

на правах рукописи

**ШАРАХОВ**  
**Игорь Валентинович**

**МОЛЕКУЛЯРНАЯ ОРГАНИЗАЦИЯ И ЭВОЛЮЦИЯ ХРОМОСОМ У  
МАЛЯРИЙНЫХ КОМАРОВ РОДА *Anopheles* (Diptera, Culicidae)**

03.01.07 – молекулярная генетика

**АВТОРЕФЕРАТ**  
диссертации на соискание ученой степени  
доктора биологических наук

Новосибирск – 2017

Работа выполнена в лаборатории экологии, генетики и охраны окружающей среды ФГАОУ ВО «Национальный исследовательский Томский государственный университет».

Научный консультант: **Стегний Владимир Николаевич**, доктор биологических наук, профессор, заведующий лабораторией экологии, генетики и охраны окружающей среды, ФГАОУ ВО «Национальный исследовательский Томский государственный университет», г. Томск.

Официальные оппоненты: **Бугров Александр Геннадьевич**, доктор биологических наук, профессор, ведущий научный сотрудник лаборатории филогении и фауногенеза, ФГБУН Институт систематики и экологии животных СО РАН г. Новосибирск.

**Гундерина Лариса Ивановна**, доктор биологических наук, ведущий научный сотрудник сектора эволюционной геномики хирономид, ФГБНУ «ФИЦ Институт цитологии и генетики СО РАН» г. Новосибирск.

**Калмыкова Алла Ивановна**, доктор биологических наук, заведующая лабораторией исследования геномных повторов эукариот, ФГБУН Институт молекулярной генетики РАН, г. Москва.

Ведущее учреждение: ФГБУН Институт общей генетики им. Н. И. Вавилова РАН, г. Москва

Защита состоится: «13» декабря 2017 г. в 10 часов на заседании диссертационного совета Д 003.074.01 по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук при ФГБУН Институт молекулярной и клеточной биологии СО РАН (630090, Новосибирск, пр. академика Лаврентьева, 8/2, тел. (383)-373-02-49, e-mail: [ovant@mcb.nsc.ru](mailto:ovant@mcb.nsc.ru)

С диссертацией можно ознакомиться в библиотеке и на сайте Института молекулярной и клеточной биологии СО РАН [www.mcb.nsc.ru](http://www.mcb.nsc.ru)  
Автореферат разослан «\_\_\_» \_\_\_\_\_ 2017 г.

Ученый секретарь  
диссертационного совета,  
кандидат биологических наук

Антоненко О.В.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

### Актуальность исследования

Инфекционные и паразитарные заболевания, переносимые комарами, являются одной из самых серьезных проблем мирового здравоохранения. Только заболеваемость малярией, согласно данным ВОЗ, составляет 300-500 млн. человек в год. Из них – до 1 млн. со смертельным исходом. Устойчивость к инсектицидам и лекарственным антималярийным препаратам продолжает возникать и активно распространяться, и эта ситуация усугубляется медленным прогрессом при внедрении альтернативных подходов. Именно поэтому основные надежды на успешный и долговременный контроль малярии и подобных заболеваний связаны с геномными исследованиями видов-переносчиков.

До сих пор нет ясности в вопросе, почему одни виды или популяции комаров участвуют в передаче малярии, тогда как другие – не имеют медицинского значения. Известно, что способность переносить малярию определяется многими факторами, в том числе поведением, иммунитетом, и жизненным циклом комаров. Можно предположить, что различия в геномной пластичности у разных видов или групп комаров влияют на эти ключевые признаки, определяя способность быть носителем и передавать патоген человеку. Геномная эра предоставляет прекрасные возможности для понимания молекулярно-генетических основ адаптации и эволюции малярийных комаров и их взаимоотношений с хозяевами и патогенами. Полногеномные данные становятся важным инструментом исследований, а наличие полных хромосом у *Anopheles* благоприятствует созданию цитогенетических и физических геномных карт высокого разрешения. Такое картирование играет ключевую роль в работах по функциональной аннотации генетических данных и проведению комплексного сравнительно-геномного анализа на межвидовом и более высоких уровнях. Физические карты играют важную роль и в поиске генных локусов, отвечающих за резистентность комаров к воздействию инсектицидов, экологическую адаптацию и степень восприимчивости к патогену. Данный подход становится особо актуальным в свете того, что африканский комплекс малярийных комаров *An. gambiae* представлен группой морфологически неразличимых видов-сиблингов (Coetzee et al., 2013), которые, тем не менее, характеризуются индивидуальными особенностями адаптации, поведения и восприимчивости к переносимому паразиту *Plasmodium falciparum* (White, 1974).

В настоящее время является общепризнанным тот факт, что функциональная характеристика любого генома должна включать эпигенетическую составляющую. Эпигенетические события и модификации хроматина активно исследуются как у человека ([www.epigenome.org](http://www.epigenome.org)), так и у малярийного плазмодия (Ponts et al., 2010; Salcedo-Amaya et al., 2009). В триаде хозяин-переносчик-патоген только для *Anopheles* пока отсутствуют систематические эпигеномные исследования. В то же время, взаимодействия между патогеном и переносчиком всегда довольно пластичны и динамичны, что предполагает эпигенетическую регуляцию (Gómez-Díaz et al., 2012). С этих позиций представляется особо интересным исследование молекулярной организации гетерохроматина малярийных комаров. В первую очередь, гетерохроматиновых генов и генов, расположенных на эугетерохроматиновой границе, которые могут быть подвержены эпигенетической регуляции со стороны гетерохроматина. Известно, что гетерохроматиновая Y-хромосома контролирует ряд базовых функций, включая иммунный ответ, определение пола и фертильность самцов (Lemos et al., 2010). Однако, из-за обогатенности высоко-повторенными последовательностями, до сих пор не удалось исследовать эту «терра инкогнита» у малярийных комаров, в том числе на наличие генов. Выявление таких генов и определение молекулярной структуры Y-хромосомы было бы крайне полезно для реализации пол-специфичного контроля переносчиков малярии. Исследования на дрозофиле показали, что с гетерохроматином также связаны процессы регуляции генома, в которые вовлечены малые некодирующие РНК класса пиРНК (Piwi-interacting RNA), играющие важную роль в защите целостности генома от повреждающего действия транспозонов в клетках зародышевой линии (Brennecke et al., 2007) и в процессе распада материнской мРНК у эмбрионов (Rouget et al., 2010). Генетическая регуляция фертильности и эмбрионального развития могут быть использованы для снижения репродуктивного потенциала природных популяциях комаров, а пиРНК могут стать основой этой генно-инженерной стратегии. В связи с этим важно изучить пиРНК и производящие ее геномные локусы у малярийных комаров.

Анализ эволюции хромосом является перспективным подходом для выявления функционально-важных участков генома. У малярийных комаров хромосомные инверсии часто неслучайно расположены по

хромосомным плечам и связаны с эпидемиологически важными адаптациями и, возможно, видообразованием (Ayala et al., 2017). Исследование специфичных генных комбинаций, сохраняющихся в процессе эволюции в пределах полиморфных инверсий, способствует поиску генов, отвечающих за фенотипическую пластичность малярийных комаров. Сравнительные геномные исследования малярийных комаров позволяют, изучив хромосомную эволюцию разных видов, с высокой степенью достоверности установить как филогенетические связи между ними, так и направление эволюции. Также появляется возможность проводить исследования молекулярной организации ДНК в точках разрывов хромосомных перестроек, что также необходимо для понимания популяционно-генетической структуры видов малярийных комаров и филогенетических взаимоотношений между ними. В свою очередь, это необходимо, чтобы вплотную подойти к пониманию механизмов, отвечающих за различия между видами, касающиеся устойчивости к инфицированию малярийным плазмодием и выбора объекта для питания. В целом же, развитие представлений о молекулярной структуре, функциях и эволюции хромосом малярийных комаров дает исследователям новое понимание механизмов работы генома и широкий спектр возможностей для разработки новых подходов к контролю опасных для человека заболеваний.

### **Цель и задачи исследования**

**Целью работы** было комплексное изучение молекулярной организации и эволюции хромосом эпидемиологически значимых видов малярийных комаров из подродов *Cellia*, *Anopheles* и *Nyssorhynchus*, а также выяснение филогенетических взаимоотношений между видами комплекса *An. gambiae*.

### **В связи с этим были поставлены следующие задачи:**

1. Используя возможности политенных хромосом провести цитогенетическое и физическое картирование генов *An. gambiae*, расположенных на эу-гетерохроматиновой границе, и изучить молекулярную структуру гетерохроматина у данного вида.

2. Исследовать локализацию пиРНК кластеров в геноме *An. gambiae*, сравнить с организацией пиРНК кластеров у *Drosophila melanogaster* и *Aedes aegypti*.

3. Исследовать геномную структуру и эволюцию Y-хромосомы в комплексе видов *An. gambiae*.

4. Разработать цитогенетические карты и провести физическое картирование генов у *An. gambiae*, *An. funestus*, *An. stephensi* и *An. nili*.

5. Изучить характер, скорость, молекулярные механизмы и функциональную значимость хромосомных перестроек у малярийных комаров.

6. На основе аутомсомных и X-хромосомных инверсий реконструировать филогенетические взаимоотношения между видами-сиблингами комплекса *An. gambiae*.

### **Научная новизна работы**

Следует считать новыми следующие результаты, полученные в данной работе: впервые проведено разностороннее изучение молекулярной организации и эволюции хромосом у малярийных комаров рода *Anopheles*. Созданы новые исследовательские инструменты – цитогенетические и физические карты политенных хромосом высокого разрешения для эпидемиологически значимых видов малярийных комаров подрода *Cellia*. Изучена молекулярная организация гетерохроматина у африканского малярийного комара *An. gambiae* и его хромосомная локализация. Впервые у малярийных комаров исследованы локализация и молекулярная структура пиРНК кластеров, а также геномная организация и эволюция Y-хромосомы. Впервые выявлены закономерности хромосомной эволюции у малярийных комаров и продемонстрированы существенные различия в эволюционной динамике между X-хромосомой и аутомсомами. Предложен новый надежный подход для реконструкции филогенетических взаимоотношений внутри комплексов близкородственных видов, основанный на выяснении порядка генов в точках инверсионных разрывов относительно «внешней» группы видов.

### **Практическая значимость работы**

Полученная информация о молекулярных особенностях организации хромосом малярийного комара, в первую очередь, данные о гетерогенности и лабильности гетерохроматиновой части генома, высокой эволюционной динамике Y-хромосомы, организации кластеров пиРНК и связанного с ними каскада генов, позволит планировать направленную работу по сокращению репродуктивного потенциала самок комаров и создания стерильных самцов для успешного контроля популяций переносчиков. Знание о характере и скорости хромосомной эволюции у *Anopheles* будет способствовать работе по проектированию и внедрению модифицированных комаров в природные популяции для успешного

контроля переносимых заболеваний. Выяснение родственных отношений и направления эволюции в комплексе видов *An. gambiae* позволит выявить конкретные генетические изменения, отвечающие за выбор комаром хозяина и адаптацию к изменениям в окружающей среде. Обнаруженный параллелизм по наборам генов в полиморфных адаптивных инверсиях хромосомы 2 (2R) у разных видов *Anopheles* дает начало перспективному направлению по поиску генных кластеров, играющих важную роль в эволюционной адаптации малярийных комаров.

### **Положения, выносимые на защиту**

1. Геном *An. gambiae* имеет крупные блоки интеркалярного гетерохроматина, обладающие гетерогенной структурой. Гетерохроматин обогащен генами, участвующими в регуляции биологических процессов и организации хромосом.

2. У *An. gambiae*, в сравнении с *D. melanogaster*, произошло распространение пиРНК кластеров из гетерохроматина в эухроматин, где значительная часть пиРНК продуцируется генами, важными для развития и размножения малярийных комаров.

3. Молекулярная организация Y-хромосомы кардинально меняется даже у близких видов комаров. X- и Y-хромосомы у *Anopheles* обмениваются генетическим материалом.

