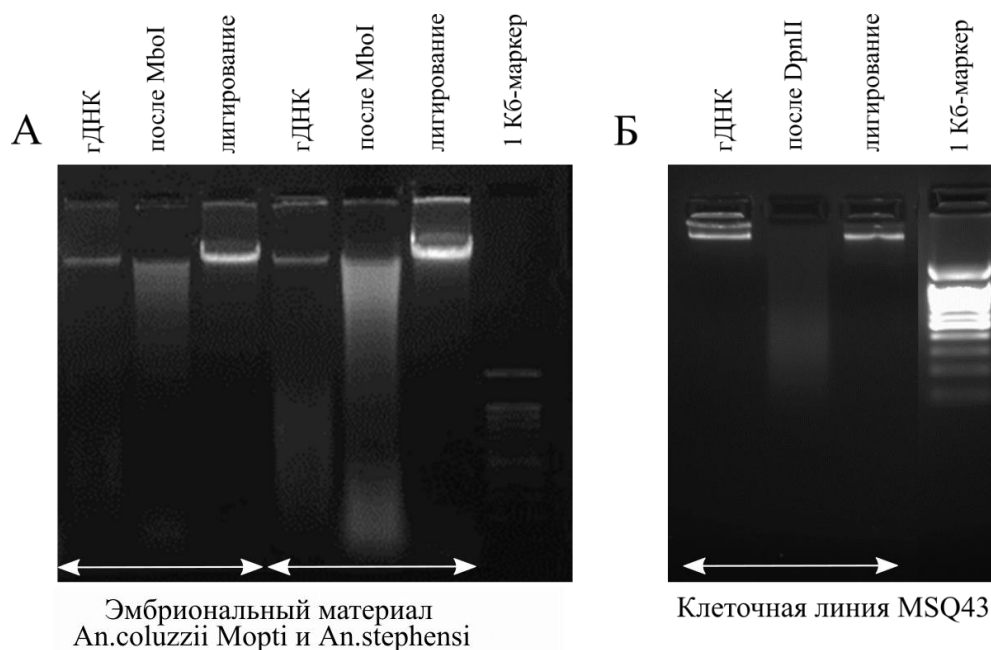


Рисунок 11. Первичная оценка эффективности прохождения основных молекулярных этапов Ni-C-протокола А) для эмбрионального материала и Б) для клеточной линии. В случае успешного эксперимента мы наблюдали единый высокомолекулярный бэнд в кармане для K_0 (гДНК), расщепление на градиент более мелких фрагментов для K_1 (после MboI/DpnII), снова высокомолекулярный бэнд в случае успешного лигирования концов (лигирование).

2.2.4 Полимеразная цепная реакция с градиентом температур и без

Для тестирования оптимальной температуры отжига олигонуклеотидных праймеров для проверки качества сборки библиотек мы проводили ПЦР с градиентом температур. Для этого мы определяли температуру отжига олигонуклеотидных праймеров согласно приложению NebOligocalculator и программировали амплификатор на 5 градусов ниже и выше рекомендуемой с шагом в 1.5-2 градуса. Проводили ПЦР с высокопроизводительной Q5-полимеразой. На 25 мкл реакции использовали 5 мкл 5X Q5-буфера, по 1,25 мкл 10 мМ каждого из праймеров, 0,5 мкл 10 мМ dNTP, 0,5 единиц активности Q5-полимеразы (NEB), ~500 нг ДНК. Условия реакции подбирали в соответствии с рекомендациями производителя относительно нуклеотидного состава праймеров и размера ожидаемого ПЦР-продукта. Для анализа полученных результатов использовали гель-электрофорез ДНК в 1,5-3% агарозных гелях.

2.2.5. Гель-электрофорез в агарозном геле

Гель-электрофорез ДНК проводили в 1,5-3% агарозных гелях, содержащих буфер ТАЕ (Tris-Acetate-EDTA) или ТВЕ (Tris-borate-EDTA) с 10 мкг/мл бромистого этидия, при напряжении 5 В/см. Для определения длины фрагментов ДНК, в один из карманов геля

добавляли 10 мкл маркера необходимой длины. Смешивали 200-300 нг ДНК исследуемого образца и 5 мкл буфера для нанесения, содержащего краситель бромфеноловый синий, либо коммерческий аналог (NEB, USA). Для фрагментов длиной менее 1000 н.п. использовался маркер с шагом 100 н.п. (Invitrogen), для более длинных фрагментов ДНК – с шагом 500-1000 н.п. (Invitrogen). ДНК детектировали с помощью трансиллюминатора (ChemIDoc, BioRad) в УФ-свете (302 нм).

2.2.6 Иммуноокрашивание на антитела

Для проведения экспериментов по иммунопреципитации хроматина было необходимо проверить функциональность выбранных антител на нашем объекте исследования. Поскольку эффективность работы большинства антител среди насекомых показана на клетках *Drosophila*, нам было необходимо убедиться в пригодности выбранных антител для использования на комарах рода *Anopheles*. Мы провели два теста – иммуноокрашивание (для всех используемых антител) и вестерн-блоттинг (для двух гистоновых модификаций – H3K27Ac и H3K27me3). Для проведения иммуноокрашивания адгезионную культуру клеток MSQ43 растили на покровном стекле до требуемой плотности (~80%), фиксировали в 3%-м PFA в течение 30 минут, затем отмывали трижды в фосфатно-солевом растворе по 5 минут. Переносили стекла в чистые лунки. Далее, для пенетрации клеточной мембраны добавляли 0.2% раствор Triton X-100 в PBS (фосфатно-солевой буфер), инкубировали 15 минут. Затем добавляли раствор 3-5% БСА в PBS без антител для снижения неспецифического связывания, оставляли на 30 минут при 4°C. Разводили целевые антитела в рекомендуемом соотношении (1:500 – 1:1000) в буфере для инкубации (1% БСА, 0.1% Tween20, PBS) и добавляли к клеткам, оставляли на 4°C в течение ночи. На следующий день проводили отмывку препаратов в буфере 0.2% Tween20/PBS три раза по 5 минут на комнатной температуре. Добавляли вторичные антитела в рекомендуемом разведении (1:200 – 1:500) в буфере для инкубации (1% БСА, 0.1% Tween20, PBS) и оставляли на 2 часа на комнатной температуре. Проводили отмывку препаратов в буфере 0.2% Tween20/PBS три раза по 5 минут. Проводили окрашивание ядер раствором DAPI в течение 2-10 минут, переносили в чистые лунки, затем отмывали клетки дважды в PBS, один раз – в MQ воде, подсушивали на воздухе, наносили раствор для длительного хранения препаратов и фиксировали на предметном стекле. Проводили детекцию сигнала с помощью конфокального микроскопа Zeiss LSM780 (ЦКП, ИЦиГ СО РАН). Результаты иммуноокрашивания приведены на Рисунке 12, А – Г, а также Рисунке 13, А.

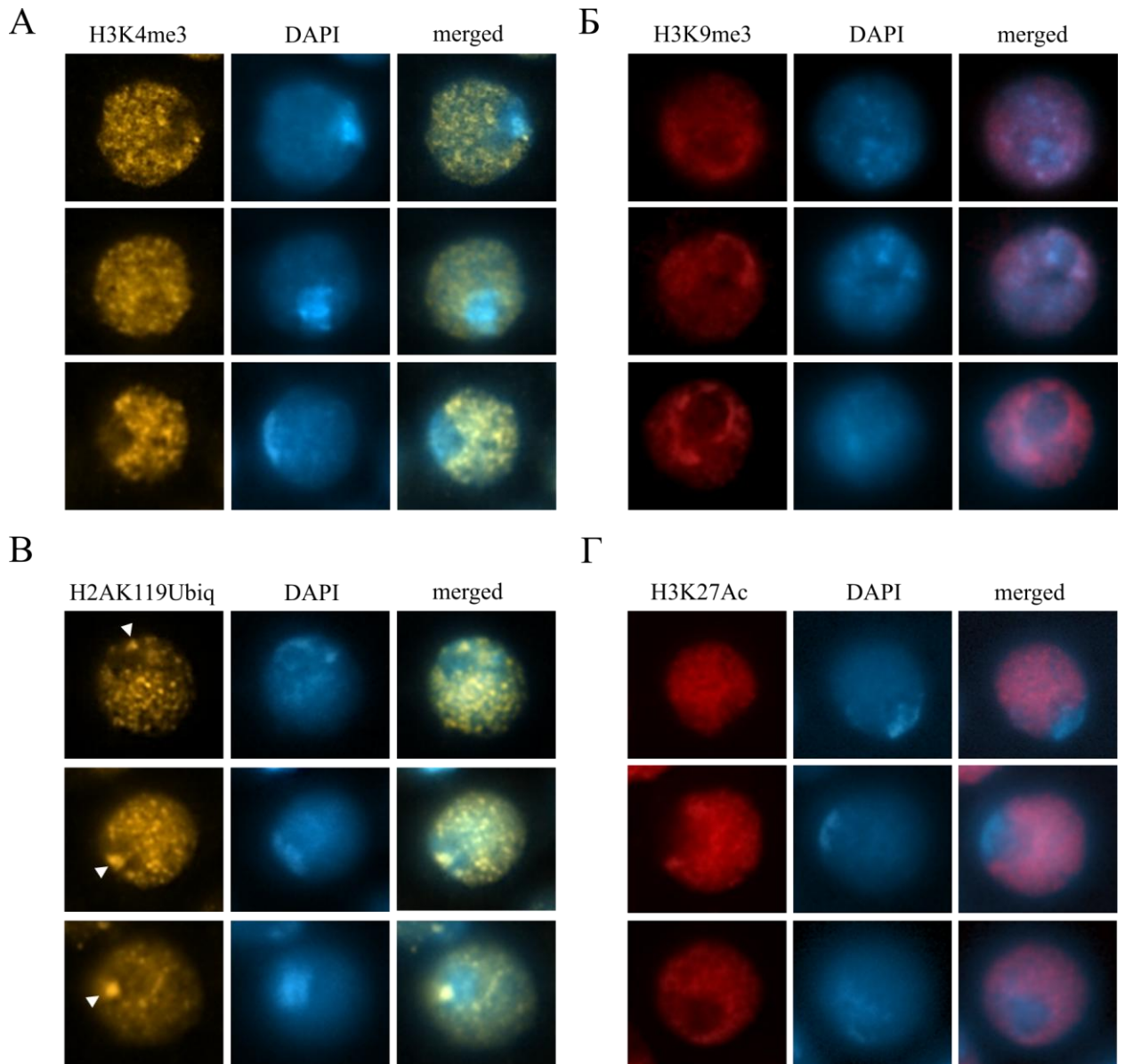


Рисунок 12. Иммуноокрашивание за антитела против гистоновых модификаций: А) H3K4me3, Б) H3K9me3, В) H2AK119Ubiq, Г) H3K27me3.

2.2.7 Выделение белков из тканей и проведение вестерн блот – анализа

Белковая фракция была выделена из тканей эмбрионов и взрослых особей в буфере RIPA (150 mM NaCl, 1% NP-40, 0.5% деоксихолат натрия, 0.1% SDS, 50 mM Tris-HCl, ингибиторы протеаз), концентрации определялись с использованием спектрофотометра Nanodrop в режиме Bradford и BCA. Затем, мы проводили разведение образцов до унифицированных значений 1-5 мг/мл и, смешав в соотношении 1:1 с загрузочным буфером (4% SDS, 10% 2-меркаптоэтанол, 20% глицерол, 0.004% бромфеноловый синий, 0.125 M Tris-HCl pH=6.8) и проведя денатурацию на 95°C в течение 5 минут, наносили на предварительно залитый полиакриламидный гель 7,5-12%. Гель-электрофорез белков проводили в полиакриламидных гелях в растворе на основе буфера TBE(Tris-borate-EDTA)

при постоянном охлаждении в течение 60-80 минут при настройках источника 180 В, 400 мА. Далее, проводили трансфер белков на активированную в растворе метанола нитроцеллюлозную мембрану в течение 60 минут при настройках источника 100 В, 350 мА. После полного переноса проводили инкубацию с первичными антителами на 4С в течение ночи с постоянным перемешиванием в разведении 1:1000, предварительно заблокировав неспецифические белковые связывания 5% БСА. На следующий день инкубировали с буфером TBS (0.2 М Tris base, 1.5 М NaCl) дважды, с TBST (0.2 М Tris base, 1.5 М NaCl, 0.1% Tween 20) дважды. Затем добавляли разведение вторичных антител 1:2000, конъюгированных с пероксидазой хрена (HRP), и инкубировали 2 часа при комнатной температуре с постоянным перемешиванием. Проведя последовательные раунды отмывки в TBS и TBST, детектировали сигнал при помощи набора SuperSignal™ West Pico (ThermoFisher, 34079) и системы гель-документирования Gel Doc XR+ (BioRad). Пример вестерн-блот-анализа для гистоновой метки H3K27me3 приведен на Рисунке 13, Б.

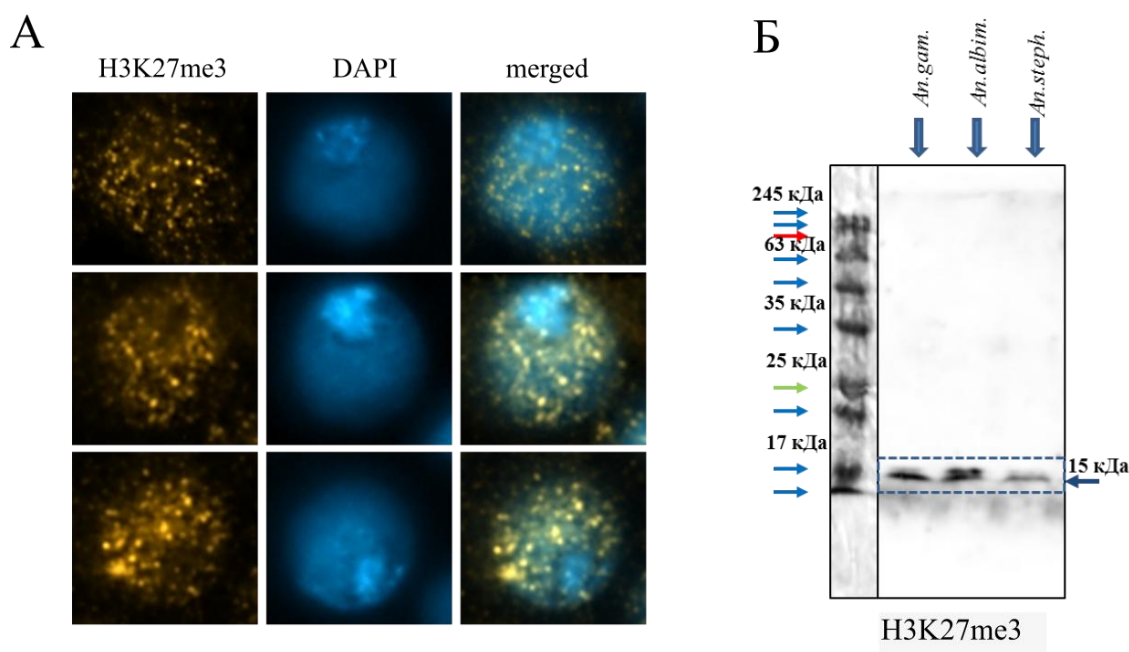


Рисунок 13. Результаты валидации антител против гистоновой модификации H3K27me3 при помощи А) вестерн-блот-анализа и Б) иммуноокрашивания.

2.2.8 Иммунопреципитация хроматина (ChIP)

Для иммунопреципитации хроматина клетки линии MSQ43 культивировали при оптимальной плотности до необходимого количества, промывали в PBS, снимали раствором TrypleExpress (ThermoFischer, USA), пересчитывали в камере Горяева и переносили в пробирки для фиксации по 5 млн. клеток на один ChIP-эксперимент. Проводили сшивку белков с хроматином в 2%-м растворе параформальдегида,

инактивировали добавлением глицина до конечной концентрации 0.25M, отмывали клеточный осадок в фосфатном буфере и замораживали до начала эксперимента по иммунопреципитации. После разморозки ресуспендировали осадок в лизирующем буфере (10 mM Tris pH 7.5, 1 mM ЭДТА, 0.5% SDS, 0.1% диоксихолат натрия, 1% Triton X-100, протеазные ингибиторы 1X) и инкубировали 20 минут на льду при периодическом перемешивании. Обработывали клетки ультразвуком на приборе Bandelin Sonopulse в режиме 30 сек ON/90 сек OFF. Затем удаляли обломки клеток и клеточных ядер центрифугированием на 18 000 g в течение 20 минут, переносили супернатант, содержащий фрагменты ДНК, в новые пробирки и отбирали небольшой объем для контроля количества хроматина и эффективности соникации. На указанных параметрах основной пик длин фрагментов приходился на 200-300 п.н.. В реакции иммунопреципитации использовали 10-30 мкг хроматина. Для снижения неспецифического связывания к хроматину добавляли 20 мкл магнитных шариков, связанных с иммуноглобулином А, и инкубировали при 4°C с постоянным перемешиванием. В это же время инкубировали целевые антитела с магнитными шариками, связанными с иммуноглобулином А, при 4° с постоянным перемешиванием. Через 2 часа отмывали белковые комплексы в фосфатном буфере от излишка антител. В тот же момент на магнитном штативе отбирали шарики и переносили подготовленный хроматин в пробирки с белковыми комплексами, оставляли иммунопреципитацию хроматина на ночь при 4°C с постоянным перемешиванием. Часть хроматина (1/10 - 1/20) до объединения с антителами забирали в качестве контроля (далее - инпут) и замораживали на -20°C до момента добавления 1% SDS в ТЕ. На следующий день проводили последовательные отмывки по 10 минут при постоянном перемешивании в следующих буферных растворах: 1) лизис-буфер (10 mM Tris pH 7.5, 1 mM ЭДТА, 0.1% SDS, 0.1% диоксихолат натрия, 1% Triton X-100, протеазные ингибиторы 1X) - дважды; 2) лизис-буфер с высоким содержанием соли (0.3M NaCl, 10 mM Tris pH 7.5, 1 mM ЭДТА, 0.1% SDS, 0.1% диоксихолат натрия, 1% Triton X-100, протеазные ингибиторы 1X) - дважды; 3) буфер с литием (0.25M LiCl, 0.5% IGEPAL-630, 0.5% диоксихолат натрия) - дважды; 4) буфер ТЕ с тритоном (10 mM Tris pH 7.5, 1 mM ЭДТА, 1% Triton X-100) – один раз; 5) буфер ТЕ (10 mM Tris pH 7.5, 1 mM ЭДТА) – один раз. После чего элюировали целевые молекулы и снимали ДНК-белковые сшивки в буфере ТЕ с 1%-м содержанием SDS при 65°C в течение ночи (с данного этапа в протоколе два образца – иммунопреципитированный хроматин и инпут). На следующий день последовательно добавляли РНКазу А (инкубация 1 час при 37°C с постоянным перемешиванием), протеиназу К (инкубация 2 часа при 55°C с постоянным перемешиванием), отбирали шарики на магнитном стенде и выделяли ДНК при помощи специализированных колонок

(Zymo Research, D5205). Оценивали концентрацию при помощи спектрофотометра Qubit с использованием набора Qubit HS DNA. Переходили к подготовке полученных библиотек к NGS-секвенированию (раздел 2.2.10).

2.2.9 Выделение РНК из тканей (очистка от рибосомальной фракции) для анализа транскриптома

Выделение тотальной РНК из комаров проводили с использованием Monarch Total RNA Miniprep Kit согласно рекомендациям производителя набора с небольшими модификациями. В одну реплику эксперимента брали 1000-1500 комариных эмбрионов, предварительно обработанных 50%-м водным раствором хлорной извести. Всего получали три реплики на один биологический вид. Материал гомогенизировали в 800 мкл лизирующего буфера, после чего инкубировали в течение 10 минут без протеиназы К, а затем дополнительные 5 минут с протеиназой. Обработку дезоксирибонуклеазой (ДНКой) увеличивали до 20 минут, элюцию РНК проводили в 50 мкл MQ H₂O. После выделения проводили оценку количества материала на спектрофотометре Nanodrop в режиме для РНК, в дальнейшие реакции использовали 1 мкг. Для обогащения фракцией матричных РНК использовали специализированный набор NEBNext® Poly(A) mRNA Magnetic Isolation Module, далее переходили к этапу подготовки библиотек на секвенирование.

2.2.10 Подготовка ДНКовых и РНКовых библиотек для проведения секвенирования нового поколения на платформе Illumina

После финального шага протоколов и выделения ДНК/РНК, мы приступали к подготовке полученных библиотек к NGS-секвенированию. Мы использовали наборы олигонуклеотидных праймеров NEBNext® Multiplex Oligos for Illumina® и набор для приготовления библиотек NEBNext® Ultra™ II DNA Library Prep Kit for Illumina®, либо NEBNext® Ultra™ II RNA Library Prep Kit for Illumina® согласно рекомендациям производителя. На заключительном шаге проводили ПЦР-амплификацию собранных обогащенных библиотек с минимально возможным количеством циклов амплификации.

2.2.11 Выделение геномной ДНК из тканей

Для подготовки меченных флуоресцентных проб было необходимо качественно выделить геномную ДНК. Для этой цели мы использовали набор Monarch® Genomic DNA Purification Kit (NEB, T3010). В качестве источника материала мы использовали стадию куколок или личинок 4го возраста. На первом этапе проводили гомогенизацию тканей в лизирующем буфере при помощи пластикового гомогенизатора. Оставляли полученную суспензию для полного лизиса на 2-3 часа на 55°C с интенсивным перемешиванием. Далее

следовали приведенному протоколу и рекомендациям производителя. Растворили полученный осадок геномной ДНК в 100 мкл буфера для элюции. Конечную концентрацию измеряли на спектрофотометре NanoDrop 2000.

2.2.12 Полимеразная цепная реакция (ПЦР) с Q5-полимеразой

Для мечения геномных районов повторенных последовательностей, таких как пробы 72A, 190A, 53B, 53C и 18S мы проводили ПЦР высокоэффективной полимеразой Q5 по протоколу, рекомендованному производителем полимеразы. Добавленные компоненты реакции приведены в Таблице 2.

Таблица 2. Компоненты ПЦР с Q5-полимеразой.

КОМПОНЕНТ	25 мкл	50 мкл	ФИНАЛЬНАЯ КОНЦЕНТРАЦИЯ
5X Q5 буфер	5 мкл	10 мкл	1X
10 мМ дНТФ (смесь нуклеотидов)	0.5 мкл	1 мкл	200 мкМ
10 мкМ F-праймер	1.25 мкл	2.5 мкл	0.5 мкМ
10 мкМ R-праймер	1.25 мкл	2.5 мкл	0.5 мкМ
ДНК	вариативно	вариативно	< 1,000 нг
Q5-полимераза	0.25 мкл	0.5 мкл	0.02 Ю/мкл
5X Q5 энхансер	5 мкл	10 мкл	1X
mQ H ₂ O	to 25 мкл	to 50 мкл	

Температуру на каждом из этапов реакции использовали согласно рекомендациям производителя полимеразы. Температуру отжига варьировали согласно рекомендациям программы Neb Tm calculator [331] и производителя олигонуклеотидных праймеров.

Затем ПЦР-продукты очищали при помощи набора для очистки Monarch™ PCR and DNA Cleanup Kit (NEB, T1030), измеряли концентрацию на спектрофотометре NanoDrop 2000. В одну реакцию ник-трансляции добавляли ~500 нг очищенного продукта.

2.2.13 Подготовка специфических флуоресцентных проб с помощью избирательного мечения участков ДНК методом ник-трансляции

Метод ник-трансляции позволяет получать качественно меченные флуоресцентные пробы, которые дают яркий сигнал в FISH-экспериментах, но требует предварительного подбора условий на начальный размер матрицы и требуемую длину конечных фрагментов.

Поскольку мы использовали ПЦР-продукты в качестве матрицы, их размер не превышал нескольких тысяч пар оснований (1-1.5 Кб в случае пробы на экзоны). На первом этапе мы смешивали все компоненты в одну реакцию, как описано в Таблице 3. Аккуратно пипетировали и оставляли на инкубацию при 15°C в течение 1.5-2.5 часов (в зависимости от исходной длины продукта и продуктивности работы ферментов) без перемешивания. Затем переносили пробирки в лед, отбирали 1/10 часть и наносили на агарозный гель для проверки эффективности реакции. Если размер полученных фрагментов составлял 200-800 п.н. (пример на Рисунке 14, А), мы добавляли 0.5 М ЭДТА для остановки реакции, ацетат натрия и 3 объема 96% этанола. Если результат не был удовлетворительным, продолжали инкубацию. Далее переходили к высадению и подготовке меченных проб к гибридизации.

Таблица 3. Компоненты ПЦР с Q5-полимеразой.

КОМПОНЕНТ	50 мкл	ФИНАЛЬНАЯ КОНЦЕНТРАЦИЯ
ДНК-матрица (500 нг)	X мкл	-
ДНК-полимераза I	1 мкл	0.2 Ю/мкл
буфер для полимеразы	2.5 мкл	0.5 мкМ
10X смесь из нуклеотидов (0.5 мМ АТФ, 0.5 мМ ЦТФ, 0.5 мМ ГТФ, 0.15 мМ ТТФ)	5 мкл	1X
БСА, 0.5 мг/мл	5	0.05 мг/мл
ДНКаза I (разведение до 0.02 Ю/мкл)	2 мкл*	0.008 – 0.016 Ю/мкл
флуоресцентно меченный нуклеотид дУТФ-Су3/дУТФ-Су3	10 мкл	-
mQ H ₂ O	to 50 мкл	

*количество единиц на реакцию необходимо оптимизировать в конкретном эксперименте.

2.2.14 Мечение участков ДНК методом случайных праймеров

Для мечения уникальных последовательностей мы также применяли метод случайных праймеров, где смесь из 6-буквенных одноцепочечных олигонуклеотидов случайного состава (Random Hexamer Primer, Invitrogen) добавляли в реакцию с фрагментом Кленова, миксом нуклеотидов, меченным нуклеотидом и ДНКовой матрицей. В качестве матрицы в реакции выступали уникальные ПЦР-продукты, полученные ранее в

реакции с Q5-полимеразой и очищенные с использованием набора Monarch™ PCR and DNA Cleanup Kit (NEB, T1030). Пример электрофореграммы ПЦР-продуктов до очистки приведен на Рисунке 14, Б.

В результате отжига праймеров в случайные места на матрице происходил синтез фрагментов, которые несли вставку меченой буквы. Далее, действие фермента ингибировали добавлением 0.5 М раствора ЭДТА до финальной концентрации 0.01 М, а полученный продукт высаждали добавлением спирта.

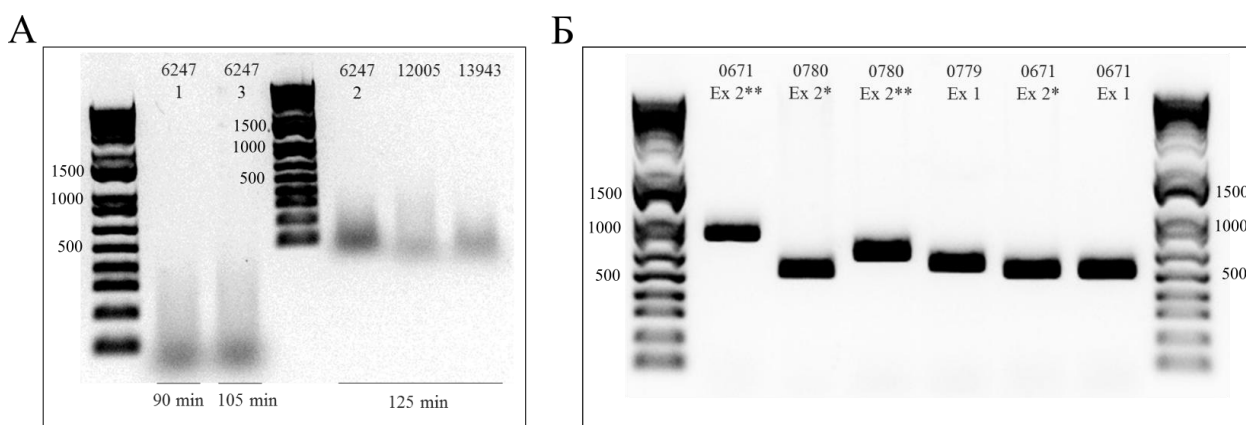


Рисунок 14. Мечение флуоресцентных проб для FISH-экспериментов при помощи двух подходов. А) Оценка эффективности ник-трансляции для подготовки проб для валидации дальней хроматиновой петли у *An.atroparvus*; Б) подготовка ПЦР-продуктов для мечения проб для валидации дальней хроматиновой петли у *An.stephensi* методом случайных праймеров.

2.2.15 Высаждение и подготовка флуоресцентных проб к гибридизации

После получения меченных флуоресцентной меткой специфичных ДНК-проб методом ник-трансляции или методом случайных праймеров, пробы высаждались из раствора добавлением трех объемов 96% спирта. Затем откручивали раствор в течение 20 минут при 18 000g, спирт удаляли, остатки упаривали в сушильной центрифуге при 40С в течение 20 минут и проверяли наличие сухого цветного осадка на дне пробирки. Далее, полученный осадок растворяли в 20-40 мкл буфера для гибридизации (1.2xSSC (цитратно-солевой буфер), 10% декстран сульфата, 60% деионизированного формамида) и оставляли в защищенном от света месте на 37°С с небольшим перемешиванием до полного растворения. В случае необходимости хранили пробы при температуре -20 – 80°С.

2.3 FISH-эксперименты и кариотипирование

2.3.1 Диссекция и фиксация отдельных органов насекомых для проведения 3D-FISH

Диссекция проводилась на предварительно обездвиженных холодом взрослых особях или личинках 3го возраста с использованием препаровальных игл и фиксирующих пинцетов. Все манипуляции осуществляли на холодовом столике в небольшом количестве PBS. В случае диссекции яичников брюшко самки отделяли по 4му с хвоста сегменту, удаляли хвост, очищали яичники от мальпигиевых и кровеносных сосудов и других сторонних тканей, переносили в свежий фосфатный буфер с добавлением Tween 20 до концентрации 0.01% (буфер PBST), а затем помещали в раствор фиксатора (4% параформальдегид в фосфатном буфере).

Для некоторых видов, например, *An. albimanus*, политенные хромосомы слюнных желез имеют более четкую исчерченность и структуру по сравнению с хромосомами из питающих клеток яичников. Поэтому в случае диссекции слюнных желез у личинки 3го возраста отделялась голова, затем две слюнные железы, симметрично расположенные по бокам в районе шеи, отделяли от сопутствующих тканей, переносили в свежий фосфатный буфер, который затем меняли на фиксатор (4% параформальдегид в фосфатном буфере).

Длительность фиксации составляла 30-60 минут с постоянным перемешиванием, после чего образцы трижды отмывали в охлажденном фосфатном буфере и переходили к следующему этапу.

2.3.2 Подготовка образцов к гибридизации

Далее в течение 20 минут инкубировали образцы в присутствии РНКазы А в концентрации 20 мкг/мкл на 37°C. Затем, проведя три отмывки в фосфатно-солевом буфере, инкубировали образцы в растворе 1% Triton X-100/0.1 М HCl в течение 20 минут на комнатной температуре для подготовки клеточных мембран к гибридизации. Далее, трижды отмыв образцы, проводили денатурацию ДНК в растворе 50% формамида в 2xSSC при 75°C в течение получаса. Помещали пробирки в лед и промывали трижды предварительно охлажденным PBS.

2.3.3 Протокол 3D-FISH (флуоресцентная иммуногибридизация)

Смешивали флуоресцентные пробы в желаемой комбинации и денатурировали при 90°C в течение 5 минут, помещали в лед и затем оставляли на 42°C на 30 минут. Добавляли 30 мкл нагретого микса из проб в пробирку с зафиксированным органом, предварительно

помещенного в 30 мкл гибридизационного буфера. Оставляли в защищенном от света месте при 42°C на 16 часов. На следующий день трижды отмывали образец в растворе 50% формамида в 2xSSC на 42°C в течение 5 минут, затем трижды в растворе 2xSSC на 42°C в течение 5 минут. Промывали образец в фосфатно-солевом буфере, добавляли в пробирку каплю защитного реагента ProLong™ Gold Antifade Mountant с DAPI (ThermoFisher #P36931), переносили орган на подготовленное стекло, сверху накрывали покровным стеклом таким образом, чтобы образец сохранял объемную структуру. Детектировали сигнал на лазерном сканирующем микроскопе Zeiss LSM 880 (Carl Zeiss AG, Oberkochen, Germany).

2.3.4 Приготовление давленных препаратов политенных хромосом, выделенных из яичников и слюнных желез

Для приготовления препаратов политенных хромосом мы следовали протоколу, предложенному в работе [332] с небольшими изменениями. Через 24-30 часов после второго или третьего кормления самок (третья стадия развития яичников по Кристоферу согласно [333]), яичники выделяли согласно описанной в пункте 2.3.1 стратегии и обрабатывали фиксирующим раствором (метанол и ледяная уксусная кислота в соотношении 3:1). После инкубации при комнатной температуре в течение 24 часов, образцы переносили в морозильную камеру и хранили на -20°C нескольких месяцев. Пробирки с образцами доставали из заморозки, отделяли индивидуальные фолликулы на предметном стекле с использованием препаровальных игл в небольшом количестве фиксирующего буфера, переносили в каплю 50% раствора пропионовой кислоты, разрушали оболочки фолликулов, позволяя хромосомам выйти из ядра. На один препарат политенных хромосом использовали 3-5 фолликулов. Во время инкубации в пропионовой кислоте отделяли сторонние ткани, сосуды, убирали жировые капли. Сменив раствор пропионовой кислоты на свежий распластывали политенные хромосомы по стеклу, накрыв покровным стеклом и попеременно надавливая на препарат. Степень распластывания контролировали под световым микроскопом. Выдерживали готовые препараты в течение 2-3 часов во влажной камере на 4°C, проводили быструю заморозку давленных препаратов в азоте, скалывали покровное стекло, помещали предметные стекла в охлажденный 50% этиловый спирт и оставляли на ночь. На следующий день проводили дегидратацию препаратов в серии последовательных разведений этилового спирта по возрастанию концентрации (50%, 70%, 90%, 96%), инкубируя в каждом по 5 минут. После 96% спирта препараты подсушивали на воздухе и оценивали качество полученных политенных

хромосом. Успешные препараты помещали в папку для хранения препаратов и использовали в FISH-экспериментах течение 3х месяцев.

2.3.5 Протокол 2D-FISH (флуоресцентная гибридизация)

Протокол проведения 2D FISH эксперимента был выбран из методического пособия под редакцией И.В.Шарахова [333] согласно разделу 4.2.1.7, Глава 4. Препараты политенных хромосом хорошего качества подготавливали к эксперименту инкубацией в однократном фосфатно-солевом буфере в течение 20 минут. Далее, их переносили в раствор 4% формалина в фосфатно-солевом буфере и выдерживали в течение 1-2 минут. Далее переносили препараты в 50% этанол и последовательно дегидрировали их в серии спиртов (70%, 80%, 90%, 96%), выдерживая в каждом растворе по 5 минут. Подсушивали препараты на воздухе, проверяли сохранность хромосом под бинокуляром.

Наносили микс из флуоресцентных проб, по 250 нг каждой из меток, на подготовленный препарат хромосом, накрывали покровным стеклом и герметично запечатывали термоклеем. Оставляли в гибридизационной камере на 42°C на ночь с предварительной денатурацией ДНК на 70-75°C в течение 10 минут. Отмывали препараты дважды в 1xSSC в течение 20 минут при 42°C, один раз в 1xSSC в течение 20 минут при комнатной температуре, один раз в PBS в течение 10 минут при комнатной температуре, один раз в воде, добавляли каплю защитного реагента ProLong™ Gold Antifade Mountant с DAPI (ThermoFisher #P36931). Проводили детекцию флуоресцентного сигнала на лазерном сканирующем микроскопе Zeiss LSM 710 (Carl Zeiss Microimaging GmbH, Oberkochen, Germany).

