

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ ХИМИЧЕСКОЙ БИОЛОГИИ И ФУНДАМЕНТАЛЬНОЙ МЕДИЦИНЫ
СИБИРСКОГО ОТДЕЛЕНИЯ РОССИЙСКОЙ АКАДЕМИИ НАУК

На правах рукописи

Кичигин Илья Геннадьевич

НАУЧНЫЙ ДОКЛАД
об основных результатах выполненной
научно-квалификационной работы

Изучение состава и эволюции
половых хромосом чешуйчатых

Направление подготовки

06.06.01 Биологические науки

Направленность

03.01.07 - Молекулярная генетика

Новосибирск - 2018

Общая характеристика работы

Актуальность

Раздельнополость обнаружена у подавляющего большинства позвоночных. В разных таксонах могут быть совершенно разные механизмы определения пола. Программа дифференциации клеток зародышевого пути неизменна среди позвоночных, разнообразие наблюдается на уровне генов запускающих каскад определения пола. Пол может определяться как хромосомами (системы XY и ZW), так и температурой окружающей среды (TSD). Лучше всего изучена консервативная система XY хромосом плацентарных млекопитающих. Хорошо охарактеризована ZW система птиц. Всё чаще появляются работы, описывающие разнообразные системы амфибий и костистых рыб. Системы определения пола рептилий, а в частности отряда чешуйчатых (Squamata), изучены не достаточно хорошо. У чешуйчатых присутствует широкий спектр систем определения пола, включающий как XY и ZW системы, так и TSD. Интересна также и скорость эволюции половых систем — даже у близкородственных видов могут встречаться разные системы. Ещё одна уникальная особенность этого таксона — частые случаи слияния аутосом с половыми хромосомами.

Половые системы чешуйчатых в основном изучены на цитогенетическом уровне. Почти ничего не известно о геномном составе хромосом, а так же о структуре повторённых последовательностей. В век повсеместных проектов секвенирования геномов рептилии оказались незаслуженно обделены исследованиями на уровне секвенированных последовательностей ДНК. Не так давно была получена частичная сборка полного генома каролинского анолиса *Anolis carolinensis*. Макрохромосомы были почти собраны, тогда как микрохромосомы остались на уровне групп сцепления и скаффолдов. Половую хромосому не удалось собрать, только обозначить одну из групп сцепления, как группу с половой хромосомы.

В отделе разнообразия и эволюции животных ИМКБ СО РАН собрана коллекция хромосомспецифичных библиотек ДНК некоторых рептилий, в том числе и половых хромосом. Высокопроизводительное секвенирование отдельных библиотек может расширить наше понимание о механизмах определения пола чешуйчатых. Однако анализ данных высокопроизводительного секвенирования - не стандартная процедура, присутствует ряд трудностей, связанных с контаминацией полногеномной ДНК изучаемого вида и человека, необходимостью работать с межвидовыми выравниваниями и т.д.

Цель

Целью данной работы является исследования генетического состава половых хромосом каролинского анолиса *Anolis carolinensis*, коричневого анолиса *Anolis sagrei* и плоскохвостого домового геккона *Hemidactylus platyurus* с применением методов анализа данных высокопроизводительного секвенирования (Illumina Miseq) ДНК отдельных хромосом.

Задачи:

1. Адаптировать метод поиска уникальных геномных районов в данных высокопроизводительного секвенирования Illumina под специфику неполных сборок геномов рептилий
2. Применить адаптированный метод для описания состава половых хромосом каролинского анолиса *Anolis carolinensis*, коричневого анолиса *Anolis sagrei* и плоскохвостого домового геккона *Hemidactylus platyurus*.
3. Описать различие в составе повторённых последовательностей между аутосомами и половыми хромосомами этих видов.
4. Проанализировать особенности эволюции половых хромосом исследованных видов.

Научная новизна и практическая ценность

Недавно разработанный метод поиска уникальных геномных районов в данных высокопроизводительного секвенирования отдельных хромосом был впервые применён для улучшения незавершённой сборки и последующего анализа структуры хромосом. Применимость данного аспекта метода была продемонстрирована на собранной аутосоме 6 каролинского анолиса. Известная последовательность хромосомы была разделена случайным образом на псевдо-скаффолды случайного размера 10 раз. После обработки проверяемым методом удавалось собирать (90±3)% начальной аутосомы 6. Таким образом было доказано, что метод применим для экспериментальных данных.

В ходе данной работы удалось присвоить 189 млн. п.н. неопределённых контигов к микрохромосомам каролинского анолиса. Хромосома 13 каролинского анолиса была определена как половая хромосома. Было показано что половая хромосома коричневого анолиса состоит из групп сцепления гомологичных хромосом X, 9, 12 и 18 каролинского анолиса. Выявлено, что половая хромосома плоскохвостого домового геккона гомологична хромосоме 9 каролинского анолиса. В ходе анализа литературных данных и результатов этого и других проектов были обнаружены консервативные группы сцепления, независимо связанные с полом у разных таксонов рептилий. Проведено сравнение повторённых последовательностей между половыми хромосомами и аутосомами у исследованных видов.