4. Хромосомная эволюция малярийных комаров связана с перетасовкой порядка генов путем парацентрических инверсий и полноплечевых транслокаций. При этом половая X-хромосома значительно чаще меняет свою структуру, по сравнению с аутосомами, несмотря на малочисленность полиморфных инверсий. Механизм формирования инверсии 2La включал инсерции повторяющейся ДНК и формирование псевдогенов в точках разрыва.

5. Аутосомные плечи различаются по своей толерантности к генерации точек инверсионных разрывов. Адаптивные полиморфные инверсии в 2R плече у отдалённых видов *Anopheles* имеют сходный генный состав.

6. Порядок генов в точках инверсионных разрывов с учетом внешней группы является новым надежным признаком, позволяющим реконструировать филогенетические взаимоотношения внутри комплекса близкородственных видов. *An. merus* и *An. gambiae* являются видами, наиболее близкими к предковому виду в комплексе *An. gambiae*.

## **Апробация работы**

Основные результаты были доложены более чем на 20 международных конференциях, симпозиумах и конгрессах в России, США, Греции, Франции, Италии, Мали, Южной Кореи, Великобритании и Китае, таких как International Conference on *Drosophila* Heterochromatin (Губбио, Италия, 2005, 2011), Annual Meeting of American Society of Tropical Medicine and Hygiene (Филадельфия, США, 2007, 2011), Joint ISCB Africa ASBCB Conference on Bioinformatics of Infectious Diseases (Бамако, Мали, 2009), EMBO meeting «Molecular and Population Biology of Mosquitoes and Other Disease Vectors» (Колимбари, Греция, 2005, 2009, 2011, 2013, 2015), EMBO Conference «Nuclear structure and dynamics» (Л'Иль-сюр-ла-Сорг, Франция, 2011), The American Genetic Association Annual Symposium (Дурам, США, 2012), The International conference «Genomics Epidemiology of Malaria» (Хинкстон, Великобритания, 2012, 2014), CSHL Asia symposium on Evolutionary Genetics and Genomics (Сужоу, Китай, 2014), The 9th Annual Arthropod Genomics Symposium (Манхеттен, Канзас, США, 2015), Международная конференция «Хромосома» (Новосибирск, Россия, 2009, 2012, 2015), Санкт-Петербургский симпозиум по биологии систем и биоинформатики (Санкт-Петербург, Россия, 2016).

## **Вклад автора**

Все основные результаты, вошедшие в диссертацию, получены автором лично или под его руководством, как в индивидуальных, так и коллективных исследованиях. Все выводы и теоретические обобщения сделаны автором самостоятельно. Часть результатов была получена в ходе сотрудничества с другими лабораториями. Ф. Джордж и М.В. Шарахова внесли вклад в картирование генома и характеризацию гетерохроматина *An. gambiae* под руководством автора. И.В. Брусенцова провела иммуноокрашивание хромосом *An. gambiae*. Анализ обогащения терминов генных онтологий был проведен К.Д. Смитом. С.К. Леман разработал и использовал новые Байесовские статистические модели. Дж.А. Бэйли проанализировал сегментные дубликации. Х.М. Тубио определил распределение LTR ретротранспозонов. С. Дженсен под руководством Ш. Вари провела идентификацию пиРНК и картирование пиРНК в генах. Б. Холл, Дж. Ту, Ф. Папатанос обнаружили кандидатные гены в Y-хромосоме и провели анализ данных PacBio у *An. gambiae*. В. Тимошевский и А. Шарма провели гибридизацию *in situ* на Y-хромосоме у

видов комплекса *An. gambiae* под руководством автора. А. Ща, М.В. Шарахова, А. Пери под руководством автора, а так же М. Унгер, Э.М. Баричева и Д. Карагодин принимали участие в молекулярном картировании хромосом *An. stephensi*. Антонио Нкондио, Фред Симард и Кирил Ндо провели сбор комаров *An. nili* и *An. melas* в Африке. М.В. Шарахова и А. Ща под руководством автора построили цитогенетическую и физическую карту *An. nili*. Х. Жианг и Дж. Ту провели сборку генома *An. stephensi*. Г. Ян и Н. Безански руководили проектом по картированию *An. funestus*. О.Г. Грушко и О. Брагинец участвовали в картировании кДНК и микросателлитов у *An. funestus* под руководством автора. М. Фонтаин, Н. Безански и М. Хан провели молекулярный анализ эволюции Y-хромосомы и геномов в комплексе *An. gambiae*. Д.Е. Нифси, Р.М. Вотерхаус, Н. Безански внесли главный вклад в организацию секвенирования и анализ геномов 16 видов *Anopheles*. Г.Н. Артемов, А.Н. Науменко, М.В. Шарахова под руководством автора, а так же В.Н. Стегний прокартировали геном *An. atroparvus*. Х. Жианг провела биоинформатический анализ и визуализацию хромосомных перестроек при помощи R программы genoPlotR под руководством автора. М. Камали участвовала в клонировании и картировании точек разрывов инверсий у *An. gambiae* и *An. merus*. Автор лично осуществлял планирование, координировал эксперименты и проводил анализ полученных результатов в рамках данной работы.

### **Структура и объем работы**

Диссертация состоит из шести глав: «Обзор литературы», «Материалы и методы», «Молекулярная структура и эволюция гетерохроматина у африканского комплекса малярийных комаров *An. gambiae*», «Картирование генома основных переносчиков малярии Африки и Азии: *An. gambiae*, *An. funestus*, *An. nili* и *An. stephensi*», «Хромосомная эволюция у малярийных комаров из подродов *Cellia*, *Anopheles* и *Nyssorhynchus*» и «Инверсионная филогения в комплексе видов *An. gambiae*». Главы с 3 по 6 включают результаты и обсуждение собственных исследований по четырем основным направлениям изучения малярийных комаров: организация гетерохроматина, физическое и цитогенетическое картирование, хромосомная эволюция и филогения. Объем диссертации составляет 344 машинописные страницы, включая 78 рисунков, 49 таблиц, в том числе, в приложении. Список литературы содержит 383 ссылки.

## Публикации по теме диссертации

Общее число работ по теме диссертации, включая сборники трудов конференций, составляет 59, из них 28 статей в рецензируемых международных и отечественных журналах (26 статей из списка, рекомендованного Перечнем ВАК РФ) и 1 глава в книге.

## СОДЕРЖАНИЕ РАБОТЫ

**Глава 1 «Особенности организации и эволюции генома малярийных комаров»** является обзором литературы по теме работы и содержит информацию о видах малярийных комаров, работа с которыми проводилась в рамках исследования. А также о цитологическом и физическом картировании разных видов. Представлен подраздел, включающий работы, посвященные хромосомной пластичности и ее роли в адаптации малярийных комаров. Показаны проблемы и трудности установления филогенетических связей между близкородственными видами. Описаны исследования гетерохроматина и некоторые механизмы эпигенетической регуляции генома.

### **Глава 2 «Материалы и методы исследования».**

В работе использовали разнообразные виды, колонии и лабораторные линии малярийных комаров. Комары популяций африканского вида *An. funestus* были собраны лично автором в ходе экспедиционных работ. Всего было собрано в природе или взято из лабораторной культуры и обработано: 380 особей *An. albimanus*, 60 особей *An. arabiensis*, 560 особей *An. atroparvus*, 950 особей *An. gambiae*, 1200 особей *An. funestus*, 50 особей *An. melas*, 200 особей *An. merus*, 750 особей *An. nili*, 80 особей *An. quadriannuatus*, 800 особей *An. stephensi*.

Препараты политенных и митотических хромосом готовили по стандартным протоколам. Также был применен и усовершенствован метод получения уплощенных политенных хромосом, обеспечивающий быстрое и качественное физическое картирование геномов малярийных комаров. Соответствующая методическая статья входит в список работ автора.

Флюоресцентную гибридизацию *in situ* и иммуноокрашивание политенных хромосом проводили по стандартным протоколам с некоторыми модификациями. Модификации, в первую очередь, были связаны с избытком жира на препаратах хромосом, получаемых из питающих клеток яичников малярийных комаров.

Микродиссекцию Y-хромосомы, создание библиотек кДНК, выделение РНК и создание библиотек малых РНК, скрининг библиотек,

амплификацию ДНК, секвенирование ДНК по Сэнгеру, полногеномное секвенирование проводили согласно протоколам производителей использованных реакционных наборов и соответствующего оборудования. Сопутствующий биоинформатический анализ проводили с помощью различных программ и программных пакетов: BLASTN, CLUSTALX 1.81, SEQMAN, PRIMER3, TEALIGN, визуализировали в genoPlotR.

Информация об ортологах была получена из базы данных OrthoDB (<http://people.csail.mit.edu/waterhouse/AGCC/Orthology/>). ID генов *An. gambiae* были получены на основании аннотации из базы данных VectorBase (<http://www.vectorbase.org/>).

Плотность генов и степень охвата транспозонами анализировали с помощью программ Biomart (Haider et al., 2009) и RepeatMasker (Smit, Hubley, Green, 2004), соответственно. Микро и минисателлиты анализировали с помощью Tandem Repeats Finder (Benson, 1999).

При идентификации транспозонов, ассоциированных с пиРНК, использовали программу BowTie2, базу данных экзонов BioMART (Guberman et al., 2011). Онтологию генов определяли с помощью DAVID.

Оценку количества хромосомных инверсий между видами проводили с помощью программы Genome Rearrangements in Mouse and Man (GRIMM) (Tesler, 2002). Для подсчета инверсионных расстояний использовали программы Multiple Genome Rearrangements (MGR) (Bourque, Pevzner, 2002), GRIMM и Sorting Permutation by Reversals and block-INterchanGes (SPRING) (Lin et al., 2006).

Для полногеномного сравнения скорости перестроек у четырех видов малярийных комаров и группы из 8 видов дрозофилы было использовано два независимых групповых t-теста (STATISTICA 10.0).

Все отсеквенированные в процессе работы последовательности ДНК были помещены в National Center for Biotechnology Information short read archive ([www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi](http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi)) и в базу данных GenBank database (<http://www.ncbi.nlm.nih.gov/Genbank/>).

**Глава 3 «Молекулярная структура и эволюция гетерохроматина у африканского комплекса малярийных комаров *An. gambiae*»** посвящена работам по геномному картированию и изучению молекулярной структуры гетерохроматина (ГХ) у «модельного» вида малярийного комара.

## ***Геномное картирование и молекулярная структура гетерохроматина у *An. gambiae****

Набор политенных хромосом питающих клеток яичника у самок малярийных комаров *An. gambiae* включает пять хромосомных плеч: четыре аутомсомных плеча 2R, 2L, 3R, 3L и X хромосому. В ходе работы были описаны разные типы ГХ, в том числе, с характерной морфологией, не выявленной ранее в ГХ *D. melanogaster*: прицентромерный (ПЦ), диффузный интеркалярный (ИГХд) и компактный интеркалярный (ИГХк).

Локализация ГХ на цитологической карте позволила приступить к поиску гетерохроматин-эухроматиновых (ГХ-ЭХ) границ. Приблизительные координаты были определены на основании геномной позиции ВАС и кДНК клонов из интересующих нас районов (Holt et al., 2002). Поскольку ГХ районы имели недостаточное покрытие маркерами, с помощью ПЦР были амплифицированы дополнительные генные фрагменты и использованы в качестве ДНК проб для физического картирования. С помощью флуоресцентной гибридизации *in situ* (FISH) были выявлены ПЦР продукты, локализованные наиболее близко к предполагаемой ГХ-ЭХ границе каждого ГХ района каждого хромосомного плеча. Это позволило установить положение границы, основываясь на самых дальних от центра ГХ и ЭХ маркерах. Основываясь на этих границах, примерно 16,6 Мб были отнесены к ГХ в текущей сборке генома *An. gambiae*. По итогам работы картированная часть ГХ составила 6,4% от ~260 Мб генома (Holt et al., 2002).