2.3.6 Карiotипирование клеточной линии MSQ43

Для анализа карiotипа проводили фиксацию клеток культуры MSQ43 при плотности рассадки 70-80%. Сначала клетки обрабатывали раствором коллемеида (ROCHE, Germany) с финальной концентрацией 0.1 мкг/мл в питательной среде, инкубировали в течение 60-90 минут. Затем снимали клетки с пластиковой поверхности при помощи 0.25% раствора трипсина, переносили клеточную суспензию в коническую пробирку и осаждали центрифугированием при 1,3 RPM в течение 7 минут. Удаляли супернатант, инкубировали клеточный осадок в гипотоническом растворе 0.56 % KCl при 37°C в течение 25 минут. Добавляли 100-150 мкл свежего фиксирующего раствора (метанол и ледяная уксусная кислота в соотношении 3:1) и осаждали клетки центрифугированием. Удаляли супернатант и добавляли 2-3 мл свежего, предварительно охлажденного, фиксирующего раствора, не ресуспендировали клетки. Инкубировали клеточный осадок на льду в течение 20 минут,

после чего аккуратно пипетировали и осаждали клетки центрифугированием. Повторяли процесс фиксации дважды, инкубируя клетки на льду в течение 10 минут. Наконец осаждали клеточный осадок и ресуспендировали в небольшом количестве фиксирующего буфера (200-300 мкл), при правильной концентрации наблюдая опалесценцию клеточной суспензии. Готовили препараты метафазных пластинок, раскапывая суспензию на охлажденные предметные стекла, инкубировали надо льдом в течение 2-3 минут, окончательно подсушивали в потоке теплого воздуха. Предварительно оценивали целостность и количество метафазных пластинок. Проводили окрашивание хромосом красителем DAPI, после чего делали снимки метафазных хромосом на микроскопе Carl Zeiss AxioPlan 2 (Jena, Germany) в ЦКП ИЦиГ СО РАН. Для анализа кариотипа клеточной линии MSQ43 выполняли подсчет 100 метафаз на препарат, анализировали 4 предметных стекла (Рисунок 8, В - Е).

2.4 Сборка геномов

Используя результаты Hi-C-эксперимента и данные физического картирования, мы собрали *de novo* три генома до хромосомного уровня (*An. coluzzii Mopti*, *An. merus*, *An. stephensi*) и улучшили существующие хромосомные сборки для *An. albimanus* и *An. atroparvus*. В качестве референсных геномов были использованы предыдущие геномные сборки [297,334].

Исходно, в программном обеспечении Juicer происходило выравнивание прочтений на референсный геном [100] и автоматическая полировка полученных геномных сборок в программе 3D-DNA [101]. Далее, мы проводили визуальную инспекцию, необходимую ручную корректировку ошибок в сборках и реконструкцию положения контигов и скаффолдов в программе Juicebox в режиме Assembly mode [335] в соответствии с данными о пространственных контактах, а также сопоставление полученных сборок с результатами физического картирования, описанными ранее [297,334,336–338]. Взаиморасположение хромосомных плеч задавалось данными физического картирования. Созданные таким образом геномы затем служили референсными при финальном запуске программы Juicer, где первичные Hi-C-прочтения еще раз выравнивались на обновленный геном. В результате такого алгоритма действий мы получили тепловые карты пространственных контактов и сборки геномов хромосомного уровня высокого качества для пяти представителей рода *Anopheles* (более наглядно данный подход представлен на рисунке 15).

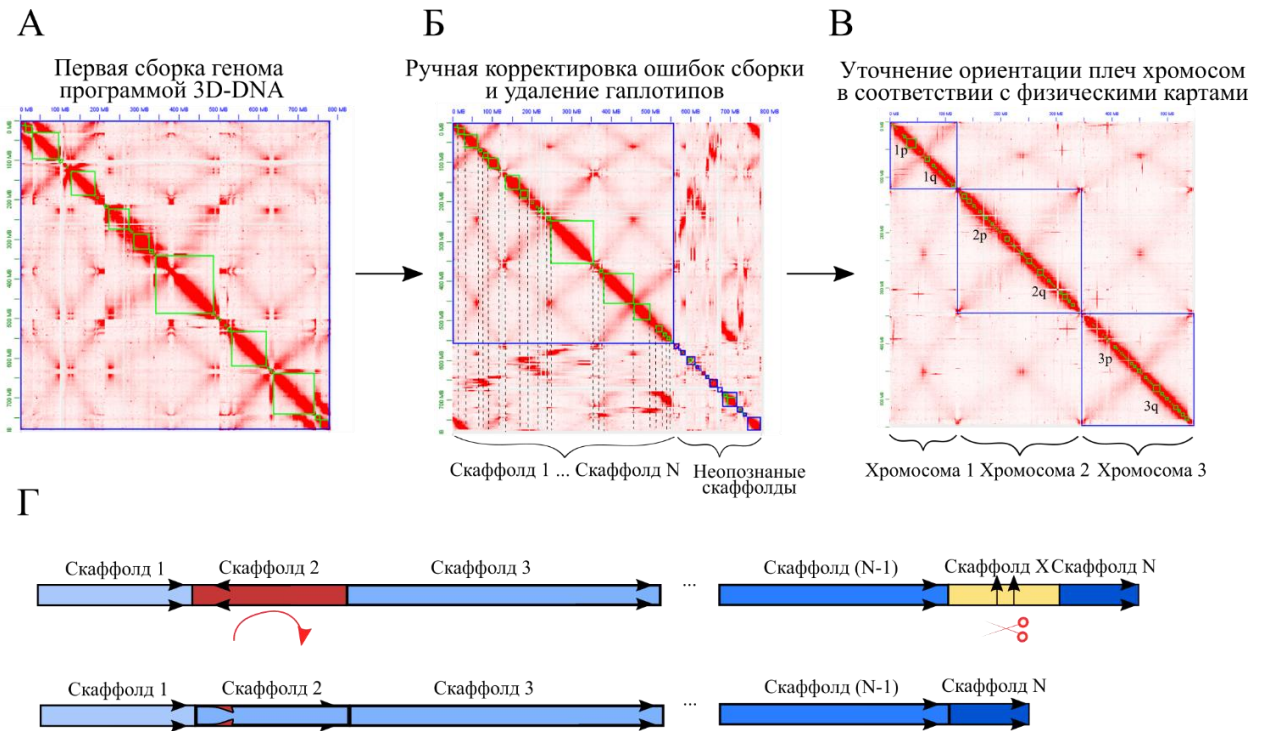


Рисунок 15. Алгоритм сборки генома на основе данных Hi-C. А. Черновой вариант сборки после запуска программы 3D-DNA. Б. Ручная корректировка ошибок в стыковке скаффолдов и удаление гаплотипов. В. Уточнение ориентации хромосомных плеч в соответствии с физическими картами. Г. Схема корректировки геномной сборки в программе Juicebox. «Федеральный

ГЛАВА 3. РЕЗУЛЬТАТЫ

3.1 Оптимизация протокола Hi-C для эмбрионов комаров рода *Anopheles*. Оценка качества и характеристика полученных библиотек

Для исследования трехмерной организации геномов анофелесов, мы провели Hi-C-эксперимент для каждого из пяти видов комаров, выбранных в исследование. За основу были взяты протоколы *in situ* Hi-C из [119] и Hi-C-метода согласно [4]. Краткая схема эксперимента, для которого мы использовали 15-18-часовых эмбрионов комаров, изображена на рисунке 16, А-В. Поскольку размер кладок отличается в зависимости от возраста самок и порядка кормления, мы брали в эксперимент яйца, отложенные после 2го или 3го кормлений, получая в сумме несколько тысяч комариных эмбрионов. Спустя 2-3 дня после кормления кровью, мы оставляли в клетке емкость с водой, покрытую фильтровальной бумагой, для сбора эмбрионов (Рисунок 7, А). На следующий день извлекали их из инкубатора и переходили к Hi-C-эксперименту. В силу специфики работы с комариными эмбрионами мы внесли несколько технических модификаций в оригинальный Hi-C-протокол.

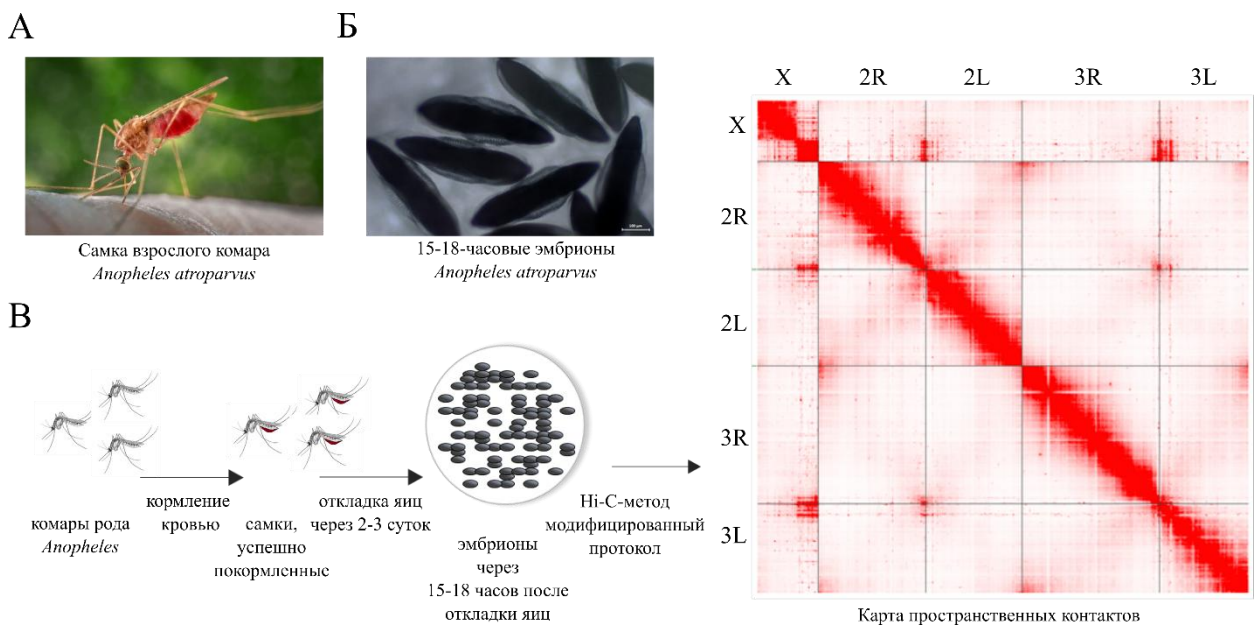


Рисунок 16. Дизайн Hi-C-эксперимента. А и Б – женская взрослая особь и эмбрионы *Anopheles atroparvus*. В. Последовательность действий в Hi-C эксперименте с эмбрионами малярийных комаров. Изображение взрослой самки *An.atroparvus* заимствовано с [284].

Первое важное изменение – это предобработка комариных яиц 3%-м раствором гипохлорита натрия для разрушения поверхностных мембран экзохориона (Рисунок 17, А-В), чтобы обеспечить лучшую доступность тканей эмбриона во время фиксации

параформальдегидом. 15-18-часовые эмбрионы помещались в 3%-раствор гипохлорита натрия на 8-10 минут, с постоянным наблюдением за целостностью оболочек под бинокляром. В момент, когда ~80% яиц подверглись химическому воздействию (осветленные пятна покрыли поверхность эмбриона достаточно равномерно, но содержимое оставалось внутри), мы трижды отмывали яйца в чистой дистиллированной воде, удаляя остатки жидкости с помощью фильтровальной бумаги. Затем переходили к следующему этапу.

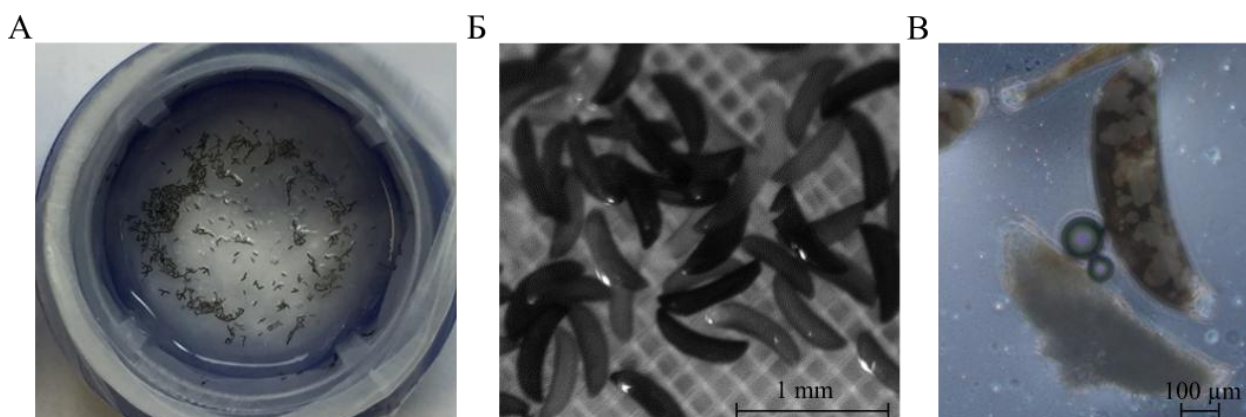


Рисунок 17. Предобработка комариных яиц 50%-раствором хлорной извести для разрушения поверхностных мембран экзохориона: А) эмбрионы до фиксации; Б и В) в процессе обработки появляется осветление защитных покровов.

Второе изменение - это совмещение этапов фиксации и лизиса клеток. Мы подготовили свежий буфер для фиксации, перенесли и гомогенизировали эмбрионы в 2 мл-гомогенизаторе Даунса, последовательно пропустили через 100, 60, 10 мкм нейлоновый фильтр, инкубируя в фиксирующем буфере в течение 25 минут. Затем инактивировали параформальдегид добавлением глицина до конечной концентрации 0.4 М. Отмыли ядра в буфере без PFA дважды, заморозили осадок быстрой заморозкой в азоте или на сухом льду и хранили на -80°C до начала эксперимента. Далее следовали протоколу, который подробно описан в главе Материалы и методы. В результате мы получили подготовленные для полногеномного секвенирования Hi-C-библиотеки в двух репликах для каждого из 5 видов. Первичная качественная оценка эффективности прохождения основных молекулярных этапов Hi-C-протокола была проведена с использованием гель-электрофореза в агарозном геле, результаты которой приведены на Рисунке 11, Б и В (раздел Материалы и Методы).

Количество прочтений, полученных после NGS-секвенирования собранных библиотек, с указанием доли уникальных и Hi-C-прочтений (содержащих сайт узнавания

MboI в середине фрагмента, но не расположенных линейно друг за другом по геному) представлено в Таблице 4.

Таблица 4. Статистика прочтений и первичной обработки в программном обеспечении Juicer.

* стандартом в данном случае мы считали показатели из протокола Hi-C ENCODE для категории “pass”, проценты указаны для уникальных прочтений (второй показатель процентов в скобках, выделен полужирным шрифтом)

Вид	<i>An. albimanus</i> AalbS4	<i>An. atroparvus</i> AatrE4	<i>An. coluzzii</i> <i>Mopti</i> AcolN2	<i>An. merus</i> эмбрионы AmerM5	<i>An. merus</i> взрослые AmerM5	<i>An. stephensi</i> AsteI4	MSQ43	Стандарт*
Σ прочтений	67,591,619	217,895,615	231,961,292	188,700,705	129,643,770	253,760,093	88,422,839	-
Уникальные прочтения	56,129,336 (83.04%)	150,572,935 (69.10%)	165,637,626 (71.41%)	114,668,816 (60.77%)	77,371,708 (59.68%)	135,257,540 (53.30%)	60,102,741 (67.97%)	-
Hi-C прочтения	52,985,219 (78.39% / 94.40%)	136,465,823 (62.63% / 90.63%)	125,215,908 (53.98% / 75.60%)	86,933,619 (46.07% / 75.81%)	60,823,869 (46.92% / 78.61%)	112,540,526 (44.35% / 83.20%)	48,153,488 (54.46% / 80.12%)	>50%
Межхромосомные контакты	6,862,999 (10.15% / 12.23%)	24,255,276 (11.13% / 16.11%)	21,566,658 (9.30% / 13.02%)	29,961,382 (15.88% / 26.13%)	6,270,919 (4.84% / 8.10%)	18,456,987 (7.27% / 13.65%)	7,840,863 (8.87% / 13.05%)	<40%
Внутрихромосомные контакты	46,122,220 (68.24% / 82.17%)	112,210,547 (51.50% / 74.52%)	103,649,250 (44.68% / 62.58%)	56,972,237 (30.19% / 49.68%)	54,552,950 (42.08% / 70.51%)	94,083,539 (37.08% / 69.56%)	40,312,625 (45.59% / 67.07%)	>40%
Близкие взаимодействия (<20 Кб)	13,468,625 (19.93% / 24.00%)	25,525,319 (11.71% / 16.95%)	29,262,843 (12.62% / 17.67%)	13,002,707 (6.89% / 11.34%)	16,850,740 (13.00% / 21.78%)	37,468,994 (14.77% / 27.70%)	17,433,250 (19.72% / 29.01%)	<30%
Дальние взаимодействия (>20 Кб)	32,653,284 (48.31% / 58.18%)	86,684,768 (39.78% / 57.57%)	74,385,329 (32.07% / 44.91%)	43,969,011 (23.30% / 38.34%)	37,702,135 (29.08% / 48.73%)	56,613,127 (22.31% / 41.86%)	22,879,329 (25.87% / 38.07%)	>35%

Оценивая качество полученных библиотек, мы ориентировались на стандарты, принятые Геномным центром Hi-C ENCODE в 2018 году [339]. По количеству выровненных на геном Hi-C-прочтений - тех фрагментов, которые содержат сайт узнавания рестриктазы и выравниваются на отдаленные, линейно не сближенные, геномные локусы - библиотеки для эмбрионов *An. albimanus* и *An. atroparvus* обладают наиболее высокими показателями и составляют 78% и 62% соответственно. Процент межхромосомных, а также внутрихромосомных ближних и внутрихромосомных дальних взаимодействий удовлетворяет принятым высоким стандартам у всех собранных библиотек, но наиболее успешен для эмбрионов *An. albimanus* и *An. atroparvus*. Ознакомиться с детальной статистикой библиотек можно в таблице в Приложении 2.

3.2 Улучшение существующих сборок геномов для пяти видов комаров рода *Anopheles*

По результатам NGS-секвенирования были получены первичные Hi-C-прочтения собранных библиотек. Далее, в результате алгоритма действий, описанного в главе Материалы и методы, мы выравнивали полученные прочтения на референсные контиги, объединили отдельные контиги в скаффолды, а скаффолды - в хромосомные плечи (X, 2R, 2L, 3R, 3L), которые затем были ориентированы в хромосомы на основании их пространственных контактов и данных по физическому картированию [297,334,336–338]. Множественные ошибки в сборках были устранены, а вариации гаплотипов перенесены в отдельный скаффолд. Качество сборок было подтверждено алгоритмом BUSCO, суммарной длиной фрагментов генома, собранных в хромосомы, а также показателем N50 (Таблица 5). Таким образом, мы получили тепловые карты пространственных контактов и обновленные сборки геномов хромосомного уровня высокого качества для пяти представителей рода *Anopheles* (Рисунок 18, А - Д). Последние были отправлены в международную базу данных NCBI (BioProject: PRJNA660041, геномные сборки: JADFFJ000000000, JADGIR000000000, JADGIQ000000000, JADFFN000000000, JADFFO000000000).

Стоит подчеркнуть, что геном AgamP4 PEST, долгое время считавшийся эталонным для представителей комплекса *An.gambiae* (*An. coluzzii*, *An. gambiae*, *An. arabiensis*, *An. quadriannulatus*, *An. melas*, *An. merus*, *An. bwambae*) и выбранный в качестве референса для *An. coluzzii*, содержал множество ошибок в состыковке контигов внутри крупных скаффолдов (>50), а также неверно ориентированных фрагментов. По этой причине для уточнения положения контигов мы использовали данные PacBio-секвенирования -

технологии прочтения длинных геномных фрагментов, недавно полученные нашими коллегами для другой колонии вида *An. coluzzii* - Ngouso [287], являющейся ближайшим родственником *An. gambiae* [340]. Мы использовали скаффолды, полученные на основе данных секвенирования PacBio, в качестве черновой версии генома, на которую картировали полученные нами прочтения Hi-C. Таким образом, мы смогли ориентировать и выстроить PacBio-скаффолды согласно их положению на хромосомах. В результате геномная реконструкция на основе Hi-C и PacBio-данных оказалась значительно более аккуратной и приемлемой для дальнейшего использования по сравнению со сборкой PEST, реконструированной при помощи Hi-C без использования длинных прочтений.

В случае *An. merus* в нашем распоряжении также оказались данные PacBio, предоставленные нашими коллегами из университета Вирджинии, что позволило выровнять данные Hi-C на длинные прочтения и повысить качество геномной сборки. Для остальных геномов были использованы существующие на тот момент сборки высшего уровня, хромосомного для *An. albimanus* и *An. atroparvus* [297,334] и скаффолдного для *An. stephensi* [341]. Таким образом, геномы для *An. coluzzii*, *An. stephensi* и *An. merus* были собраны *de novo*, а для *An. albimanus* и *An. atroparvus* существенно улучшены. В приведенной ниже Таблице 5 представлены сравнительные характеристики для полученных геномныхборок.

Таблица 5. Характеристики полученных геномных сборок по сравнению с предыдущими геномами.

* не учитывалась митохондриальная ДНК и скаффолды короче 200 пн.

** анализ BUSCO был выполнен сотрудниками лаборатории Роберта Уотерхауса (Лозанна, Швейцария)

Вид	Название генома	Общая длина скаффолдов*	Длина генома в хромосомах	N50 для скаффолдов	BUSCO**
<i>An. albimanus</i>	AalbS4/ AalbS2	170,170,859/ 170,336,140	169,445,844/ 167,376,415	5,915,397/ 9,735,467	96.5
<i>An. atroparvus</i>	AatrE4/ AatrE3	220,078,729/ 224,290,125	216,747,066/ 200,912,972	688,000/ 9,206,694	99.1
<i>An. coluzzii</i>	AcolN2/ AcolN1	245,361,604/ 251,414,185	231,601,574/ -	3,835,000/ 3,468,756	96.4
<i>An. merus</i>	AmerM5/ AmerM4	300,125,122/ 300,704,392	234,366,716/ -	3,903,996/ 2,729,089	95.8
<i>An. stephensi</i>	AsteI4/ AsteI2	220,673,503/ 221,324,304	196,394,606/ -	980,283/ 1,631,802	97.7

На Рисунке 18 показана визуализация карт пространственных контактов, полученных методом Hi-C, снимки политенных хромосом, а также приведен характерный размер генома для пяти представителей рода *Anopheles*. Так, финальные версии генома состоят из 245 млн.п.н. для *An.coluzzii*, 300 млн.п.н. для *An.merus*, 221 млн.п.н. для *An.stephensi*, 220 млн.п.н. для *An.atroparvus*, 170 млн.п.н. для *An.albimanus*, соответственно. Значения показателя BUSCO варьируют в диапазоне 95.8-99.1% для наших пяти видов, тогда как общепринятым стандартом геномной сборки хорошего качества считаются показатели со значением >95% [342].

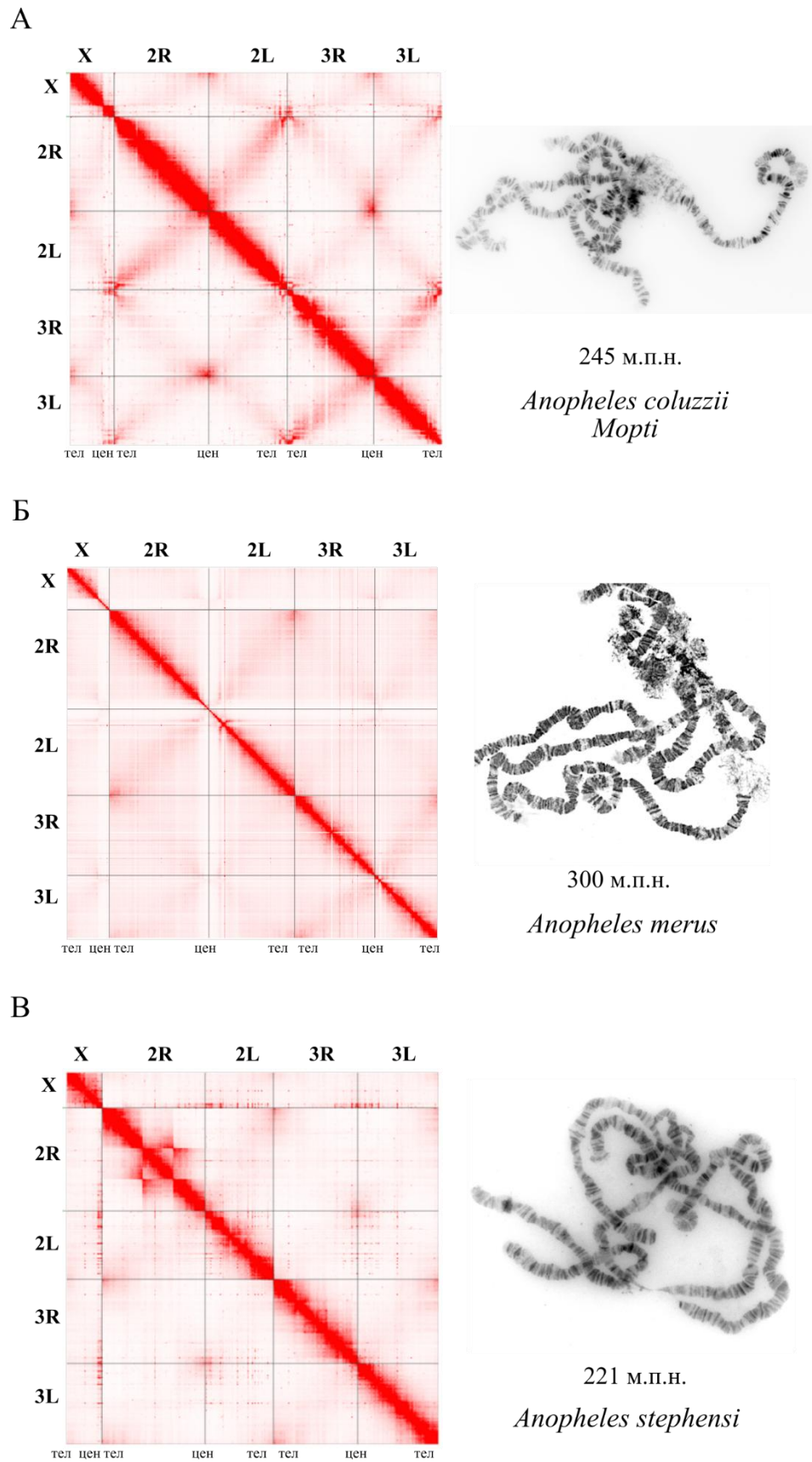


Рисунок 18. Визуализация полученных карт пространственных контактов, снимки политенных хромосом, а также характерный размер генома для пяти представителей рода *Anopheles*. А. *An. coluzzii* Mopti; Б. *An. merus*, изображение политенных хромосом адаптировано из [343]; В. *An. stephensi*. * Цветовой бегунок выставлен на значение 2400 на всех представленных картах контактов.

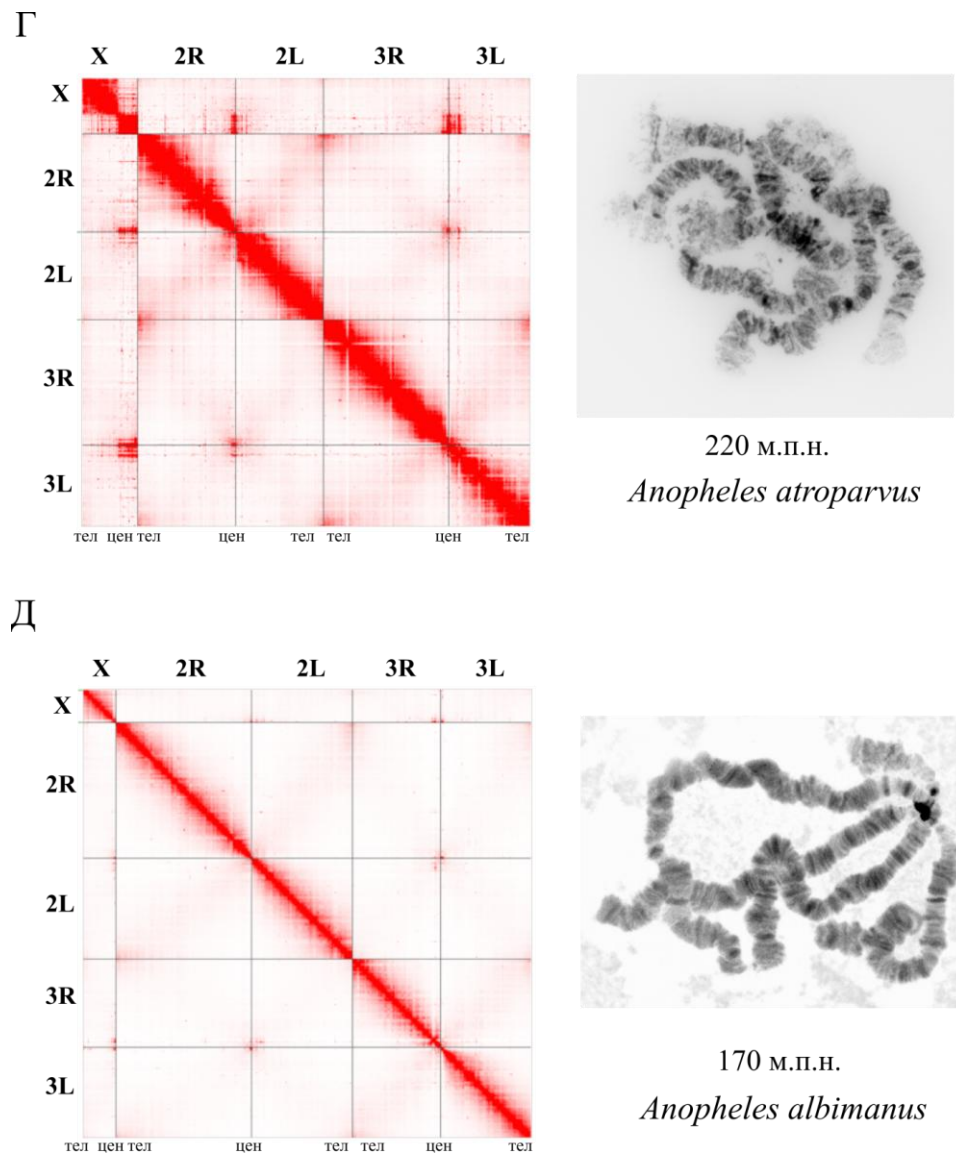


Рисунок 18 (продолжение). Визуализация полученных карт пространственных контактов, снимки политенных хромосом, а также характерный размер генома для пяти представителей рода *Anopheles*. Г. *An. atroparvus*; Д. *An. albimanus*. * Цветовой бегунок выставлен на значение 2400 на всех представленных картах контактов.

3.3 Детекция ранее известных и новых хромосомных перестроек с помощью данных Hi-C. Полиморфные и фиксированные инверсии у представителей родов *Anopheles* и *Aedes*

Малярийные комары представляют удобную биологическую модель для изучения инверсионных полиморфизмов, так как известно, что анофелесы способны накапливать большое количество хромосомных перестроек. По данным Аяла и коллег в 2014 году было известно о существовании 49 инверсионных полиморфизмов среди 8 видов малярийных комаров, которые оказывают влияние на экологическую и адаптивную приспособленность их носителей [293]. Поскольку метод Hi-C дает информацию о пространственных контактах всего генома, хромосомные перестройки, и, в частности, инверсии, очень заметны на

тепловых картах и имеют характерный рисунок/паттерн в виде бабочки (Рисунок 19, А). Ранее была описана возможность идентификации сбалансированных инверсионных перестроек с использованием карт пространственных контактов, где точки разрывов определяли по пересечению перпендикуляров, опущенных из центра инверсионной “бабочки”, с главной диагональю [344] (пример показан на Рисунке 19, А, Д, З, И). Используя предложенную стратегию, мы охарактеризовали инверсионные полиморфизмы, специфичные для малярийных комаров в диссертационном исследовании, а также для нескольких колоний *Aedes aegypti*.

3.3.1 Характеристика инверсий у *Anopheles*

Среди пяти видов *Anopheles* мы идентифицировали 4 инверсии, детектируемые на тепловых картах пространственных контактов и имеющих размер в диапазоне от 2.8 Мб до 16 Мб (Таблица 6, Рисунок 19, А, Г, Ж, З).

Таблица 6. Обнаруженные полиморфные инверсии и их координаты среди пяти видов *Anopheles*. * - обозначение инверсионных перестроек, описанных впервые в данном исследовании.

Вид	Колония	Наименование	Геномные координаты	Размер
<i>An. stephensi</i>	Indian	2Rb	2R: 21,140,000 - 37,180,000	16 Мб
<i>An. atroparvus</i>	EBRO	2L1*	2L: 2,980,000 - 13,940,000	11 Мб
<i>An. coluzzii</i>	MOPTI	2Ru	2R: 30,940,000 - 35,340,000	4.4 Мб
<i>An. merus</i>	MAF	2Rt*	2R: 60,175,000 - 62,930,000	2.8 Мб

3.3.1.1 Полиморфная 2Rb-инверсия у *An. stephensi*

Согласно обзорному исследованию Аяла и коллег, для *An. stephensi* описано 16 инверсионных перестроек, имеющих влияние на жизнеспособность комаров в вариативных природных условиях, а также при обработке инсектицидами [293]. Детально анализируя карту пространственных контактов, полученную для лабораторной колонии индийского подвида *An. stephensi*, мы обнаружили паттерн–бабочку на участке 2R-плеча (Рисунок 19, А), характерный для инверсионных структур [344]. Отложив два перпендикуляра в направлении основной диагонали, мы получили точки разрывов для данной инверсии с разрешением +/- 5 Кб и определили координаты перестройки как 2R: 21,140,000 - 37,180,000. Сопоставив популяционные данные по ранее описанным инверсиям для *An.*

stephensi, мы соотнесли обнаруженную инверсию с ранее известной инверсией 2Rb [337,345,346], которая имеет характерный размер 16 Мб и является наиболее распространенной инверсионной перестройкой у *An. stephensi*. По данным Колюцци и коллег, полученных еще в 1973 году, частота встречаемости инверсий в популяциях существенно варьирует в зависимости от региона [345] и отличается между урбанизированными и дикими природными популяциями [346]. Для носителей 2Rb-инверсии была отмечена разница в продолжительности периода активности по сравнению с диким типом [347], а также чувствительности к альфаметрину в экспериментах на инбредных линиях по устойчивости к различным инсектицидам [348]. Стоит подчеркнуть полиморфную структуру данной перестройки внутри лабораторной колонии (Рисунок 19, А - Г).

3.3.1.2 Новая 2L-инверсия у *An. atroparvus*

An. atroparvus относится к видам с низким инверсионным полиморфизмом. В середине прошлого века для данного комара были описаны несколько инверсий, расположенных на хромосоме 3 и на длинном плече хромосомы 2 (2R), предположительно ассоциированных с устойчивостью к ДДТ и дильдрину [349,350]. В природных популяциях с невысокой частотой встречается единственная инверсионная перестройка на 3L-плече, на участке 34a-38c [351], хотя последние данные для Ставропольского края свидетельствуют о частотах $33,3 \pm 9,1\%$ [352], а для региона Калмыкия - от $20,6 \pm 7,0\%$ до $59,1 \pm 10,7\%$ [353]. Стоит отметить, что в Калмыкии инверсия наблюдается как в гомозиготном, так и гетерозиготном состояниях [353]. Поскольку для *An. atroparvus* публикаций по инверсионным полиморфизмам не так много, стоит указать, что у близкородственного вида *An. messeae* из того же комплекса видов-сиблингов *An. maculipennis*, для которого характерно большое количество инверсионных перестроек в природных популяциях, инверсии описаны для всех хромосом, кроме короткого плеча второй хромосомы (2L). Карта пространственных контактов хроматина, построенная нами для эмбрионов *An. atroparvus* колонии EBRO, показала наличие выраженного инверсионного полиморфизма на хромосомном плече 2L. Точная локация перестройки была определена как 2L: 2,980,000 - 13,940,000, а размер составил 11 Мб. Попытка искусственного вращения инвертированного участка в приложении Juicebox не повлияла на положение характерного паттерна-бабочки, что свидетельствует о наличии геномных фрагментов как в нормальной ориентации, так и в инвертированной. Анализ цитологических препаратов политенных хромосом, полученных от нескольких особей внутри одной колонии EBRO (*An. atroparvus*), позволил выявить особей с инвертированной ориентацией участка 2L и без, что

свидетельствует о полиморфном характере наблюдаемой инверсии и согласуется с нашими выводами на основе анализа тепловых карт пространственных взаимодействий. Данная инверсия была выявлена и описана впервые в представленной диссертационной работе (Рисунок 19, Д - Е).