Был проведен анализ молекулярных различий между разными типами хроматина. С помощью Байесовской статистической модели удалось показать, что ГХ и ЭХ различаются по плотности распространения генов, покрытию транспозонами (ТЭ) и сегментными дупликациями (рис. 1). ПЦГХ имел самую высокую плотность покрытия ретроэлементами и тандемными повторами, тогда как ИГХ был обогащен сегментными дупликациями. ИГХд был обогащен транспозонами, минисателлитами и сателлитами в сравнении с ИГХк. ИГХк был обогащен соло-LTR-транспозонами.

Чтобы охарактеризовать гены из ГХ *An. gambiae*, были использованы «генные термины» биоинформатического проекта генной онтологии (Gene Ontology или GO) (Ashburner et al., 2000). Оказалось, что ГХ аккумулирует гены, важные для процессов регуляции генной экспрессии и организации хроматина.

## Организация локусов пиРНК в геноме малярийного комара

Для анализа сходства и различия в пиРНК-опосредованных регуляторных путях у дрозофилы и малярийного комара была выделена пиРНК из тканей яичников самок *An. gambiae* и исследована локализация пиРНК кластеров на хромосомах.

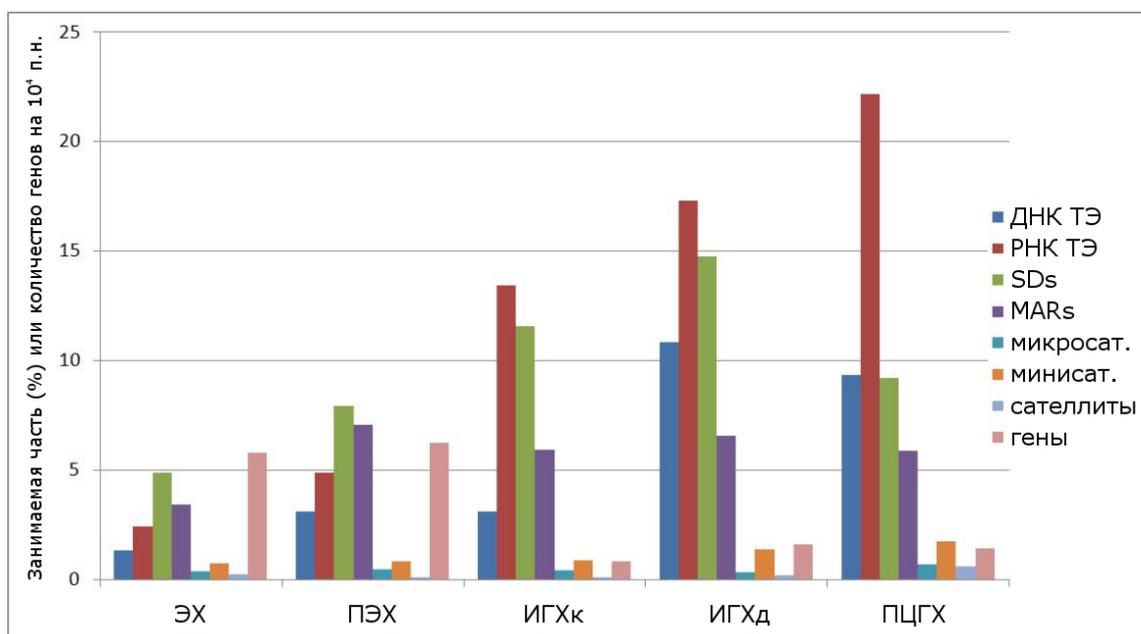


Рис. 1 Усредненное распределение плотности генов, транспозонов и повторов в различных типах хроматина *An. gambiae*. ТЭ – транспозоны, SD - короткие дупликации, MARs (matrix associated regions) – районы, связанные с ядерным матриксом.

В первую очередь мы создали базу данных, секвенировав малые РНК *An. gambiae* с помощью Illumina Small RNA TruSeq technology. Были использованы аналогичные данные, полученные ранее для *D. melanogaster* (Brennecke et al., 2008) и комара *Ae. aegypti* (Arensburger et al., 2011). Описание и аннотации малых РНК трех видов представлены на рис. 2.

Последовательности 24-29 п.н. длиной из созданной нами базы данных были картированы на 7080 аннотированных транспозонов надкласса *Hehrada* из RepBase (Bao, Kojima, Kohany, 2015). Чтобы идентифицировать пиРНК, связанные с транспозонами, использовали два подхода: "консенсусное" картирование и "перекрывающееся" картирование. Оказалось, что у дрозофилы 81,6% пиРНК происходит из транспозонов, тогда как у малярийного комара только 39,4%.

При дальнейшем анализе в геноме *An. gambiae* было выявлено 187 пиРНК кластеров и 155 пиРНК кластеров - в геноме *D. melanogaster*.

Минимум 24% картированной пиРНК соответствовало последовательностям транспозонов в геноме комара, и 65% - в геноме дрозофилы. У обоих видов пиРНК, в первую очередь, связывается с ретротранспозонами. В отличие от дрозофилы, у *An. gambiae* основная часть кластеров пиРНК подвергается однонаправленной транскрипции.

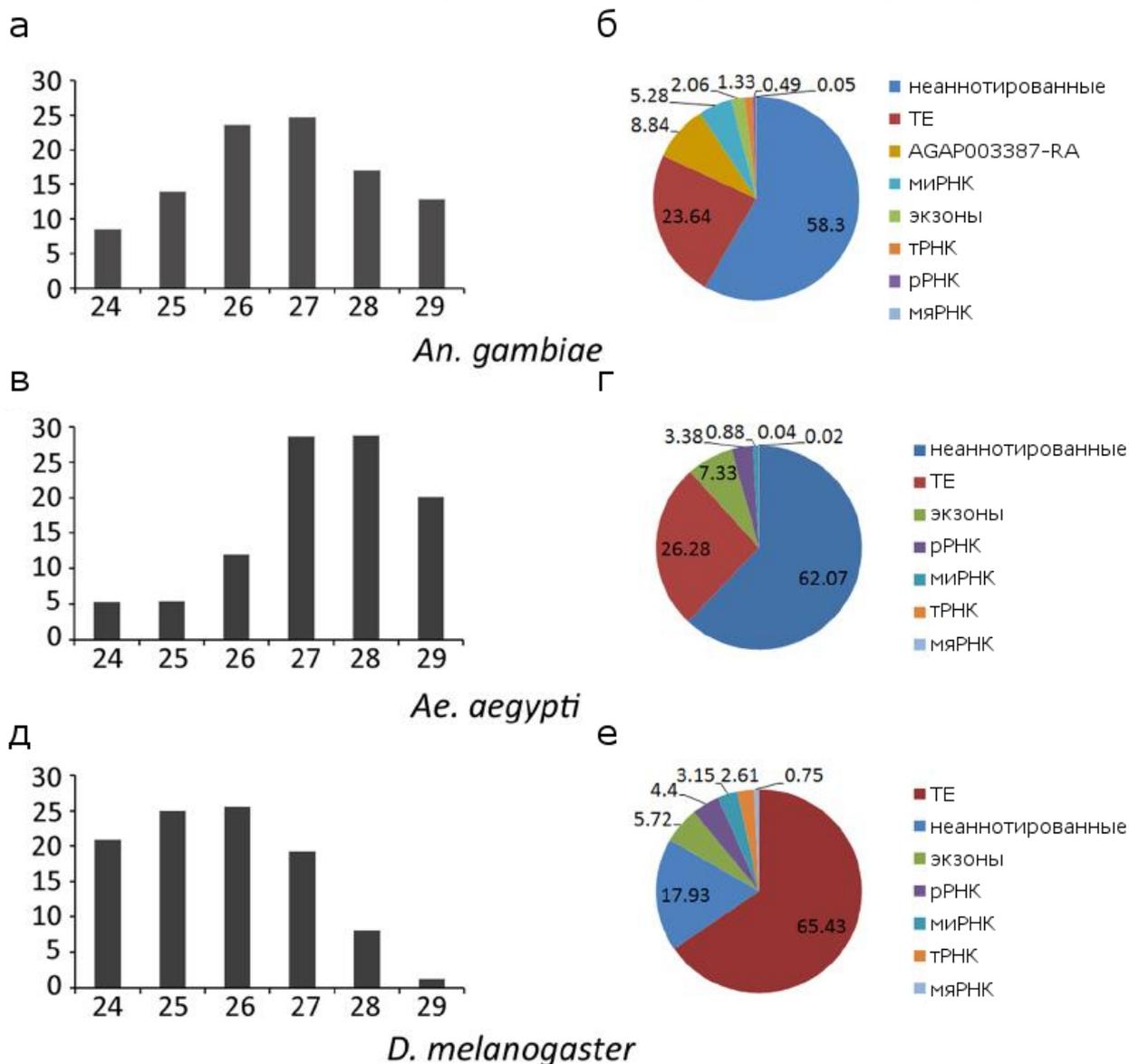


Рис. 2. Распределение по размерам и аннотация малых РНК, имеющих размеры, сходные с пиРНК (24-29п.н) в нашей базе данных *An. gambiae*, *Ae. aegypti* и *D. melanogaster*. А, В, Д показывают распределение длин малых РНК в отсеквенированной библиотеке для трех видов (по оси X – длина в п.н; по оси Y – процент от всех малых РНК указанной длины). Б, Г, Е – аннотации этих малых РНК для всех трех видов.

Анализ геномной локализации кластеров пиРНК показал, что у малярийного комара только около 30% уникальных пиРНК продуцируются кластерами, расположенными в ПЦГХ. 29.3% –

кластерами, расположенными в ЭХ, тогда как у дрозофилы только 6.6%. Другое интересное отличие заключается в том, что у дрозофилы вне кластеров продуцируется 22% пиРНК, тогда как у комара – 43.2%. В целом, имеет место явный сдвиг транскрипции пиРНК из ПЦГХ (у дрозофилы) в ЭХ (у комара). Значительная часть популяции пиРНК комара также связана с ИГХд. У *Ae. aegypti* основная часть пиРНК кластеров сосредоточена по ЭХ (рис. 3).

Около 11% популяции пиРНК у *An. gambiae* связано с геной транскрипцией, что существенно больше, чем у дрозофилы (6%). Анализ пиРНК обогащенных генов (GO) показал, что большая их часть связана с размножением и развитием малярийного комара.

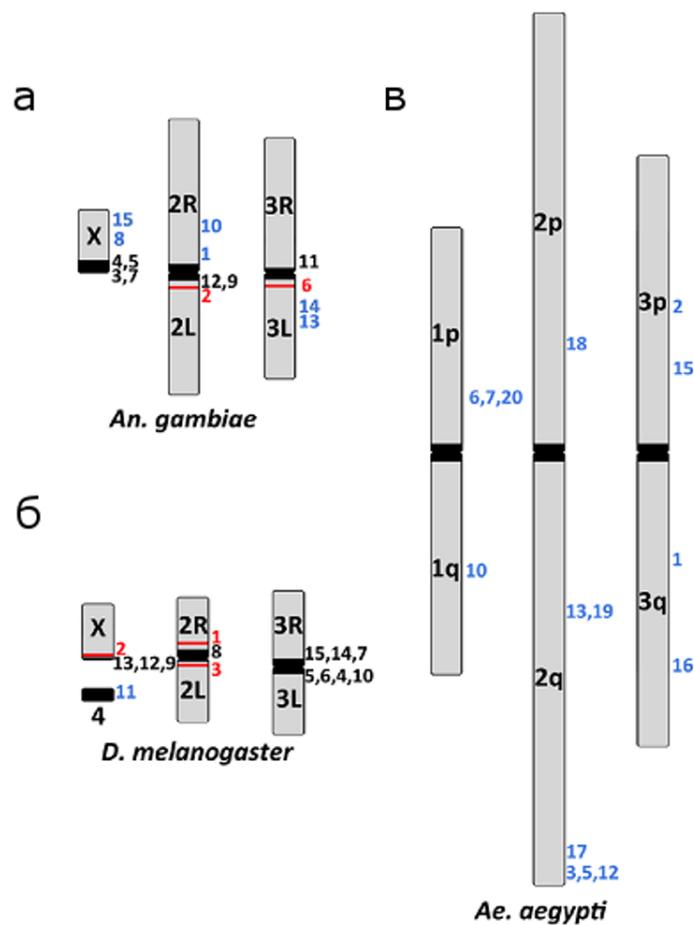


Рис. 3 Хромосомное распределение 15 топовых пиРНК кластеров у трех представителей двукрылых: а) *An. gambiae*, б) *D. melanogaster*, в) *Ae. aegypti*. Серым цветом окрашены районы эухроматина, черным – прицентромерного гетерохроматина, красным – интеркалярного гетерохроматина. Номера кластеров присвоены согласно убыванию числа уникальных пиРНК.