3.3.1.3 Полиморфная 2Ru-инверсия у *An. coluzzii*

Для комаров комплекса *An. gambiae*, куда входит *An. coluzzii*, инверсионные полиморфизмы описаны наиболее полно, поскольку специализация отдельных подвидов комплекса, вероятно, происходила в результате возникновения и закрепления инверсионных полиморфизмов [306,354]. Так, у представителей комплекса наиболее известная перестройка – инверсия 2La - влияет на приспособленность комаров к засушливым условиям [301,302,355,356], термоустойчивость личинок 4го возраста [357,358], повышенную устойчивость гетерозигот к инсектицидам, в частности блокирующим рецепторы аминокислотной кислоты (фипронила и дильдрин) [359], а также на способность к переносу малярийного плазмодия [360]. Группа инверсий в локусе 2R у представителей комплекса *An. gambiae* - 2Rb, -c, -d, -j, -u и -s -, также вовлечена в различные изменения биохимических свойств у своих носителей (из обзора [293]). Анализируя карту пространственных контактов для *An. coluzzii* колонии МОПТИ мы детектировали инверсию в локусе 2R-хромосомного плеча, которая по своему положению очень напоминала перестройку 2Ru, выявленную ранее и ассоциированную с географическим положением и влажностью климата [361]. Согласно результатам нашего эксперимента, инверсия 2Ru имеет координаты 2R: 30,940,000 - 35,340,000, что составляет 4.4 Мб в длину, и представляет полиморфный тип хромосомной перестройки (Рисунок 19, З). Выраженность инверсии небольшая, сопоставима по интенсивности с перестройкой у *An. atroparvus* на 2L (Рисунок 19, Д).

3.3.1.4 Новая 2Rt-инверсия у *An. merus*

В прицентромерном локусе длинного плеча 2й хромосомы у *An. merus* мы обнаружили небольшую инверсионную перестройку размером 2.8 Мб и геномными координатами - 2R: 60,175,000 - 62,930,000. Упоминание о данной инверсии не встречается в литературных данных. Выраженность инверсионной структуры на карте пространственных контактов довольно слабая, что может свидетельствовать о низкой частоте встречаемости в популяции (Рисунок 19, И). Но, поскольку инвертированный регион примыкает к центромерной области, где качество сборки существенно снижено по сравнению со средним по геному, данная находка требует подтверждения другим независимым методом, например, флуоресцентной гибридизацией *in situ*.

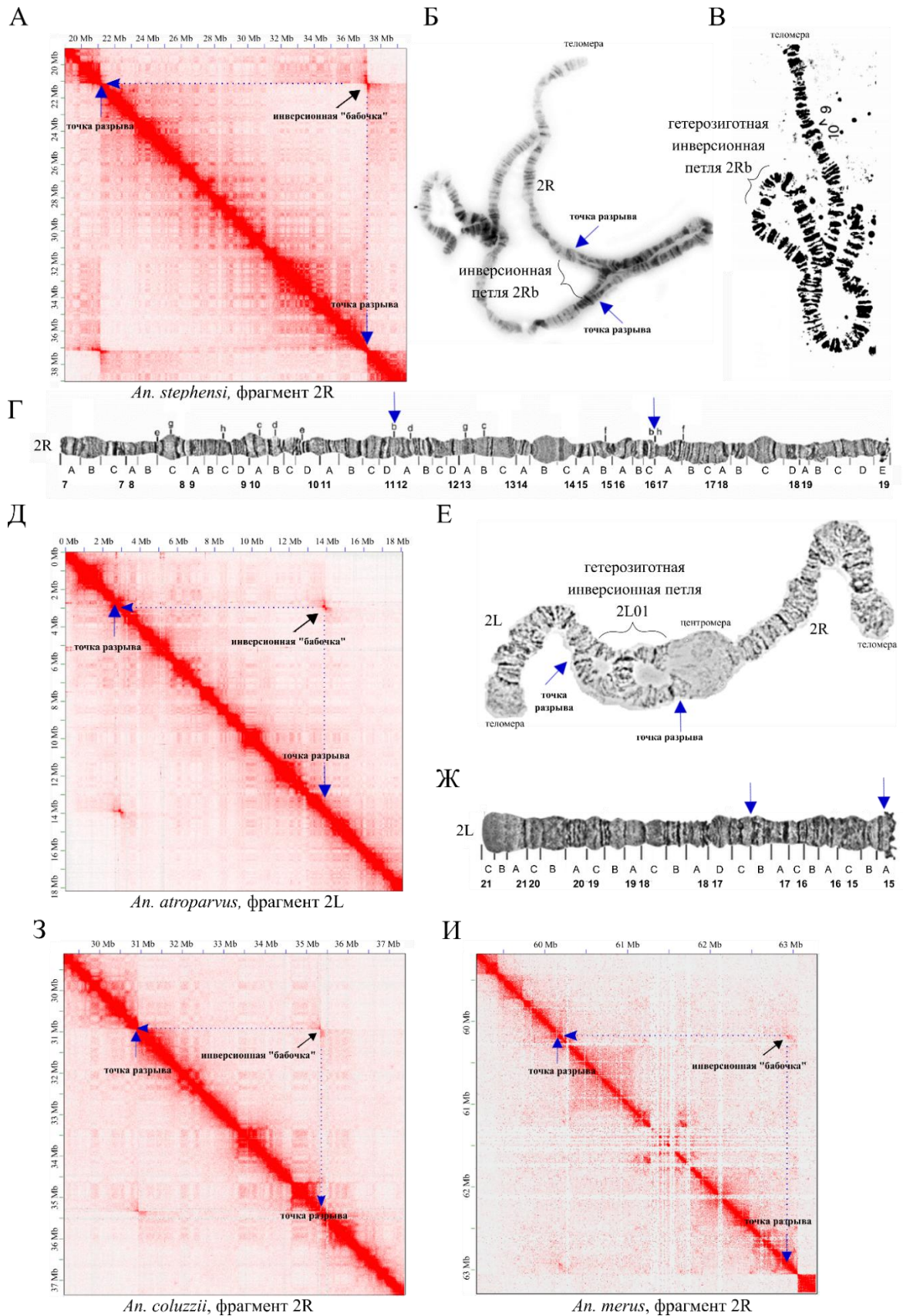


Рисунок 19. А, Д, З, И – визуализация инверсий у малярийных комаров, выявленных при помощи метода Hi-C. Б, Е – цитологические препараты политенных хромосом с демонстрацией инверсионной петли. В – изображение инверсионной петли 2Rb, адаптированное из работы Coluzzi et al., 1973. Г и Ж – физические карты политенных хромосом с указанием инверсионных точек разрыва.

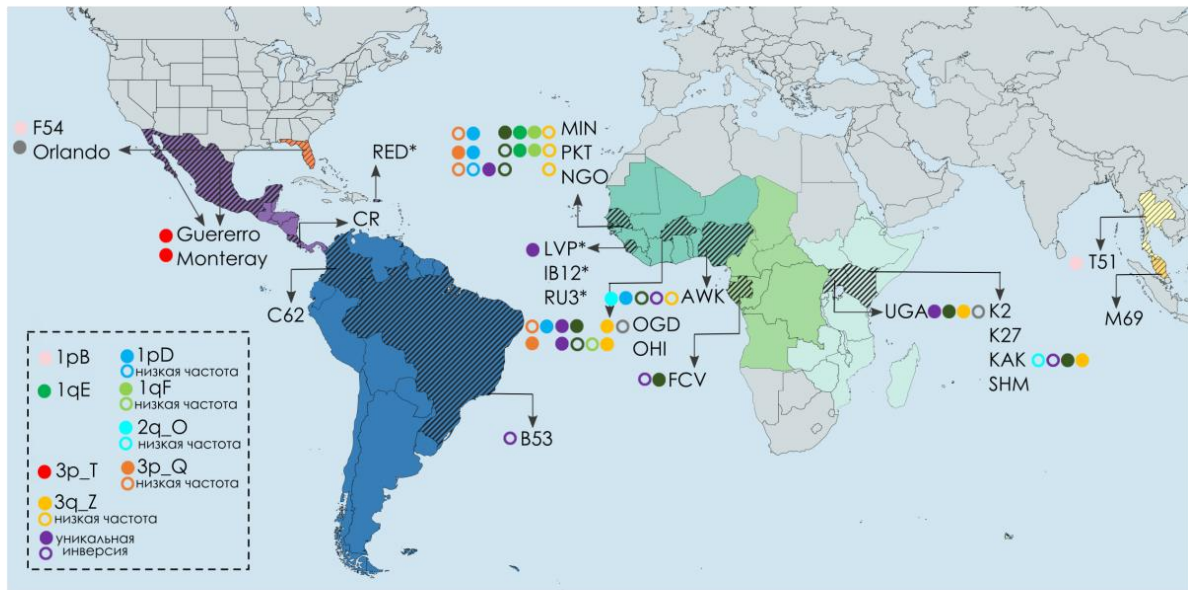
3.3.2 *Полиморфные и фиксированные инверсии у различных представителей рода Aedes. Инверсионный полиморфизм как фактор формирования пищевых и экологических предпочтений*

Как было подробно описано в Литературном обзоре, у комаров рода *Anopheles* инверсионные перестройки являются драйверами адаптационных процессов. Предположив, что инверсионный полиморфизм может играть важную роль и у представителей рода *Aedes*, мы проанализировали 25 карт пространственных контактов, полученных для 12 колоний с Африканского континента, семи колоний Нового Света (Бразилия, Колумбия, Коста Рика, США, Мексика), двух азиатских и четырех лабораторных колоний, имеющих разные пищевые и экологические предпочтения (Африканские колонии показаны на Рисунке 20, А). В результате мы обнаружили 25 хромосомных инверсий, часть из которых оказалась уникальной, а часть – повторяющейся у нескольких колоний (Таблица 7). Размер описанных хромосомных перестроек варьировал в диапазоне от 0,985 Мб до 54,775 Мб, встречались как полиморфные, так и фиксированные варианты. Наибольшее количество инверсий наблюдалось у западноафриканских колоний: OGD (8), ОН1 (7), РКТ (6), MIN (6), NGO (5) и АWK (5), из которых OGD и NGO – антропофильны. В некоторых случаях на месте инверсионной “бабочки” мы видели точечные фокусы пространственных контактов, при этом их основания полностью совпадали с границами инверсий, наблюдаемых в других колониях. Поскольку частоты инверсионных полиморфизмов могут варьировать в зависимости от сезона и конкретной популяции насекомых [345], мы предполагаем, что низкочастотные перестройки на тепловых картах контактов могут иметь вид точечных взаимодействий. В Таблице 7 эти случаи отмечены буквой К – контакт. Для колоний MIN и Uganda в нашем распоряжении оказались Hi-C-данные для эмбрионов и для взрослых комаров. Стоит отметить, что материал, полученный с эмбриональных фиксаций, имеет более четкий инверсионный паттерн на картах контактов по сравнению с картами от взрослых особей.

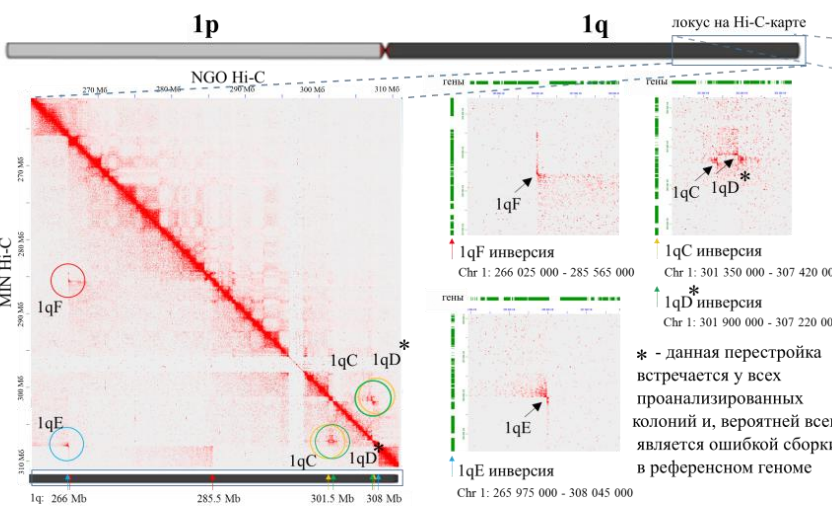
Наиболее интересными оказались инверсионные перестройки на q-плече 1й хромосомы: 1qE, 1qF, 1qC, 1qD, 1qH. Дело в том, что в работах [362] и [363] были определены геномные локусы, предположительно ответственные за выбор человека в качестве источника питания (Рисунок 20, Б). Согласно схеме, приведенной в вышеуказанной статье, на участке ~240-305 Мб 1й хромосомы расположен крупный генный кластер, связанный с распознаванием запахов, в том числе человека, и ассоциированный со сменой типа питания. Наши результаты Hi-C-эксперимента выявили, что точки разрывов инверсий 1qE, 1qF, 1qC, 1qD и 1qH попадают как раз в этот геномный участок (Рисунок 20, Б и В). 1qE и 1qF являются уникальными для колоний MIN и РКТ, 1qD была обнаружена у

OGD, PKT, NGO и MIN, 1qH – у OGD, а инверсия 1qD – у всех 25 исследованных колоний (предположительно – ошибка сборки). Специфичной для мексиканских колоний оказалась довольно крупная инверсия на 3p-плече размером 16,65 Мб. Африканские колонии, за исключением FCV, SHM, K2 и K27, являются носителями небольшой инверсии на 3q-плече размером 5,2 Мб. У западноафриканских колоний (MIN, NGO, PKT, OGD и OHI) была обнаружена инверсионная перестройка 3pQ размером 9.05 Мб.

А



Б



В

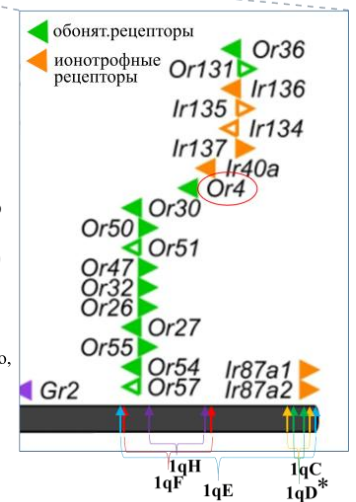


Рисунок 20. Инверсии у представителей рода *Aedes*. А. География распространения колоний *Ae. aegypti*, для которых были построены карты пространственных контактов и охарактеризованы инверсионные полиморфизмы. Цветом отмечены регионы, где были собраны образцы; штриховкой отмечены страны сбора. Наличие перестроек обозначено цветными кругами напротив наименования колонии. Б. Инверсии 1qC, 1qD, 1qF и 1qE на картах пространственных контактов, построенных в сравнении для колоний MIN (под основной диагональю) и NGO (над основной диагональю). В. Пересечение инверсий 1qC, 1qD, 1qF и 1qE с кластером обонятельных и ионотрофных рецепторов на 1q-плече, рисунок частично адаптирован из Matthews et al., 2018.

контактах (Рисунок 27, А, межцентромерные контакты выделены серыми овалами, а межтеломерные - желтыми). Кроме того, мы наблюдали появление слабовыраженных дополнительных диагоналей, перпендикулярных основной и пересекающих ее в центромерных областях и по краям хромосом. Мы считаем, что наблюдаемый факт является следствием взаимодействия плеч хромосом и целых хромосом (на Рисунке 27, А, перпендикулярные диагонали выделены красными и синими пунктирными прямоугольниками).

Для проверки существования данной особенности организации ядра мы применили независимый подход флуоресцентной *in situ* гибридизации (FISH) на фиксациях яичников малярийных комаров, выделенных через сутки после кормления. Данный выбор был обусловлен наличием нескольких клеточных типов (классические интерфазные ядра фолликулярного слоя, питающие клетки с политенными хромосомами с разной степенью политении), а также легкостью выделения органа и приготовления препаратов. Структура яичника комара, а также строение вторичного фолликула представлены на Рисунке 21, панель А и Б, соответственно. Избирательно пометив последовательности аутосомных и X-хромосомных центромер (пробы Aste72A and Aste190A - наиболее распространенные повторенные последовательности в центромерном гетерохроматине половых хромосом и аутосом *An. stephensi*, соответственно, согласно данным из [341]), мы провели гибридизацию проб методом 3D FISH, с сохранением структуры ткани яичника, после чего проверили локализацию меток на сканирующем конфокальном микроскопе Zeiss LSM 880 (Carl Zeiss AG, Oberkochen, Germany). Результаты эксперимента для комара *An. stephensi* представлены на Рисунке 21, панель В, где визуализируется колокализация центромерных проб, характерная как для интерфазных ядер фолликулярного эпителия, так и для политенных ядер. Наглядная схема данной конфигурации приведена на Рисунке 27, Б.

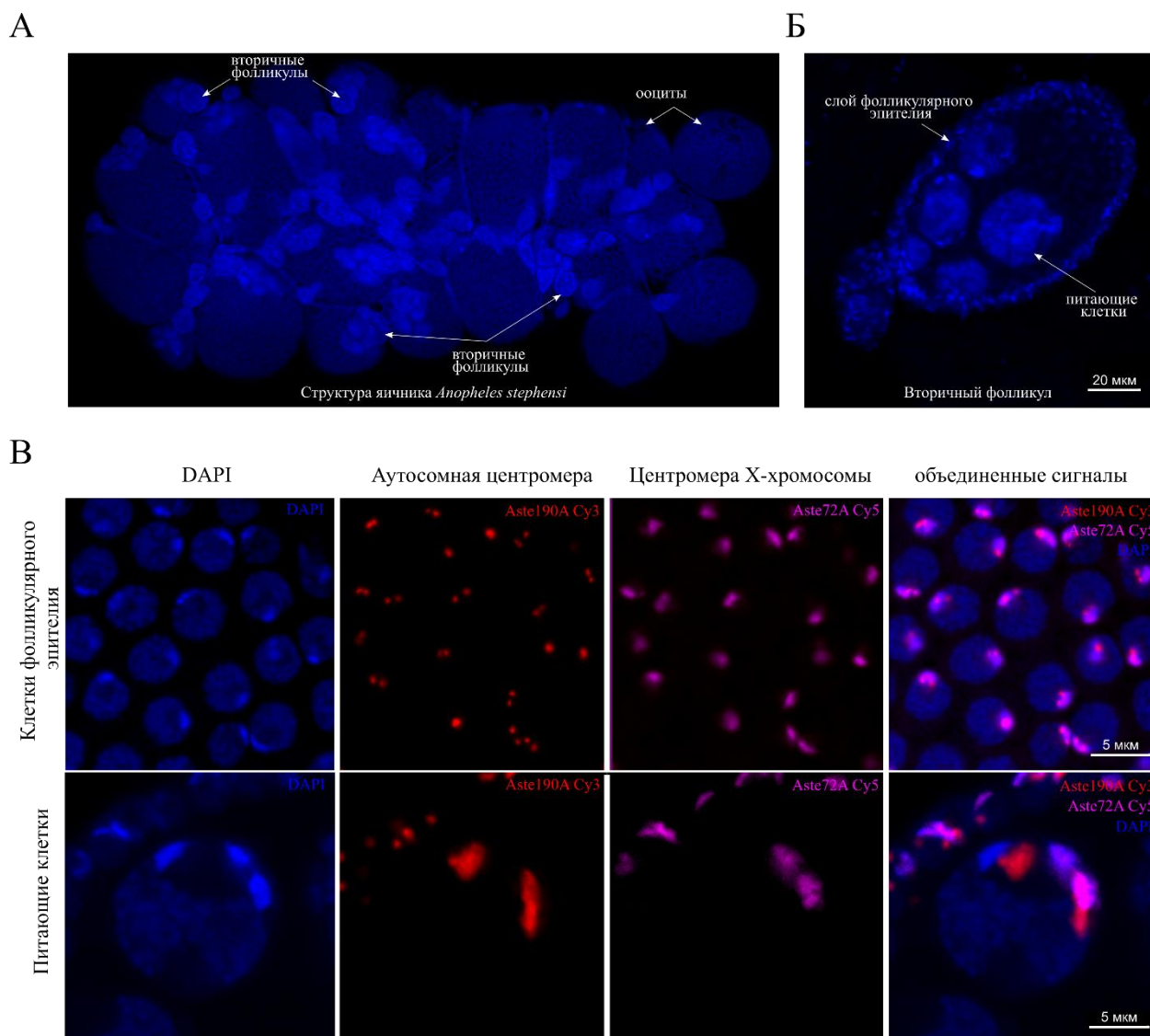


Рисунок 21. Rab1-конфигурация хроматина в интерфазном ядре у комаров рода *Anopheles*. А и Б. Структура яичника и вторичного фолликула малярийного комара *An.stephensi*; В. Колокализация центромерных проб в клетках фолликулярного эпителия (верхняя панель) и в питающих клетках яичника в 3D FISH-эксперименте.

3.4.2 Характеристика выделенных структур: компартменты

Детальный анализ карт пространственных контактов позволил нам выявить ключевые структуры в организации хроматина у малярийных комаров. На разрешении карт 10-100 Кб мы наблюдали типичный “клетчатый” паттерн, характерный для активных и неактивных компартментов. Для корректной интерпретации обнаруженных структур мы провели иммунопреципитацию хроматина с последующим секвенированием обогащенных фрагментов (ChIP-seq) по методу, описанному в разделе 2.2.8 на линии клеток MSQ43, полученной из личинок первого возраста комаров вида *An. stephensi*, для пяти гистоновых модификаций: H3K27Ac, H3K4me3, H3K9me3, H2AK119Ubiq и H3K27me3. Кроме того, нами был проведен анализ транскриптома эмбрионального материала пяти видов

анофелесов, полученного с той же стадии развития, что и фиксации на Hi-C-эксперимент. В результате, деление на компартменты, наблюдаемое на Hi-C-карте, хорошо соответствовало активности транскрипции и эпигенетическим модификациям гистонов (Рисунок 22, А). А-компартменты оказались обогащены метками активного хроматина – H3K27Ac и H3K4me3, а В-компартменты – репрессивными метками гетерохроматина и комплекса белков группы поликомб (первого или второго типа), такими как H3K9me3, H2AK119Ubiq и H3K27me3, (Рисунок 22, Б).

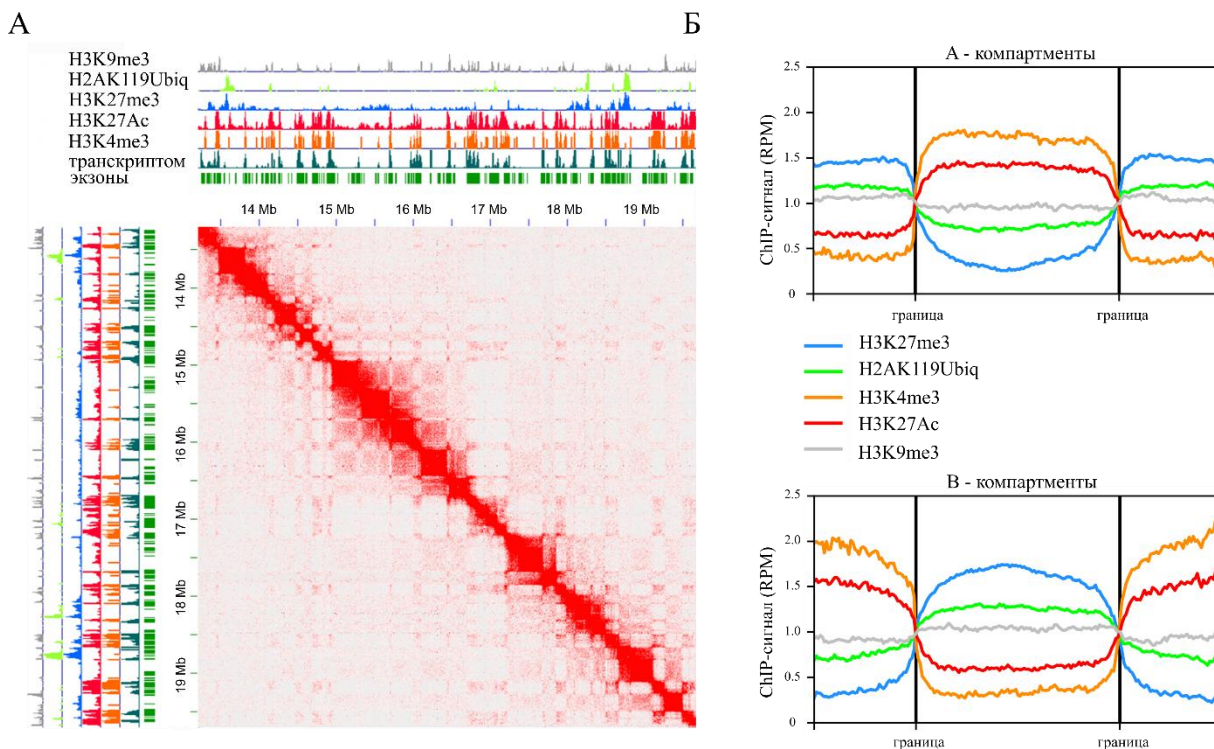


Рисунок 22. Соответствие клеточных компартментов гистоновым модификациям на примере клеточной линии MSQ43, *An.stephensi*. А) Пересечение рисунка ядерных компартментов с транскрипционными данными/активностью транскрипции и эпигенетическими модификациями гистонов; Б) Обогащение А- и В-компартментов метками активного хроматина и гетерохроматина, соответственно.

3.4.3 Характеристика выделенных структур: домены и хроматиновые петли

На более глубоком разрешении карт, в диапазоне 1-10 Кб, мы обнаружили структуры хроматиновых доменов (ТАДов или компартментных доменов), имеющие характерную форму треугольников вблизи основной диагонали (Рисунок 23, А; Рисунок 27, Г), и петель, точечных кластеров более интенсивных взаимодействий в вершинах, либо внутри доменов (Рисунок 23, Б и В; Рисунок 27, Д). Размер выявленных ТАДов варьировал в диапазоне от 10 до 1 640 Кб, где средний размер (median length) составил 130 Кб для *An.atroparvus*, 135 Кб для *An.albimanus* и *An.stephensi*, 145 Кб для *An.coluzzii* и 185 Кб для *An.merus*, соответственно.

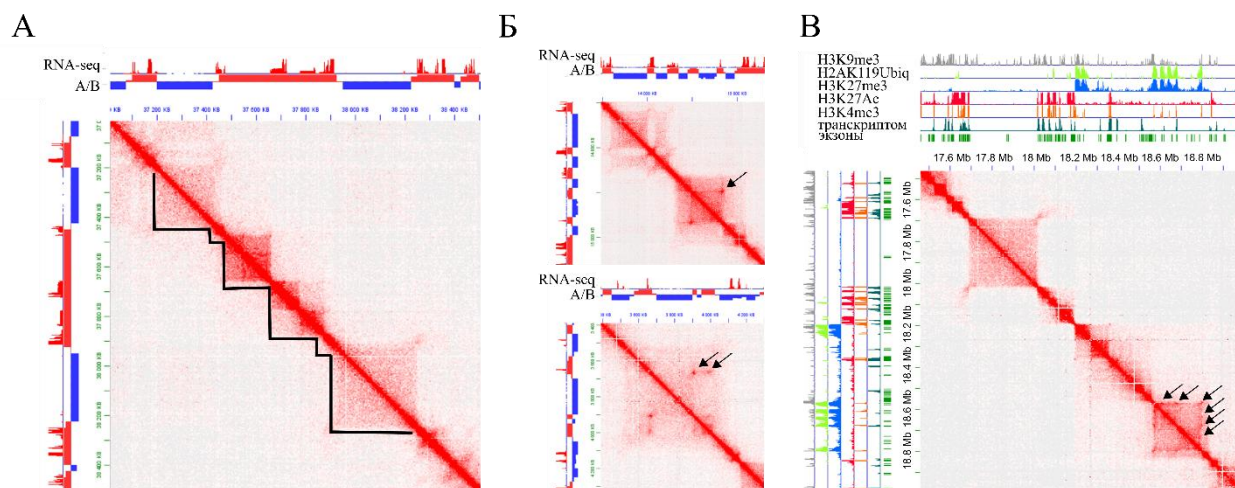


Рисунок 23. Характеристика выделенных структур: А) показаны домены и Б) петли у *An.albimanus* и В) хроматиновые петли и их пересечение с пиками H2AK119Ubiq и H3K27me3 на линии MSQ43. Петли отмечены черными стрелками.

Следует отметить линейную зависимость средней протяженности доменов от размера генома – закономерность, также наблюдаемую у представителей других таксономических групп (Schloissnig et al., 2021). В среднем на геном малярийных комаров было выявлено ~1200 доменов. Индивидуальные характеристики доменов, выделенные для каждого генома по отдельности, указаны в Таблице 8.

Таблица 8. Характеристики доменов, выделенные для каждого генома по отдельности. * Размер генома новой сборки.

Параметр	<i>An. albimanus</i>	<i>An. atroparvus</i>	<i>An. stephensi</i>	<i>An. coluzzii</i>	<i>An. merus</i>
Минимальная длина	25 000	20 000	15 000	10 000	10 000
Максимальная длина	660 000	870 000	650 000	1 055 000	1 640 000
Медианный размер	135 000	130 000	135 000	145 000	185 000
Средний размер	155 133	151 749	155 902	173 393	236 575
Количество доменов	1 084	1 442	1 252	1 326	984
Размер генома*, млн п.н.	170	220	221	245	300

Кроме доменов мы обнаружили более мелкие структуры – хроматиновые петли, которые представлены на карте контактов как отдельные фокусы интенсивных взаимодействий. Как правило, данные структуры носят регуляторный характер, опосредуя энхансер-промоторные взаимодействия. Сравнив данные по расположению гистоновых

меток комплексов белков группы поликомб первого и второго типа (H2AK119Ubiq - PRC1 и H3K27me3 - PRC2), мы обнаружили их обогащение в основаниях некоторых хроматиновых петель (Рисунок 23, В). Данный тип петель был ранее замечен у *Drosophila melanogaster* и вовлечен в репрессию некоторых геномных локусов и замолкание экспрессии [238]. В другой работе было отмечено, что зачастую формирование репрессионных хроматиновых петель сопряжено с регуляцией генов развития [237].

3.4.4 Свойства и особенности эухроматина и гетерохроматина на тепловой карте пространственных контактов

Как было описано выше, размер хроматиновых доменов у малярийных комаров варьирует в пределах 1,5 Мб. Сопоставив данные транскриптомного анализа с топологическим рисунком на карте Hi-C, мы заметили, что наиболее крупные блоки гетерохроматина совпадают с В-компартаментами или неактивными ТАДами, тогда как активные домены представлены структурами меньшего размера. На рисунке 24 видно преобладание активных участков хроматина среди маленьких доменов и наоборот – неактивных среди крупных.

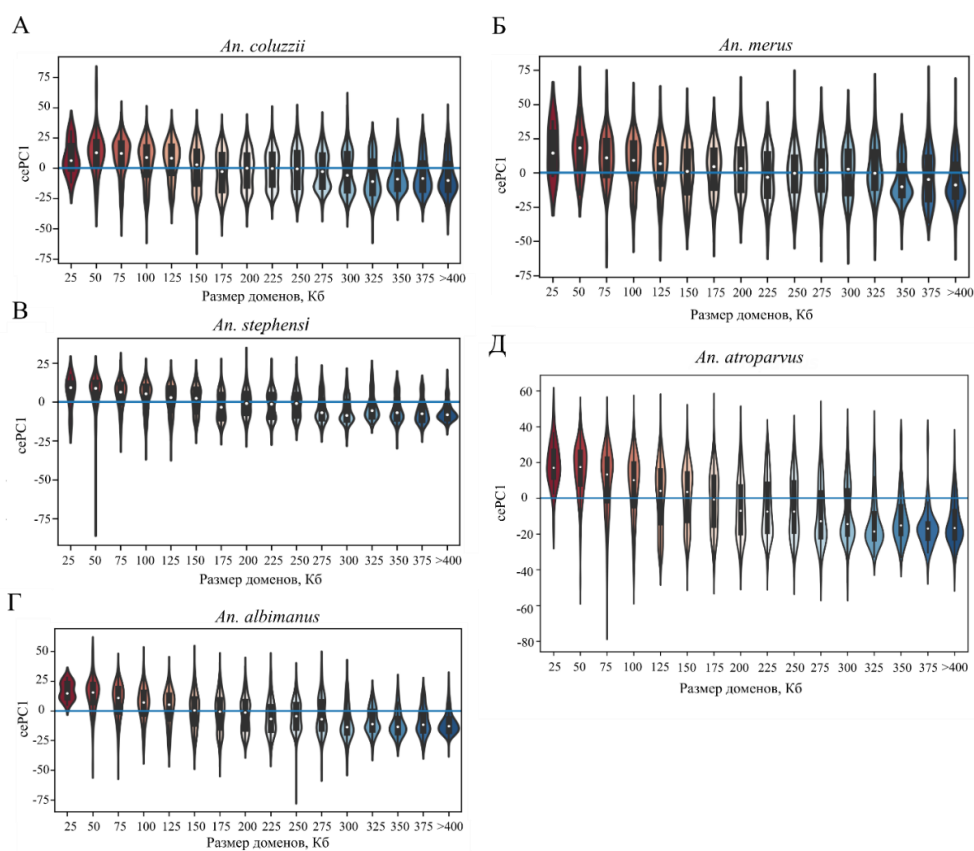


Рисунок 24. Распределение доменов по размеру в зависимости от принадлежности к типу компартамента. Заметно преобладание крупных ТАДов среди В-компартаментов и мелких среди А-компартаментов. По оси ординат отложена величина первой главной компоненты, полученная алгоритмом АВСЕ [364] и отражающая принадлежность к А-, либо В-компарменту; по оси абсцисс указан размер доменов.

Кроме того, расположение цитологических структур политенных хромосом, представляющих протяженные блоки интеркалллярного гетерохроматина, описанные ранее [337], совпадает с расположением протяженных локусов В-компарментов на картах пространственных контактов (Рисунок 25, А – 3; координаты крупных гетерохроматиновых локусов согласно Hi-C-картам приведены в таблице Приложения 3). Мы сопоставили координаты самых крупных доменов с повышенной инсуляцией у *An. merus* с цитогенетическими картами и обнаружили их хорошее соответствие. На панели А отмечено расположение гетерохроматиновых областей (PH и IH - прицентромерный и интеркалллярный гетерохроматин) относительно координат плеча хромосомы на цитологической карте *An. gambiae*. Д и Е панели отображают увеличенную область гетерохроматина на 2R и 2L, а Ж и З – на 3R и 3L, а верхний трек на панелях Б - 3 иллюстрирует полное отсутствие активных транскриптов в гетерохроматиновых локусах.