### Геномная структура и эволюция Y-хромосомы в комплексе *An. gambiae*

Хромосома Y у малярийных комаров является полностью

гетерохроматиновой, но, в отличие от дрозофилы, показывает частичную гомологию с X хромосомой и играет роль в определении пола у самцов. Чтобы исследовать организацию нерекомбинированной Y хромосомы в комплексе видов *An. gambiae* использовали современные возможности метода одномолекулярного геномного секвенирования. В итоге, 246 Мб генома были отнесены к последовательностям Y-хромосомы и вошли в базу данных, обозначенную как Ydb. Основные результаты, полученные автором в рамках данного исследования, были следующие: 1. Y-хромосома у *An. gambiae* практически полностью состоит из небольшого числа массивов амплифицированных тандемных повторов и ретротранспозонов. Сателлитная ДНК составляет примерно 49% всех отсеквенированных последовательностей, отнесенных к Y-хромосоме, однако представлена только 6 различными сателлитными мономерами, из которых на два - *AgY477* и *AgY373* - приходится до 93% всех сателлитных последовательностей. Другая часть – 43.5% - представлена транспозонами 8 типов, преимущественно ретротранспозонами. 2. Между близкими видами комплекса *An. gambiae* (время расхождения которых оценивается примерно в 2 млн. лет) количество и типы повторенной ДНК на Y подвергались быстрым эволюционным изменениям. Так *AgY477* и *AgY373* - наиболее многочисленные сателлиты на нерекомбинированной Y-хромосоме *An. gambiae* - отсутствуют или не показывают половой специфичности у видов-сблингов. Физическое картирование подтверждает высокую динамику последовательностей Y-хромосомы в пределах комплекса (рис. 4). 3. Некоторые из этих повторов могут вступать в рекомбинацию со сходными повторами на X-хромосоме.

### **Заключение к главе 3**

Морфологическое описание и детальное физическое картирование расширило наши представления о ГХ части генома *An. gambiae*. 16.6 Мб картированного ГХ стали доступны для дальнейших описаний и исследований. ГХ оказался обогащен генами, кодирующими белки, которые могут быть вовлечены в эпигенетическую регуляцию хроматина. Идентификация большого пула пиРНК, продуцируемых генами, связанными с развитием и размножением комара, показывает, что пиРНК-регуляторный путь может играть важную роль в соответствующих процессах. Это представляет интерес в плане возможного контроля репродукции комаров. Благодаря новой методологии впервые удалось исследовать организацию «terra incognita» генома –

гетерохроматиновой Y-хромосомы. Оказалось, что она сформирована массово амплифицированными тандемными повторами всего нескольких видов сателлитов и транспозонов. Эти повторы эволюционируют очень быстро. Даже между близкими видами комаров выявлены принципиальные различия. Данное исследование станет основой дальнейшего изучения биологии и эволюции Y-хромосомы у разных видов.

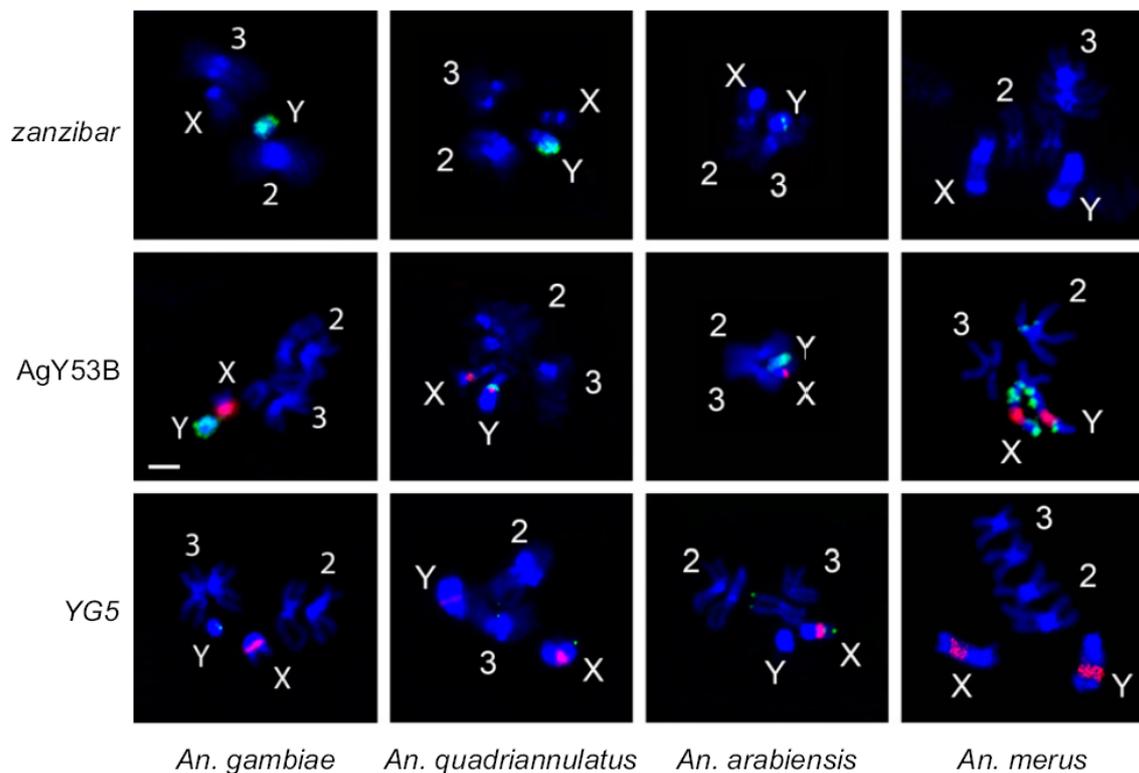


Рис. 4 Физическое картирование подтверждает структурную динамику Y-хромосомы в комплексе видов *An. gambiae*. Представлены результаты FISH ретротранспозона *zanzibar*, сателлита *AgY53B*, и Y-сцепленного гена *YG5* (зеленый сигнал). Гибридизация была сделана на хромосомах имагинальных дисков, за исключением хромосом *An. merus* (*AgY53B*), полученных из семенников. Пробы 18S рДНК (красный сигнал) были использованы на всех слайдах, за исключением гибридизации с *zanzibar*. Хромосомы окрашены DAPI (синий). Масштабная линейка для всех изображений 2 мкм

**Глава 4 «Картирование генома основных переносчиков малярии Африки и Азии»** включает результаты многолетней работы по цитогенетическому и физическому картированию ряда видов малярийных комаров подрода *Cellia*.

#### **Организация хромосом у малярийного комара *An. funestus***

В ходе работы была составлена цитогенетическая карта политенных

хромосом питающих клеток яичников этого вида и локализованы на ней точки инверсионных разрывов. Также была получена физическая карта микросателлитных ДНК-маркеров, пригодная для популяционных исследований.

С помощью FISH было проведено физическое картирование 157 кДНК *An. funestus* на политенных хромосомах этого вида. Из них 116 имели уникальную локализацию в геноме. Эти же последовательности были *in silico* картированы в геноме *An. gambiae*. Анализ результатов показал, что имела место реципрокная полноплечевая транслокация между 2L и 3R.

### ***Цитогенетическая карта высокого разрешения An. gambiae***

Ранее на дрозофиле был апробирован метод приготовления препаратов в условиях повышенного давления на хромосомы (Novikov et al., 2007). В работе этот метод был модифицирован и использован, чтобы создать новую цитогенетическую карту высокого разрешения для *An. gambiae*. Чтобы привязать хромосомные районы к геномной последовательности были определены геномные координаты для 302 маркеров, включающих ВАС, кДНК и ПЦР-амплифицированные генные фрагменты. Т.о., была получена физическая карта с разрешением, в среднем, 0.76 Мб (рис. 5).

### ***Карты хромосом азиатского переносчика малярии An. stephensi***

Была построена стандартная цитогенетическая фотокарта *An. stephensi*. Используя полученную карту, в ходе многолетней работы на хромосомах этого вида были картированы различные маркерные последовательности ДНК, включая ВАС-клоны и кДНК. После того, как геном *An. stephensi* был полностью секвенирован (Jiang et al., 2014), было проведено физическое картирование 227 проб ДНК. Этого оказалось достаточно, чтобы привязать 86 скэффолдов к уникальным позициям на политенных хромосомах. 86 геномных скэффолда суммарно составили 137.14 Мб (62%) собранного генома.

Также для этого вида была построена интегрированная карта микросателлитных маркеров и точек инверсионных разрывов, пригодная для популяционных исследований.

### ***Цитогенетическая карта малярийного комара An. nili***

Данный вид африканских малярийных комаров, относящийся к мажорным переносчикам, является малоизученным. Однако представляет особый интерес в связи с уникальной экологической адаптацией (Ayala et

al., 2009) и, несомненно, требует всестороннего изучения. В ходе работы была построена стандартная цитогенетическая карта политенных хромосом для данного вида и описаны маркерные районы для всех хромосомных плеч. ДНК-маркеры были прокартированы в плечах 2R и 3L у *An. nili*.

#### Заключение к главе 4

В главе представлены результаты хромосомного картирования четырех видов рода *Anopheles*, относящихся к основным переносчиками малярии в Африканском и Азиатском регионах. Цитогенетические карты лежат в основе исследований в области таксономии, систематики и популяционной генетики малярийных комаров. Локализация ДНК-маркерных последовательностей на политенных хромосомах по отношению к полиморфным хромосомным перестройкам необходима для правильной интерпретации популяционно-генетической структуры видов.

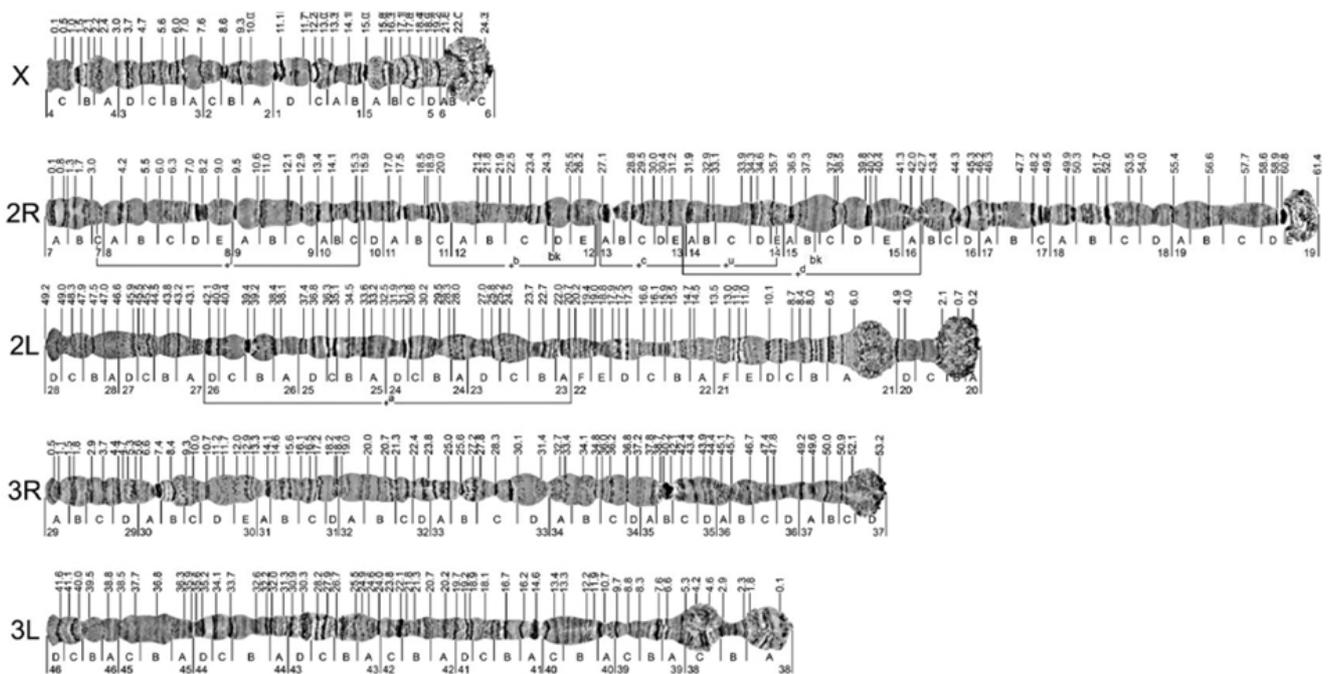


Рис. 5. Цитогенетическая фотокарта политенных хромосом питающих клеток яичника *An. gambiae*. Геномные координаты показаны над названиями хромосомных плеч. Районы и подрайоны показаны цифрами и буквами под хромосомами.