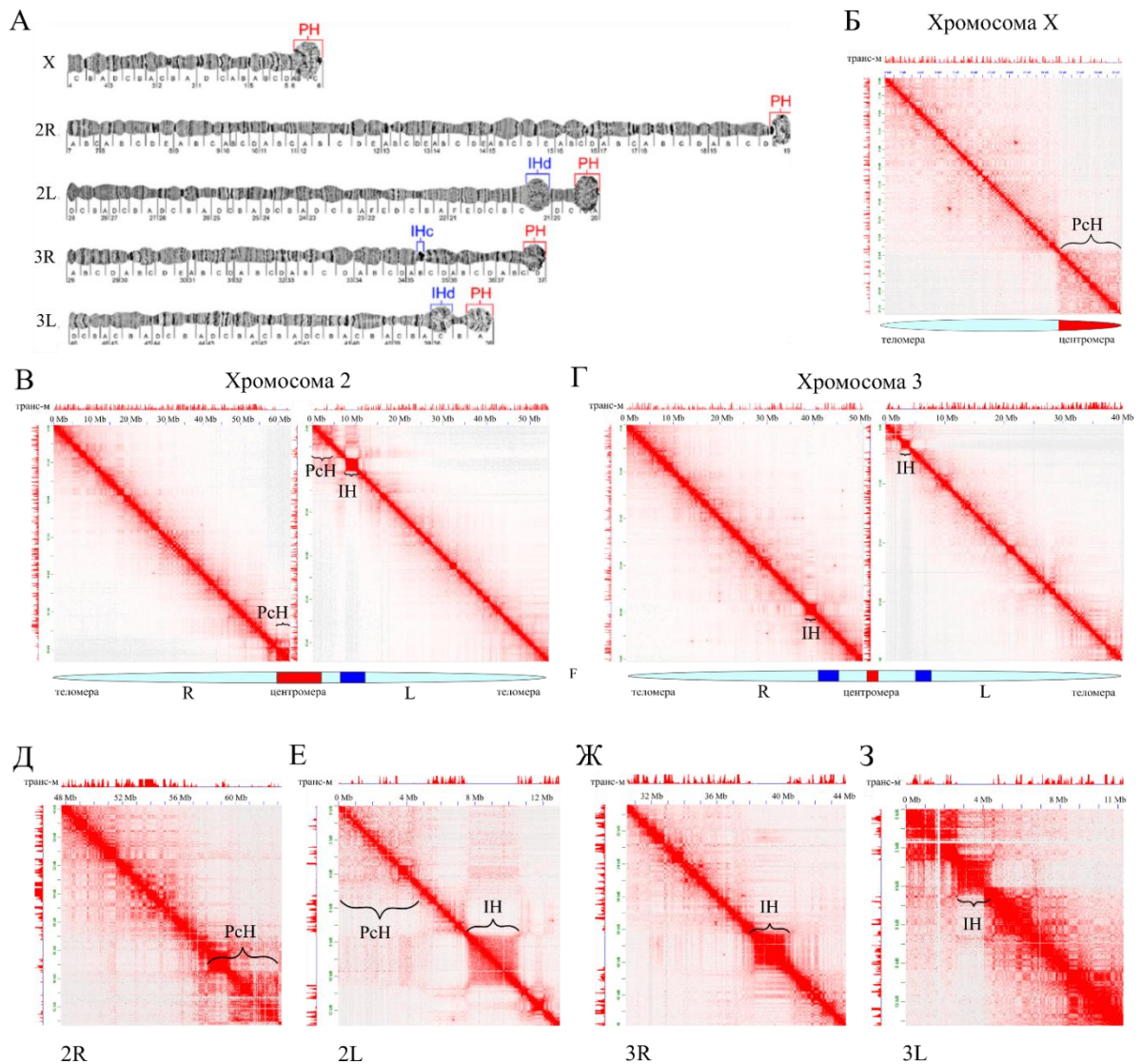


Рисунок 25. Прицентромерный (PcH) и интеркалярный (IH) гетерохроматин на карте пространственных контактов у *An. merus*. А. Цитогенетическая карта приведена для *An. gambiae* и адаптирована из Sharakhova et al., 2010. Паттерн Hi-C-контактов на хромосоме X, на хромосоме 2 и на хромосоме 3 показан на панелях Б, В и Г, соответственно. Д и Е панели отображают увеличенную область гетерохроматина на 2R и 2L, а Ж и З – на 3R и 3L.

Более того, оказалось, что блоки прицентромерного гетерохроматина на X-хромосоме, описанные ранее на физической карте, имеют характерную структуру на картах пространственных контактов у всех пяти видов малярийных комаров. Хорошо заметна инсуляция гетерохроматина от остальной части хромосомы X (Рисунок 26, А - Е). Стоит отметить, что, несмотря на гетерохроматиновый статус, мы наблюдали формирования нескольких ТАДов (компаратментных доменов) внутри прицентромерного гетерохроматинового локуса. Также примечательно, что в нашей выборке размер гетерохроматина на X пропорционален длине X-хромосомы малярийного комара (Таблица 9). *An. atroparvus* несколько выпадает из наблюдаемой закономерности, что может быть связано с несовершенной сборкой гетерохроматиновых локусов.

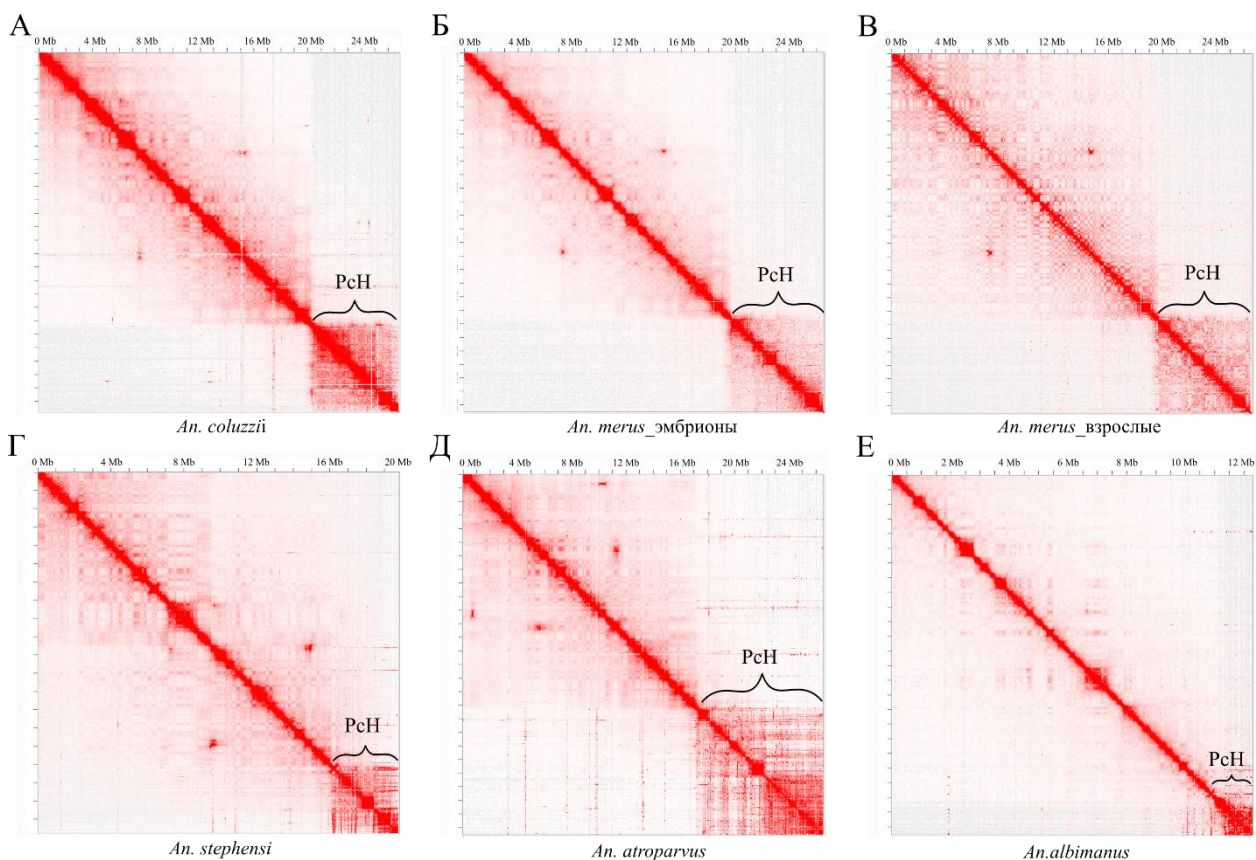


Рисунок 26. А–Б, Г–Е - гетерохроматиновый локус на хромосоме X у эмбрионов 5 видов малярийных комаров, В - у взрослых особей *An. merus*. PcH – прицентромерный гетерохроматин.

Таблица 9. Соответствие размера гетерохроматина на карте пространственных контактов длине X-хромосомы у малярийных комаров. **An.atroparvus* несколько выпадает из наблюдаемой закономерности.

Вид	Размер генома в хромосомах, п.н.	Размер X-хромосомы, Мб	Размер гетерохроматина, Мб
<i>An.atroparvus</i> *	216,747,066	26.5	9.5
<i>An.merus</i>	234,366,716	26.8	6.59
<i>An.coluzzii</i>	231,601,574	26.2	6.54
<i>An.stephensi</i>	196,394,606	19.7	3.44
<i>An.albimanus</i>	169,445,844	12.5	1.5

Таким образом, хроматин в интерфазном ядре у малярийных комаров организован следующим образом: отдельные хромосомы уложены в соответствии с конфигурацией по Раблю, когда центромеры объединены, а плечи хромосом вытянуты по всей длине параллельно друг другу; также в пространстве ядра наблюдается кластеризация активных и неактивных компартментов, соответствующих характерным модификациям гистонов и транскриптомным данным; на более детальном разрешении карт контактов детектируется уровень доменов (ТАДов) – треугольников вблизи основной диагонали, инсулирующих входящие в их состав элементы генома, и хроматиновых петель, часть из которых представлена регуляторными взаимодействиями промоторов и их активных энхансеров, а часть - интенсивными кластерами взаимодействий на крайне больших расстояниях в линейном геноме, о которых будет подробно рассказано в следующей главе. Описанные свойства хроматина проиллюстрированы на Рисунке 27.

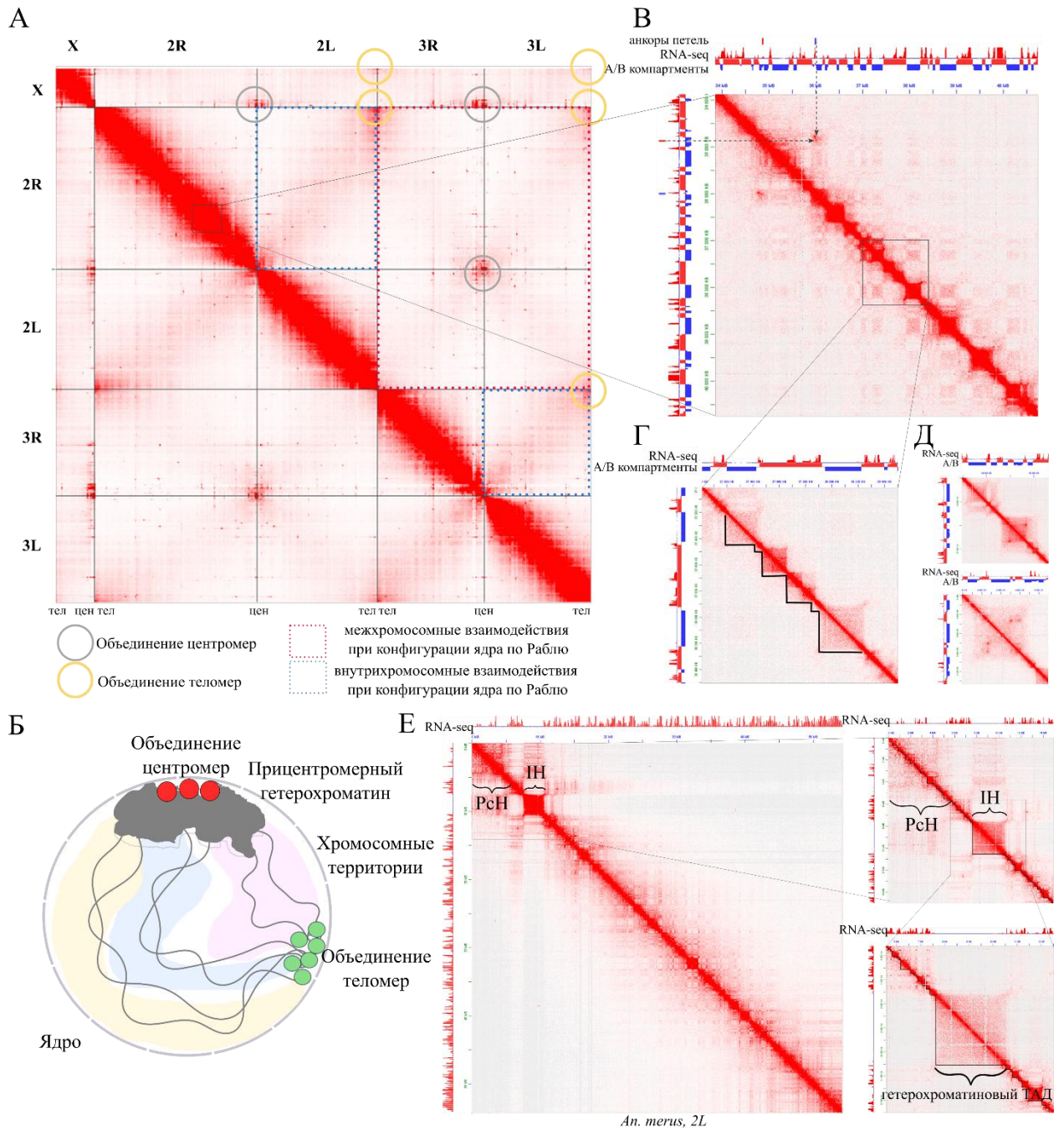


Рисунок 27. Основные консервативные принципы 3D-организации генома комаров рода *Anopheles*. А. Карта пространственных контактов в полногеномном разрешении – пояснения в тексте раздела 3.3.1. Б. Модель организации хроматина в ядре у малярийных комаров. В-Д. Показаны уровни активных и неактивных компартментов, доменов и хроматиновых петель. Е. Крупные блоки прицентромерного и интеркалярного гетерохроматина.

3.5 Интенсивные дальние взаимодействия хроматина, специфичные для комаров *Anopheles*. Их визуализация на картах пространственных контактов и подтверждение методом FISH

Помимо консервативных элементов 3D-архитектуры генома, анализ данных Hi-C выявил пары локусов (далее - основания хроматиновых петель или анкеры), разделенных значительным геномным расстоянием (от нескольких миллионов до нескольких десятков миллионов пар нуклеотидов) и демонстрирующих повышенный уровень пространственных контактов друг с другом, в десятки раз превосходящий ожидаемые. У каждого вида было выявлено от трех до девяти дальних петель, координаты всех обнаруженных петель приведены в Таблице из Приложения 4. Расстояние между основаниями петель варьирует, но максимально у *An.atroparvus* - 31 Мб, что составляет половину хромосомного плеча 3R. Наиболее выраженными оказались два взаимодействия - на хромосоме X и хромосомном плече 2R (3R у *An.atroparvus*), что в терминах хромосомных элементов соответствует элементам 1 и 2 (нумерация используется согласно номенклатуре, указанной в [12]). Далее в тексте этим двум петлям будут соответствовать наименования X-петля и A-петля, соответственно (Рисунки 30 и 31).

Стоит отметить, что по своей структуре на картах пространственных контактов дальние хроматиновые петли напоминают регуляторные взаимодействия хроматина, описанные в постмитотических нейронах обонятельных рецепторов у мышей [1]. А именно, у малярийных комаров на пересечении анкеров мы наблюдаем довольно протяженный выраженный компартмент с основанием в несколько сотен тыс.п.н, внутри которого выделяются точечные фокусы более интенсивных взаимодействий размером в несколько тыс.п.н. (Рисунок 29, В; Рисунок 30, Е и Ж; Рисунок 31, Е).

Наличие данных формирований у видов, разделенных в эволюции расстоянием в 80-100 млн. лет, натолкнуло нас на мысль о консервативности и функциональной значимости петель. Сперва мы сравнили генный состав в основаниях X-петли у всех видов в исследовании и обнаружили семь генов, ортологичных у четырех из пяти видов малярийных комаров. Два гена из этого списка имели ортологов и у *An. albimanus*, но были расположены в пределах 1 Мб от оснований петли. Примечательно, что функции некоторых из генов-ортологов связаны с возникновением устойчивости к инсектицидам (Cytochrome P450 CYP9K1 согласно [365]), а также с поддержанием иммунной реакции у малярийных комаров (prostaglandin reductase 1 согласно [366]). В основаниях A-петли мы идентифицировали три гена, ортологичных для всех пяти видов, и два гена-ортолога для четырех видов в исследовании. Подробные результаты сравнения приведены в Таблице из

Приложения 5. Выполненное М.А. Нуриддиновым эволюционное сравнение нуклеотидных последовательностей петель, найденных на Hi-C-картах разных видов, показало, что локусы, лежащие в их основаниях, расположены в синтенных участках даже у видов, дивергировавших порядка 100 млн лет назад. Подробные результаты эволюционного сравнения изложены в его диссертационной работе.

Дальние хроматиновые взаимодействия наблюдаются не только на эмбриональной стадии, когда согласно [240,367] активен транскрипционный фактор Zelda (формирующий один из типов хроматиновых петель у *Drosophila melanogaster*), но и у взрослых комаров, что было независимо подтверждено в экспериментах по флуоресцентной *in situ* гибридизации (Рисунок 31, Ж), а также на клеточной линии *An. stephensi* MSQ43 (Рисунок 28, А и Б). Можно заметить, что интенсивность взаимодействий А-петли в клеточной линии MSQ43 значительно снижена, в то время как X-петля выглядит лишь немного слабее по сравнению с эмбриональными данными.

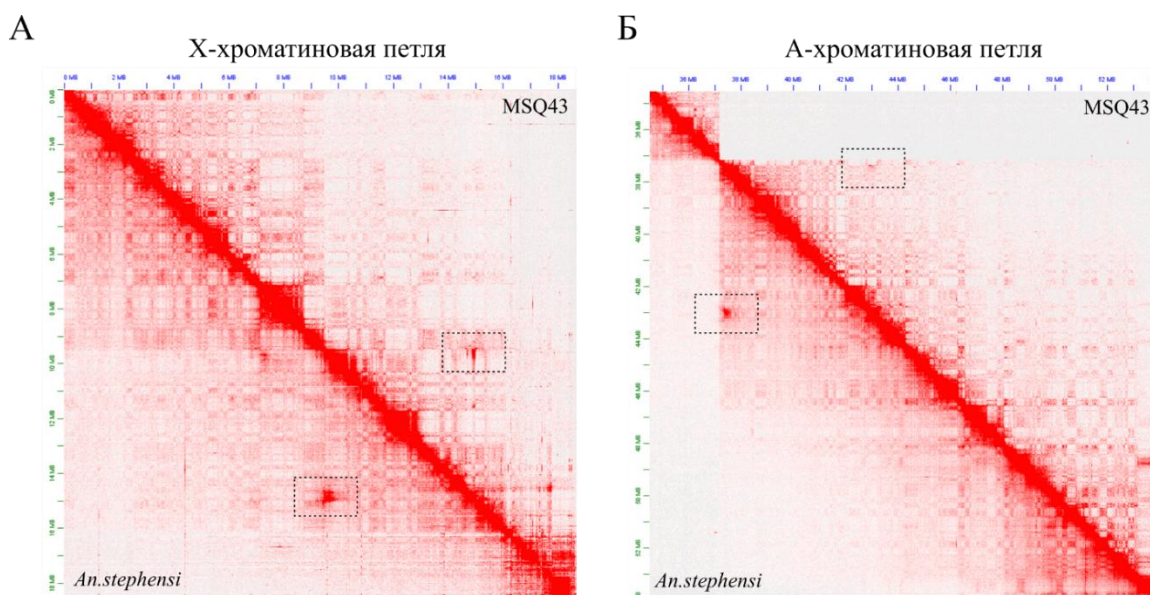


Рисунок 28. Сравнение карт пространственных контактов, полученных для эмбрионов комаров *An. stephensi* и клеточной линии MSQ43, в районе X-хроматиновой петли (панель А) и А-хроматиновой петли (панель Б).

Пространственные контакты между удаленными участками хроматина были ранее показаны для генома дрозофилы и ассоциированы с взаимодействиями либо белков группы поликомб, либо участками активного хроматина [154,368]. С целью выяснить эпигенетический ландшафт в основаниях петель мы провели серию экспериментов по иммунопреципитации хроматина на линии MSQ43. Однако, результат ChIP-seq-эксперимента за гистоновые метки H2AK119Ubiq и H3K27me3, связанные с

репрессированием хроматина комплексами белков группы поликомб PRC1 и PRC2, а также за гетерохроматиновую метку H3K9me3 не выявил уникального обогащения в основаниях дальних петель, обнаруженных нами у комаров рода *Anopheles* (Рисунок 29, А - Г). Несмотря на тот факт, что для некоторых анкоров нами было показано присутствие характерных для белков группы поликомб эпигенетических меток, геномные районы с идентичными характеристиками не вовлекались в формирования дальних хроматиновых петель (Рисунок 29, А, показаны черными стрелками). Кроме того, согласно результатам экспериментов на эмбрионах комаров по РНК-секвенированию, а также иммунопреципитации за метки активного хроматина (H3K27Ac, H3K4me3), в основаниях петель мы не увидели высоко экспрессирующихся транскриптов, или же гистоновых меток активной транскрипции.

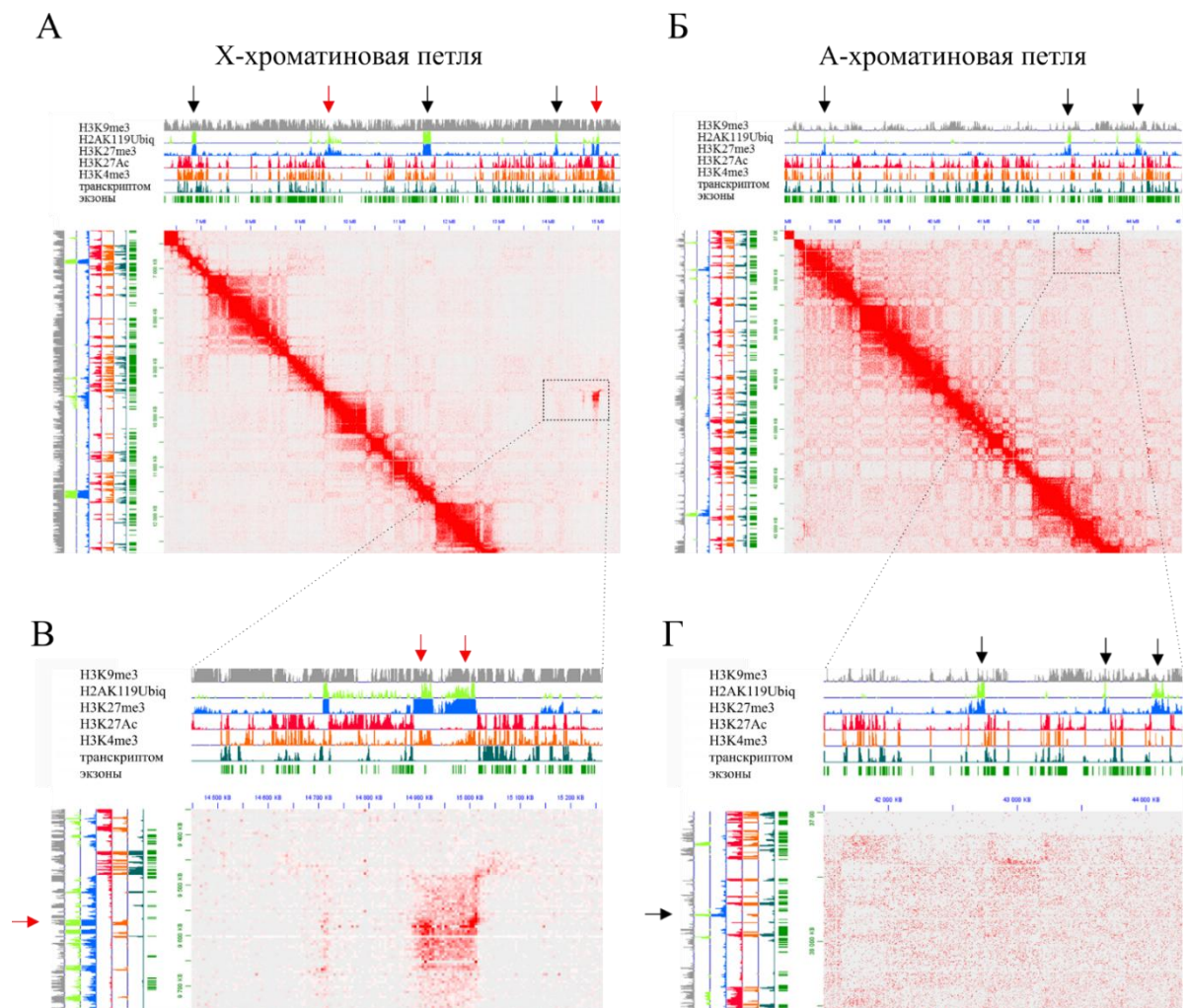


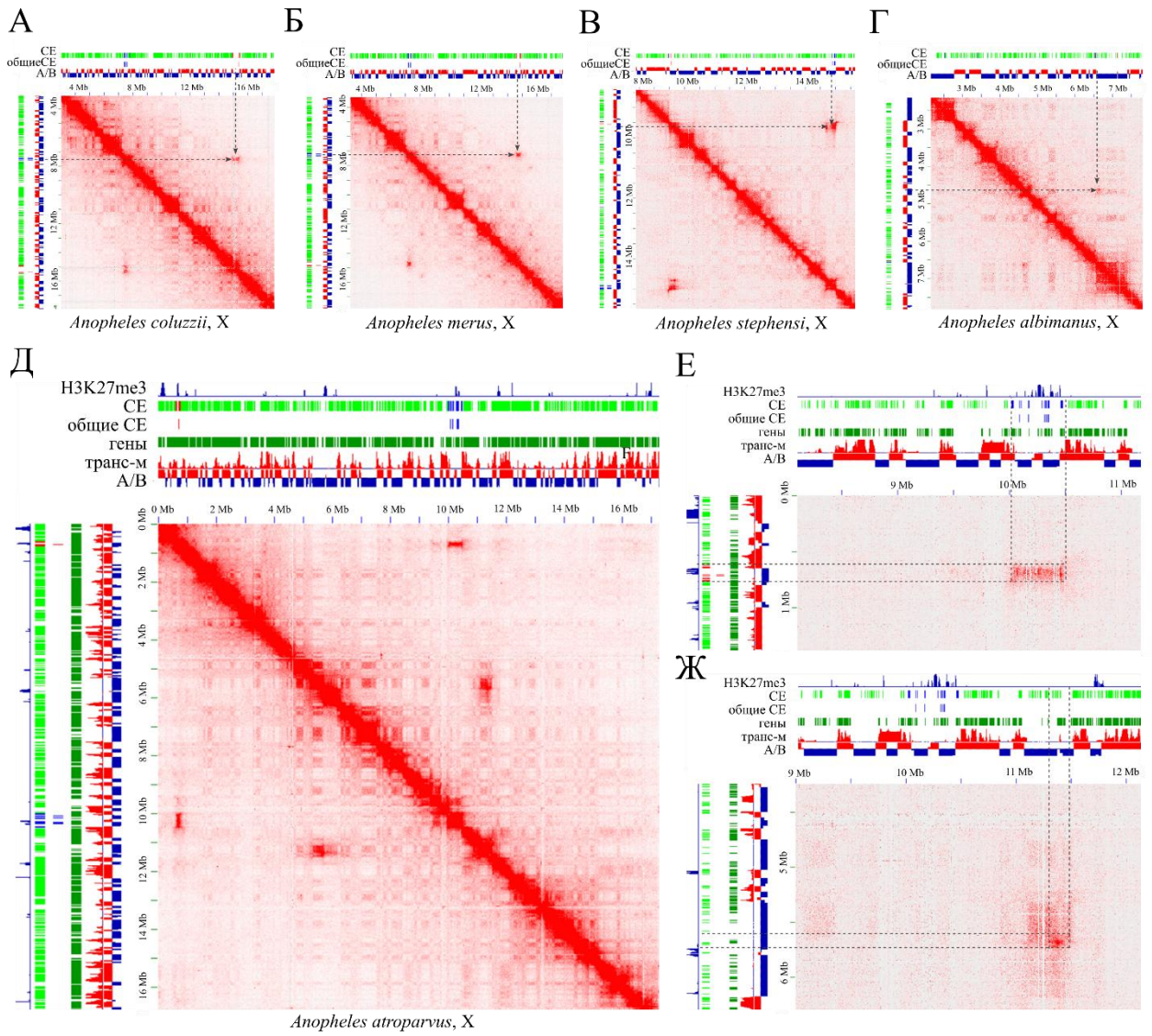
Рисунок 29. Распределение гистоновых меток в анкерах дальних хроматиновых петель в клеточной линии MSQ43 комара *An. stephensi*. А и В – петля на X-хромосоме; Б и Г – петля на 2R. На панелях А и Б представлено разрешение 10 Кб, В – 1 Кб, Г – 5 Кб соответственно. Красными стрелками обозначены перекрывающиеся с анкерами петель пики H2AK119Ubiq, черными - пики H2AK119Ubiq без перекрытия.

В 3D-FISH-экспериментах мы обнаружили, что колокализация сигнала А-петель у *An.stephensi* и *An.atroparvus* в клетках фолликулярного эпителия и высокополитенизированных ядрах питающих клеток происходит в 8-15% случаев, тогда как в низкополитенизированных ядрах гораздо чаще и составляет 39-70%. В случае Х-петли сигналы пересекались в 97% ядер клеток фолликулярного эпителия у *An.atroparvus*, но гораздо реже совпадали в высокополитенизированных ядрах питающих клеток (6.7%) (Рисунки 30, 3 и Рисунок 31, Ж). Эти результаты хорошо согласуются с Hi-C-данными, полученными на эмбрионах всех пяти видов и взрослых особях (для комаров рода *An.merus*): в обоих случаях мы наблюдали Х- и А-петели, что свидетельствует о формировании этих структур в нескольких различных типах клеток.

Результаты 2D FISH-экспериментов для трех видов малярийных комаров показали колокализацию анкоров дальних хроматиновых петель во всех интерфазных ядрах фолликулярного эпителия, тогда как в высокополитенизированных хромосомах питающих клеток мы наблюдали отдаленное положение проб в соответствии с координатами в линейном геноме (Рисунок 32, А-В).

По нашему мнению, противоречивые данные для интерфазных ядер могут возникать из-за разницы в первичной фиксации материала. Так, фиксация клеток для 3D FISH проходила в растворе 4% PFA, что идентично фиксациям ядер для Hi-C-экспериментов. В методе 2D FISH при распластывании препарата на стекле используется обработка пропионовой и уксусной кислотами с последующей фиксацией материала в 50% формамиде. Разница же в результатах для интерфазных и высокополитенизированных ядер может быть связана со сниженной «гибкостью» последних в результате высокой копииности и, как следствие, физического блока на изгиб [234]. Последовательности и координаты праймеров для подготовки флуоресцентных проб указаны в Приложении 6, а полные результаты статистического подсчета приведены в Приложении 7.

Полученные нами данные позволяют утверждать: 1) дальние хроматиновые петли в геномах комаров рода *Anopheles* действительно присутствуют, что было подтверждено двумя независимыми методами – Hi-C и FISH; 2) наблюдаемые структуры не стадиоспецифичны и могут являться одним из вариантов взаимодействий, опосредованных белковыми комплексами PRC1/PRC2, в сочетании с особым эпигенетическим ландшафтом; 3) петли формируются за счет неопisanного ранее механизма и, вероятно, связаны с исключительной биологической функцией.



3

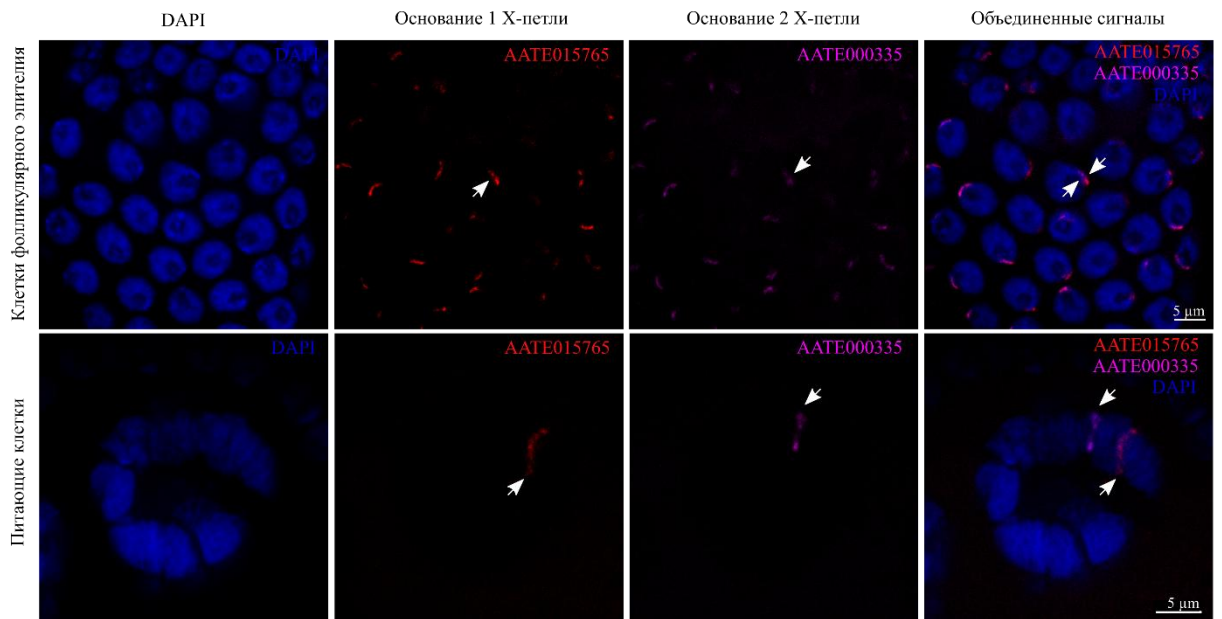


Рисунок 30. Дальняя хроматиновая X-петля. А – Д – петля представлена у всех пяти видов. Е – увеличенное изображение петли у *An.atroparvus*. Ж. Визуализация X-петли методом 3D-FISH в нескольких типах клеток.

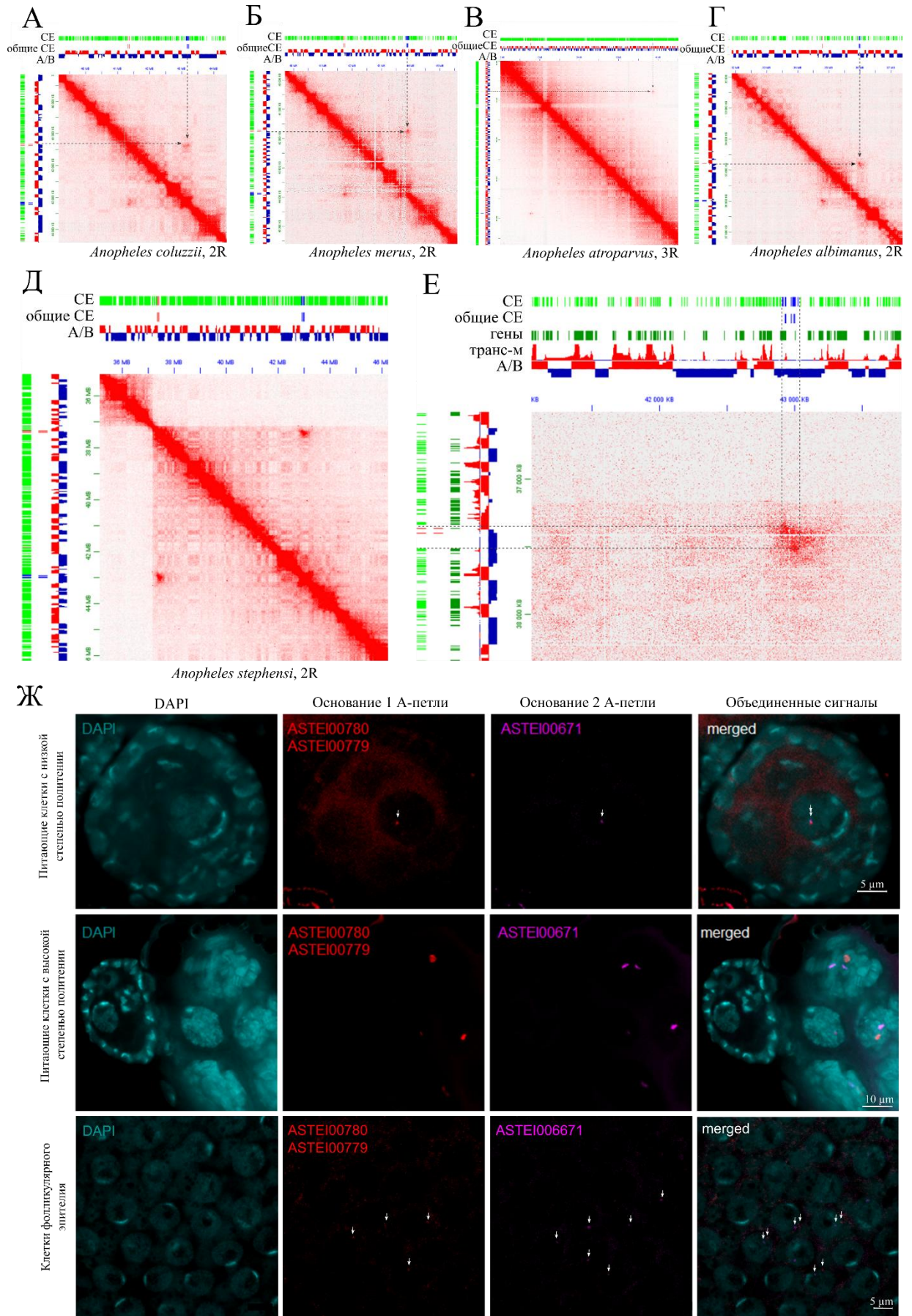


Рисунок 31. Дальняя хроматиновая А-петля. А – Д – петля представлена у всех пяти видов. Е – увеличенное изображение петли у *An.stephensi*. Ж. Визуализация А-петли методом 3D-FISH в нескольких типах клеток.