Физическое картирование позволяет работать над идентификацией локусов, отвечающих за устойчивость комаров к инсектицидам и восприимчивость к малярийному плазмодию. Различные виды картирования дополняют друг друга и, используемые совместно с

данными геномного секвенирования, позволяют комплексно подойти к изучению организации и эволюции генома малярийных комаров. Сравнительное геномное картирование легло в основу изучения характера и механизмов хромосомных перестроек у изучаемых видов (глава 5) и филогенетических связей между ними (глава 6).

### **Глава 5 «Хромосомная эволюция у малярийных комаров из подродов *Cellia*, *Anopheles* и *Nyssorhynchus*»**

Хромосомная изменчивость является базовым элементом эволюционного процесса. У Diptera отмечен очень высокий уровень хромосомного полиморфизма, в первую очередь определяемый парацентрическими инверсиями (Krimbas, Powell, 1992). Инверсии подавляют рекомбинацию и стабилизируют положение адаптивных аллелей или регуляторных комбинаций, играя основную роль в экологической дифференциации популяций и способствуя видообразованию. Нельзя отвергать и роль инверсий в системной реорганизации генома на основе «эффекта положения» в процессах адаптации и видообразования (Стегний, 1984). В случае малярийных комаров было показано, что быстрая адаптация к экологическим изменениям (в том числе связанным с человеческой деятельностью), коррелирует с высокой степенью инверсионного полиморфизма (Ayala et al., 2014).

#### ***Роль геномного ландшафта в эволюционной динамике хромосомных инверсий***

У видов комплекса *An. gambiae* 18 из 31 общих полиморфных инверсий располагаются на плече 2R. И только 2 полиморфные инверсии локализируются на X-хромосоме. Из 82 полиморфных инверсий у *An. gambiae*, 67 локализируются на 2R, а 15 на 2L, 3R и 3L вместе взятых и ни одной на X-хромосоме. В ходе исследования был проведен сравнительный анализ организации геномов у трех видов малярийных комаров подрода *Cellia* - *An. stephensi*, *An. funestus* и *An. gambiae*. Для этого использовали 231 маркерную последовательность ДНК с известной локализацией на хромосомах *An. stephensi*. Анализ подтвердил, что парацентрические инверсии и полноплечевые транслокации являются основными видами перестроек в подроде *Cellia*. Также было подсчитано минимально возможное число инверсий между *An. gambiae* и *An. stephensi*. Было показано, что X-хромосома имеет самую высокую частоту фиксации инверсий, а 2R плечо эволюционирует быстрее других аутосомных плеч.

Самый высокий темп эволюции наблюдался у X-хромосомы, несмотря на отсутствие на ней полиморфных инверсий у всех трех исследуемых видов.

Анализ геномного ландшафта у *An. gambiae* обнаружил, что каждое хромосомное плечо обладает индивидуальными молекулярными особенностями. X-хромосома характеризуется самой высокой плотностью транспозонов и самым высоким покрытием всех видов сателлитов. 2R плечо имеет наивысшую плотность генов и сегментных дупликаций и наименьшую плотность транспозонов, как и наименьшее покрытие минисателлитами и MAR, которые сконцентрированы на остальных аутосомных плечах. Довольно низкая концентрация транспозонов наблюдалась в ЭХ, и основные пики приходились на ИГХ и ПЦГХ. Распространение генов имело противоположный паттерн. MAR были сконцентрированы в ПЦ районах всех плеч и также присутствовали в ЭХ 2L, 3R, 3L. Мы обнаружили район максимальной плотности сегментных дупликаций, приходящийся на проксимальную половину плеча 2R. Именно в этом месте расположена зона, богатая точками разрывов инверсий (Pombi et al., 2008).

На всех хромосомах *An. stephensi* плотность транспозонов меньше, чем у *An. gambiae*, при этом, на X-хромосоме плотность транспозонов более чем в 2 раза выше, чем на аутосомах. Плотность простых повторов в эухроматине *An. gambiae* в несколько раз выше, чем у *An. stephensi*. Наивысшая плотность простых повторов у обоих видов выявляется на X-хромосоме (рис. 6).

### **Консервативность и перетасовывание генов в ходе эволюции малярийных комаров.**

Геномное исследование синтении и изменения порядка генов у *An. gambiae* и *An. funestus* показало, что порядок генов у этих видов сильно изменен по длине всех хромосомных плеч. Однако были найдены и отдельные сегменты с консервативным порядком генов. Согласно проведенным расчетам, с момента расхождения видов произошла фиксация, по крайней мере, 70 инверсий. Т.о., предполагаемая скорость возникновения инверсий у малярийных комаров оказалась одной из самых высоких среди всех исследованных ранее видов эукариот, включая *Drosophila* (González, Ranz, Ruiz, 2002).

Чтобы исследовать синтению и число хромосомных инверсий между *An. gambiae* и *An. stephensi*, использовали данные по хромосомной локализации 6448 генов-ортологов. 2L хромосомное плечо у *An. gambiae*

является гомологом 3L плеча у *An. stephensi* (Sharakhova et al., 2011). 66 блоков синтении было определено на X-хромосоме. 104 и 64 блока на 2R и 2L (3L у *An. stephensi*). 104 и 42 – на 3R и 3L (2L у *An. stephensi*). Таким образом, X-хромосома демонстрирует наибольшее количество перестроек на единицу длины.

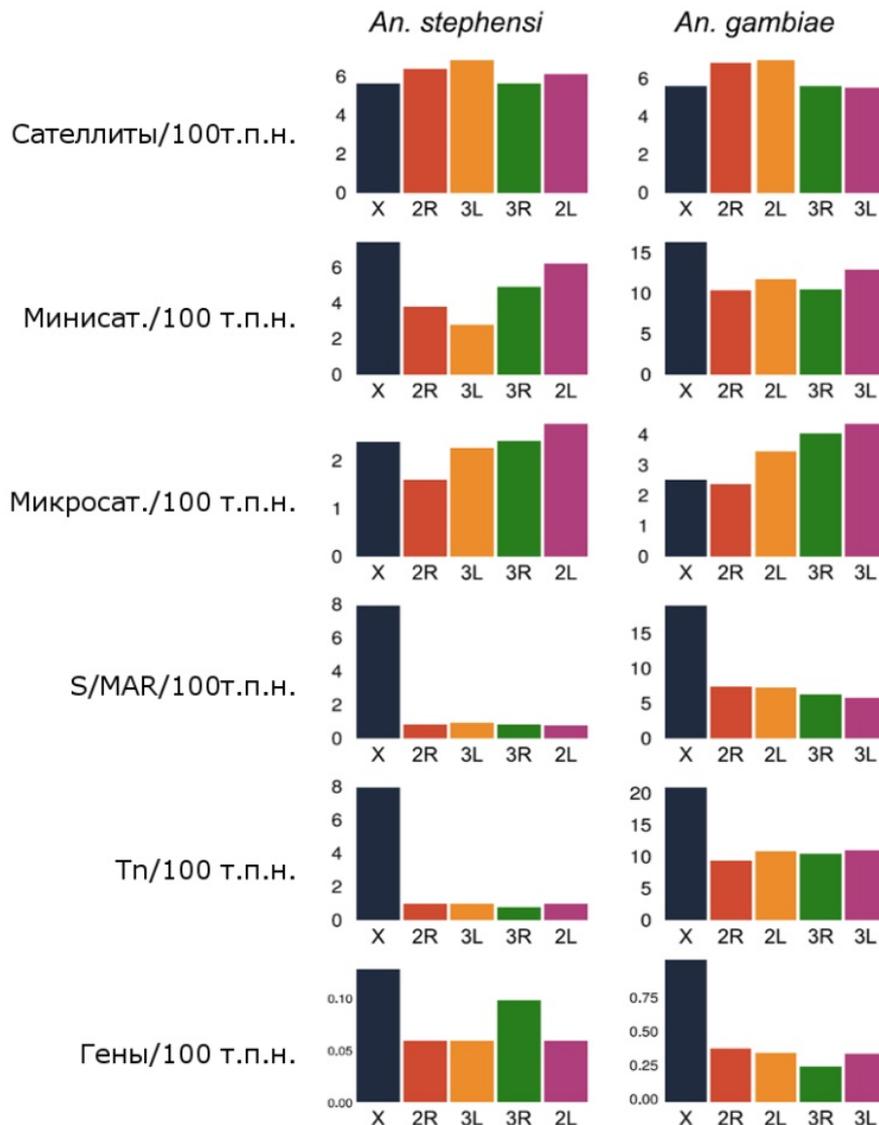


Рис. 6. Сравнительный анализ плотности расположения генов, транспозонов, MAR и сателлитов на 100 Кб генома у двух видов малярийных комаров

В рамках геномного исследования 16 видов малярийных комаров (Neafsey et al., 2015) появилась возможность провести сравнительный анализ разных видов малярийных комаров Америки, Евразии и Африки. Используя разработанные нами физические карты для *An. stephensi*, *An. funestus* и карту *An. albimanus* (Cornel, Collins, 2000). А также, проведя дополнительное хромосомное картирование для *An. albimanus* и *An.*

*atroparvus*, мы получили возможность исследовать паттерн хромосомной эволюции малярийных комаров на примере пяти видов: *An. gambiae*, *An. stephensi*, *An. funestus* (подрод *Cellia*), *An. atroparvus* (подрод *Anopheles*) и *An. albimanus* (подрод *Nyssorhynchus*). У *An. gambiae* плечи были обозначены как хромосомные элементы 1 (X), 2+3 (2R+2L), 4+5 (3R+3L). Соответствие для других видов было следующее: *An. funestus* 1, 2+4, 3+5; *An. stephensi* 1, 2+5, 3+4; *An. atroparvus* 1, 4+3, 2+5 и *An. albimanus* 1, 2+4, 5+3. Каждый из видов индивидуально сравнили с *An. gambiae*. Сравнительная позиция генов в каждой паре видов была использована для определения консервативных синтенных блоков. Были проанализированы два параметра: ориентация и порядок ортологичных генов. Порядок генов определялся относительно генома *An. gambiae*. Исходя из позиции 3908 однокопийных ортологичных генов было определено 253 консервативных блока синтении и 130 инверсий между *An. stephensi* и *An. gambiae*, 253 блока и 145 инверсий между *An. funestus* и *An. gambiae*, 340 блоков и 195 инверсий между *An. atroparvus* и *An. gambiae*, 903 блока и 589 инверсий между *An. albimanus* и *An. gambiae*. Синтения на уровне целого плеча оказалась весьма консервативной, несмотря на полноплечевые транслокации. В отличие от дрозофилы, эти хромосомные элементы переходят с хромосомы на хромосому путем их полной транслокации и не показывают слияний или «растворения» одних элементов в других.