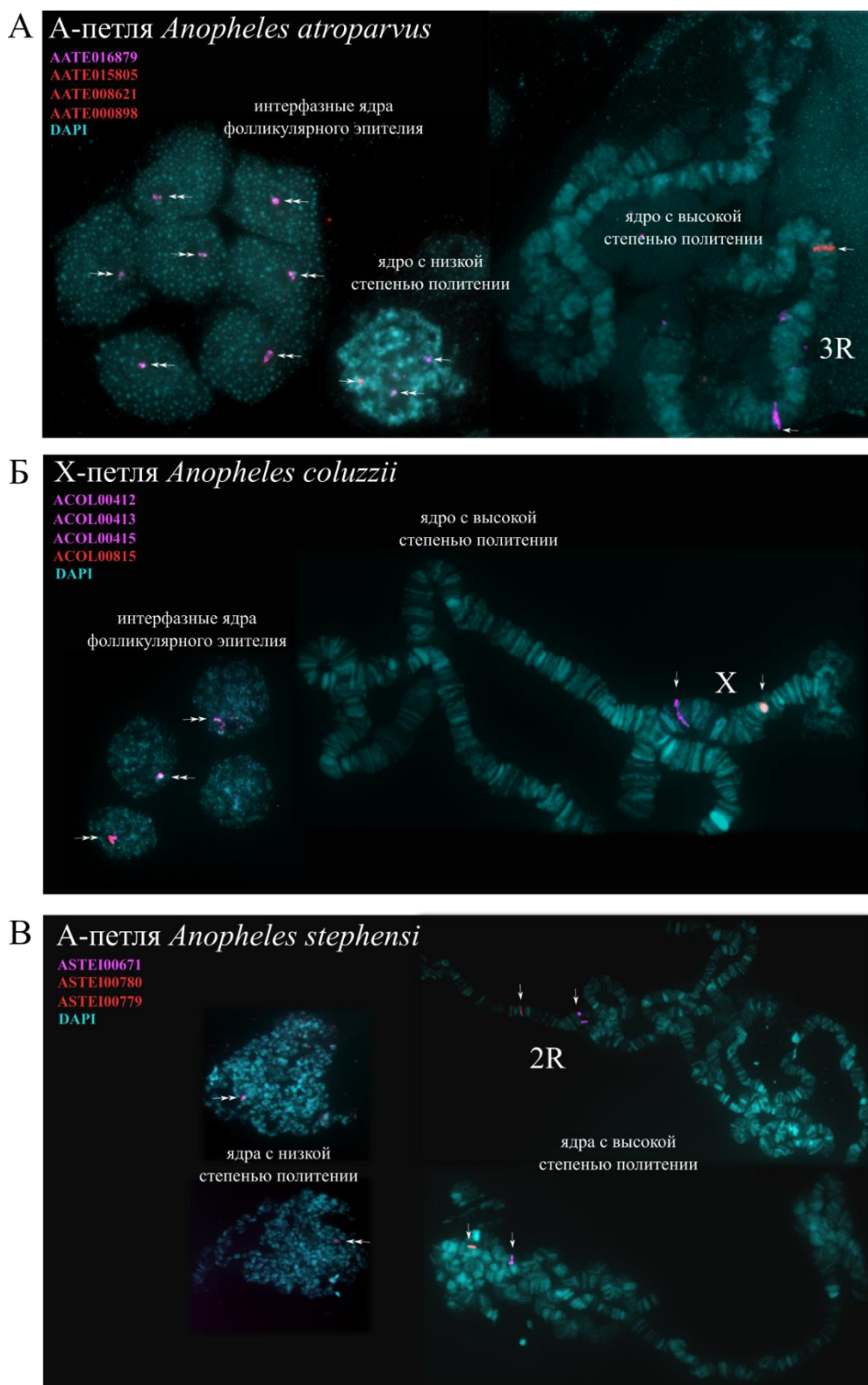


Рисунок 32. Дальние хроматиновые петли в 2D FISH. А. А-петля у *An. atroparvus* на 3R-плече, заметно пересечение сигналов в интерфазных ядрах, частичное перекрытие в низкополитенизированных ядрах и разделение сигналов согласно геномным координатам в высокополитенизированных ядрах; Б. X-петля у *An. coluzzii* на X-хромосоме, заметно пересечение сигналов в интерфазных ядрах и разделение сигналов согласно геномным координатам в высокополитенизированных ядрах; В. А-петля у *An. stephensi*, частичное перекрытие в низкополитенизированных ядрах и разделение сигналов согласно геномным координатам в высокополитенизированных ядрах.

3.6 Сравнение выявленных принципов укладки хроматина у малярийных комаров с другими таксонами

Чтобы понять, как полученные нами результаты согласуются с существующими представлениями о пространственной организации генома в ядре, мы оценили зависимость вероятности пространственных контактов от геномного расстояния ($P(S)$) и сравнили с данными для других таксономических групп (Рисунок 33, А). Ранее было показано, что частота контактов падает с увеличением расстояния согласно степенному закону [66,222,369,370], что оказалось справедливо и для генома малярийных комаров. Для выяснения особенностей представленной зависимости, мы взяли производную от вероятности контактов ($P(S)$) и построили дополнительный график в логарифмических координатах (Рисунок 33, Б). Оказалось, что скорость падения вероятности контактов не одинакова с увеличением расстояния, и график можно разделить на две фазы. Первая фаза графика у малярийных комаров представляет кривую в форме U и покрывает расстояния от 10 Кб до 1 Мб. Значение производной медленно падает и достигает локального минимума в значении -1 (200-500 Кб), после чего начинает возрастать и достигает первоначального значения на участке 1 Мб. Затем следует вторая фаза с локальным максимумом в значении -0,6, что соответствует геномным расстояниям 2-3 Мб, после чего кривая быстро опускается к значениям ниже -1. Наблюдаемая форма графика очень схожа с данными для других насекомых, включая рода *Aedes* и *Drosophila*, с той лишь разницей, что с увеличением размера генома происходит смещение (увеличение) геномных координат.

В данных для позвоночных мы наблюдаем несколько иной паттерн: изгиб кривой в первой фазе существенно более выражен по сравнению с насекомыми, а точка минимума производной наблюдается на геномном расстоянии 800 Кб для курицы, 1-1,5 Мб для млекопитающих, 4-5 Мб в случае индуцибельного дегрона по белку WAPL, а при деплеции по белку RAD21 провал на графике отсутствует вовсе. Напомним, что RAD21 входит в состав когезинового комплекса, а биологической функцией WAPL является снятие этого комплекса с молекулы ДНК в процессе выпетливания хроматина [205,207,210].

Мы предполагаем, что биологический смысл первой фазы отражает формирования ТАДов, что подтверждается характерными значениями геномных расстояний в точке минимума [364]. Для более крупных геномов, например, аксолотля, это предположение также справедливо [109], и таксоноспецифичный размер ТАДов может быть определен исходя из графика зависимости вероятности контактов от геномного расстояния. Стоит отметить также слабую выраженность ТАДов на графиках для насекомых по сравнению с ТАДАми позвоночных, что может косвенно указывать на силу этих формирований.

Максимум второй фазы попадает на значение 6-8 Мб для куриных интерфазных ядер и на 15-20 Мб для ядер млекопитающих, после чего в обоих случаях наблюдается резкое падение кривой. Отметим, что падение происходит к значениям, лежащим выше минимума первой фазы. Биологическая интерпретация второй фазы графика представляет более сложную задачу. Наличие наклона кривой во второй фазе не зависит от присутствия известных архитектурных белков, как-то: когезиновый комплекс (RAD21), конденсиновый комплекс (CAP-H2, SMC2), WAPL и NIPBL, - а также от формирований ТАДов, но падение существенно более резко и пролонгировано в случае метафазных хромосом (Рисунок 33, Б для *Gallus gallus*, желтая кривая). Кроме того, наблюдаемый паттерн кривой сохраняется у различных типов клеток, фаз клеточного цикла и хромосом альтернативных длин, что подтверждает его универсальную природу.

По нашему предположению, резкость/угол падения наблюдаемого пика может отражать особую форму хромосомных территорий, существующих в интерфазе. Если считать, что молекула хромосомы имеет форму эллипса или цилиндра, то ее радиус будет определять характерный объем сферы, внутри которого контакты не затруднены «вытянутостью» хромосомной территории. Вероятность контактов регионов, расположенных в разных частях по оси эллипса, будет в этом случае снижена по сравнению с идеальной сферой. Основываясь на этом предположении, мы считаем, что в интерфазных ядрах хромосомы у разных видов организованы в цилиндрические структуры, минимальный радиус которых специфичен для насекомых, птиц и млекопитающих. Действительно, свидетельства вытянутой сферической формы хромосомных территорий были получены в экспериментах на млекопитающих [371,372] и комарах [373]. В работе Фляймера и коллег авторы утверждают, что характерный изгиб графика $P(S)$ формируется в результате нескольких причин: плотности упаковки хроматина в ядре, стадии клеточного цикла, а также «клеточной памяти» о предыдущем состоянии генома (то состояние крайней компактизации, в котором находился геном) [92]. Мы предполагаем, что феномен «клеточной памяти» как раз может отвечать за угол наклона кривой во второй фазе на графике производной $P(S)$ (Рисунок 33, Б).

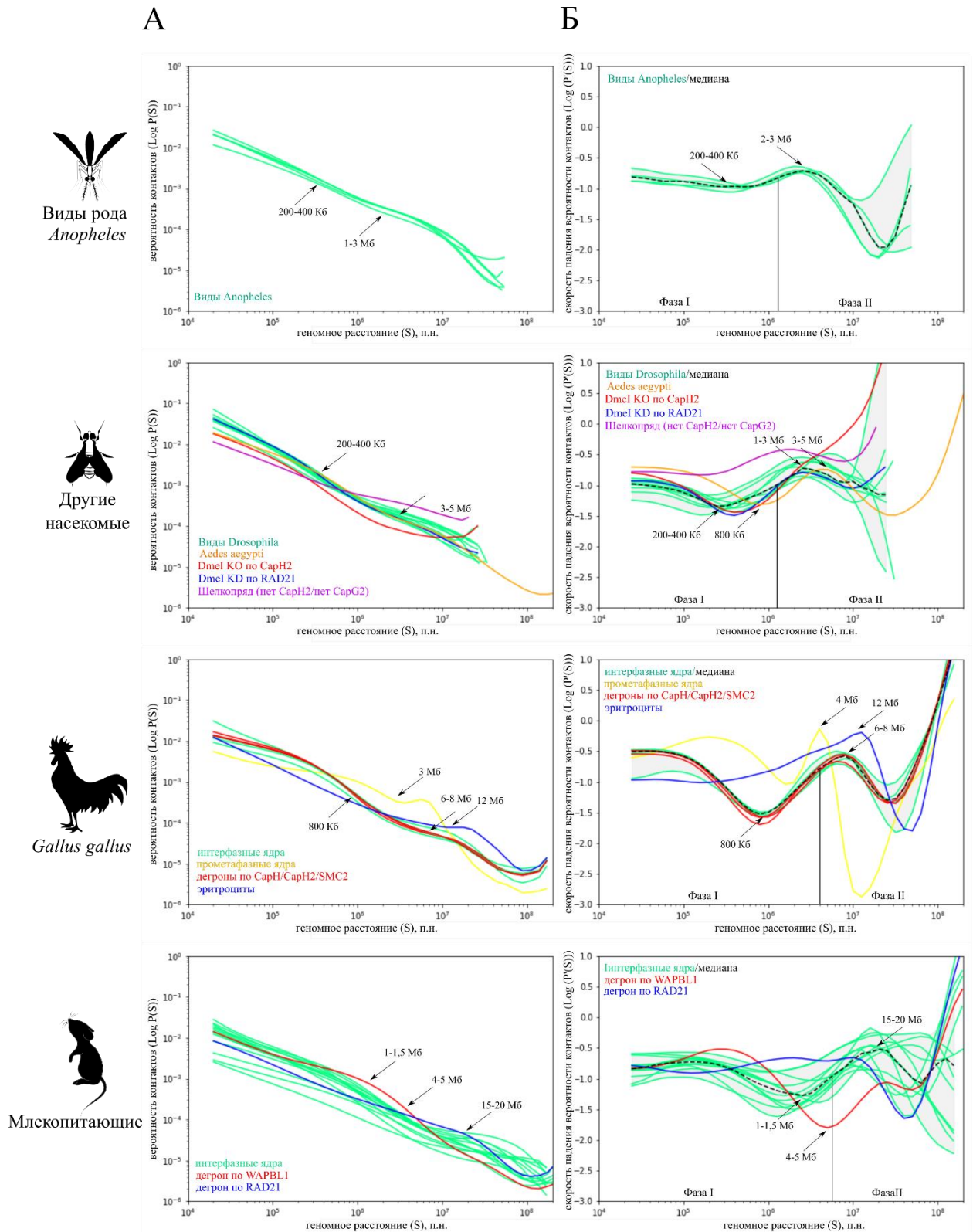


Рисунок 33. Частота пространственных контактов изменяется неравномерно с геномным расстоянием. А. График зависимости вероятности контактов от геномного расстояния (P(S)) для малярийных комаров, других насекомых, куриц и млекопитающих в координатах log-log; Б. Скорость изменения графика P(S) (панель А) как функция от геномного расстояния. Черный пунктир изображает усредненное значение по всем представленным типам клеток (зеленые кривые).

ГЛАВА 4. ОБСУЖДЕНИЕ.

4.1 Улучшение существующих сборок геномов для пяти видов комаров рода *Anopheles*

На сегодняшний день особое внимание уделяется расшифровке и изучению геномов патогенных организмов и их переносчиков для разработки генетических методов борьбы с вредителями, а также выяснению фундаментальных молекулярных механизмов, определяющих их активное распространение среди человеческой популяции. Технология 3C – метод захвата конформации хроматина и его модификация Hi-C [6,72] дает возможность с высокой точностью и хорошим разрешением собирать/создавать геномы живых организмов *de novo*, улучшать существующие геномные сборки, состоящие из мелких скаффолдов и содержащих неточности в сборке [101,374,375]. Используя результаты Hi-C-эксперимента и данные физического картирования, мы собрали *de novo* три генома до хромосомного уровня (*An. coluzzii* Mopti, *An. merus*, *An. stephensi*) и улучшили существующие хромосомные сборки для *An. albimanus* и *An. atroparvus*. В качестве референсных геномов были использованы предыдущие геномные сборки [297,334]. Геномная реконструкция на основе Hi-C и PacBio-данных оказалась значительно более аккуратной и приемлемой для дальнейшего использования по сравнению со сборкой PEST, реконструированной при помощи Hi-C без использования длинных прочтений.

В Таблице 5 можно заметить, что показатель N50 для скаффолдов, свидетельствующий о качестве геномной сборки с точки зрения её непрерывности, несколько снизился для некоторых видов (*An. albimanus*, *An. atroparvus*, *An. stephensi*). Данный результат является следствием корректировки структуры скаффолдов в режиме *assembly mode* в программе Juicebox [335]. Для того, чтобы внести изменение (перевернуть участок, встроить или удалить фрагмент внутри скаффолда), необходимо внести разрыв (или несколько) в существующий скаффолд, что ведет к уменьшению его суммарной длины. Таким образом, повышение точности/аккуратности сборки с использованием Hi-C-технологии неизбежно приводит к снижению размера крупных скаффолдов, что отражается на финальном значении N50.

4.2 Технология Hi-C для выявления полиморфных и фиксированных инверсий у представителей рода *Anopheles* и *Aedes*

Предполагается, что именно хромосомные aberrации, закрепленные в популяции, обеспечивают столь быструю адаптацию к новым условиям обитания и географическим

ареалам, могут влиять на поведение комаров, их пищевые предпочтения и способность к переносу патогенов [263,293,298,300,360,376]. Адаптивное значение хромосомных инверсий подтверждается корреляцией между частотами их встречаемости и рядом фенотипических признаков, которые приобретают их носители [293]. Так, например, носители некоторых вариантов обладают повышенной устойчивостью к засухе в первые дни после превращения в имаго [302] или выраженным адаптивным ответом на смену сезонов [377,378]. Поскольку геномные районы внутри инвертированного участка в гетерозиготном состоянии оказываются практически инсулированы от рекомбинации и сторонней интрогрессии, это позволяет закреплять выгодные генетические варианты в популяции [379]. Активная обработка больших природных площадей соединениями против насекомых в середине прошлого века и, как следствие, появление устойчивости некоторых популяций комаров к инсектицидам натолкнуло исследователей на мысль о влиянии хромосомных перестроек на возникновение данных свойств [349]. По этой причине популяционные исследования инверсионного полиморфизма имеют важный прикладной аспект, а идентификация хромосомных перестроек в природных популяциях комаров с помощью Hi-C-технологии является хорошей альтернативой/дополнением к существующим методам кариотипирования.

Комары рода *Aedes* считаются векторами-переносчиками вирусных заболеваний, вызывающих лихорадку денге, чикунгуньи, жёлтую лихорадку, лихорадку Западного Нила, лихорадку Зика и некоторые другие. Активно распространяясь по африканскому континенту и за его пределы, представители *Ae. aegypti* заняли различные экологические ниши, сформировав индивидуальные предпочтения в выборе питания и мест обитания. Именно переход некоторых колоний *Ae. aegypti* от питания на животных к человеку повлиял на динамичное распространение многочисленных вирусных заболеваний среди людей. Причины сменившихся пищевых предпочтений неясны. Внутри вида *Aedes aegypti* принято выделять подвиды *Ae. aegypti formosus* и *Ae. aegypti aegypti*, где первый преимущественно обитает в Африке и тяготеет к зоофильному типу питания, а второй успешно распространился на другие континенты и практически полностью перешел на питание человеческой кровью. Несколько лет назад было показано, что смена типа питания с животных на человека совпадает с изменениями в работе обонятельных рецепторов, и, в частности, *Or4*, который отвечает за распознавание химических компонент, формирующих запах человека [362]. Наши коллеги из лаборатории Линди МакБрайд провели масштабное исследование пищевых предпочтений представителей *Ae. aegypti*, собранных из 27 мест обитания в Тропической Африке, где сайты отлова широко варьировали по показателям

близости людей (плотность человеческой популяции) и сезонности осадков (контрастность и длительность сухих и дождливых периодов). Согласно результатам исследования, в 83% случаев указанные параметры являются ключевыми в определении пищевых привычек *Ae. aegypti* [363]. Кроме того, было проведено генотипирование 375 особей, выявившее однонуклеотидные полиморфизмы (SNP) и геномные локусы, изменения в которых предположительно влияют на выбор источника пищи [363]. Оказалось, колонии *Ae. aegypti*, отловленные в Восточной и Центральной Африке – NGO и OGD, относящиеся к подвиду *Ae. aegypti formosus*, демонстрируют выраженное предпочтение к питанию человеческой кровью. В наш Hi-C-эксперимент вошли 12 колоний с Африканского континента, исследованных Линди МакБрайд, а также семи колоний Нового Света, двух колоний из Азии и четырех лабораторных колоний. Нам удалось показать, что часть обнаруженных инверсий пересекается с кластером обонятельных рецепторов, изменения в функционировании которых предположительно ответственны за переход от питания на животных к человеку. А именно, инверсионные перестройки 1qE, 1qF, 1qC, 1qD и 1qH, расположенные на длинном плече (q-плече) 1й хромосомы, содержат крупный генный кластер, включающий и ген *Or4* [362].

В свете новых данных о расположении инверсий интересным представляется тот факт, что хромосомное плечо 2R у анофелесов, на котором было обнаружено большинство хромосомных aberrаций [293], соответствует 1q и 3p плечам у комаров рода *Aedes*, где также наблюдается подавляющее число перестроек (хромосомный элемент e2 в терминологии Мюллера [305]). Можно предположить, что «генетическое содержание» данного хромосомного элемента способствует накоплению инверсионных полиморфизмов даже у представителей родов, эволюционно разделенных расстоянием в 150-200 млн лет.

4.3 Особенности пространственных контактов хроматина у малярийных комаров

Поскольку в геноме комаров рода *Anopheles* присутствуют огромные гетерохроматизированные районы, наблюдается выраженная компартиментализация эухроматина и гетерохроматина, а также разделение на транскрипционно активные и неактивные участки, подтвержденное данными РНК-секвенирования (наши данные). Наличие интеркалярного и прицентромерного гетерохроматина, формирующих сеть взаимодействующих между собой, но крайне инсулированных от остального генома регионов, сказывается на общей организации генома и влияет на алгоритмы выделения пространственных геномных структур. Схожие проблемы встречаются при работе с геномами некоторых червей (бескишечные турбеллярии *Hofstenia miamia*,

неопубликованные данные), дрожжей [380], земноводных [109] и растений [381,382], характерной особенностью которых является ориентация хромосом по Раблю. Интерфазные хромосомы у данных организмов формируют взаимодействия гетерохроматиновых областей центромер с ламиной и между собой, что, вероятно, происходит по механизму фазирования/фазового разделения хроматина. Известно три типичных признака конфигурации по Раблю: 1) хроматиновые взаимодействия центромер, 2) пространственные контакты теломерных районов и 3) взаимодействия хромосомных плеч. При этом степень выраженности конфигурации может отличаться, с преобладанием или отсутствием некоторых признаков. Совсем недавно была проведена масштабная работа по изучению особенностей пространственной организации генома у 24 видов эукариот, где авторы выяснили, что хромосомные территории и конфигурация по Раблю представляют два крайних состояния организации генома, которые сменяли друг друга на протяжении эволюции эукариот, исчезая и возникая вновь [22]. Более того, им удалось связать степень интенсивности взаимодействий гомологичных хромосом (наличия синапсиса гомологов или его отсутствие), за которые обычно отвечает комплекс белков конденсина II, с принадлежностью к тому или иному состоянию в интерфазе. У части исследованных видов (дрожжи, тихоходки, плодовые мушки, комары рода *Aedes*) отсутствие некоторых субъединиц конденсина II совпало с организацией хромосом по Раблю, что в свете эволюции представляется как взаимосвязанные события. Эксперименты по деградации субъединицы CAP-G2 на клетках человека, которые в норме обладают классическими хромосомными территориями, элегантно подтвердили предположение о ключевом участии комплекса белков конденсина II в пространственной организации генома. В работе утверждается, что реорганизация хроматина конденсином II происходит на М-стадии клеточного цикла и в короткий промежуток после нее [22]. Авторы предполагают существование непосредственной связи между длиной хромосом в митозе и типом пространственных структур в интерфазе. Согласно этой гипотезе хроматин более коротких хромосом в процессе деконденсации приводит к образованию хромосомных территорий, тогда как длинные хромосомы переходят в ориентацию по Раблю. В другом исследовании были представлены свидетельства того, что ориентированные по Раблю хромосомы менее подвержены перемешиванию хроматина, что не противоречит данным работы Хоенкампа [383].

Ранее было отмечено, что доля генома, покрытая транскрипционно менее активными В-компартаментами, существенно выше, чем А-компартаментами, а также что размеры В-компарментов превосходят размеры А [153,154,227,231]. Данная закономерность

прослеживалась в нескольких исследованиях по изучению организации хроматина у дрозофилид [225]. Мы в своей работе наблюдали аналогичную картину, что свидетельствует о схожей организации генома у двукрылых.

Необходимо подчеркнуть, что размер ТАДов у живых организмов имеет положительную корреляцию с размером генома. Так, сравнивая данные, полученные для представителей малярийных комаров, мы определили средний размер в 135 Кб, при этом наблюдали небольшие флуктуации в среднем значении, соответствующие разнице в длинах геномов. Известно, что характерный размер ТАДов для представителей рода *Drosophila* в среднем соответствует 100 Кб при среднем размере генома в 200 Мб [153,154]. Данные, полученные для *Aedes aegypti*, у которых средний размер ТАДов составляет 700 Кб при среднем размере генома в 1500 Мб, также отлично подчиняются предполагаемой закономерности [102]. У млекопитающих, соответственно, ТАДы имеют средний размер 0.9-1 Мб [144,145]. У аксолотля, размер генома которого 32 млрд.п.н., средний размер ТАДа составляет 9 Мб [109].

Примечательно, в некоторых работах утверждается, что на более глубоком разрешении карт контактов (увеличенном покрытии генома) можно увидеть структуры меньшего масштаба внутри ТАДов, где размер таких субдоменов составляет десятки тыс.п.н. [231]. В недавнем исследовании французской группы с использованием микроскопии суперразрешения были идентифицированы наноструктуры внутри ТАДов в эмбриональных стволовых клетках, названные хроматиновыми нанодоменами (CND). Структуры не изменялись при деплеции белка CTCF или когезина, их количество внутри ТАДа возрастало при удалении RAD21, а глобальное увеличение ацетилирования приводило к разрушению наноструктур, что, вероятно, свидетельствует о когезин-независимом механизме образования последних. Авторы исследования выдвигают гипотезу, что наблюдаемый уровень организации хроматина может быть универсальным для всех позвоночных и представлять аналог хроматиновых доменов у насекомых, так как их размер схож с наблюдаемыми у *Drosophila melanogaster* [48].

4.4 Уникальные хроматиновые петли

3D-контакты между удаленными участками генома или дальние петли хроматина были ранее показаны у различных живых организмов. Среди них выделяют когезин-опосредованные клеточно-специфичные регуляторные взаимодействия промоторов и их энхансеров, а также взаимодействия на более высоком уровне, связанные с регуляцией

экспрессии целых групп генов и формируемые когезин-независимыми механизмами. Касательно первого типа взаимодействий было выяснено, что только 28% всех энхансер-промоторных контактов в человеческом геноме являются переменными между клеточными типами и связаны с функциональной специализацией/спецификацией клеток [384]. Было также отмечено, что вариативность в структуре петель перекрывается с изменениями в статусе хроматина, а интенсивность пространственных контактов имеет позитивную корреляцию с активностью экспрессии.

У дрозофилы хроматиновые петли могут быть ассоциированы с активной экспрессией или репрессией участков хроматина. Еще в 2007 профессором Кавалли был показан интересный феномен - “chromosome kissing” или “поцелуй хромосом” [385], когда отдаленные друг от друга участки хромосомы контактируют в пространстве ядра в результате общих регуляторных процессов [386]. В новейших геномных работах предполагается, что формируемые в пространстве регуляторные кластеры взаимодействий (энхансерные хабы) образуются в эмбриональном развитии раньше ТАДов и могут переключаться на противоположные активности в зависимости от стадии в онтогенезе [387]. При этом образование регуляторной энхансерно-промоторной петли не гарантирует активацию целевого гена, но лишь обеспечивает платформу для посадки транскрипционных факторов, в зависимости от клеточной функции которых происходит активация или репрессия [388,389]. Предположив, что наблюдаемые у анофелесов петли носят регуляторный характер, отсутствие высокой экспрессии на эмбриональной стадии может быть объяснено сменой экспрессионного паттерна при переходе на другую стадию развития, где данная регуляторная петля необходима. Кроме того, поскольку на 15-18 часовой стадии клеточная дифференцировка в тканях эмбриона уже подходит к завершению [330], могут возникать тканеспецифичные особенности экспрессии, которые теряются на фоне данных со смешанной популяции клеток. По этой причине было бы интересно изучить тканеспецифичную картину пространственных контактов совместно с активностью генов у взрослых анофелесов.

Известно, что у *Drosophila* многие хроматиновые петли ассоциированы с активностью именно белков группы поликомб. Линейные расстояния между участками генома, контактирующими в пространстве ядра благодаря комплексам PRC1 и PRC2 могут почти в 100 раз превосходить расстояния между регуляторными элементами, связанными по когезин-зависимому механизму [390]. Взаимодействия, контролируемые белками группы поликомб, мультивалентны (включают несколько геномных локусов), отличны от ТАДов и, как правило, связаны с плотной компактизацией участков генома и репрессией

транскрипции. Экспрессия данных генов тонко регулирует клеточно-специфичную дифференцировку, а их абберрантная активность может приводить к нарушениям в развитии организма [390,391]. При этом утверждается, что субъединицы канонического PRC1-комплекса способны формировать клеточные конденсаты и необходимы для инициации фазового разделения хроматина, тогда как поддержание разделения возможно без участия белков группы поликомб [392]. Было показано, что в результате молекулярных взаимодействий CBX2-субъединиц, входящих в состав комплекса группы поликомб, осуществляется фазирование белковых молекул и, как следствие, разделение хроматина по механизму фазовой сепарации [393,394]. Точечные мутации в Cbx2-локусе приводят к исчезновению поликомбных кластеров в ядре, декомпактизации плотных гетерохроматиновых районов *in vitro*, а также дефектам развития у мышей [393].

Участки генома, распознаваемые комплексами поликомб, как правило, содержат характерные гистоновые модификации - H3K27me3 и H2AK119Ubiq. Стоит отметить существование хроматиновых петель, поддерживаемых белками группы поликомб, но лишенных характерных эпигенетических меток - не канонические нео-PRC1-локусы [395]. Данные локусы обогащены активной меткой H3K27Ac и достаточно высоко экспрессируются. Таким образом, комплекс поликомб может выполнять роль как репрессора (канонический), так и активатора экспрессии (неканонический), независимо от транскрипционного фактора CTCF [390].

В анкерах некоторых дальних петель, обнаруженных у анофелесов, нами было выявлено присутствие характерных для белков группы поликомб эпигенетических меток H3K27me3 и H2AK119Ubiq, но в то же время многочисленные геномные районы с идентичными характеристиками не вовлекались в формирования дальних контактов. Во всех основаниях отсутствовали метки активного хроматина и высоко экспрессирующиеся транскрипты, таким образом исключая вероятность дальних петель быть не каноническими нео-PRC1-локусами. Тем не менее, дальние хроматиновые петли могут являться одним из вариантов взаимодействий, опосредованных белковыми комплексами PRC1/PRC2, в сочетании с особым эпигенетическим ландшафтом или дополнительными белковыми факторами.

На картах пространственных контактов дальние петли структурно схожи с регуляторными взаимодействиями хроматина, описанными в постмитотических нейронах обонятельных рецепторов у мышей – “греческими островками” [1]. В обоих случаях наблюдается протяженный компартмент с выраженными точечными фокусами интенсивных взаимодействий (Рисунок 26, В, Рисунок 30, Е-Ж, Рисунок 31, Е). В случае с

обонятельными нейронами хроматиновые взаимодействия, вовлекая в регуляторную сеть более 1000 генов с 18 хромосом, выполняют роль ключевых медиаторов, задающих спецификацию конкретного нейрона. Формирование функционального компартмента происходит предположительно по механизму фазовой сепарации и объединению участков со схожими характеристиками, определяемыми транскрипционными факторами (LHX2, EBF и LDB1) и метками гетерохроматина.

Предполагается, что клетки с высокоспецифичной функцией могут существенно модифицировать базовые правила организации хроматина в ядре. Палочки в сетчатке – один из примеров, когда, вопреки ключевым механизмам, происходит реорганизация положений эу- и гетерохроматина относительно ламины для выполнения биологической задачи собирательной линзы в данном типе клеток [396]. Другой необычный случай – это ядра клеток слюнных желез или питающих клеток яичников у некоторых насекомых, содержащие гигантские политенные хромосомы, функцией которых является синтез больших количеств белка. В них совершенно отсутствуют структуры компартментов и дальние контакты хроматина, вероятно в результате физического блока на изгиб [234]. Описанные выше случаи, включая “греческие островки”, считаются примерами высокоспециализированных посмитотических формирований.

В своей работе, анализируя результаты экспериментов с флуоресцентным маркированием анкоров дальних петель, мы обнаружили колокализацию оснований петель в пространстве ядра не во всех клетках даже одного типа, что может свидетельствовать о пространственно-временной избирательности данного процесса. В интерфазных ядрах фолликулярного эпителия мы наблюдаем разные варианты положения анкоров: в небольшой доле клеток наблюдаются петлевые структуры, но в большинстве ядер мы их не видим. В высокополитенизированных ядрах основания петель не сближены никогда, что согласуется с отсутствием и других геномных структур (ТАДов и компартментов) в данном типе клеток [234]. Тогда как в ядрах с низкой степенью политении еще сохраняется физическая мобильность и гибкость хромосом и типичная пространственная структура ядра, поэтому мы встречаем колокализацию петель в ядрах данного типа. Эволюционная консервативность дальних хроматиновых петель и специфичность формируемых ими взаимодействий говорит о том, что они, с высокой вероятностью, задаются уникальными и эволюционно-консервативными нуклеотидными последовательностями.

4.5 Заключение

Среди насекомых пространственная организация ядра наиболее детально исследована у представителей рода *Drosophila*. Геномика других групп насекомых, имеющих важное фундаментальное и биомедицинское значение, только начинает развиваться и остается малоизученной. Одним из таких примеров являются анофелесы или малярийные комары, ключевые переносчики малярийного плазмодия и других опасных патогенов животных и человека. Среди ~450 видов анофелесов только ~60 видов способны создавать благоприятные ниши для существования и переноса паразитов, а также гораздо более эффективно адаптироваться к условиям обитания. Устойчивость к лекарственным препаратам продолжает возникать, а анофелесные комары – в связи с изменяющимися климатическими условиями - активно распространяться за обычные ареалы своего обитания. Именно поэтому основные надежды на успешный и долговременный контроль малярии и подобных заболеваний связаны с геномными исследованиями видов-переносчиков.

В представленной диссертационной работе была проведена сборка геномов до хромосомного уровня, а также подробно охарактеризованы фундаментальные принципы пространственной организации хроматина нескольких представителей комаров рода *Anopheles*: *An. coluzzii*, *An. merus*, *An. stephensi*, *An. atroparvus* и *An. albimanus*, представляющих достаточно протяженную ветвь на эволюционном дереве (от 0.5 миллионов лет до 100). В ходе исследования 3D-организации геномов пяти видов комаров рода *Anopheles* методом Hi-C, который позволяет оценивать частоты взаимодействий различных районов по геному с точностью до сотен нуклеотидов, мы показали основные уровни пространственной организации хроматина у анофелесов. Так, мы описали A/B-компарменты, разделяющие геном на активные и неактивные области и хорошо совпадающие с характерными гистоновыми модификациями; домены со средним размером в 135 Кб; несколько типов хроматиновых петель, часть из которых обусловлена взаимодействием белков группы поликомба; также продемонстрировали конфигурацию хромосом по Раблю, как на картах пространственных контактов, так и в 3D-FISH-экспериментах. Кроме того, мы обнаружили несколько крайне интенсивных взаимодействий между участками хроматина, удаленными на огромное расстояние друг от друга (десятки миллионов пар оснований). Обнаруженные взаимодействия присутствуют на Hi-C-картах как у эмбрионов, так и у взрослых комаров, а FISH-эксперименты наглядно подтверждают их наличие по крайней мере в нескольких типах клеток взрослых особей.

Кроме того, в диссертационной работе метод Hi-C был впервые применен для поиска хромосомных перестроек у комаров родов *Anopheles* и *Aedes*. Инверсионный полиморфизм давно является краеугольным камнем в исследованиях комаров с большими геномами и низким качеством препаратов политенных хромосом (например, *Culex* и *Aedes*). На примере малярийных комаров нами было выяснено, что технология Hi-C представляет удобный молекулярный инструмент для поиска хромосомных перестроек у насекомых. У анофелесов мы продемонстрировали четыре инверсионные перестройки, две из которых ранее не были известны (*An. merus* и *An. atroparvus*). Две самые крупные были подтверждены на препаратах политенных хромосом и детектировались в виде инверсионных петель (*An. stephensi* и *An. atroparvus*). Мы успешно применили анализ Hi-C-данных для 25 колоний комаров рода *Aedes* (полученных от представителей разных частей Африки, Южной и Центральной Америки, Азии и нескольких лабораторных колоний) и выявили точки разрывов хромосомных перестроек с точностью до 5 кб. Мы показали, что инверсионные перестройки на 1q-плече – 1qE, 1qF, 1qC, 1qD и 1qH – пересекаются с кластером генов обонятельных рецепторов, что, согласно работам [362] и [363], может иметь ключевое значение в пищевом переходе комаров рода *Aedes* от животных к человеку. Мы выявили несколько специфичных для западно-африканских и мексиканских колоний перестроек, а также показали, что больше всего инверсий встречается в африканских популяциях.

Мы сравнили пространственную организацию генома малярийных комаров с данными по другим насекомым и позвоночным и показали, что основные уровни организации хроматина (конфигурация хромосом по Раблю, компартменты, домены, хроматиновые петли) присутствуют у всех указанных таксонов и частота контактов в пространстве ядра подчиняется экспоненциальной зависимости. Уникальной структурой у анофелесов являются не описанные ранее дальние хроматиновые петли, молекулярные механизмы формирования и биологическую функцию которых еще предстоит исследовать.

ВЫВОДЫ

1. Полученные при помощи технологии Hi-C геномные сборки пяти видов комаров рода *Anopheles*: *An. coluzzii*, *An. merus*, *An. stephensi*, *An. atroparvus*, *An. albimanus* характеризуются высокой континуальностью и соответствуют данным физического картирования хромосом, что свидетельствует об их высоком качестве. Полученные сборки геномов позволяют исследовать механизмы эпигенетической регуляции комаров рода *Anopheles* на хромосомном уровне.
2. Границы точек разрыва, установленные на основе карт трехмерных контактов хроматина с разрешением 5 тысяч пар оснований, хорошо согласуются с цитогенетическими данными, полученными ранее, что свидетельствует в пользу возможности установления границ геномных инверсий на основе методов захвата конформации хромосом.
3. Компартиментализация эу- и гетерохроматина в совокупности с организацией по Раблю хромосомных территорий определяют паттерн дистальных контактов ДНК в геномах комаров рода *Anopheles*.
4. Геномы комаров рода *Anopheles* подразделяются на топологически ассоциированные домены со средним размером около 135 тысяч пар оснований, причем блоки прицентромерного, прителомерного и интеркалярного гетерохроматина выделяются в отдельные, инсулированные от окружения крупные домены, что позволяет идентифицировать их на картах пространственных контактов хроматина.
5. Детальное изучение карт трехмерных контактов хроматина позволило идентифицировать петли нескольких типов:
 - многочисленные короткие петли (менее 1 млн. пар оснований), ассоциированные с активностью белков группы поликомб;
 - значительно более редкие дальние (более 1 млн. пар оснований) петли, также ассоциированные с активностью белков группы поликомб и
 - немногочисленные дальние петли (более 1 млн. пар оснований), не ассоциированные с активностью белков группы поликомбом.
 Участки генома, задействованные в образовании петель последнего типа оказались консервативны среди исследованных комаров рода *Anopheles*. В то же время, формирование петель такого типа не было описано ранее у других животных, что позволяет предположить уникальный для комаров рода *Anopheles*, ранее не описанный механизм образования хроматиновых петель.

6. Геномы представителей рода *Anopheles* и *Drosophila* демонстрируют сходные принципы укладки хроматина: экспоненциальную зависимость частоты контактов хроматина от расстояния; соответствие хроматиновых доменов компартментам; наличие конфигурации хромосом по Раблю; а также наличие хроматиновых петель, образуемых белками группы поликомб.

СПИСОК ЛИТЕРАТУРЫ

1. Monahan K., Horta A., Lomvardas S. LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice // *Nature*. - 2019. - Vol. 565. - No 7740. - P. 448–453.
2. Dekker J. Capturing Chromosome Conformation // *Science*. - 2002. - Vol. 295. - No 5558. - P. 1306–1311.
3. Plessis L. du, Skunca N., Dessimoz C. The what, where, how and why of gene ontology--a primer for bioinformaticians // *Brief Bioinform*. - 2011. - Vol. 12. - No 6. - P. 723–735.
4. Comet I., Schuettengruber B., Sexton T., et al. A chromatin insulator driving three-dimensional Polycomb response element (PRE) contacts and Polycomb association with the chromatin fiber // *Proceedings of the National Academy of Sciences*. - 2011. - Vol. 108. - No 6. - P. 2294–2299.
5. Lieberman-Aiden E., Berkum N. L. van, Williams L., et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome // *Science*. - 2009. - Vol. 326. - No 5950. - P. 289–293.
6. Berkum N. L. van, Lieberman-Aiden E., Williams L., et al. Hi-C: A Method to Study the Three-dimensional Architecture of Genomes. // *JoVE*. - 2010. - No 39. - P. 1869.
7. Bystricky K., Laroche T., Houwe G. van, et al. Chromosome looping in yeast // *Journal of Cell Biology*. - 2005. - Vol. 168. - No 3. - P. 375–387.
8. Shaw P. The architecture of interphase chromosomes and nucleolar transcription sites in plants // *Journal of Structural Biology*. - 2002. - Vol. 140. - No 1–3. - P. 31–38.
9. Coetzee M., Hunt R. H., Wilkerson R., et al. *Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex // *Zootaxa*. - 2013. - Vol. 3619. - No 3. - P. 246–274.
10. <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021>.
11. Budiansky S. Creatures of Our Own Making // *Science*. - 2002. - Vol. 298. - No 5591. - P. 80–86.
12. Neafsey D. E., Waterhouse R. M., Abai M. R., et al. Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes // *Science*. - 2015. - Vol. 347. - No 6217. - P. 1258522.
13. Straub T. Heterochromatin Dynamics // *PLoS Biol*. - 2003. - Vol. 1. - No 1. - P. e14.
14. Zink D., Cremer T., Saffrich R., et al. Structure and dynamics of human interphase chromosome territories in vivo // *Human Genetics*. - 1998. - Vol. 102. - No 2. - P. 241–251.
15. Pecinka A., Schubert V., Meister A., et al. Chromosome territory arrangement and homologous pairing in nuclei of *Arabidopsis thaliana* are predominantly random except for NOR-bearing chromosomes // *Chromosoma*. - 2004. - Vol. 113. - No 5. - P. 258–269.
16. Stadler S., Schnapp V., Mayer R., et al. The architecture of chicken chromosome territories changes during differentiation // *BMC Cell Biol*. - 2004. - Vol. 5. - No 1. - P. 44.
17. Misteli T. Beyond the Sequence: Cellular Organization of Genome Function // *Cell*. - 2007. - Vol. 128. - No 4. - P. 787–800.
18. Rosin L. F., Nguyen S. C., Joyce E. F. Condensin II drives large-scale folding and spatial partitioning of interphase chromosomes in *Drosophila* nuclei // *PLoS Genet*. - 2018. - Vol. 14. - No 7. - P. e1007393.
19. Vandriel R., Fransz P. Nuclear architecture and genome functioning in plants and animals: what can we learn from both? // *Experimental Cell Research*. - 2004. - Vol. 296. - No 1. - P. 86–90.
20. Cowan C. R., Carlton P. M., Cande W. Z. The Polar Arrangement of Telomeres in Interphase and Meiosis. Rabl Organization and the Bouquet // *Plant Physiology*. - 2001. - Vol. 125. - No 2. - P. 532–538.

21. Bauer C. R., Hartl T. A., Bosco G. Condensin II Promotes the Formation of Chromosome Territories by Inducing Axial Compaction of Polyploid Interphase Chromosomes // *PLoS Genet.* - 2012. - Vol. 8. - No 8. - P. e1002873.
22. Hoencamp C., Dudchenko O., Elbatsh A. M. O., et al. 3D genomics across the tree of life reveals condensin II as a determinant of architecture type // *Science.* - 2021. - Vol. 372. - No 6545. - P. 984–989.
23. Hochstrasser M., Sedat J. W. Three-dimensional organization of *Drosophila melanogaster* interphase nuclei. I. Tissue-specific aspects of polytene nuclear architecture. // *The Journal of Cell Biology.* - 1987. - Vol. 104. - No 6. - P. 1455–1470.
24. Hiraoka Y., Dernburg A. F., Parmelee S. J., et al. The onset of homologous chromosome pairing during *Drosophila melanogaster* embryogenesis. // *Journal of Cell Biology.* - 1993. - Vol. 120. - No 3. - P. 591–600.
25. Marshall W. F., Dernburg A. F., Harmon B., et al. Specific interactions of chromatin with the nuclear envelope: positional determination within the nucleus in *Drosophila melanogaster*. // *MBoC.* - 1996. - Vol. 7. - No 5. - P. 825–842.
26. Vazquez J., Belmont A. S., Sedat J. W. The Dynamics of Homologous Chromosome Pairing during Male *Drosophila* Meiosis // *Current Biology.* - 2002. - Vol. 12. - No 17. - P. 1473–1483.
27. Dej K. J., Spradling A. C. The endocycle controls nurse cell polytene chromosome structure during *Drosophila* oogenesis // *Development.* - 1999. - Vol. 126. - No 2. - P. 293–303.
28. Cremer T., Cremer C., Schneider T., et al. Analysis of chromosome positions in the interphase nucleus of Chinese hamster cells by laser-UV-microirradiation experiments // *Hum Genet.* - 1982. - Vol. 62. - No 3. - P. 201–209.
29. Haaf T., Schmid M. Chromosome topology in mammalian interphase nuclei // *Experimental Cell Research.* - 1991. - Vol. 192. - No 2. - P. 325–332.
30. Branco M. R., Pombo A. Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations // *PLoS Biol.* - 2006. - Vol. 4. - No 5. - P. e138.
31. Cremer T., Cremer M. Chromosome Territories // *Cold Spring Harbor Perspectives in Biology.* - 2010. - Vol. 2. - No 3. - P. a003889–a003889.
32. Solovei I., Kreysing M., Lanctôt C., et al. Nuclear Architecture of Rod Photoreceptor Cells Adapts to Vision in Mammalian Evolution // *Cell.* - 2009. - Vol. 137. - No 2. - P. 356–368.
33. Solovei I., Wang A. S., Thanisch K., et al. LBR and Lamin A/C Sequentially Tether Peripheral Heterochromatin and Inversely Regulate Differentiation // *Cell.* - 2013. - Vol. 152. - No 3. - P. 584–598.
34. Solovei I., Thanisch K., Feodorova Y. How to rule the nucleus: divide et impera // *Current Opinion in Cell Biology.* - 2016. - Vol. 40. - P. 47–59.
35. Pickersgill H., Kalverda B., Wit E. de, et al. Characterization of the *Drosophila melanogaster* genome at the nuclear lamina // *Nat Genet.* - 2006. - Vol. 38. - No 9. - P. 1005–1014.
36. Guelen L., Pagie L., Brasset E., et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions // *Nature.* - 2008. - Vol. 453. - No 7197. - P. 948–951.
37. Peric-Hupkes D., Meuleman W., Pagie L., et al. Molecular Maps of the Reorganization of Genome-Nuclear Lamina Interactions during Differentiation // *Molecular Cell.* - 2010. - Vol. 38. - No 4. - P. 603–613.
38. Koningsbruggen S. van, Gierliński M., Schofield P., et al. High-Resolution Whole-Genome Sequencing Reveals That Specific Chromatin Domains from Most Human Chromosomes Associate with Nucleoli // *MBoC.* - 2010. - Vol. 21. - No 21. - P. 3735–3748.
39. Németh A., Conesa A., Santoyo-Lopez J., et al. Initial Genomics of the Human Nucleolus // *PLoS Genet.* - 2010. - Vol. 6. - No 3. - P. e1000889.

40. Pederson T. The Nucleolus // *Cold Spring Harbor Perspectives in Biology*. - 2011. - Vol. 3. - No 3. - P. a000638–a000638.
41. Hall L. L., Carone D. M., Gomez A. V., et al. Stable C0T-1 Repeat RNA Is Abundant and Is Associated with Euchromatic Interphase Chromosomes // *Cell*. - 2014. - Vol. 156. - No 5. - P. 907–919.
42. Langer-Safer P. R., Levine M., Ward D. C. Immunological method for mapping genes on *Drosophila* polytene chromosomes. // *Proceedings of the National Academy of Sciences*. - 1982. - Vol. 79. - No 14. - P. 4381–4385.
43. Brown K. E., Guest S. S., Smale S. T., et al. Association of Transcriptionally Silent Genes with Ikaros Complexes at Centromeric Heterochromatin // *Cell*. - 1997. - Vol. 91. - No 6. - P. 845–854.
44. Brown K. E., Baxter J., Graf D., et al. Dynamic Repositioning of Genes in the Nucleus of Lymphocytes Preparing for Cell Division // *Molecular Cell*. - 1999. - Vol. 3. - No 2. - P. 207–217.
45. Zink D., Amaral M. D., Englmann A., et al. Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei // *Journal of Cell Biology*. - 2004. - Vol. 166. - No 6. - P. 815–825.
46. Lakadamyali M., Cosma M. P. Advanced microscopy methods for visualizing chromatin structure // *FEBS Letters*. - 2015. - Vol. 589. - No 20PartA. - P. 3023–3030.
47. Finn E. H., Pegoraro G., Brandão H. B., et al. Extensive Heterogeneity and Intrinsic Variation in Spatial Genome Organization // *Cell*. - 2019. - Vol. 176. - No 6. - P. 1502–1515.e10.
48. Szabo Q., Donjon A., Jerković I., et al. Regulation of single-cell genome organization into TADs and chromatin nanodomains // *Nat Genet*. - 2020. - Vol. 52. - No 11. - P. 1151–1157.
49. Schermelleh L., Carlton P. M., Haase S., et al. Subdiffraction Multicolor Imaging of the Nuclear Periphery with 3D Structured Illumination Microscopy // *Science*. - 2008. - Vol. 320. - No 5881. - P. 1332–1336.
50. Cogger V. C., McNerney G. P., Nyunt T., et al. Three-dimensional structured illumination microscopy of liver sinusoidal endothelial cell fenestrations // *Journal of Structural Biology*. - 2010. - Vol. 171. - No 3. - P. 382–388.
51. Sparmann A., Lohuizen M. van. Polycomb silencers control cell fate, development and cancer // *Nat Rev Cancer*. - 2006. - Vol. 6. - No 11. - P. 846–856.
52. Hediger F., Gasser S. M. Heterochromatin protein 1: don't judge the book by its cover! // *Current Opinion in Genetics & Development*. - 2006. - Vol. 16. - No 2. - P. 143–150.
53. Wit E. de, Greil F., Steensel B. van. High-Resolution Mapping Reveals Links of HP1 with Active and Inactive Chromatin Components // *PLoS Genet*. - 2007. - Vol. 3. - No 3. - P. e38.
54. Filion G. J., Bommel J. G. van, Braunschweig U., et al. Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in *Drosophila* Cells // *Cell*. - 2010. - Vol. 143. - No 2. - P. 212–224.
55. Zhimulev I. F., Zykova T. Yu., Goncharov F. P., et al. Genetic Organization of Interphase Chromosome Bands and Interbands in *Drosophila melanogaster* // *PLoS ONE*. - 2014. - Vol. 9. - No 7. - P. e101631.
56. Boldyreva L., Goncharov F., Demakova O., et al. Protein and Genetic Composition of Four Chromatin Types in *Drosophila melanogaster* Cell Lines // *CG*. - 2017. - Vol. 18. - No 2. - P. 214–226.
57. Sequeira-Mendes J., Aragüez I., Peiró R., et al. The Functional Topography of the *Arabidopsis* Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States // *The Plant Cell*. - 2014. - Vol. 26. - No 6. - P. 2351–2366.
58. Sotelo-Silveira M., Chávez Montes R. A., Sotelo-Silveira J. R., et al. Entering the Next Dimension: Plant Genomes in 3D // *Trends in Plant Science*. - 2018. - Vol. 23. - No 7. - P. 598–612.
59. Vergara Z., Gutierrez C. Emerging roles of chromatin in the maintenance of genome organization and function in plants // *Genome Biol*. - 2017. - Vol. 18. - No 1. - P. 96.

60. Grosberg A. Yu., Nechaev S. K., Shakhnovich E. I. The role of topological constraints in the kinetics of collapse of macromolecules // *J. Phys. France*. - 1988. - Vol. 49. - No 12. - P. 2095–2100.
61. Grosberg A., Rabin Y., Havlin S., et al. Crumpled Globule Model of the Three-Dimensional Structure of DNA // *Europhys. Lett.* - 1993. - Vol. 23. - No 5. - P. 373–378.
62. Fudenberg G., Mirny L. A. Higher-order chromatin structure: bridging physics and biology // *Current Opinion in Genetics & Development*. - 2012. - Vol. 22. - No 2. - P. 115–124.
63. Gibcus J. H., Samejima K., Goloborodko A., et al. A pathway for mitotic chromosome formation // *Science*. - 2018. - Vol. 359. - No 6376. - P. eaao6135.
64. Grosberg A. I., Khokhlov A. R., Gennes P.-G. de. Giant molecules: here, there, and everywhere. Hackensack, N.J: World Scientific, 2011. Вып. 2nd ed. 322 P.
65. Tark-Dame M., Driel R. van, Heermann D. W. Chromatin folding – from biology to polymer models and back // *Journal of Cell Science*. - 2011. - Vol. 124. - No 6. - P. 839–845.
66. Mirny L. A. The fractal globule as a model of chromatin architecture in the cell // *Chromosome Res.* - 2011. - Vol. 19. - No 1. - P. 37–51.
67. Dekker J., Steensel B. van. The Spatial Architecture of Chromosomes // *Handbook of Systems Biology*. : Elsevier, 2013. P. 137–151.
68. Yokota H., Engh G. van den, Hearst J. E., et al. Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human G0/G1 interphase nucleus. // *Journal of Cell Biology*. - 1995. - Vol. 130. - No 6. - P. 1239–1249.
69. Dekker J., Marti-Renom M. A., Mirny L. A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. // *Nature reviews. Genetics*. - 2013. - Vol. 14. - No 6. - P. 390–403.
70. Zhao Z., Tavoosidana G., Sjölander M., et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. // *Nature genetics*. - 2006. - Vol. 38. - No 11. - P. 1341–1347.
71. Dostie J., Richmond T. A., Arnaout R. A., et al. Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements // *Genome Research*. - 2006. - Vol. 16. - No 10. - P. 1299–1309.
72. Lieberman-aiden E., Berkum N. L. Van, Williams L., et al. Comprehensive Mapping of Long-Range Interactions Revels Folding Principles of the Human Genome // *Science (New York, N.Y.)*. - 2009. - Vol. 326. - No October. - P. 289–294.
73. Berkum N. L. van, Lieberman-Aiden E., Williams L., et al. Hi-C: a method to study the three-dimensional architecture of genomes. // *Journal of visualized experiments : JoVE*. - 2010. - Vol. 6. - No 39. - P. 1869.
74. Ramani V., Cusanovich D. A., Hause R. J., et al. Mapping 3D genome architecture through in situ DNase Hi-C // *Nat Protoc*. - 2016. - Vol. 11. - No 11. - P. 2104–2121.
75. Ramani V., Deng X., Qiu R., et al. Sci-Hi-C: A single-cell Hi-C method for mapping 3D genome organization in large number of single cells // *Methods*. - 2020. - Vol. 170. - P. 61–68.
76. Wit E. de, Laat W. de. A decade of 3C technologies: insights into nuclear organization // *Genes & Development*. - 2012. - Vol. 26. - No 1. - P. 11–24.
77. Ma W., Ay F., Lee C., et al. Using DNase Hi-C techniques to map global and local three-dimensional genome architecture at high resolution // *Methods*. - 2018. - Vol. 142. - P. 59–73.
78. Gridina M., Mozheiko E., Valeev E., et al. A cookbook for DNase Hi-C // *Epigenetics & Chromatin*. - 2021. - Vol. 14. - No 1. - P. 15.

79. Hsieh T.-H. S., Cattoglio C., Slobodyanyuk E., et al. Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding // *Molecular Cell*. - 2020. - Vol. 78. - No 3. - P. 539- - 553.e8.
80. Krietenstein N., Abraham S., Venev S. V., et al. Ultrastructural Details of Mammalian Chromosome Architecture // *Molecular Cell*. - 2020. - Vol. 78. - No 3. - P. 554- - 565.e7.
81. Jäger R., Migliorini G., Henrion M., et al. Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci // *Nat Commun*. - 2015. - Vol. 6. - No 1. - P. 6178.
82. Mifsud B., Tavares-Cadete F., Young A. N., et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C // *Nat Genet*. - 2015. - Vol. 47. - No 6. - P. 598–606.
83. Baxter J. S., Leavy O. C., Dryden N. H., et al. Capture Hi-C identifies putative target genes at 33 breast cancer risk loci // *Nat Commun*. - 2018. - Vol. 9. - No 1. - P. 1028.
84. Fullwood M. J., Liu M. H., Pan Y. F., et al. An oestrogen-receptor- α -bound human chromatin interactome // *Nature*. - 2009. - Vol. 462. - No 7269. - P. 58–64.
85. Li G., Fullwood M. J., Xu H., et al. ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing // *Genome Biol*. - 2010. - Vol. 11. - No 2. - P. R22.
86. Li X., Luo O. J., Wang P., et al. Long-read ChIA-PET for base-pair-resolution mapping of haplotype-specific chromatin interactions // *Nat Protoc*. - 2017. - Vol. 12. - No 5. - P. 899–915.
87. Li G., Ruan X., Auerbach R. K., et al. Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation // *Cell*. - 2012. - Vol. 148. - No 1–2. - P. 84–98.
88. Bertolini J. A., Favaro R., Zhu Y., et al. Mapping the Global Chromatin Connectivity Network for Sox2 Function in Neural Stem Cell Maintenance // *Cell Stem Cell*. - 2019. - Vol. 24. - No 3. - P. 462- - 476.e6.
89. Wang P., Feng Y., Zhu K., et al. *In situ* Chromatin Interaction Analysis Using Paired-End Tag Sequencing // *Current Protocols*. - 2021. - Vol. 1. - No 8.
90. Fang R., Yu M., Li G., et al. Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq // *Cell Res*. - 2016. - Vol. 26. - No 12. - P. 1345–1348.
91. Mumbach M. R., Rubin A. J., Flynn R. A., et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture // *Nat Methods*. - 2016. - Vol. 13. - No 11. - P. 919–922.
92. Flyamer I. M., Gassler J., Imakaev M., et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition // *Nature*. - 2017. - Vol. 544. - No 7648. - P. 110–114.
93. Nagano T., Lubling Y., Stevens T. J., et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure // *Nature*. - 2013. - Vol. 502. - No 7469. - P. 59–64.
94. Tan L., Xing D., Chang C.-H., et al. Three-dimensional genome structures of single diploid human cells // *Science*. - 2018. - Vol. 361. - No 6405. - P. 924–928.
95. Ranisavljevic N., Borensztein M., Ancelin K. Understanding Chromosome Structure During Early Mouse Development by a Single-Cell Hi-C Analysis // *Epigenetic Reprogramming During Mouse Embryogenesis Methods in Molecular Biology*. / под ред. K. Ancelin, M. Borensztein. New York, NY: Springer US, 2021. P. 283–293.
96. Stevens T. J., Lando D., Basu S., et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C // *Nature*. - 2017. - Vol. 544. - No 7648. - P. 59–64.
97. Kim H.-J., Yardımcı G. G., Bonora G., et al. Capturing cell type-specific chromatin compartment patterns by applying topic modeling to single-cell Hi-C data // *PLoS Comput Biol*. - 2020. - Vol. 16. - No 9. - P. e1008173.
98. Belaghzal H., Dekker J., Gibcus J. H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation // *Methods*. - 2017. - Vol. 123. - P. 56–65.

99. Servant N., Varoquaux N., Lajoie B. R., et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing // *Genome Biol.* - 2015. - Vol. 16. - P. 259.
100. Durand N. C., Shamim M. S., Machol I., et al. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments // *Cell Systems.* - 2016. - Vol. 3. - No 1. - P. 95–98.
101. Dudchenko O., Batra S. S., Omer A. D., et al. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds // *Science.* - 2017. - Vol. 356. - No 6333. - P. 92–95.
102. Dudchenko O., Shamim M. S., Batra S. S., et al. The Juicebox Assembly Tools module facilitates *de novo* assembly of mammalian genomes with chromosome-length scaffolds for under \$1000. : *Genomics*, 2018.
103. Biello R., Singh A., Godfrey C. J., et al. A chromosome-level genome assembly of the woolly apple aphid, *Eriosoma lanigerum* Hausmann (Hemiptera: Aphididae) // *Mol Ecol Resour.* - 2021. - Vol. 21. - No 1. - P. 316–326.
104. Davidson P. L., Guo H., Wang L., et al. Chromosomal-Level Genome Assembly of the Sea Urchin *Lytechinus variegatus* Substantially Improves Functional Genomic Analyses // *Genome Biology and Evolution.* - 2020. - Vol. 12. - No 7. - P. 1080–1086.
105. Garg V., Dudchenko O., Wang J., et al. Chromosome-length genome assemblies of six legume species provide insights into genome organization, evolution, and agronomic traits for crop improvement // *Journal of Advanced Research.* - 2021. - P. S2090123221002058.
106. Liang Q., Li H., Li S., et al. The genome assembly and annotation of yellowhorn (*Xanthoceras sorbifolium* Bunge) // *GigaScience.* - 2019. - Vol. 8. - No 6. - P. giz071.
107. Liu C., Zhang Y., Ren Y., et al. The genome of the golden apple snail *Pomacea canaliculata* provides insight into stress tolerance and invasive adaptation // *GigaScience.* - 2018. - Vol. 7. - No 9.
108. Ma Z., Zhang Y., Wu L., et al. High-quality genome assembly and resequencing of modern cotton cultivars provide resources for crop improvement // *Nat Genet.* - 2021. - Vol. 53. - No 9. - P. 1385–1391.
109. Schloissnig S., Kawaguchi A., Nowoshilow S., et al. The giant axolotl genome uncovers the evolution, scaling, and transcriptional control of complex gene loci // *Proc Natl Acad Sci USA.* - 2021. - Vol. 118. - No 15. - P. e2017176118.
110. Schreiber M., Mascher M., Wright J., et al. A Genome Assembly of the Barley ‘Transformation Reference’ Cultivar Golden Promise // *G3 Genes|Genomes|Genetics.* - 2020. - Vol. 10. - No 6. - P. 1823–1827.
111. Shingate P., Ravi V., Prasad A., et al. Chromosome-level genome assembly of the coastal horseshoe crab (*Tachypleus gigas*) // *Mol Ecol Resour.* - 2020. - Vol. 20. - No 6. - P. 1748–1760.
112. Wang L., Wu J., Liu X., et al. A high-quality genome assembly for the endangered golden snub-nosed monkey (*Rhinopithecus roxellana*) // *GigaScience.* - 2019. - Vol. 8. - No 8. - P. giz098.
113. <https://www.illumina.com/systems/sequencing-platforms/miseq/specifications.html>.
114. Turchaninova M. A., Davydov A., Britanova O. V., et al. High-quality full-length immunoglobulin profiling with unique molecular barcoding // *Nat Protoc.* - 2016. - Vol. 11. - No 9. - P. 1599–1616.
115. Fox E. J., Reid-Bayliss K. S., Emond M. J., et al. Accuracy of Next Generation Sequencing Platforms // *Next Gener Seq Appl.* - 2014. - Vol. 1. - P. 1000106.
116. Rang F. J., Kloosterman W. P., Ridder J. de. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy // *Genome Biol.* - 2018. - Vol. 19. - No 1. - P. 90.
117. Mahmoud M., Zywicki M., Twardowski T., et al. Efficiency of PacBio long read correction by 2nd generation Illumina sequencing // *Genomics.* - 2019. - Vol. 111. - No 1. - P. 43–49.
118. Dixon J. R., Jung I., Selvaraj S., et al. Chromatin architecture reorganization during stem cell differentiation // *Nature.* - 2015. - Vol. 518. - No 7539. - P. 331–336.

119. Rao S. S. P., Huntley M. H., Durand N. C., et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping // *Cell*. - 2014. - Vol. 159. - No 7. - P. 1665–1680.
120. Bonev B., Mendelson Cohen N., Szabo Q., et al. Multiscale 3D Genome Rewiring during Mouse Neural Development // *Cell*. - 2017. - Vol. 171. - No 3. - P. 557- - 572.e24.
121. Luperchio T. R., Wong X., Reddy K. L. Genome regulation at the peripheral zone: lamina associated domains in development and disease // *Current Opinion in Genetics & Development*. - 2014. - Vol. 25. - P. 50–61.
122. Wong X., Luperchio T. R., Reddy K. L. NET gains and losses: the role of changing nuclear envelope proteomes in genome regulation // *Current Opinion in Cell Biology*. - 2014. - Vol. 28. - P. 105–120.
123. Tomita S., Abdalla M. O. A., Fujiwara S., et al. Roles of long noncoding RNAs in chromosome domains: Roles of long noncoding RNAs in chromosome domains // *WIREs RNA*. - 2017. - Vol. 8. - No 2. - P. e1384.
124. Camozzi D., Capanni C., Cenni V., et al. Diverse lamin-dependent mechanisms interact to control chromatin dynamics: Focus on laminopathies // *Nucleus*. - 2014. - Vol. 5. - No 5. - P. 427–440.
125. Steensel B. van, Henikoff S. Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase // *Nat Biotechnol*. - 2000. - Vol. 18. - No 4. - P. 424–428.
126. Bommel J. G. van, Pagie L., Braunschweig U., et al. The Insulator Protein SU(HW) Fine-Tunes Nuclear Lamina Interactions of the *Drosophila* Genome // *PLoS ONE*. - 2010. - Vol. 5. - No 11. - P. e15013.
127. Chen S., Luperchio T. R., Wong X., et al. A Lamina-Associated Domain Border Governs Nuclear Lamina Interactions, Transcription, and Recombination of the *Tcrb* Locus // *Cell Reports*. - 2018. - Vol. 25. - No 7. - P. 1729- - 1740.e6.
128. Peric-Hupkes D., Steensel B. van. Role of the Nuclear Lamina in Genome Organization and Gene Expression // *Cold Spring Harbor Symposia on Quantitative Biology*. - 2010. - Vol. 75. - No 0. - P. 517–524.
129. Wen B., Wu H., Shinkai Y., et al. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells // *Nat Genet*. - 2009. - Vol. 41. - No 2. - P. 246–250.
130. Brunet A., Forsberg F., Fan Q., et al. Nuclear Lamin B1 Interactions With Chromatin During the Circadian Cycle Are Uncoupled From Periodic Gene Expression // *Front. Genet*. - 2019. - Vol. 10. - P. 917.
131. Meuleman W., Peric-Hupkes D., Kind J., et al. Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence // *Genome Research*. - 2013. - Vol. 23. - No 2. - P. 270–280.
132. Forsberg F., Brunet A., Ali T. M. L., et al. Interplay of lamin A and lamin B LADs on the radial positioning of chromatin // *Nucleus*. - 2019. - Vol. 10. - No 1. - P. 7–20.
133. Pascual-Reguant L., Blanco E., Galan S., et al. Lamin B1 mapping reveals the existence of dynamic and functional euchromatin lamin B1 domains // *Nat Commun*. - 2018. - Vol. 9. - No 1. - P. 3420.
134. Falk M., Feodorova Y., Naumova N., et al. Heterochromatin drives compartmentalization of inverted and conventional nuclei // *Nature*. - 2019. - Vol. 570. - No 7761. - P. 395–399.
135. Vigouroux C., Bonne G. Laminopathies: One Gene, Two Proteins, Five Diseases... // *Nuclear Envelope Dynamics in Embryos and Somatic Cells* / под ред. P. Collas. Boston, MA: Springer US, 2002. P. 153–172.
136. Worman H. J., Bonne G. “Laminopathies”: A wide spectrum of human diseases // *Experimental Cell Research*. - 2007. - Vol. 313. - No 10. - P. 2121–2133.
137. Kang S., Yoon M.-H., Park B.-J. Laminopathies; Mutations on single gene and various human genetic diseases // *BMB Rep*. - 2018. - Vol. 51. - No 7. - P. 327–337.
138. O’Keefe R. T., Henderson S. C., Spector D. L. Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences. // *Journal of Cell Biology*. - 1992. - Vol. 116. - No 5. - P. 1095–1110.

139. Farkash-Amar S., Lipson D., Polten A., et al. Global organization of replication time zones of the mouse genome // *Genome Research*. - 2008. - Vol. 18. - No 10. - P. 1562–1570.
140. Hiratani I., Ryba T., Itoh M., et al. Global Reorganization of Replication Domains During Embryonic Stem Cell Differentiation // *PLoS Biol.* - 2008. - Vol. 6. - No 10. - P. e245.
141. MacAlpine D. M. Coordination of replication and transcription along a *Drosophila* chromosome // *Genes & Development*. - 2004. - Vol. 18. - No 24. - P. 3094–3105.
142. Pope B. D., Ryba T., Dileep V., et al. Topologically associating domains are stable units of replication-timing regulation // *Nature*. - 2014. - Vol. 515. - No 7527. - P. 402–405.
143. Yaffe E., Farkash-Amar S., Polten A., et al. Comparative Analysis of DNA Replication Timing Reveals Conserved Large-Scale Chromosomal Architecture // *PLoS Genet.* - 2010. - Vol. 6. - No 7. - P. e1001011.
144. Dixon J. R., Selvaraj S., Yue F., et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions // *Nature*. - 2012. - Vol. 485. - No 7398. - P. 376–380.
145. Nora E. P., Lajoie B. R., Schulz E. G., et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre // *Nature*. - 2012. - Vol. 485. - No 7398. - P. 381–385.
146. Ali T., Renkawitz R., Bartkuhn M. Insulators and domains of gene expression // *Current Opinion in Genetics & Development*. - 2016. - Vol. 37. - P. 17–26.
147. Guo Y., Xu Q., Canzio D., et al. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function // *Cell*. - 2015. - Vol. 162. - No 4. - P. 900–910.
148. de Wit E., Vos E. S. M., Holwerda S. J. B., et al. CTCF Binding Polarity Determines Chromatin Looping // *Molecular Cell*. - 2015. - Vol. 60. - No 4. - P. 676–684.
149. Vietri Rudan M., Barrington C., Henderson S., et al. Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture // *Cell Reports*. - 2015. - Vol. 10. - No 8. - P. 1297–1309.
150. Seitan V. C., Faure A. J., Zhan Y., et al. Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments // *Genome Research*. - 2013. - Vol. 23. - No 12. - P. 2066–2077.
151. Zuin J., Dixon J. R., Reijden M. I. J. A. van der, et al. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells // *Proc Natl Acad Sci USA*. - 2014. - Vol. 111. - No 3. - P. 996–1001.
152. Yang R., Kerschner J. L., Gosalia N., et al. Differential contribution of *cis*-regulatory elements to higher order chromatin structure and expression of the *CFTR* locus // *Nucleic Acids Res.* - 2016. - Vol. 44. - No 7. - P. 3082–3094.
153. Hou C., Li L., Qin Z. S., et al. Gene Density, Transcription, and Insulators Contribute to the Partition of the *Drosophila* Genome into Physical Domains // *Molecular Cell*. - 2012. - Vol. 48. - No 3. - P. 471–484.
154. Sexton T., Yaffe E., Kenigsberg E., et al. Three-Dimensional Folding and Functional Organization Principles of the *Drosophila* Genome // *Cell*. - 2012. - Vol. 148. - No 3. - P. 458–472.
155. Dong P., Tu X., Chu P.-Y., et al. 3D Chromatin Architecture of Large Plant Genomes Determined by Local A/B Compartments // *Molecular Plant*. - 2017. - Vol. 10. - No 12. - P. 1497–1509.
156. Szabo Q., Jost D., Chang J.-M., et al. TADs are 3D structural units of higher-order chromosome organization in *Drosophila* // *Sci. Adv.* - 2018. - Vol. 4. - No 2. - P. eaar8082.
157. Ryba T., Hiratani I., Lu J., et al. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types // *Genome Research*. - 2010. - Vol. 20. - No 6. - P. 761–770.
158. Ulianov S. V., Khrameeva E. E., Gavrilov A. A., et al. Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains // *Genome Res.* - 2016. - Vol. 26. - No 1. - P. 70–84.