### **Высокая частота перестроек на X-хромосоме малярийных комаров**

Чтобы сравнить разницу между скоростью эволюции X-хромосомы и аутосом у *An. stephensi*, *An. funestus*, *An. atroparvus* и *An. albimanus* использовали два независимых статистических теста. Получили достоверно значимые различия  $t(18)=12.527$ ,  $p < 0.0000001$ , со средним значением 0.126 и 0.041, соответственно для X и аутосом (рис. 7). Далее была проведена оценка частоты хромосомных перестроек у малярийных комаров в сравнении с дрозофилами. Число инверсий между *D. melanogaster* и 8 другими видами рода было ранее опубликовано (von Grotthuss, Ashburner, Ranz, 2010). В целом, частота перестроек у дрозофилы выше, чем у малярийных комаров. Однако соотношение частот перестроек на половой хромосоме комаров по отношению к аутосомам существенно выше, чем у дрозофил. Это может свидетельствовать о важной роли перестроек на X-хромосоме в видообразовании малярийных комаров.

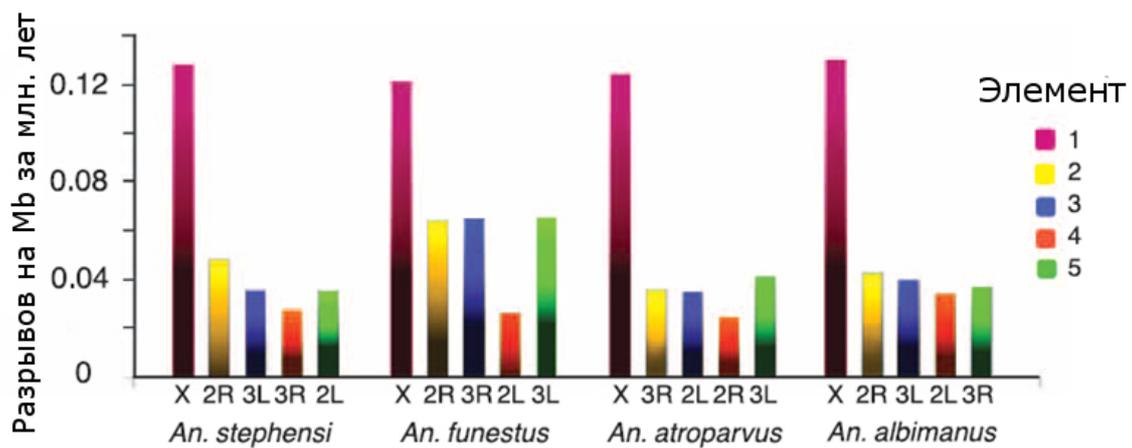


Рис. 7. У малярийных комаров X-хромосома показывает наиболее высокую степень перестроек в сравнении с аутосомами. Несмотря на отсутствие полиморфных инверсий на этой хромосоме у всех исследованных видов.

### **Молекулярные механизмы происхождения некоторых инверсий в комплексе видов *Anopheles gambiae***

Инверсия 2La является высоко полиморфной и широко распространенной в популяциях *An. gambiae*. Высокая частота данной инверсии выявляется в засушливой части ареала, есть также сезонная цикличность: от дождливого к засушливому сезону частота встречаемости инверсии растет (Powell et al., 1999). *An. gambiae* – единственный вид комплекса, у которого данная инверсия полиморфна. *An. bwambae*, *An. melas*, *An. quadriannulatus* А и В – мономорфны по инверсии 2L<sup>+</sup><sub>a</sub>, предположительно предковой или стандартной. *An. arabiensis* и *An. merus* имеют мономорфную инверсию, идентичную на цитологическом уровне 2La.

Чтобы пролить свет на происхождение 2La инверсии, были клонированы и секвенированы фрагменты ДНК из точек разрыва инверсии у двух изолятов *An. gambiae* и двух видов-сблингов *An. arabiensis* и *An. merus*. Также были секвенированы фрагменты из зоны точек разрыва альтернативной инверсии 2L<sup>+</sup><sub>a</sub> у *An. melas* и *An. quadriannulatus*. Было показано, что молекулярная организация точек разрыва 2La инверсии одинакова у всех трех видов: *An. gambiae*, *An. arabiensis* и *An. merus*. Различия между альтернативными перестройками заключались в фрагменте ДНК 750 п.н., представленном одной копией в проксимальной точке разрыва 2La инверсии и дупликацией в дистальной точке разрыва 2L<sup>+</sup><sub>a</sub> инверсии. Возникновение инверсии 2L<sup>+</sup><sub>a</sub> не могло произойти по простому сценарию, предполагающему «cut-and-paste»

механизм, включающий только два разрыва на хромосоме. Мы предложили модель, включающую три разрыва и обе гомологичные хромосомы (рис. 8). Доступные сиквенсы из точек разрыва инверсии  $2L^{+a}$  у *An. melas* и *An. quadriannulatus* подтверждают монофилитическое происхождение этой перестройки у комплекса *An. gambiae*. И этот факт имел решающее значение в изменении представлений о филогенетической истории видов комплекса.

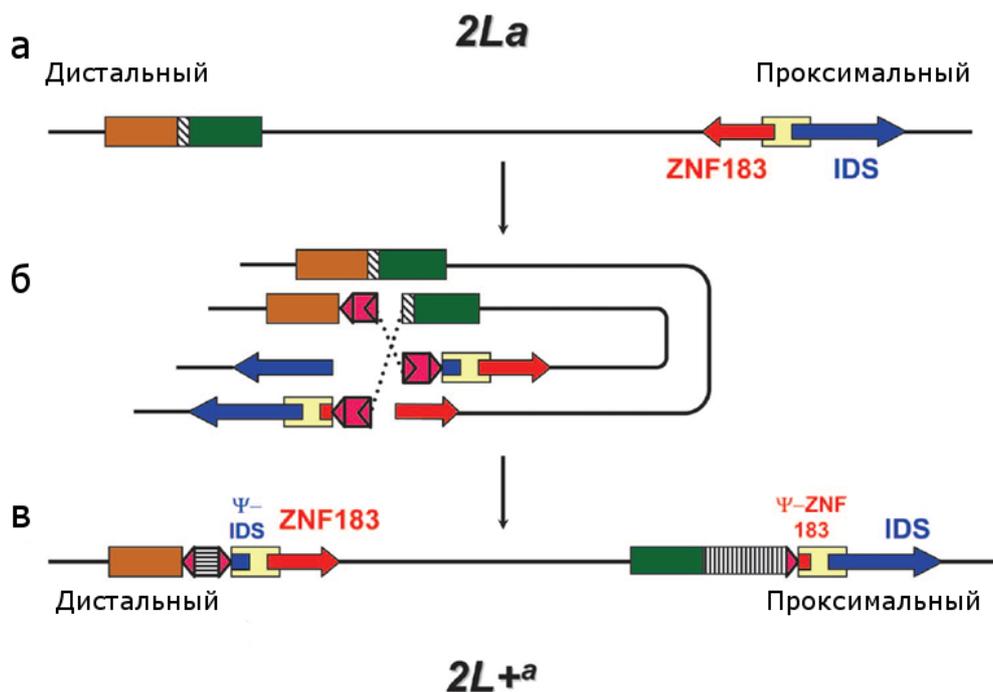


Рис. 8. Модель образования  $2L^{+a}$  из  $2L^a$ . Гомологичные последовательности представлены в виде прямоугольников и текста одного цвета. Красные и голубые горизонтальные стрелки и их обрезанные аналоги соответствуют генам или псевдогенам ( $\Psi$ ), которые потенциально кодируют ZNF183 и IDS белки. Желтый прямоугольник маркирует 750 п.н. фрагмент в  $2L^a$ , представленный в обратной ориентации в обеих точках разрыва  $2L^{+a}$ . Заштрихованный прямоугольник показывает комплекс дефектных транспозонов и повторенной ДНК. Треугольник соответствует терминальному фрагменту повторенных элементов, непосредственно прилегающему к точке разрыва. (а) предковый порядок  $2L^a$ ; (б) Два проксимальных разрыва на двух гомологах: один в районе гена *ZNF183*, второй – гена *IDS*. Три копии повторенных элементов (показаны красным) могли сгенерировать эти разрывы или могли быть связаны с предыдущими разрывами; (в) возникшая перестройка  $2L^{+a}$  с неповрежденными генами и инсерцией повторенных элементов на обеих точках разрыва. ДНК последовательность, гомологичная фрагменту в 750 п.н. здесь представлена дважды, в противоположной ориентации на обеих точках разрыва.

### **Функциональный аспект хромосомных перестроек**

Многочисленные наблюдения позволяют предположить, что гены, расположенные на 2R и 2L хромосомных плечах *An. gambiae* и их гомологов у *An. funestus* и *An. stephensi*, могут быть связаны с селективным ответом на давление окружающей среды (Mahmood, Sakai, 1984; Coluzzi et al., 2002; Cohuet et al., 2005). Чтобы изучить этот вопрос мы провели исследование синтении и сравнение порядка расположения генов на хромосомах у трех видов подрода *An. gambiae*, *An. funestus* и *An. stephensi*.

Наличие сходных групп генов в полиморфных инверсиях гомологичных хромосомных плеч у всех трех видов могло указывать, что происходит естественный отбор комбинаций адаптивных генов в сходных условиях окружающей среды. Мы тестировали наличие или отсутствие физически и *in silico* картированных кДНК и ВАС клонов (содержащих гены) относительно общих полиморфных инверсий изучаемых видов.

Исходя из предположения о том, что гены распределяются случайным образом относительно полиморфных инверсий и друг друга, мы должны были бы выявить случайное распределение идентичных маркеров по хромосомам. Однако была обнаружена неслучайная кластеризация маркеров в пределах полиморфных инверсий у разных видов. На рис. 9 представлены графики термокарт, которые демонстрируют различия между наблюдаемым и ожидаемым числом общих маркеров в каждой инверсии *An. gambiae* и *An. stephensi*, *An. gambiae* и *An. funestus*. Схожий *p-value* был рассчитан при симуляции распределения методом Monte Carlo. Полученные результаты указывают на то, что некоторые полиморфные инверсии, по крайней мере, на плече 2R *An. gambiae*, не случайно содержат генные комбинации, общие с инверсиями *An. stephensi* и *An. funestus*.

Анализ консервативных и разрушенных генных блоков выявил, что плечо 2R характеризуется наивысшей концентрацией разрушенных блоков на единицу длины. Напротив, 3R плечо у *An. gambiae* и его гомологи у других видов показывает сохранность всех выявленных блоков между всеми тремя видами. Следовательно, 3R имеет эволюционно консервативный набор точек инверсионных разрывов и обладает самой низкой толерантностью к возникновению новых инверсионных разрывов.

## Заключение к главе 5

Сравнительный анализ представителей *Anopheles* из разных серий подрода *Cellia*: *Pyretophorus* (*An. gambiae*), *Myzomyia* (*An. funestus*) и *Neocellia* (*An. stephensi*) и из подрода *Anopheles* (*An. atroparvus*) и подрода *Nyssorhynchus* (*An. albimanus*) показал, что хромосомная эволюция у малярийных комаров шла путем парацентрических инверсий и полноплечевых транслокаций. При этом геномный ландшафт и эволюционная история индивидуальных хромосомных плеч существенно различаются.