159. Weddington N., Stuy A., Hiratani I., et al. ReplicationDomain: a visualization tool and comparative database for genome-wide replication timing data // *BMC Bioinformatics*. - 2008. - Vol. 9. - No 1. - P. 530.
160. Phillips-Cremins J. E., Sauria M. E. G., Sanyal A., et al. Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment // *Cell*. - 2013. - Vol. 153. - No 6. - P. 1281–1295.
161. Rao S. S. P., Huntley M. H., Durand N. C., et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping // *Cell*. - 2014. - Vol. 159. - No 7. - P. 1665–1680.
162. Phillips-cremins J. E., Sauria M. E. G., Sanyal A., et al. Genomes During Lineage Commitment // 2014. - Vol. 153. - No 6. - P. 1281–1295.
163. Haarhuis J. H. I., Weide R. H. van der, Blomen V. A., et al. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension // *Cell*. - 2017. - Vol. 169. - No 4. - P. 693- - 707.e14.
164. Cattoni D. I., Cardozo Gizzi A. M., Georgieva M., et al. Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions // *Nat Commun*. - 2017. - Vol. 8. - No 1. - P. 1753.
165. Dermitzakis E. T., Clark A. G. Differential Selection After Duplication in Mammalian Developmental Genes // *Molecular Biology and Evolution*. - 2001. - Vol. 18. - No 4. - P. 557–562.
166. Borneman A. R., Gianoulis T. A., Zhang Z. D., et al. Divergence of Transcription Factor Binding Sites Across Related Yeast Species // *Science*. - 2007. - Vol. 317. - No 5839. - P. 815–819.
167. The ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project // *Nature*. - 2007. - Vol. 447. - No 7146. - P. 799–816.
168. Schmidt D., Wilson M. D., Ballester B., et al. Five-Vertebrate ChIP-seq Reveals the Evolutionary Dynamics of Transcription Factor Binding // *Science*. - 2010. - Vol. 328. - No 5981. - P. 1036–1040.
169. Harmston N., Ing-Simmons E., Tan G., et al. Topologically associating domains are ancient features that coincide with Metazoan clusters of extreme noncoding conservation // *Nat Commun*. - 2017. - Vol. 8. - No 1. - P. 441.
170. Rao S. S. P., Huang S.-C., Glenn St Hilaire B., et al. Cohesin Loss Eliminates All Loop Domains // *Cell*. - 2017. - Vol. 171. - No 2. - P. 305- - 320.e24.
171. Sanborn A. L., Rao S. S. P., Huang S.-C., et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes // *Proc Natl Acad Sci USA*. - 2015. - Vol. 112. - No 47. - P. E6456–E6465.
172. Schuijers J., Manteiga J. C., Weintraub A. S., et al. Transcriptional Dysregulation of MYC Reveals Common Enhancer-Docking Mechanism // *Cell Reports*. - 2018. - Vol. 23. - No 2. - P. 349–360.
173. Nora E. P., Goloborodko A., Valton A.-L., et al. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization // *Cell*. - 2017. - Vol. 169. - No 5. - P. 930- - 944.e22.
174. Thiecke M. J., Wutz G., Muhar M., et al. Cohesin-Dependent and -Independent Mechanisms Mediate Chromosomal Contacts between Promoters and Enhancers // *Cell Reports*. - 2020. - Vol. 32. - No 3. - P. 107929.
175. Hariharan N., Kelley D. E., Perry R. P. Delta, a transcription factor that binds to downstream elements in several polymerase II promoters, is a functionally versatile zinc finger protein. // *Proceedings of the National Academy of Sciences*. - 1991. - Vol. 88. - No 21. - P. 9799–9803.
176. Park K., Atchison M. L. Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta), that binds to the immunoglobulin kappa 3' enhancer and the immunoglobulin heavy-chain mu E1 site. // *Proceedings of the National Academy of Sciences*. - 1991. - Vol. 88. - No 21. - P. 9804–9808.
177. Shi Y., Seto E., Chang L.-S., et al. Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein // *Cell*. - 1991. - Vol. 67. - No 2. - P. 377–388.

178. Beagan J. A., Duong M. T., Titus K. R., et al. YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment // *Genome Res.* - 2017. - Vol. 27. - No 7. - P. 1139–1152.
179. Weintraub A. S., Li C. H., Zamudio A. V., et al. YY1 Is a Structural Regulator of Enhancer-Promoter Loops // *Cell.* - 2017. - Vol. 171. - No 7. - P. 1573- - 1588.e28.
180. Verheul T. C. J., Hijfte L. van, Perenthaler E., et al. The Why of YY1: Mechanisms of Transcriptional Regulation by Yin Yang 1 // *Front. Cell Dev. Biol.* - 2020. - Vol. 8. - P. 592164.
181. Zhang X., Jeong M., Huang X., et al. Large DNA Methylation Nadirs Anchor Chromatin Loops Maintaining Hematopoietic Stem Cell Identity // *Molecular Cell.* - 2020. - Vol. 78. - No 3. - P. 506- - 521.e6.
182. Long H. K., Sims D., Heger A., et al. Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates // *eLife.* - 2013. - Vol. 2. - P. e00348.
183. Xie W., Schultz M. D., Lister R., et al. Epigenomic Analysis of Multilineage Differentiation of Human Embryonic Stem Cells // *Cell.* - 2013. - Vol. 153. - No 5. - P. 1134–1148.
184. Jeong M., Sun D., Luo M., et al. Large conserved domains of low DNA methylation maintained by Dnmt3a // *Nat Genet.* - 2014. - Vol. 46. - No 1. - P. 17–23.
185. Padeken J., Heun P. Nucleolus and nuclear periphery: Velcro for heterochromatin // *Current Opinion in Cell Biology.* - 2014. - Vol. 28. - P. 54–60.
186. Chu C., Zhang Q. C., da Rocha S. T., et al. Systematic Discovery of Xist RNA Binding Proteins // *Cell.* - 2015. - Vol. 161. - No 2. - P. 404–416.
187. McHugh C. A., Chen C.-K., Chow A., et al. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3 // *Nature.* - 2015. - Vol. 521. - No 7551. - P. 232–236.
188. Sun L., Goff L. A., Trapnell C., et al. Long noncoding RNAs regulate adipogenesis // *Proc Natl Acad Sci USA.* - 2013. - Vol. 110. - No 9. - P. 3387–3392.
189. Hacisuleyman E., Goff L. A., Trapnell C., et al. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre // *Nat Struct Mol Biol.* - 2014. - Vol. 21. - No 2. - P. 198–206.
190. Yang F., Deng X., Ma W., et al. The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation // *Genome Biol.* - 2015. - Vol. 16. - No 1. - P. 52.
191. Barutcu A. R., Maass P. G., Lewandowski J. P., et al. A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus // *Nat Commun.* - 2018. - Vol. 9. - No 1. - P. 1444.
192. Schwarzer W., Abdennur N., Goloborodko A., et al. Two independent modes of chromatin organization revealed by cohesin removal // *Nature.* - 2017. - Vol. 551. - No 7678. - P. 51–56.
193. Hyman A. A., Weber C. A., Jülicher F. Liquid-Liquid Phase Separation in Biology // *Annu. Rev. Cell Dev. Biol.* - 2014. - Vol. 30. - No 1. - P. 39–58.
194. Li P., Banjade S., Cheng H.-C., et al. Phase transitions in the assembly of multivalent signalling proteins // *Nature.* - 2012. - Vol. 483. - No 7389. - P. 336–340.
195. Erdel F., Rippe K. Formation of Chromatin Subcompartments by Phase Separation // *Biophysical Journal.* - 2018. - Vol. 114. - No 10. - P. 2262–2270.
196. Larson A. G., Elnatan D., Keenen M. M., et al. Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin // *Nature.* - 2017. - Vol. 547. - No 7662. - P. 236–240.
197. Strom A. R., Emelyanov A. V., Mir M., et al. Phase separation drives heterochromatin domain formation // *Nature.* - 2017. - Vol. 547. - No 7662. - P. 241–245.

198. Nasmyth K. Disseminating the Genome: Joining, Resolving, and Separating Sister Chromatids During Mitosis and Meiosis // *Annu. Rev. Genet.* - 2001. - Vol. 35. - No 1. - P. 673–745.
199. Alipour E., Marko J. F. Self-organization of domain structures by DNA-loop-extruding enzymes // *Nucleic Acids Research.* - 2012. - Vol. 40. - No 22. - P. 11202–11212.
200. Fudenberg G., Abdennur N., Imakaev M., et al. Emerging Evidence of Chromosome Folding by Loop Extrusion // *Cold Spring Harb Symp Quant Biol.* - 2017. - Vol. 82. - P. 45–55.
201. Haering C. H., Farcas A.-M., Arumugam P., et al. The cohesin ring concatenates sister DNA molecules // *Nature.* - 2008. - Vol. 454. - No 7202. - P. 297–301.
202. Kim J.-S., He X., Liu J., et al. Systematic proteomics of endogenous human cohesin reveals an interaction with diverse splicing factors and RNA-binding proteins required for mitotic progression // *Journal of Biological Chemistry.* - 2019. - Vol. 294. - No 22. - P. 8760–8772.
203. Kong M., Cutts E. E., Pan D., et al. Human Condensin I and II Drive Extensive ATP-Dependent Compaction of Nucleosome-Bound DNA // *Molecular Cell.* - 2020. - Vol. 79. - No 1. - P. 99- - 114.e9.
204. Pradhan B., Barth R., Kim E., et al. SMC complexes can traverse physical roadblocks bigger than their ring size. : *Biophysics*, 2021.
205. Wutz G., Várnai C., Nagasaka K., et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins // *EMBO J.* - 2017. - Vol. 36. - No 24. - P. 3573–3599.
206. Nanni L., Ceri S., Logie C. Spatial patterns of CTCF sites define the anatomy of TADs and their boundaries // *Genome Biol.* - 2020. - Vol. 21. - No 1. - P. 197.
207. Kueng S., Hegemann B., Peters B. H., et al. Wapl Controls the Dynamic Association of Cohesin with Chromatin // *Cell.* - 2006. - Vol. 127. - No 5. - P. 955–967.
208. Huis in 't Veld P. J., Herzog F., Ladurner R., et al. Characterization of a DNA exit gate in the human cohesin ring // *Science.* - 2014. - Vol. 346. - No 6212. - P. 968–972.
209. Briño-Enríquez M. A., Moak S. L., Toledo M., et al. Cohesin Removal along the Chromosome Arms during the First Meiotic Division Depends on a NEK1-PP1 γ -WAPL Axis in the Mouse // *Cell Reports.* - 2016. - Vol. 17. - No 4. - P. 977–986.
210. Tedeschi A., Wutz G., Huet S., et al. Wapl is an essential regulator of chromatin structure and chromosome segregation // *Nature.* - 2013. - Vol. 501. - No 7468. - P. 564–568.
211. Luppino J. M., Park D. S., Nguyen S. C., et al. Cohesin promotes stochastic domain intermingling to ensure proper regulation of boundary-proximal genes // *Nat Genet.* - 2020. - Vol. 52. - No 8. - P. 840–848.
212. Ganji M., Shaltiel I. A., Bisht S., et al. Real-time imaging of DNA loop extrusion by condensin // *Science.* - 2018. - Vol. 360. - No 6384. - P. 102–105.
213. Vian L., Pękowska A., Rao S. S. P., et al. The Energetics and Physiological Impact of Cohesin Extrusion // *Cell.* - 2018. - Vol. 173. - No 5. - P. 1165- - 1178.e20.
214. Rowley M. J., Lyu X., Rana V., et al. Condensin II Counteracts Cohesin and RNA Polymerase II in the Establishment of 3D Chromatin Organization // *Cell Reports.* - 2019. - Vol. 26. - No 11. - P. 2890- - 2903.e3.
215. Busslinger G. A., Stocsits R. R., Lelij P. van der, et al. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl // *Nature.* - 2017. - Vol. 544. - No 7651. - P. 503–507.
216. Stigler J., Çamdere G. Ö., Koshland D. E., et al. Single-Molecule Imaging Reveals a Collapsed Conformational State for DNA-Bound Cohesin // *Cell Reports.* - 2016. - Vol. 15. - No 5. - P. 988–998.

217. Brandão H. B., Paul P., Berg A. A. van den, et al. RNA polymerases as moving barriers to condensin loop extrusion // *Proc Natl Acad Sci USA*. - 2019. - Vol. 116. - No 41. - P. 20489–20499.
218. Li Y., Haarhuis J. H. I., Sedeño Cacciatore Á., et al. The structural basis for cohesin–CTCF-anchored loops // *Nature*. - 2020. - Vol. 578. - No 7795. - P. 472–476.
219. Goloborodko A., Marko J. F., Mirny L. A. Chromosome Compaction by Active Loop Extrusion // *Biophysical Journal*. - 2016. - Vol. 110. - No 10. - P. 2162–2168.
220. Mirny L. A., Imakaev M., Abdennur N. Two major mechanisms of chromosome organization // *Current Opinion in Cell Biology*. - 2019. - Vol. 58. - P. 142–152.
221. Naumova N., Imakaev M., Fudenberg G., et al. Organization of the Mitotic Chromosome // *Science*. - 2013. - Vol. 342. - No 6161. - P. 948–953.
222. Ryzhkova A., Taskina A., Khabarova A., et al. Erythrocytes 3D genome organization in vertebrates // *Sci Rep*. - 2021. - Vol. 11. - No 1. - P. 4414.
223. Rowley M. J., Corces V. G. The three-dimensional genome: principles and roles of long-distance interactions // *Current Opinion in Cell Biology*. - 2016. - Vol. 40. - P. 8–14.
224. Rowley M. J., Nichols M. H., Lyu X., et al. Evolutionarily Conserved Principles Predict 3D Chromatin Organization // *Molecular Cell*. - 2017. - Vol. 67. - No 5. - P. 837–852.e7.
225. Szabo Q., Bantignies F., Cavalli G. Principles of genome folding into topologically associating domains // *Sci. Adv*. - 2019. - Vol. 5. - No 4. - P. eaaw1668.
226. Cubeñas-Potts C., Rowley M. J., Lyu X., et al. Different enhancer classes in *Drosophila* bind distinct architectural proteins and mediate unique chromatin interactions and 3D architecture // *Nucleic Acids Research*. - 2017. - Vol. 45. - No 4. - P. 1714–1730.
227. Ramírez F., Bhardwaj V., Arrigoni L., et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies // *Nat Commun*. - 2018. - Vol. 9. - No 1. - P. 189.
228. Maksimenko O., Bartkuhn M., Stakhov V., et al. Two new insulator proteins, Pita and ZIPIC, target CP190 to chromatin // *Genome Res*. - 2015. - Vol. 25. - No 1. - P. 89–99.
229. Van Bortle K., Ramos E., Takenaka N., et al. *Drosophila* CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains // *Genome Research*. - 2012. - Vol. 22. - No 11. - P. 2176–2187.
230. El-Sharnouby S., Fischer B., Magbanua J. P., et al. Regions of very low H3K27me3 partition the *Drosophila* genome into topological domains // *PLoS ONE*. - 2017. - Vol. 12. - No 3. - P. e0172725.
231. Wang Q., Sun Q., Czajkowsky D. M., et al. Sub-kb Hi-C in *D. melanogaster* reveals conserved characteristics of TADs between insect and mammalian cells // *Nat Commun*. - 2018. - Vol. 9. - No 1. - P. 188.
232. Schoborg T. A., Labrador M. The Phylogenetic Distribution of Non-CTCF Insulator Proteins Is Limited to Insects and Reveals that BEAF-32 Is *Drosophila* Lineage Specific // *J Mol Evol*. - 2010. - Vol. 70. - No 1. - P. 74–84.
233. Van Bortle K., Nichols M. H., Li L., et al. Insulator function and topological domain border strength scale with architectural protein occupancy // *Genome Biol*. - 2014. - Vol. 15. - No 5. - P. R82.
234. Eagen K. P., Hartl T. A., Kornberg R. D. Stable Chromosome Condensation Revealed by Chromosome Conformation Capture // *Cell*. - 2015. - Vol. 163. - No 4. - P. 934–946.
235. Chathoth K. T., Zabet N. R. Chromatin architecture reorganization during neuronal cell differentiation in *Drosophila* genome // *Genome Res*. - 2019. - Vol. 29. - No 4. - P. 613–625.
236. Kaushal A., Mohana G., Dorier J., et al. CTCF loss has limited effects on global genome architecture in *Drosophila* despite critical regulatory functions // *Nat Commun*. - 2021. - Vol. 12. - No 1. - P. 1011.

237. Eagen K. P., Aiden E. L., Kornberg R. D. Polycomb-mediated chromatin loops revealed by a subkilobase-resolution chromatin interaction map // *Proc Natl Acad Sci USA*. - 2017. - Vol. 114. - No 33. - P. 8764–8769.
238. Ogiyama Y., Schuettengruber B., Papadopoulos G. L., et al. Polycomb-Dependent Chromatin Looping Contributes to Gene Silencing during *Drosophila* Development // *Molecular Cell*. - 2018. - Vol. 71. - No 1. - P. 73- - 88.e5.
239. King T. D., Leonard C. J., Cooper J. C., et al. Recurrent Losses and Rapid Evolution of the Condensin II Complex in Insects // *Molecular Biology and Evolution*. - 2019. - Vol. 36. - No 10. - P. 2195–2204.
240. Dufourt J., Trullo A., Hunter J., et al. Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs // *Nat Commun*. - 2018. - Vol. 9. - No 1. - P. 5194.
241. Lewis E. B. The Theory and Application of a New Method of Detecting Chromosomal Rearrangements in *Drosophila melanogaster* // *The American Naturalist*. - 1954. - Vol. 88. - No 841. - P. 225–239.
242. Gelbart W. M., Wu C.-T. INTERACTIONS OF ZESTE MUTATIONS WITH LOCI EXHIBITING TRANSVECTION EFFECTS IN *DROSOPHILA MELANOGASTER* // *Genetics*. - 1982. - Vol. 102. - No 2. - P. 179–189.
243. Gelbart W. M. Synapsis-dependent allelic complementation at the decapentaplegic gene complex in *Drosophila melanogaster*. // *Proceedings of the National Academy of Sciences*. - 1982. - Vol. 79. - No 8. - P. 2636–2640.
244. Geyer P. K., Green M. M., Corces V. G. Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome: the molecular basis of transvection in *Drosophila* // *EMBO J*. - 1990. - Vol. 9. - No 7. - P. 2247–2256.
245. Koeman J. M., Russell R. C., Tan M.-H., et al. Somatic Pairing of Chromosome 19 in Renal Oncocytoma Is Associated with Deregulated ELGN2-Mediated Oxygen-Sensing Response // *PLoS Genet*. - 2008. - Vol. 4. - No 9. - P. e1000176.
246. Duncan I. W. Transvection Effects in *Drosophila* // *Annu. Rev. Genet*. - 2002. - Vol. 36. - No 1. - P. 521–556.
247. LaSalle J. M., Lalande M. Homologous Association of Oppositely Imprinted Chromosomal Domains // *Science*. - 1996. - Vol. 272. - No 5262. - P. 725–728.
248. Rodriguez J. D., Myrick D. A., Falcatori I., et al. A Model for Epigenetic Inhibition via Transvection in the Mouse // *Genetics*. - 2017. - Vol. 207. - No 1. - P. 129–138.
249. Apostolou E., Thanos D. Virus Infection Induces NF- κ B-Dependent Interchromosomal Associations Mediating Monoallelic IFN- β Gene Expression // *Cell*. - 2008. - Vol. 134. - No 1. - P. 85–96.
250. Ling J. Q. CTCF Mediates Interchromosomal Colocalization Between *Igf2/H19* and *Wsb1/Nf1* // *Science*. - 2006. - Vol. 312. - No 5771. - P. 269–272.
251. Simonis M., Klous P., Splinter E., et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C) // *Nat Genet*. - 2006. - Vol. 38. - No 11. - P. 1348–1354.
252. Spilianakis C. G., Lalioti M. D., Town T., et al. Interchromosomal associations between alternatively expressed loci // *Nature*. - 2005. - Vol. 435. - No 7042. - P. 637–645.
253. Joyce E. F., Erceg J., Wu C. Pairing and anti-pairing: a balancing act in the diploid genome // *Current Opinion in Genetics & Development*. - 2016. - Vol. 37. - P. 119–128.
254. Lim B., Heist T., Levine M., et al. Visualization of Transvection in Living *Drosophila* Embryos // *Molecular Cell*. - 2018. - Vol. 70. - No 2. - P. 287- - 296.e6.
255. Fukaya T., Lim B., Levine M. Enhancer Control of Transcriptional Bursting // *Cell*. - 2016. - Vol. 166. - No 2. - P. 358–368.

256. AlHaj Abed J., Erceg J., Goloborodko A., et al. Highly structured homolog pairing reflects functional organization of the *Drosophila* genome // *Nat Commun.* - 2019. - Vol. 10. - No 1. - P. 4485.
257. Erceg J., AlHaj Abed J., Goloborodko A., et al. The genome-wide multi-layered architecture of chromosome pairing in early *Drosophila* embryos // *Nat Commun.* - 2019. - Vol. 10. - No 1. - P. 4486.
258. Hnisz D., Shrinivas K., Young R. A., et al. A Phase Separation Model for Transcriptional Control // *Cell.* - 2017. - Vol. 169. - No 1. - P. 13–23.
259. Mellert D. J., Truman J. W. Transvection Is Common Throughout the *Drosophila* Genome // *Genetics.* - 2012. - Vol. 191. - No 4. - P. 1129–1141.
260. Hartl T. A., Smith H. F., Bosco G. Chromosome Alignment and Transvection Are Antagonized by Condensin II // *Science.* - 2008. - Vol. 322. - No 5906. - P. 1384–1387.
261. Kolesnikova T. D. Banding Pattern of Polytene Chromosomes as a Representation of Universal Principles of Chromatin Organization into Topological Domains // *Biochemistry Moscow.* - 2018. - Vol. 83. - No 4. - P. 338–349.
262. Harbach R. E. The Phylogeny and Classification of Anopheles // *Anopheles mosquitoes - New insights into malaria vectors / под ред. S. Manguin. : InTech, 2013.*
263. Coluzzi M. A Polytene Chromosome Analysis of the *Anopheles gambiae* Species Complex // *Science.* - 2002. - Vol. 298. - No 5597. - P. 1415–1418.
264. Sallum M. A. M., Schultz T. R., Foster P. G., et al. Phylogeny of Anophelinae (Diptera: Culicidae) based on nuclear ribosomal and mitochondrial DNA sequences: Molecular phylogeny of Anophelinae // *Systematic Entomology.* - 2002. - Vol. 27. - No 3. - P. 361–382.
265. Hebert P. D. N., Cywinska A., Ball S. L., et al. Biological identifications through DNA barcodes // *Proc. R. Soc. Lond. B.* - 2003. - Vol. 270. - No 1512. - P. 313–321.
266. Zomuanpui R., Ringnghe L., Brindha S., et al. ITS2 characterization and *Anopheles* species identification of the subgenus *Cellia* // *Acta Tropica.* - 2013. - Vol. 125. - No 3. - P. 309–319.
267. Lorenz C., Alves J. M. P., Foster P. G., et al. Phylogeny and temporal diversification of mosquitoes (Diptera: Culicidae) with an emphasis on the Neotropical fauna // *Syst Entomol.* - 2021. - Vol. 46. - No 4. - P. 798–811.
268. Poinar G., Zavortink T. J., Brown A. *Priscoculex burmanicus* n. gen. et sp. (Diptera: Culicidae: Anophelinae) from mid-Cretaceous Myanmar amber // *Historical Biology.* - 2020. - Vol. 32. - No 9. - P. 1157–1162.
269. Borkent A., Grimaldi D. A. The Earliest Fossil Mosquito (Diptera: Culicidae), in Mid-Cretaceous Burmese Amber // *an.* - 2004. - Vol. 97. - No 5. - P. 882–888.
270. Poinar G., Zavortnik, Pike, et al. *Paleoculicis minutus* (Diptera: Culicidae) n. gen., n. sp., from Cretaceous Canadian amber, with a summary of described fossil mosquitoes // *ACTA GEOLOGICA HISPANICA.* - 2000. - Vol. 35. - No 1–2. - P. 119–128.
271. Zavortink T. J., Poinar G. O. *Anopheles (Nyssorhynchus) dominicanus* sp. n. (Diptera: Culicidae) from Dominican Amber // *an.* - 2000. - Vol. 93. - No 6. - P. 1230–1235.
272. Statz G. Neue Dipteren (Nematocera) aus dem Oberoligozän von Rott. III. Familie: Limnobiidae (Stelzmücken), IV. Familie: Tipulidae (Schnaken), V. Familie: Culicidae (Stechmücken) // *Palaeontographica Abteilung A.* - 1944. - Vol. A095. - No 3–6. - P. 93–121.
273. Brown W. M., George M., Wilson A. C. Rapid evolution of animal mitochondrial DNA. // *Proceedings of the National Academy of Sciences.* - 1979. - Vol. 76. - No 4. - P. 1967–1971.
274. Liang J., Sharakhov I. V. Premeiotic and meiotic failures lead to hybrid male sterility in the *Anopheles gambiae* complex // *Proc. R. Soc. B.* - 2019. - Vol. 286. - No 1906. - P. 20191080.

275. Liang J., Hodge J. M., Sharakhov I. V. Asymmetric Phenotypes of Sterile Hybrid Males From Reciprocal Crosses Between Species of the *Anopheles gambiae* Complex // *Front. Ecol. Evol.* - 2021. - Vol. 9. - P. 660207.
276. Danks H. V. The range of insect dormancy responses // *Eur. J. Entomol.* - 2002. - Vol. 99. - No 2. - P. 127–142.
277. Diniz D. F. A., Albuquerque C. M. R. de, Oliva L. O., et al. Diapause and quiescence: dormancy mechanisms that contribute to the geographical expansion of mosquitoes and their evolutionary success // *Parasites Vectors.* - 2017. - Vol. 10. - No 1. - P. 310.
278. Ruzzante L., Reijnders M. J. M. F., Waterhouse R. M. Of Genes and Genomes: Mosquito Evolution and Diversity // *Trends in Parasitology.* - 2019. - Vol. 35. - No 1. - P. 32–51.
279. Sinka M. E. Global Distribution of the Dominant Vector Species of Malaria // *Anopheles mosquitoes - New insights into malaria vectors / под ред. S. Manguin.* : InTech, 2013.
280. Pantoja-Sánchez H., Gomez S., Velez V., et al. Precopulatory acoustic interactions of the New World malaria vector *Anopheles albimanus* (Diptera: Culicidae) // *Parasites Vectors.* - 2019. - Vol. 12. - No 1. - P. 386.
281. Gómez G. F., Márquez E. J., Gutiérrez L. A., et al. Geometric morphometric analysis of Colombian *Anopheles albimanus* (Diptera: Culicidae) reveals significant effect of environmental factors on wing traits and presence of a metapopulation // *Acta Tropica.* - 2014. - Vol. 135. - P. 75–85.
282. Loyola E. G., González-Cerón L., Rodríguez M. H., et al. *Anopheles albimanus* (Diptera: Culicidae) Host Selection Patterns in Three Ecological Areas of the Coastal Plains of Chiapas, Southern Mexico // *Journal of Medical Entomology.* - 1993. - Vol. 30. - No 3. - P. 518–523.
283. Escobar D., Ascencio K., Ortiz A., et al. Blood Meal Sources of *Anopheles* spp. in Malaria Endemic Areas of Honduras // *Insects.* - 2020. - Vol. 11. - No 7. - P. 450.
284. <https://phil.cdc.gov/>.
285. Becker N., Petrić D., Zgomba M., et al. Systematics // *Mosquitoes and Their Control.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. P. 3–8.
286. Dorn P. L., Justi S., Krafsur E. S., et al. Genetics of Major Insect Vectors // *Genetics and Evolution of Infectious Diseases.* : Elsevier, 2017. P. 341–382.
287. Kingan S. B., Urban J., Lambert C. C., et al. A high-quality genome assembly from a single, field-collected spotted lanternfly (*Lycorma delicatula*) using the PacBio Sequel II system // *GigaScience.* - 2019. - Vol. 8. - No 10. - P. giz122.
288. Mwangangi J. M., Mbogo C. M., Orindi B. O., et al. Shifts in malaria vector species composition and transmission dynamics along the Kenyan coast over the past 20 years // *Malar J.* - 2013. - Vol. 12. - No 1. - P. 13.
289. Bartilol B., Omedo I., Mbogo C., et al. Bionomics and ecology of *Anopheles merus* along the East and Southern Africa coast // *Parasites Vectors.* - 2021. - Vol. 14. - No 1. - P. 84.
290. Sinka M. E., Bangs M. J., Manguin S., et al. The dominant *Anopheles* vectors of human malaria in the Asia-Pacific region: occurrence data, distribution maps and bionomic précis // *Parasites Vectors.* - 2011. - Vol. 4. - No 1. - P. 89.
291. Schneider I. Establishment of three diploid cell lines of *Anopheles stephensi* (Diptera: Culicidae) // *J Cell Biol.* - 1969. - Vol. 42. - No 2. - P. 603–606.
292. Dana A. N., Hong Y. S., Kern M. K., et al. Gene expression patterns associated with blood-feeding in the malaria mosquito *Anopheles gambiae* // *BMC Genomics.* - 2005. - Vol. 6. - No 1. - P. 5.
293. Ayala D., Ullastres A., González J. Adaptation through chromosomal inversions in *Anopheles* // *Front. Genet.* - 2014. - Vol. 5.

294. Dobzhansky T. Genetics of natural populations; experiments on chromosomes of *Drosophila pseudoobscura* from different geographic regions // *Genetics*. - 1948. - Vol. 33. - No 6. - P. 588–602.
295. Powell J. R. Progress and prospects in evolutionary biology: the *Drosophila* model. New York: Oxford University Press, 1997.
296. Berthelot C., Muffato M., Abecassis J., et al. The 3D Organization of Chromatin Explains Evolutionary Fragile Genomic Regions // *Cell Reports*. - 2015. - Vol. 10. - No 11. - P. 1913–1924.
297. Artemov G. N., Bondarenko S. M., Naumenko A. N., et al. Partial-arm translocations in evolution of malaria mosquitoes revealed by high-coverage physical mapping of the *Anopheles atroparvus* genome // *BMC Genomics*. - 2018. - Vol. 19. - No 1. - P. 278.
298. Torre A. della, Costantini C., Besansky N. J., et al. Speciation Within *Anopheles gambiae* -- the Glass Is Half Full // *Science*. - 2002. - Vol. 298. - No 5591. - P. 115–117.
299. Sharakhova M. V., Xia A., Leman S. C., et al. Arm-specific dynamics of chromosome evolution in malaria mosquitoes // *BMC Evol Biol*. - 2011. - Vol. 11. - No 1. - P. 91.
300. Gildenhart M., Rono E. K., Diarra A., et al. Mosquito microevolution drives *Plasmodium falciparum* dynamics // *Nat Microbiol*. - 2019. - Vol. 4. - No 6. - P. 941–947.
301. Fouet C., Gray E., Besansky N. J., et al. Adaptation to Aridity in the Malaria Mosquito *Anopheles gambiae*: Chromosomal Inversion Polymorphism and Body Size Influence Resistance to Desiccation // *PLoS ONE*. - 2012. - Vol. 7. - No 4. - P. e34841.
302. Gray E. M., Rocca K. A., Costantini C., et al. Inversion 2La is associated with enhanced desiccation resistance in *Anopheles gambiae* // *Malar J*. - 2009. - Vol. 8. - No 1. - P. 215.
303. Riehle M. M. Natural Malaria Infection in *Anopheles gambiae* Is Regulated by a Single Genomic Control Region // *Science*. - 2006. - Vol. 312. - No 5773. - P. 577–579.
304. Speciation: natural processes, genetics and biodiversity / под ред. P. Michalak. New York: Nova Biomedical, 2013. 264 P.
305. Muller H. J. Bearings of the ‘*Drosophila*’ work on systematics // *The New Systematics* edited by Huxley J. Clarendon. - 1940. - P. 185–268.
306. Lee Y., Collier T. C., Sanford M. R., et al. Chromosome Inversions, Genomic Differentiation and Speciation in the African Malaria Mosquito *Anopheles gambiae* // *PLoS ONE*. - 2013. - Vol. 8. - No 3. - P. e57887.
307. Bhutkar A., Schaeffer S. W., Russo S. M., et al. Chromosomal Rearrangement Inferred From Comparisons of 12 *Drosophila* Genomes // *Genetics*. - 2008. - Vol. 179. - No 3. - P. 1657–1680.
308. Khrameeva E. E., Fudenberg G., Gelfand M. S., et al. History of chromosome rearrangements reflects the spatial organization of yeast chromosomes // *J. Bioinform. Comput. Biol*. - 2016. - Vol. 14. - No 02. - P. 1641002.
309. Cox F. E. History of the discovery of the malaria parasites and their vectors // *Parasites Vectors*. - 2010. - Vol. 3. - No 1. - P. 5.
310. White N. J. *Plasmodium knowlesi*: The Fifth Human Malaria Parasite // *Clinical Infectious Diseases*. - 2008. - Vol. 46. - No 2. - P. 172–173.
311. Sutherland C. J., Tanomsing N., Nolder D., et al. Two Nonrecombining Sympatric Forms of the Human Malaria Parasite *Plasmodium ovale* Occur Globally // *J INFECT DIS*. - 2010. - Vol. 201. - No 10. - P. 1544–1550.
312. Allison A. C. Protection Afforded by Sickle-cell Trait Against Subtertian Malarial Infection // *BMJ*. - 1954. - Vol. 1. - No 4857. - P. 290–294.
313. Nagel R. L., Fleming A. F. Genetic epidemiology of the β s gene // *Baillière’s Clinical Haematology*. - 1992. - Vol. 5. - No 2. - P. 331–365.

314. Kramer L. D., Ciota A. T. Dissecting vectorial capacity for mosquito-borne viruses // *Current Opinion in Virology*. - 2015. - Vol. 15. - P. 112–118.
315. Osta M. A. Effects of Mosquito Genes on Plasmodium Development // *Science*. - 2004. - Vol. 303. - No 5666. - P. 2030–2032.
316. Blandin S. A., Wang-Sattler R., Lamacchia M., et al. Dissecting the Genetic Basis of Resistance to Malaria Parasites in *Anopheles gambiae* // *Science*. - 2009. - Vol. 326. - No 5949. - P. 147–150.
317. Mitri C., Jacques J.-C., Thiery I., et al. Fine Pathogen Discrimination within the APL1 Gene Family Protects *Anopheles gambiae* against Human and Rodent Malaria Species // *PLoS Pathog.* - 2009. - Vol. 5. - No 9. - P. e1000576.
318. Crawford J. E., Bischoff E., Garnier T., et al. Evidence for Population-Specific Positive Selection on Immune Genes of *Anopheles gambiae* // *G3 Genes|Genomes|Genetics*. - 2012. - Vol. 2. - No 12. - P. 1505–1519.
319. Ruiz J. L., Yerbanga R. S., Lefèvre T., et al. Chromatin changes in *Anopheles gambiae* induced by *Plasmodium falciparum* infection // *Epigenetics & Chromatin*. - 2019. - Vol. 12. - No 1. - P. 5.
320. Ruiz J. L., Ranford-Cartwright L. C., Gómez-Díaz E. The regulatory genome of the malaria vector *Anopheles gambiae* : integrating chromatin accessibility and gene expression // *NAR Genomics and Bioinformatics*. - 2021. - Vol. 3. - No 1. - P. lqaa113.
321. Mitri C., Bischoff E., Eiglmeier K., et al. Gene copy number and function of the APL1 immune factor changed during *Anopheles* evolution // *Parasites Vectors*. - 2020. - Vol. 13. - No 1. - P. 18.
322. Blandin S., Shiao S.-H., Moita L. F., et al. Complement-Like Protein TEP1 Is a Determinant of Vectorial Capacity in the Malaria Vector *Anopheles gambiae* // *Cell*. - 2004. - Vol. 116. - No 5. - P. 661–670.
323. Pompon J., Levashina E. A. A New Role of the Mosquito Complement-like Cascade in Male Fertility in *Anopheles gambiae* // *PLoS Biol.* - 2015. - Vol. 13. - No 9. - P. e1002255.
324. Gómez-Díaz E., Rivero A., Chandre F., et al. Insights into the epigenomic landscape of the human malaria vector *Anopheles gambiae* // *Front. Genet.* - 2014. - Vol. 5.
325. Ruiz J. L., Gómez-Díaz E. The second life of *Plasmodium* in the mosquito host: gene regulation on the move // *Briefings in Functional Genomics*. - 2019. - Vol. 18. - No 5. - P. 313–357.
326. Ay F., Bunnik E. M., Varoquaux N., et al. Three-dimensional modeling of the *P. falciparum* genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression // *Genome Research*. - 2014. - Vol. 24. - No 6. - P. 974–988.
327. Bunnik E. M., Cook K. B., Varoquaux N., et al. Changes in genome organization of parasite-specific gene families during the *Plasmodium* transmission stages // *Nat Commun.* - 2018. - Vol. 9. - No 1. - P. 1910.
328. Bunnik E. M., Venkat A., Shao J., et al. Comparative 3D genome organization in apicomplexan parasites // *Proc Natl Acad Sci USA*. - 2019. - Vol. 116. - No 8. - P. 3183–3192.
329. Lukyanchikova V., Fishman V., Sharakhov I. In situ Hi-C for mosquito embryos. : Protocol Exchange, 2022.
330. Goltsev Y., Rezende G. L., Vranizan K., et al. Developmental and evolutionary basis for drought tolerance of the *Anopheles gambiae* embryo // *Developmental Biology*. - 2009. - Vol. 330. - No 2. - P. 462–470.
331. <https://tmcaculator.neb.com/#!/main>.
332. Sharakhova M. V., Xia A., Mcalister S. I., et al. A Standard Cytogenetic Photomap for the Mosquito *Anopheles stephensi* (Diptera: Culicidae): Application for Physical Mapping // *J Med Entomol.* - 2006. - Vol. 43. - No 5. - P. 861–866.
333. *Protocols for Cytogenetic Mapping of Arthropod Genomes* / под ред. I. V. Sharakhov. : CRC Press, 2014. 526 P.

334. Artemov G. N., Peery A. N., Jiang X., et al. The Physical Genome Mapping of *Anopheles albimanus* Corrected Scaffold Misassemblies and Identified Interarm Rearrangements in Genus *Anopheles* // *G3*. - 2017. - Vol. 7. - No 1. - P. 155–164.
335. Durand N. C., Robinson J. T., Shamim M. S., et al. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom // *Cell Systems*. - 2016. - Vol. 3. - No 1. - P. 99–101.
336. George P., Sharakhova M. V., Sharakhov I. V. High-resolution cytogenetic map for the African malaria vector *Anopheles gambiae*: Cytogenetic map for *Anopheles gambiae* // *Insect Molecular Biology*. - 2010. - Vol. 19. - No 5. - P. 675–682.
337. Sharakhova M. V., Unger M. F., Tu Z., et al. A Physical Map for an Asian Malaria Mosquito, *Anopheles stephensi* // *The American Journal of Tropical Medicine and Hygiene*. - 2010. - Vol. 83. - No 5. - P. 1023–1027.
338. Kamali M., Sharakhova M. V., Baricheva E., et al. An Integrated Chromosome Map of Microsatellite Markers and Inversion Breakpoints for an Asian Malaria Mosquito, *Anopheles stephensi* // *Journal of Heredity*. - 2011. - Vol. 102. - No 6. - P. 719–726.
339. https://www.encodeproject.org/documents/75926e4b-77aa-4959-8ca7-87efc39d79/@@download/attachment/comp_doc_7july2018_final.pdf.
340. Thawornwattana Y., Dalquen D., Yang Z. Coalescent Analysis of Phylogenomic Data Confidently Resolves the Species Relationships in the *Anopheles gambiae* Species Complex // *Molecular Biology and Evolution*. - 2018. - Vol. 35. - No 10. - P. 2512–2527.
341. Jiang X., Peery A., Hall A. B., et al. Genome analysis of a major urban malaria vector mosquito, *Anopheles stephensi* // *Genome Biol*. - 2014. - Vol. 15. - No 9. - P. 459.
342. Waterhouse R. M., Seppey M., Simão F. A., et al. Using BUSCO to Assess Insect Genomic Resources // *Insect Genomics Methods in Molecular Biology*. / под ред. S. J. Brown, M. E. Pfrender. New York, NY: Springer New York, 2019. P. 59–74.
343. Kamali M., Xia A., Tu Z., et al. A New Chromosomal Phylogeny Supports the Repeated Origin of Vectorial Capacity in Malaria Mosquitoes of the *Anopheles gambiae* Complex // *PLoS Pathog*. - 2012. - Vol. 8. - No 10. - P. e1002960.
344. Corbett-Detig R. B., Said I., Calzetta M., et al. Fine-Mapping Complex Inversion Breakpoints and Investigating Somatic Pairing in the *Anopheles gambiae* Species Complex Using Proximity-Ligation Sequencing // *Genetics*. - 2019. - Vol. 213. - No 4. - P. 1495–1511.
345. Coluzzi M., Di Deco M., Cancrini G. Chromosomal inversions in *Anopheles stephensi* // *Parassitologia*. - 1973. - Vol. 15. - No 1. - P. 129–136.
346. Mahmood F., Sakai R. K. Inversion polymorphisms in natural populations of *Anopheles stephensi* // *Can. J. Genet. Cytol*. - 1984. - Vol. 26. - No 5. - P. 538–546.
347. Jones M. D. R. Inversion polymorphism and circadian flight activity in the mosquito *Anopheles stephensi* List. (Diptera, Culicidae) // *Bull. Entomol. Res*. - 1974. - Vol. 64. - No 2. - P. 305–311.
348. Shetty N. J., Hariprasad T. P. N., Sanil D., et al. Chromosomal inversions among insecticide-resistant strains of *Anopheles stephensi* Liston, a malaria mosquito // *Parasitol Res*. - 2013. - Vol. 112. - No 11. - P. 3851–3857.
349. D'alessandro G., Lazzaro G. F., Mariani M. Effect of DDT selection pressure on the frequency of chromosomal structures in *Anopheles atroparvus* // *Bull World Health Organ*. - 1957. - Vol. 16. - No 4. - P. 859–864.
350. Mosna E., Rivosecchi L., Ascher K. R. Studies on insecticide-resistant anophelines. I. Chromosome arrangements in a dieldrin-selected strain of *Anopheles atroparvus* // *Bull World Health Organ*. - 1958. - Vol. 19. - No 2. - P. 297–301.
351. Стегний В.Н. Популяционная генетика и эволюция малярийных комаров. Томск: Изд-во Том. университета, 1991. 136 P.







352. Москаев А.В., Гордеев М.И. Видовой состав и экологические характеристики личиночных биотопов малярийных комаров юга европейской части России // Мед. паразитология и паразитар. болезни. - 2016. - No 4. - P. 31–35.
353. Перевозкин В.П., Бондарчук С.С., Гордеев М.И. Популяционно-видовая структура малярийных комаров (Diptera, Culicidae) Прикаспийской низменности и Кумо-Манычской впадины // Мед. паразитология и паразитар. болезни. - 2012. - No 1. - P. 12–17.
354. Ayala F. J., Coluzzi M. Chromosome speciation: Humans, *Drosophila*, and mosquitoes // Proceedings of the National Academy of Sciences. - 2005. - Vol. 102. - No Supplement 1. - P. 6535–6542.
355. Powell J. R., Petrarca V., Torre A. della, et al. Population structure, speciation, and introgression in the *Anopheles gambiae* complex // Parassitologia. - 1999. - Vol. 41. - No 1–3. - P. 101–113.
356. Sharakhov I. V., White B. J., Sharakhova M. V., et al. Breakpoint structure reveals the unique origin of an interspecific chromosomal inversion (2La) in the *Anopheles gambiae* complex // Proceedings of the National Academy of Sciences. - 2006. - Vol. 103. - No 16. - P. 6258–6262.
357. Rocca K. A., Gray E. M., Costantini C., et al. 2La chromosomal inversion enhances thermal tolerance of *Anopheles gambiae* larvae // Malar J. - 2009. - Vol. 8. - No 1. - P. 147.
358. Cassone B. J., Molloy M. J., Cheng C., et al. Divergent transcriptional response to thermal stress by *Anopheles gambiae* larvae carrying alternative arrangements of inversion 2La: DIVERGENT TRANSCRIPTIONAL RESPONSE TO HEAT // Molecular Ecology. - 2011. - Vol. 20. - No 12. - P. 2567–2580.
359. Brooke B. D., Hunt R. H., Coetzee M. Resistance to dieldrin + fipronil assort with chromosome inversion 2La in the malaria vector *Anopheles gambiae* // Med Vet Entomol. - 2000. - Vol. 14. - No 2. - P. 190–194.
360. Beier J. C., Petrarca V. Intraspecific Chromosomal Polymorphism in the *Anopheles Gambiae* Complex as a Factor Affecting Malaria Transmission in the Kisumu Area of Kenya // The American Journal of Tropical Medicine and Hygiene. - 1992. - Vol. 46. - No 2. - P. 229–237.
361. White B. J., Cheng C., Sangaré D., et al. The Population Genomics of Trans-Specific Inversion Polymorphisms in *Anopheles gambiae* // Genetics. - 2009. - Vol. 183. - No 1. - P. 275–288.
362. McBride C. S., Baier F., Omondi A. B., et al. Evolution of mosquito preference for humans linked to an odorant receptor // Nature. - 2014. - Vol. 515. - No 7526. - P. 222–227.
363. Rose N. H., Sylla M., Badolo A., et al. Climate and Urbanization Drive Mosquito Preference for Humans // Current Biology. - 2020. - Vol. 30. - No 18. - P. 3570–3579.e6.
364. Lukyanchikova V., Nuriddinov M., Belokopytova P., et al. *Anopheles* mosquitoes reveal new principles of 3D genome organization in insects // Nat Commun. - 2022. - Vol. 13. - No 1. - P. 1960.
365. Vontas J., Grigoraki L., Morgan J., et al. Rapid selection of a pyrethroid metabolic enzyme CYP9K1 by operational malaria control activities // Proc Natl Acad Sci USA. - 2018. - Vol. 115. - No 18. - P. 4619–4624.
366. Kim Y., Stanley D. Eicosanoid Signaling in Insect Immunology: New Genes and Unresolved Issues // Genes. - 2021. - Vol. 12. - No 2. - P. 211.
367. Liang H.-L., Nien C.-Y., Liu H.-Y., et al. The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila* // Nature. - 2008. - Vol. 456. - No 7220. - P. 400–403.
368. Li L., Lyu X., Hou C., et al. Widespread Rearrangement of 3D Chromatin Organization Underlies Polycomb-Mediated Stress-Induced Silencing // Molecular Cell. - 2015. - Vol. 58. - No 2. - P. 216–231.
369. Battulin N., Fishman V. S., Mazur A. M., et al. Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach // Genome Biol. - 2015. - Vol. 16. - P. 77.
370. Fishman V., Battulin N., Nuriddinov M., et al. 3D organization of chicken genome demonstrates evolutionary conservation of topologically associated domains and highlights unique architecture of erythrocytes' chromatin // Nucleic Acids Research. - 2019. - Vol. 47. - No 2. - P. 648–665.

371. Khalil A., Grant J. L., Caddle L. B., et al. Chromosome territories have a highly nonspherical morphology and nonrandom positioning // *Chromosome Res.* - 2007. - Vol. 15. - No 7. - P. 899–916.
372. Sehgal N., Fritz A. J., Morris K., et al. Gene density and chromosome territory shape // *Chromosoma.* - 2014. - Vol. 123. - No 5. - P. 499–513.
373. George P., Kinney N. A., Liang J., et al. Three-dimensional Organization of Polytene Chromosomes in Somatic and Germline Tissues of Malaria Mosquitoes // *Cells.* - 2020. - Vol. 9. - No 2. - P. 339.
374. Bickhart D. M., Rosen B. D., Koren S., et al. Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome // *Nat Genet.* - 2017. - Vol. 49. - No 4. - P. 643–650.
375. Hoshino A., Matsunaga T. M., Sakamoto T., et al. Hi-C Revolution: From a Snapshot of DNA–DNA Interaction in a Single Cell to Chromosome-Scale *De Novo* Genome Assembly // *CYTOLOGIA.* - 2017. - Vol. 82. - No 3. - P. 223–226.
376. Sanford M. R., Ramsay S., Cornel A. J., et al. A preliminary investigation of the relationship between water quality and *Anopheles gambiae* larval habitats in western Cameroon // *Malar J.* - 2013. - Vol. 12. - No 1. - P. 225.
377. Costantini C., Ayala D., Guelbeogo W. M., et al. Living at the edge: biogeographic patterns of habitat segregation conform to speciation by niche expansion in *Anopheles gambiae* // *BMC Ecol.* - 2009. - Vol. 9. - No 1. - P. 16.
378. Simard F., Ayala D., Kamdem G., et al. Ecological niche partitioning between *Anopheles gambiae* molecular forms in Cameroon: the ecological side of speciation // *BMC Ecol.* - 2009. - Vol. 9. - No 1. - P. 17.
379. Hoffmann A. A., Rieseberg L. H. Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? // *Annu. Rev. Ecol. Evol. Syst.* - 2008. - Vol. 39. - No 1. - P. 21–42.
380. Schalbetter S. A., Fudenberg G., Baxter J., et al. Principles of meiotic chromosome assembly revealed in *S. cerevisiae* // *Nat Commun.* - 2019. - Vol. 10. - No 1. - P. 4795.
381. Mascher M., Gundlach H., Himmelbach A., et al. A chromosome conformation capture ordered sequence of the barley genome // *Nature.* - 2017. - Vol. 544. - No 7651. - P. 427–433.
382. Concia L., Veluchamy A., Ramirez-Prado J. S., et al. Wheat chromatin architecture is organized in genome territories and transcription factories // *Genome Biol.* - 2020. - Vol. 21. - No 1. - P. 104.
383. Pouokam M., Cruz B., Burgess S., et al. The Rab1 configuration limits topological entanglement of chromosomes in budding yeast // *Sci Rep.* - 2019. - Vol. 9. - No 1. - P. 6795.
384. Grubert F., Srivas R., Spacek D. V., et al. Landscape of cohesin-mediated chromatin loops in the human genome // *Nature.* - 2020. - Vol. 583. - No 7818. - P. 737–743.
385. Cavalli G. Chromosome kissing // *Current Opinion in Genetics & Development.* - 2007. - Vol. 17. - No 5. - P. 443–450.
386. Bantignies F., Roure V., Comet I., et al. Polycomb-Dependent Regulatory Contacts between Distant Hox Loci in *Drosophila* // *Cell.* - 2011. - Vol. 144. - No 2. - P. 214–226.
387. Espinola S. M., Götz M., Bellec M., et al. Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early *Drosophila* development // *Nat Genet.* - 2021. - Vol. 53. - No 4. - P. 477–486.
388. Ing-Simmons E., Vaid R., Bing X. Y., et al. Independence of chromatin conformation and gene regulation during *Drosophila* dorsoventral patterning // *Nat Genet.* - 2021. - Vol. 53. - No 4. - P. 487–499.
389. Misteli T., Finn E. H. Chromatin architecture is a flexible foundation for gene expression // *Nat Genet.* - 2021. - Vol. 53. - No 4. - P. 426–427.

390. Boyle S., Flyamer I. M., Williamson I., et al. A central role for canonical PRC1 in shaping the 3D nuclear landscape // *Genes Dev.* - 2020. - Vol. 34. - No 13–14. - P. 931–949.
391. Kundu S., Ji F., Sunwoo H., et al. Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation // *Molecular Cell.* - 2017. - Vol. 65. - No 3. - P. 432- - 446.e5.
392. Eeftens J. M., Kapoor M., Michieletto D., et al. Polycomb condensates can promote epigenetic marks but are not required for sustained chromatin compaction // *Nat Commun.* - 2021. - Vol. 12. - No 1. - P. 5888.
393. Plys A. J., Davis C. P., Kim J., et al. Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2 // *Genes Dev.* - 2019. - Vol. 33. - No 13–14. - P. 799–813.
394. Tatavosian R., Kent S., Brown K., et al. Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation // *J Biol Chem.* - 2019. - Vol. 294. - No 5. - P. 1451–1463.
395. Loubiere V., Delest A., Thomas A., et al. Coordinate redeployment of PRC1 proteins suppresses tumor formation during *Drosophila* development // *Nat Genet.* - 2016. - Vol. 48. - No 11. - P. 1436–1442.
396. Feodorova Y., Falk M., Mirny L. A., et al. Viewing Nuclear Architecture through the Eyes of Nocturnal Mammals // *Trends in Cell Biology.* - 2020. - Vol. 30. - No 4. - P. 276–289.
397. Sinka M. E., Bangs M. J., Manguin S., et al. The dominant *Anopheles* vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic précis // *Parasites Vectors.* - 2010. - Vol. 3. - No 1. - P. 117.

ПРИЛОЖЕНИЕ 1

Сводная таблица, описывающая предпочтения пять видов малярийных комаров в источнике питания, месте кормления и месте отдыха. (+) предпочтительное поведение, (+/-) нетипичное, но встречающееся поведение, (-) нет данных. Оформлена по материалам [279,397].

Вид	Места обитания личинок	Host - Хозяин		Предпочтения в кормлении		Предпочтения в отдыхе		Другие комментарии
								
<i>An. albimanus</i>	Солнечные, солоноватые или пресные водоемы; чистая, неподвижная или проточная вода, содержащая высшие растения или водоросли	+	+		+		+	Нападает на закате/ночью
<i>An. merus</i>	Солнечные или затененные, пресные или солоноватые водоемы; чистая или мутная, стоячая вода с высшими растениями или водорослями	+	+	+/-	+	+/-	+	Несмотря на то, что является прибрежным обитателем, не подвержен зависимости от приливов и уровней солености
<i>An. coluzzii</i> <i>Mopti</i>	Солнечные (иногда затененные), мелкие водоемы; пресная (иногда солоноватая), чистая или мутная, неподвижная или проточная вода с высшими растениями и водорослями или без них; временные, часто созданные руками человека среды обитания личинок, такие как рисовые поля или затопленные районы	+	+/-	+	+	+/-	+	короткий период развития личинок и часто встречается в местах обитания личинок, связанных с деятельностью человека
<i>An. stephensi</i>	Солнечные или затененные водоемы; свежая (иногда солоноватая), чистая или мутная,	+/-	+	+	+/-	+	-	Один из немногих видов малярийных комаров,

	неподвижная (иногда текущая) вода с высшими растениями или водорослями (иногда без растительности)							распространенных на урбанизированных территориях
<i>An. atroparvus</i>	Солнечные солоноватые места обитания личинок - болота и лагуны вдоль побережья (иногда пресные: каналы, канавы, берега рек, бассейны в руслах рек и рисовые поля); прохладная вода умеренной солености, не более 10 частей на 1000; предпочтительно содержание высоконитчатых водорослей	+/-	+	+	+	+		Отдыхает и диапаузирует в сараях и вольерах для домашних животных

ПРИЛОЖЕНИЕ 2

Сводная таблица по статистике Hi-C-библиотек, характеризующая качество полученных данных.

Вид	<i>An. albimanus</i> AalbS4	<i>An. atroparvus</i> AatrE4	<i>An. coluzzii</i> <i>Mopti</i> AcolN2	<i>An. merus</i> <i>эмбрионы</i> AmerM5	<i>An. merus</i> <i>взрослые</i> AmerM5	<i>An. stephensi</i> AsteI4
Суммарное количество прочтений	67,591,619	217,895,615	231,961,292	188,700,705	129,643,770	253,760,093
Нормальные прочтения	30,787,793 (45.55%)	103,420,169 (47.46%)	108,826,067 (46.92%)	78,420,401 (41.56%)	53,633,589 (41.37%)	129,373,218 (50.98%)
Химерные прочтения	30,295,709 (44.82%)	72,384,597 (33.22%)	81,994,702 (35.35%)	53,989,004 (28.61%)	35,096,848 (27.07%)	64,754,256 (25.52%)
Химерные сомнительные прочтения	2,614,742 (3.87%)	15,313,087 (7.03%)	33,974,675 (14.65%)	16,609,540 (8.80%)	11,671,013 (9.00%)	21,361,880 (8.42%)
Некартировано	3,893,375 (5.76%)	26,777,762 (12.29%)	7,165,848 (3.09%)	39,681,760 (21.03%)	29,242,320 (22.56%)	38,270,739 (15.08%)
Прочтения, содержащие лигазный мотив	46,422,310 (68.68%)	142,902,260 (65.58%)	168,629,612 (72.70%)	125,132,781 (66.31%)	82,596,867 (63.71%)	75,068,793 (29.58%)
Выровненные (нормальные+химерные)	61,083,502 (90.37%)	175,804,766 (80.68%)	190,820,769 (82.26%)	132,409,405 (70.17%)	88,730,437 (68.44%)	194,127,474 (76.50%)
Уникальные прочтения	56,129,336 (83.04%)	150,572,935 (69.10%)	165,637,626 (71.41%)	114,668,816 (60.77%)	77,371,708 (59.68%)	135,257,540 (53.30%)
ПЦР-дубликаты	3,502,506 (5.18%)	22,928,735 (10.52%)	21,846,321 (9.42%)	15,999,379 (8.48%)	11,326,157 (8.74%)	55,601,774 (21.91%)
Оптические дубликаты	1,451,660 (2.15%)	2,303,096 (1.06%)	3,336,822 (1.44%)	1,741,210 (0.92%)	32,572 (0.03%)	3,268,160 (1.29%)
Оценка комплексности (сложности) библиотеки	487,552,124	597,237,052	740,698,381	489,080,571	317,065,109	260,265,053

Intra-fragment Reads	2,461,570 (3.64% / 4.39%)	8,205,249 (3.77% / 5.45%)	11,381,671 (4.91% / 6.87%)	5,708,656 (3.03% / 4.98%)	3,065,395 (2.36% / 3.96%)	16,939,144 (6.68% / 12.52%)
Ниже лимита MAPQ	682,547 (1.01% / 1.22%)	5,901,863 (2.71% / 3.92%)	29,040,047 (12.52% / 17.53%)	22,026,541 (11.67% / 19.21%)	13,482,444 (10.40% / 17.43%)	5,777,870 (2.28% / 4.27%)
Hi-C-контакты	52,985,219 (78.39% / 94.40%)	136,465,823 (62.63% / 90.63%)	125,215,908 (53.98% / 75.60%)	86,933,619 (46.07% / 75.81%)	60,823,869 (46.92% / 78.61%)	112,540,526 (44.35% / 83.20%)
Прочтения, содержащие лигазный мотив*	23,987,831 (35.49% / 42.74%)	60,551,709 (27.79% / 40.21%)	61,809,999 (26.65% / 37.32%)	40,425,437 (21.42% / 35.25%)	26,855,598 (20.71% / 34.71%)	27,906,559 (11.00% / 20.63%)
Межхромосомные контакты	6,862,999 (10.15% / 12.23%)	24,255,276 (11.13% / 16.11%)	21,566,658 (9.30% / 13.02%)	29,961,382 (15.88% / 26.13%)	6,270,919 (4.84% / 8.10%)	18,456,987 (7.27% / 13.65%)
Внутрихромосомные контакты	46,122,220 (68.24% / 82.17%)	112,210,547 (51.50% / 74.52%)	103,649,250 (44.68% / 62.58%)	56,972,237 (30.19% / 49.68%)	54,552,950 (42.08% / 70.51%)	94,083,539 (37.08% / 69.56%)
Близкие взаимодействия (<20 Кб)	13,468,625 (19.93% / 24.00%)	25,525,319 (11.71% / 16.95%)	29,262,843 (12.62% / 17.67%)	13,002,707 (6.89% / 11.34%)	16,850,740 (13.00% / 21.78%)	37,468,994 (14.77% / 27.70%)
Дальние взаимодействия (>20 Кб)	32,653,284 (48.31% / 58.18%)	86,684,768 (39.78% / 57.57%)	74,385,329 (32.07% / 44.91%)	43,969,011 (23.30% / 38.34%)	37,702,135 (29.08% / 48.73%)	56,613,127 (22.31% / 41.86%)

ПРИЛОЖЕНИЕ 3

Таблица расположения крупных гетерохроматиновых блоков согласно тепловой карте Hi-C.

Вид	X PсН, Mб	Координаты	2R ИИ, Mб	Координаты	2L ИИ, Mб	Координаты	3R ИИ, Mб	Координаты	3L ИИ, Mб	Координаты
<i>An. coluzzii</i>	6.59	20,290,000 - 26,880,000					2.26	38,180,000 40,440,000		
<i>An. merus</i>	6.54	19,680,000 - 26,220,000			2.99	7,570,000- 10,560,000	1.985	38,250,000 - 40,235,000	1.38	2,670,000- 4,050,000
<i>An. stephensi</i>	3.44	16,345,000 - 19,780,000								
<i>An. atroparvus</i>	9.12	17,400,000 - 26,520,000	0.94	31,275,000 - 32,210,000		-	1.02	10,900,000 - 11,915,000	1.12	16,195,000- 17,315,000
<i>An. albimanus</i>							1.32	30,030,000 - 31,350,000		

ПРИЛОЖЕНИЕ 4

Таблица данных по координатам анкоров (оснований) всех обнаруженных дальних хроматиновых петель.

Вид	Хромосома	Геномные координаты 1го анкера петли	Геномные координаты 2го анкера петли	Суммарное количество дальних хроматиновых петель на геном
<i>An. albimanus</i> AalbS4	X:	4,595,000-4,715,000	6,590,000-6,675,000	3
	2R:	34,730,000-34,905,000	35,925,000-36,070,000	
	2L:	25,600,000-25,685,000	29,255,000-29,380,000	
<i>An. atroparvus</i> AatrE4	X:	5,320,000-5,755,000	11,100,000-11,510,000	4
	X:	620,000-790,000	10,010,000-10,505,000	
	2R:	11,210,000-11,700,000	23,055,000-23,700,000	
	3R:	7,620,000-8,160,000	38,550,000-38,950,000	
<i>An. coluzzii</i> <i>Mopti</i> AcolN2	X:	3,410,000-3,535,000	6,020,000-6,120,000	4
	X:	7,370,000-7,610,000	14,950,000-15,400,000	
	2R:	41,330,000-41,445,000	43,045,000-43,225,000	
	2R:	41,125,000-41,300,000	46,435,000-46,785,000	
<i>An. merus_ эмбрионы</i> AmerM5 (A) - <i>An. merus_ взрослые</i> AmerM5	X:	7,150,000-7,385,000	14,610,000-14,900,000	
	X:	3,335,000-3,395,000	5,740,000-5,925,000	
	2R (A):	26,783,000-26,815,000	39,725,000-39,825,000	
	2R:	41,670,000-41,840,000	43,635,000-43,945,000	

	3R (A):	19,780,000-19,935,000	35,890,000-35,990,000	9
	3R (A):	19,780,000-19,935,000	37,515,000-37,740,000	
	3R (A):	29,975,000-30,090,000	44,340,000-44,470,000	
	3R (A):	13,840,000-14,040,000	46,090,000-46,170,000	
	3R (A):	17,220,000-17,330,000	45,910,000-45,940,000	
<i>An. stephensi</i> AsteI4	X:	7,215,000-7,355,000	9,670,000-9,725,000	6
	X:	9,435,000-9,815,000	14,665,000-15,130,000	
	2R:	37,345,000-37,525,000	42,900,000-43,125,000	
	3R:	10,780,000-10,910,000	38,430,000-38,580,000	
	3R:	16,075,000-16,295,000	24,265,000-24,585,000	
	3R:	16,075,000-16,295,000	27,435,000-27,610,000	

ПРИЛОЖЕНИЕ 5

Таблица генов-ортологов, расположенных в основаниях дальних хроматиновых петель на X-хромосоме.

Молекулярная функция гена	<i>An. coluzzii</i>	<i>An. stephensi</i>	<i>An. albimanus</i>	<i>An. atroparvus</i>	<i>An. merus</i>
Cytochrome P450 CYP9K1	<u>AGAP000818</u>	<u>ASTEI01072</u>	AALB003283*	<u>AATE016607</u>	<u>AMEM003066</u>
nuclear receptor subfamily 2 group E member (Tailless)	<u>AGAP000819</u>	<u>ASTEI01071</u>	AALB003284*	<u>AATE006814</u>	<u>AMEM005434</u>
ANK_REP_REGIO N domain-containing protein	<u>AGAP000413</u>	<u>ASTEI06259</u>	AALB007797*	AATE005968	<u>AMEM017290</u>
prostaglandin reductase 1	<u>AGAP000414</u>	<u>ASTEI06258</u>	AALB006758	<u>AATE002874</u>	<u>AMEM010926</u>
hydroxyisourate hydrolase activity	<u>AGAP000415</u>	<u>ASTEI06257</u>	AALB006759	<u>AATE016248</u>	<u>AMEM011098</u>
Non-specific serine/threonine protein kinase	<u>AGAP029239</u>	<u>ASTEI01297</u>	AALB006314	<u>AATE016354</u>	<u>AMEM010384</u>
integrin beta subunit	<u>AGAP000815</u>	ASTEI07790	AALB006392	<u>AATE000443</u>	<u>AMEM009509</u>
metal ion binding	<u>AGAP029244</u>	<u>ASTEI06256</u>	AALB006313	<u>AATE018169</u>	<u>AMEM014919</u>
integral component of membrane	<u>AGAP000420</u>	<u>ASTEI06255</u>	AALB006411	<u>AATE017335</u>	AMEM017711
(heparan sulfate)-glucosamine 3-sulfotransferase 3	<u>AGAP000422</u>	<u>ASTEI06254</u>	AALB006570	<u>AATE004104</u>	<u>AMEM010094</u>
Frataxin homolog, mitochondrial	<u>AGAP000813</u>	ASTEI07792	AALB009856	<u>AATE002095</u>	AMEM001133
phosphatidylinositol binding	<u>AGAP000814</u>	ASTEI07791	AALB006417	<u>AATE012817</u>	-

Таблица генов-ортологов, расположенных в основаниях дальних хроматиновых петель на 2R-хромосоме.

Молекулярная функция гена	<i>An. coluzzii</i>	<i>An. stephensi</i>	<i>An. albimanus</i>	<i>An. atroparvus</i>	<i>An. merus</i>
ubiquitin thioesterase ZRANB1	<u>AGAP003853</u>	<u>ASTEI00780</u>	<u>AALB003080</u>	<u>AATE000898</u>	<u>AMEM006320</u>
snRNA-activating protein complex subunit 3	<u>AGAP003852</u>	<u>ASTEI00779</u>	<u>AALB003081</u>	<u>AATE008621</u>	<u>AMEM002040</u>
26 proteasome complex subunit	<u>AGAP003851</u>	<u>ASTEI00778</u>	<u>AALB003082</u>	<u>AATE007015</u>	-
signal peptide	<u>AGAP013134</u>	<u>ASTEI00777</u>	<u>AALB003083</u>	<u>AATE018674</u>	<u>AMEM017317</u>
DNA-binding transcription factor activity	<u>AGAP003726</u>	<u>ASTEI00671</u>	AALB007121	<u>AATE016879</u>	<u>AMEM010172</u>
integral component of membrane	<u>AGAP003849</u>	ASTEI00776	AALB003084	<u>AATE004695</u>	<u>AMEM017419</u>

Гены-ортологи отмечены подчеркнутым жирным шрифтом, если они расположены в основаниях петель.

Гены-ортологи, расположенные вблизи оснований петель (ближе 1 МБ) отмечены звездочкой (*).

ПРИЛОЖЕНИЕ 6

Таблица данных флуоресцентных проб, разработанная для визуализации дальних хроматиновых петель.

Вид	Хром-а	Размер петли	Координаты анкоров (AsteI4, AatrE4, AcolN2)	Ген	Координаты начала и конца пробы* (AatrE3 and AgamP4)	Локация пробы	Праймер F	Праймер R
<i>An. stephensi</i>	2R	5.5 Мб	37,345,000-37,525,000	ASTEIO0671	-	экзон 1	GCCTTAAC TACACCGG TATCGACT G	ACGAGGGTG AAATTAAT TGTTCCA
					-	экзон 2	ATCGATGG TTGTTC CGATGAT	AGTAAATTC TTCGCTCCG GTAGAGC
					-	экзон 2	CTCTACCG GAGCGAAG AATTTACT C	CACCGCTCG AATAATCTT CATAGTG
			42,900,000-43,125,000	ASTEIO0780	-	экзон 2	GATCGAAC ССАААТСТ AGCAAGC	TACGTACAA AGCGAGCAG TACCATT
					-	экзон 2	GGCCAGTA ACATCATG GТАААСАА С	AGTAAATTC CAGCGACCA GTTCAAT
					-	экзон 1	CATTTACC TACCAAGA ACGGAGCA G	ATTTTGTCCT TCAGCTCGG TTAAGT
<i>An. atroparvus</i>	X	6 Мб	5,320,000-5,755,000	AATE015765	5,530,177-5,529,721	экзон 3	GGTCAACC AGTTCTGG ATAATGT	AGCCGAGGG TGAATATTT GATA
					5,529,272-5,528,832	экзон 4	CTTCGTCC CGACTATG ACTAGC	AGAAAGGGA ACGTACAGC TTGA
					5,527,780-5,527,361	экзон 7	ACCGTACT GCGTCTAT ССААСТ	TACTCGTTT GAATGTGGC ATCT
			11,100,000-11,510,000	AATE000335	10,706,948-10,707,380	экзон 5	GGCTGGTA AAGGATCT TAGCAA	TACACTGAG CTGTGGTAC ATGG
					10,707,359-10,707,788		CCATGTAC CACAGCTC AGTGTAT	GGCTCATTT GAGGAGGAC ACT
			AATE012799	10,717,128-10,716,621	экзон 18	CGACTTCC GCAGTAAT ACATGA	TCGACATCC GAAACAGTA GTTG	
			2R	12 Мб	11,210,000-11,700,000	AATE006247	11,443,983-11,449,877	интрон 1
	11,449,851-11,459,008	интрон 3					GAATAATC TATCCGCA TCCCGGTT CCC	TAGCTCAGG TCTGTAAGC GAGAАСТСА
	AATE006247/ AATE008403	11,395,959-11,399,755				Межгенный регион	TGGGTAAA GCGGATTC CATGAAGC TAC	GACCCGATG TGTAGCACT GCTTGATTA

			23,055,000- 23,700,000	AATE01 2005/ AATE00 5110	23,037,924- 23,045,339	Межгенн ый регион	TGGAGAGT TTCGTGGA CAAATGTG TGT	GCTGTTTGA GGTG TCTTTGCTCA ACAG	
				AATE01 3943/ AATE01 9726	23,079,044- 23,085,516	Межгенн ый регион/ин трон 1	GCGTTGTC CGATGGAA GATTAATT CGC	GTGCAAATG GATC TAAGGAGGC CTTGA	
			7,620,000- 8,160,000	AATE01 6879	7,634,101- 7,633,547	экзон 2	ACGTGAAG GAAGAACC GCTACAG	AATGGGTAG AACTCGTCC ATCATTT	
	7,627,402- 7,626,918	экзон 7			TCAGCTTG TTTGGGGT ATGAACTT A	CTCTACCCT ACGCTCTCT TCCTTTG			
	7,626,307- 7,625,802				CCCGTACA CATATGAA CCAAAACA T	AAGCACAGC TATTACAGT GGTACGC			
	38,550,000- 38,950,000	AATE01 5805	37,557,565- 37,558,117	экзон 2	ACAGGCTA AATATCAC TGCATCGA G	ATCCGTACA CGTCAACGT TTCTAAC			
		AATE00 8621	37,552,328- 37,552,911	экзон 1	TACAAGCC AAAAATCGA ATAAGGTC A	TCGTAATAC GGTCCTTCA TCACAGT			
		AATE00 0898	37,550,627- 37,550,119	экзон 2	GAAATCCT CTCCACCC AGAGTAAA A	TGATGTTGC TGGCTAGGT TGATATT			
	<i>An. coluzzii</i>	X	7.5 M6	7,370,000- 7,610,000	AGAP00 0412	7,386,922- 7,387,362	экзон 3	CTGCCAGA TACTGATG AAGTGC	TACGTGTCG TGTATGAGC GACT
					AGAP00 0413	7,396,883- 7,396,342	экзон 5	GATCTGAA CCTCACCA CCTACC	GTGCTAACC ATCACGATG TCCT
AGAP00 0415					7,405,016- 7,404,637	экзон 1	CACCAAAG CCGTCAAC ATCTC	GATGAACGG GTACAGGCT TTC	
14,950,000- 15,400,000				AGAP00 0815	14,937,481- 14,938,014	экзон 1	TGTCCAAC TATCAGCT CACGAC	GATCGTCTT ATCGTCCTC CATC	
					14,946,605- 14,947,129	экзон 5	ACAACGTG ATCGACGT GAAGTA	AGAAGTTGT CGCACTCGC AGTA	
					14,947,389- 14,947,796		GTGCAACG AGTTCAAG CACTG	GTGTCCCAC TTTGCCATC AT	

ПРИЛОЖЕНИЕ 7

Таблица результатов статистической обработки FISH-экспериментов.

Вид	Координаты петли	Клетки фолликулярного эпителия		Питающие клетки с низкой степенью политении		Питающие клетки с высокой степенью политении	
		Пересечение сигнала	-	Пересечение сигнала	-	Пересечение сигнала	-
<i>An. stephensi</i>	2R: 37,345,000-37,525,000; 42,900,000-43,125,000	6	68	96	61	11	58
		8.1%	91.9%	61.1%	38.9%	16.4%	83.6%
<i>An. atroparvus</i>	3R: 7,620,000-8,160,000; 38,550,000-38,950,000	13	96	3	7	-	6
		11.9%	88.1%	30%	70%	0	100%
	2R: 11,210,000-11,700,000; 23,055,000-23,700,000	8	47	-	-	-	22
		14.5%	85.5%	-	-	0	100%
	X: 5,320,000-5,755,000; 11,100,000-11,510,000	58	2	-	-	1	14
		96.7%	3.3%	-	-	6.7%	93.3%
<i>An. coluzzii</i>	X: 7,370,000-7,610,000; 14,950,000-15,400,000	1	5	-	9	1	3