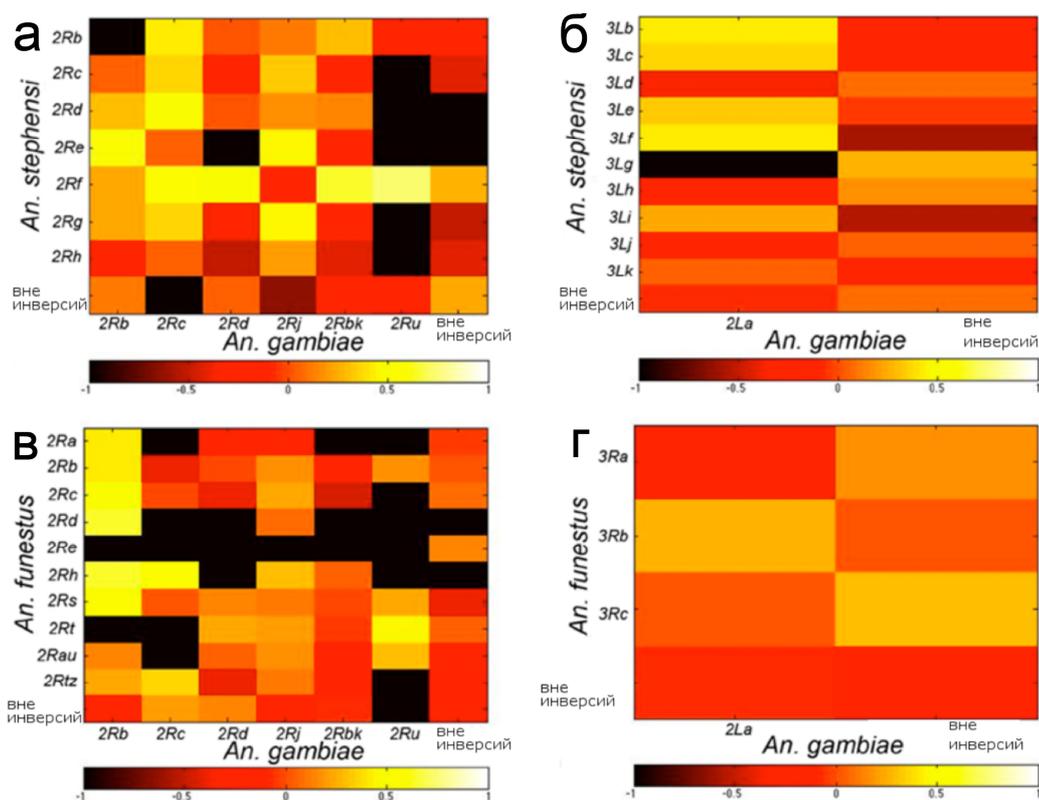


Рис. 9. Термокарты показывают обогащение (положительное и отрицательное) общими генами между инверсиями у двух видов. Цветом показано, насколько сильны отличия между наблюдаемым и расчетным (в случае, если инверсии происходили случайным образом) их количеством. Показаны инверсии на плече 2R (а) и 2L (б) у *An. gambiae* и *An. stephensi* и на 2R (в) и 2L (г) у *An. gambiae* и *An. funestus*. Светло-желтый и черный цвета показывают неслучайное присутствие или отсутствие одних и тех же маркеров в пределах инверсий. Оранжевым показаны случайные события.

Полиморфные инверсии неравномерно распространены по хромосомным плечам малярийных комаров. 2R и 2L *An. gambiae* и их гомологи у *An. stephensi* и *An. funestus* имеют наибольшую плотность генов. Ранее было показано, что 2L обогащена генами, отвечающими за

формирование кутикулы, тогда как на 2R расположены гены, вовлеченные в клеточный ответ на стрессовые условия окружающей среды. Таким образом, соответствующие полиморфные инверсии могут быть связаны с адаптацией комаров к сухому климату (Coluzzi et al., 2002). Было выдвинуто предположение, что если полиморфные инверсии у различных видов отвечают за одну и ту же экологическую адаптацию, то сходный набор генов может быть выявлен в пределах этих инверсий у разных видов. Было найдено, что некоторые полиморфные инверсии плеча 2R у всех трех видов неслучайно «ловят» сходные наборы генов. Неслучайное наличие гомологичных генов в пределах инверсии 2Rb у *An. gambiae* и инверсии 2Rh у *An. funestus* представляет особый интерес в свете экологической адаптации, связанной с этими инверсиями. (Coluzzi et al., 1979; Ayala et al., 2011).

Все основные синтенные блоки локализируются на медленно эволюционирующем плече 3R у *An. gambiae* и его гомологах у *An. funestus* и *An. stephensi*. Такая консервативность может быть следствием отрицательного отбора, направленного против разрушения определенных генных комбинаций. Кластеры генов могут иметь общие регуляторные зоны, общий паттерн экспрессии и единые эпигенетические механизмы регуляции (Ng, Wu, Zhang, 2009).

Сравнительный анализ геномов показал, что скорость изменения порядка генов на половой X-хромосоме комаров почти в три раза превышает таковую для аутосом, контрастируя с практически полным отсутствием полиморфных инверсий на X-хромосоме. Ранее было выдвинуто предположение, что гены, отвечающие за репродуктивную изоляцию у комаров, могут быть локализованы на X-хромосоме (Ayala, Coluzzi, 2005). Если подтвердится влияние гетерозиготных инверсий по X-хромосоме на жизнеспособность и размножение малярийных комаров, то их можно будет искусственно интродуцировать в природные популяции, чтобы уменьшить количество насекомых-переносчиков. Быстрое возникновение и фиксация инверсий на X-хромосоме также могут содействовать процессам видообразования у малярийных комаров, как это было показано для *Drosophila* (Machado, Haselkorn, Noor, 2007).

## **Глава 6 «Инверсионная филогения в комплексе видов *Anopheles gambiae*»**

*An. gambiae*, основной переносчик малярии в тропической Африке, принадлежит к комплексу близкородственных видов-сиблингов. Члены

комплекса характеризуются различной экологической адаптацией, поведением и способностью переносить малярийного плазмодия (Coluzzi et al., 1979, 2002; Ayala, Coluzzi, 2005). Комплекс видов *An. gambiae* представляет из себя прекрасную систему для изучения генетических основ видообразования, адаптации и отношений «паразит-хозяин».

Виды внутри комплекса можно определить на основании десяти фиксированных инверсий (Coluzzi et al., 2002; Krzywinski, Besansky, 2003). Вследствие относительно недавнего происхождения комплекса наблюдается высокий уровень сходства последовательностей в геноме и общий для разных видов полиморфизм, унаследованный от единого предка, что не позволяет однозначно реконструировать филогению: часто филогенетические построения по разным маркерам конфликтуют друг с другом (Besansky et al., 2003; Krzywinski et al., 2005). Микросателлитный полногеномный анализ выявил мозаичную геномную архитектуру в комплексе, что предполагает интрогрессию различных геномных районов между видами (Wang-Sattler et al., 2007). Филогенетический подход в рамках сравнительного геномного исследования в подобной ситуации представляет крайне высокий интерес.

### ***Филогения в комплексе An. gambiae, основанная на анализе аутосомных инверсий***

В первую очередь была протестирована возможность определения предкового порядка организации аутосом в комплексе *An. gambiae* с помощью внешней группы и комбинации биоинформатических и цитогенетических подходов и показана действенность такого подхода.

Далее было решено проанализировать порядок генов в геномных районах, включающих места разрывов хромосомных инверсий у всех видов комплекса *An. gambiae*, и сравнить их с порядком генов в этих районах у видов комаров, не относящихся к данному комплексу. Были идентифицированы гены в точках разрыва фиксированных перекрывающихся инверсий 2Ro и 2Rp у *An. merus* и гомологичные последовательности у *An. stephensi*, *Aedes aegypti* и *Culex quinquefasciatus* (все виды относятся к подсемейству Culicinae). Было показано, что «инвертированный» порядок, соответствующий 2Ro, на самом деле является предковым в изучаемом комплексе. Также было найдено, что инверсия 2La представлена у эволюционно отдаленных видов Culicinae, что также подтверждало ее предковый статус. Было сделано заключение, что клада *An. gambiae* – *An. merus* занимает базовую позицию, тогда как

*An. arabiensis* и *An. melas* – терминальную. Важно отметить, что предковый порядок хромосомных перестроек выявлен у видов – переносчиков человеческой малярии, тогда как «перестроенный» порядок – как у вида, не являющегося переносчиком – *An. quadriannulatus*, так и у вида-переносчика *An. arabiensis*.

### **Филогения в комплексе *An. gambiae*, основанная на инверсиях в X-хромосоме.**

Из 10 фиксированных инверсий в комплексе (Coluzzi et al., 2002), 5 локализируются на X-хромосоме. На основании данных по этим 5 инверсиям комплекс может быть разделен на три группы: 1) *An. merus* и *An. gambiae* с общей Хаg инверсией; 2) *An. quadriannulatus*, *An. bwambiae* и *An. melas*, со стандартным порядком генов на X и 3) *An. arabiensis* у которого обнаружена смешанная инверсия Хbcd. В части нашего исследования, входящей в работу нескольких групп ученых (Fontaine et al., 2015), были 1) определены геномные координаты точек разрывов этих 5 фиксированных инверсий. 2) С помощью «внешних», по отношению к комплексу, видов выявлены предковые и происходящие от них перестройки; 3) реконструированы филогенетические отношения между видами комплекса с помощью фиксированных инверсий на X-хромосоме в качестве маркеров; 4) проведена оценка времени дивергенции видов в комплексе на основании рассчитанной скорости эволюции X-хромосомы у *Anopheles* (Neafsey et al., 2015). Средняя скорость эволюции X-хромосомы составила  $0.126 \pm 0.004$  разрыва/Мб/млн. лет. Время расхождения между линиями *An. gambiae* и *An. arabiensis* было оценено как  $1.88 \pm 0.05$  млн. лет назад. И эта оценка очень близка к времени расхождения ( $1.85 \pm 0.47$  млн. лет), рассчитанному на основании ДНК-последовательности X-хромосом видов комплекса.

Проведенное нами исследование филогении на основе хромосомных инверсий опровергло основную массу аутомных топологий и поддержало филогению, основанную на сопоставлении последовательностей X-хромосом. Также данные инверсионной филогении о том, что *An. merus* и *An. gambiae* являются видами, эволюционно наиболее близкими к предковому, согласовываются с данными по полногеномному анализу X-сцепленных последовательностей у шести видов. Итоговая схема инверсионной филогении комплекса

представлена на рис. 10 в сравнении со схемой, существовавшей на начало наших исследований в этой области.

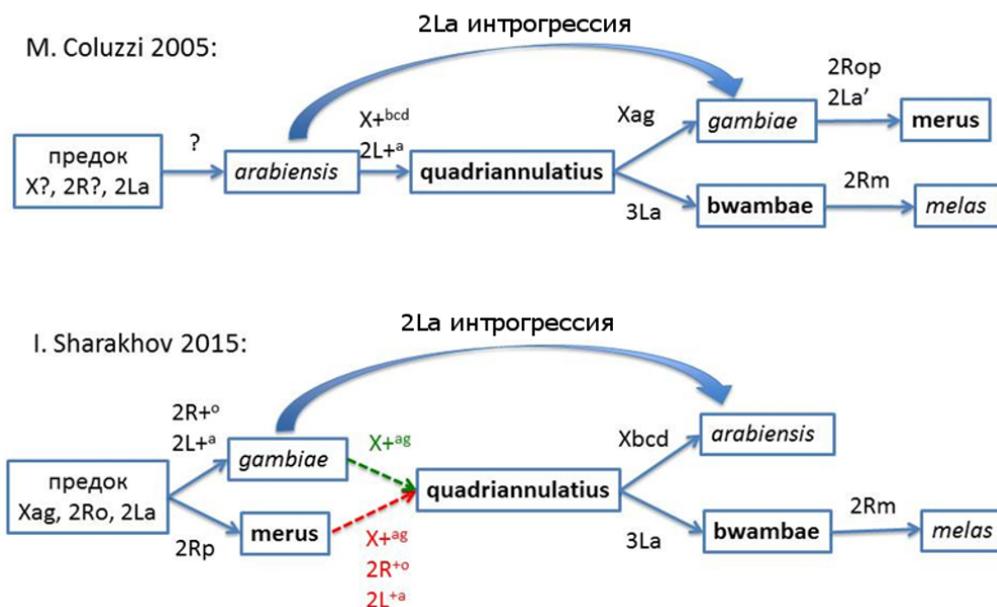


Рис. 10. Филогения видов комплекса *An. gambiae* на основе хромосомных инверсий.

### Заключение к главе 6

Использованный нами подход, основанный на анализе порядка генов в районах возникновения точек разрывов хромосомных инверсий у видов комплекса *An. gambiae* и сравнении этого порядка с таковым у «внешних» видов, кардинально изменил представление о филогении комплекса. Оказалось, что *An. merus*, минорный переносчик малярии (Рок Тсу et al., 2003), относится к видам, наиболее близким к предковым видам комплекса. *An. merus* и *An. gambiae* полагали сестринскими таксонами на основании уникального происхождения фиксированных инверсий и X-сцепленных последовательностей (Besansky et al., 2003; White et al., 2011). Согласно новой хромосомной филогении, эти два вида обладают наиболее примитивными кариотипами в исследуемом комплексе. Значит, основной переносчик малярии в Африке *An. gambiae* близок к предковым видам, тогда как кариотипы видов *An. quadriannulatus* А и В, не являющихся переносчиками малярии, произошли от кариотипа *An. gambiae*. *An. quadriannulatus* не включен в цепь передачи малярии, поскольку питается исключительно на животных (Coluzzi et al., 2002). Кариотип *An. melas* сформировался относительно недавно, но этот вид тоже является переносчиком малярии в Западной Африке (Mourou et al., 2010; Ridl et al.,

2008). Таким образом, включение человека в трофическую цепь и, как следствие, способность комаров переносить малярию, могли возникать в комплексе неоднократно.

В дальнейшем, знание о филогенетических связях между видами комплекса может быть использовано для определения конкретных генетических изменений, связанных как с выбором крови человека в качестве источника питания, так и с адаптацией отдельных видов к разным экологическим условиям.

### **Выводы:**

1. Впервые определены геномные границы между эу- и гетерохроматином в политенных хромосомах *An. gambiae*. Показано, что прицентромерный гетерохроматин обогащен РНК транспозонами, минисателлитами и сателлитами, тогда как интеркалярный гетерохроматин обогащен сегментными дупликациями. Выяснено, что гетерохроматин обогащен генами, участвующими в регуляции биологических процессов и организации хромосом.

2. Анализ локусов, продуцирующих малые некодирующие РНК класса пиРНК у *An. gambiae*, выявил следующее:

а) В сравнении с *D. melanogaster* значительно больше пиРНК кластеров малярийного комара характеризуются однонаправленной транскрипцией и располагаются в эухроматиновой части генома;

б) Существенная часть пиРНК-обогащенных генов имеет функции, связанные с размножением и эмбриональным развитием малярийного комара, что позволяет предположить участие пиРНК в эпигенетической регуляции соответствующих процессов;

в) Как у дрозофилы, так и у малярийного комара функция пиРНК в первую очередь направлена на подавление активности ретротранспозонов в клетках зародышевой линии.

3. Впервые установлено, что гетерохроматин Y-хромосомы подвергался быстрым и радикальным перестройкам во время эволюции комплекса видов *An. gambiae*. В отличие от дрозофилы, для малярийных комаров показано наличие рекомбинации между X- и Y-хромосомами.

4. Составлены новые цитогенетические и физические карты высокого разрешения с использованием высокополитенных хромосом питающих клеток яичников для эпидемиологически значимых видов малярийных комаров подрода *Cellia*: *An. gambiae*, *An. funestus*, *An. nili* и *An. stephensi*.

5. Сравнительный геномный анализ малярийных комаров из подродов

*Cellia*, *Anopheles* и *Nyssorhynchus* позволил сделать следующие важные заключения:

а) полиморфные хромосомные инверсии 2R плеча содержат сходные наборы полиморфных генов у эволюционно отдаленных видов малярийных комаров. Это открытие предполагает параллельную эволюцию геномов разных видов комаров в ходе адаптации к сходным условиям среды;

б) аутомсомные плечи различаются по своей толерантности к возникновению точек инверсионных разрывов. Все выявленные участки хромосом с совпадающим порядком генов демонстрируют консервативный характер в медленно эволюционирующем плече 3R у *An. gambiae*, *An. funestus* и *An. stephensi*; тогда как распределение точек разрывов в быстро эволюционирующем плече 2R является видоспецифичным;

в) половая X-хромосома значительно чаще меняет свою структуру, по сравнению с аутомсомами, что может свидетельствовать о важности перестроек X-хромосомы в видообразовании. Показана локализация повторяющейся ДНК в точках разрывов инверсии 2La и предложен механизм ее формирования.

6. На основании анализа хромосомных инверсий создана принципиально новая филогенетическая схема видов, входящих в комплекс *An. gambiae*. Показано, что инверсионная структура геномов *An. merus* и *An. gambiae* наиболее близка к предковому кариотипу. Выдвинуто предположение, что главные переносчики малярии, относящиеся к исследуемому африканскому комплексу, начали дивергировать от общего предка примерно  $1.88 \pm 0.05$  миллионов лет назад.

## ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

### Статьи в рецензируемых журналах:

1. **Sharakhov I.V.**, Braginets O., Mbogo C.N., Yan G. Isolation and characterization of trinucleotide microsatellites in African malaria mosquito *Anopheles funestus* // Mol Ecol Notes. - 2001.-V.1. – P. 289-292.

2. **Sharakhov I.V.**, Sharakhova M.V., Mbogo C.M., Koekemoer L., Yan G. Linear and spatial organization of polytene chromosomes of the African malaria mosquito *Anopheles funestus* //Genetics. - 2001. – V.159. – P. 211-218.

3. **Sharakhov I.V.**, Serazin A.C., Grushko O.G., Dana A. ... Besansky

N.J. Inversions and gene order shuffling in *Anopheles gambiae* and *A. funestus* // Science. - 2002. – V.298. – P. 182-185.

4. **Sharakhov I.V.**, Braginets O., Grushko O., Cohuet A., ..., Besansky N.J. A microsatellite map of the African human malaria vector *Anopheles funestus* // Journal of Heredity. - 2004. –V. 95. – P. 29-34.

5. **Sharakhov I.V.**, White B.J., Sharakhova M.V., Kayondo J., ... Besansky N.J. Breakpoint structure reveals the unique origin of an interspecific chromosomal inversion (2La) in the *Anopheles gambiae* complex // Proc Natl Acad Sci USA. - 2006. – V.103. – P. 6258-6262.

6. Sharakhova M. V., Xia A., McAlister S.I., **Sharakhov, I.V.** A standard cytogenetic photomap for the mosquito *Anopheles stephensi* (Diptera: Culicidae): application for physical mapping // J Med Entomol. - 2006. – V.43. P. 861-866.

7. Xia A., Sharakhova M.V., **Sharakhov I.V.** Reconstructing ancestral autosomal arrangements in the *Anopheles gambiae* complex // J Comp Biol. 2008. – V. 15(8). – P. 1-16.

8. George P., Sharakhova M.V., **Sharakhov I.V.** High-resolution cytogenetic map for the African malaria vector *Anopheles gambiae* // Insect Mol Biol. - 2010. – V.19 (5). - P. 675-682.

9. Sharakhova M.V., George P., Brusentsova I.V., Leman S.C., Bailey J.A., Smith C.D., **Sharakhov I.V.** Genome mapping and characterization of the *Anopheles gambiae* heterochromatin // BMC Genomics. – 2010. –V.11. – P. 459-476.

10. Sharakhova M.V., Xia A., Tu Z., Y. Shouche S., Unger M.F., **Sharakhov I.V.** A physical map for an Asian malaria mosquito, *Anopheles stephensi* // Am J Trop Med Hyg. – 2010. – V.83(5). – P. 1023–1027.

11. Xia, A., Sharakhova, M.V., Leman, S.C., Tu, Z., ..., **Sharakhov, I.V.** 2010. Genome landscape and evolutionary plasticity of chromosomes in malaria mosquitoes // PLoS One. – 2010. –V.5(5) – e10592.

12. **Шарахов И.В.**, Шарахова М.В. 2010. Хромосомная эволюция у малярийных комаров // Генетика. – 2010. –Т. 46(9). – С. 1250–1253.

13. Шарахова М.В, **Шарахов И.В.** 2010. Организация и эволюция гетерохроматина у малярийных комаров // Генетика. – 2010. – Т. 46(10). – С. 1417–1420.

14. Kamali M., Sharakhova M.V., Baricheva E., Karagodin, D., ..., **Sharakhov I.V.** 2011. An integrated chromosome map of microsatellite markers and inversion breakpoints for an Asian malaria mosquito, *Anopheles*

*stephensi* // J Hered. – 2011. –V. 102(6). – P. 719-726.

15. Sharakhova, M.V., Antonio-Nkondjio, C., Xia, A., Ndo, C., ..., **Sharakhov, I.V.** 2011. Cytogenetic map for *Anopheles nili*: application for population genetics and comparative physical mapping // Infect Genetics Evol. – 2011. –V. 11(4). – P. 746-754.

16. Sharakhova, M.V., Xia, A., Leman, S.C., **Sharakhov, I.V.** 2011. Arm-specific dynamics of chromosome evolution in malaria mosquitoes // BMC Evol Biol. – 2011. –V. 11. – P. 91.

17. Tubio J. M. C., Tojo M., Bassaganyas L., ... **Sharakhov I.V.**,..., Besansky N.J. 2011. On the distribution, abundance and occupancy of LTR retrotransposons in the genome of *Anopheles gambiae* // PLoS One. – 2011. – V. 6(1). –e16328.

18. George P., Sharakhova M.V., **Sharakhov I.V.** High-throughput physical mapping of chromosomes using automated *in situ* hybridization // J Vis Exp. – 2012. –V. 64. – P. 4007.

19. Kamali, M., Xia, A., Tu, Z., **Sharakhov, I.V.** A new chromosomal phylogeny supports the repeated origin of vectorial capacity in malaria mosquitoes of the *Anopheles gambiae* complex // PLoS Pathog. – 2012. –V. 8(10). – e1002960.

20. **Шарахов И.В.** Хромосомная филогения малярийных комаров // Цитология. – 2013. –Т. 55(4). – С. 238-240.

21. George P, Sharma A, **Sharakhov IV.** 2D and 3D chromosome painting in malaria mosquitoes // J Vis Exp. – 2014. –V. 6:(83). – e51173.

22. Jiang X, Peery A, Hall S... **Sharakhov IV,** Tu Z. Genome analysis of a major urban malaria vector mosquito, *Anopheles stephensi* // Genome Biol. – 2014. –V. 15(9). – P. 459.

23. Fontaine M.C., Pease J.B., Steele A., ..., **Sharakhov I.V.**, ..., Besansky N.J. Extensive introgression in a malaria vector species complex revealed by phylogenomics // Science. – 2015. –V. 347(6217). – P. 1258524.

24. George, P., Jensen. S., Pogorelnik, R., Lee, J., ... **Sharakhov I.V.** Increased production of piRNAs from euchromatic clusters and genes in *Anopheles gambiae* compared with *Drosophila melanogaster* // Epigenetics Chromatin. – 2015. – V. 8– P. 50.

25. Neafsey D.E., Waterhouse R.M., Abai M.R., Aganezov S.S., ..., **Sharakhov IV,** ..., Besansky NJ. Highly evolvable malaria vectors: the genomes of 16 *Anopheles* mosquitoes // Science. – 2015. –V. 347(6217). –P. 1258522.

26. **Sharakhov I.V.**, Sharakhova M.V. Heterochromatin, histone modifications, and nuclear architecture in disease vectors // *Curr Opin Insect Sci.* – 2015. –V.10. – P. 110-117.

27. Hall A.B, Papathanos P.A, Sharma A, ..., **Sharakhov I.V.**, ..., Besansky NJ. 2016. Radical remodeling of the Y chromosome in a recent radiation of malaria mosquitoes // *Proc Natl Acad Sci USA.* – 2016. –V. 113(15). –P. E2114–E2123.

28. **Sharakhov I.V.**, Artemov G.N., Sharakhova M.V. Chromosome evolution in malaria mosquitoes inferred from physically mapped genome assemblies // *J Bioinform Comput Biol.* – 2016. –V. 14(2). – P. 1630003.

#### **Монографии и главы в книгах:**

1. **Sharakhov I.V.** Chromosome plasticity, adaptation and speciation in malaria mosquitoes // 2013. Chapter 4. In: *Speciation: Natural Processes, Genetics and Biodiversity*. Michalak, P. [ed.] New York, NY: Nova Science Publishers, Inc., P. 83-118.