Материалы и методы

Методы молекулярной биологии и цитогенетики

Культуры фибробластов каролинского анолиса, коричневого анолиса и плоскохвостого домового геккона были получены из коллекций Отдела разнообразия и эволюции геномов ИМКБ СО РАН и Кэмбриджского Ресурсного Центра Сравнительной Геномики. Культивирование клеток, подготовка хромосом и хромосомный сортинг были проведены в Кэмбриджском Ресурсном Центре Сравнительной Геномики. Сортированные хромосомы амплифицировали с помощью DOP-ПЦР с праймером CCGACTCGAGNNNNNATGTGG. Высокопроизводительное секвенирование проводили на платформе Illumina MiSeq.

Методы биоинформатики

Метод анализа данных секвенирования отдельных хромосом включает в себя опознавание уникальных районов, представленных на хромосомах. В ходе адаптации метода к не полностью собранным сборкам генома, было решено использовать дополнительный параметр для определения целевых скэффолдов — гомологию к близкородственным видам, в частности гомологию к курице *Gallus gallus*. Основные процедуры были автоматизированы с помощью конвейера DOPseq_analyzer, доступным на https://github.com/ilyakichigin/DOPseq_analyzer. Анализ повторённых последовательностей проводили с помощью инструментов RepeatExplorer (<https://repeatexplorer-elixir.cerit-sc.cz/>)

Результаты и обсуждение

Характеристика хромосомных библиотек

В настоящей работе были использованы библиотеки сортированных хромосом и смесей хромосом каролинского анолиса *Anolis carolinensis* 6 (ACA6), 7-8 (ACA7-8), 9 (ACA9), 10-11 (ACA10-11), 12-13 (ACA12-13), 14 (ACA14), 15-17 (ACA15-17), 18 (ACA18); коричневого анолиса *Anolis sagrei* 6 (ASA6), X (ASAX), Y (ASAY), 8-9 (ASA8-9), 10-11 (ASA10-11), 12 (ASA12), 13-14 (ASA13-14) и плоскохвостого домового геккона *Hemidactylus platyurus* 2 (HPL2), 11-12 (HPL11-12), 13-15 (HPL13-15), 15 (HPL15), 16-17 (HPL16-17), 18-20 (HPL18-20), 19 (HPL19). Предполагалось, что other Masterpieces there will be, this is GR1/00, что у плоскохвостого домового геккона хромосома 19 является Z хромосомой. Специфичность библиотек была подтверждена флюоресцентной *in situ* гибридизацией (FISH).

Разработка метода анализа данных высокопроизводительного секвенирования хромосомспецифичных библиотек

На предыдущих этапах исследования хромосомспецифичных библиотек был разработан воспроизводимый и легко применимый метод анализа — программа-конвейер DOPseq_analyzer. Она является набором скриптов, написанных на языках программирования Python и R, автоматизирующих работу сторонних программ и проводящих дополнительные вычисления. Управление конвейером проводится с помощью конфигурационного файла с указанием всех необходимых параметров.

Для удаления последовательностей праймеров и адаптеров Illumina использовали программу cutadapt, работающую с вырожденными последовательностями ДНК. Выравнивание парных прочтений на референсный геном производили с помощью программы bwa. Для оценки и удаления загрязнения человеческой ДНК одновременно с выравниванием на целевой геном (*Anolis carolinensis*) проводили выравнивание на геном человека, прочтения с более высоким качеством картирования на геном человека, а также с низким качеством выравнивания на целевой геном отбрасывали.

Для поиска целевых скэффолдов перекрывающиеся позиции картирования прочтений интерпретировали как положения ампликонов. Ключевым параметром определения целевых скэффолдов стало расстояние между ампликонами (средние значения на целевых скэффолдах ниже, чем на нецелевых). Для автоматизации поиска целевых скэффолдов мы использовали пакет DNAscore для языка программирования R, созданный для поиска дупликаций и делеций в данных сравнительной геномной гибридизации на микрочипах. На его основе был написан скрипт region_dnascor.R, разделяющий заданные скэффолды целевого генома по средним значениям расстояний между ампликонами: низкие значения характерны для целевых скэффолдов, высокие — для нецелевых. Оказалось, что одного только расстояния между ампликонами недостаточно для точного определения состава микрохромосом. Так как хромосомспецифичные библиотеки характеризуются контаминацией полногеномной ДНК и референсным геном *Anolis carolinensis* содержит значительное количество небольших скэффолдов, скэффолды с промежуточным значением среднего расстояния между ампликонами невозможно было однозначно отнести к целевым или нецелевым. Для разрешения подобных случаев ввели дополнительный параметр — гомологию скэффолда хромосомам курицы *Gallus gallus* и человека *Homo sapiens*. Было показано, что микрохромосомы чешуйчатых консервативно гомологичны небольшому количеству хромосом курицы, таким образом скэффолды с одной микрохромосомы анолиса гомологичны одной, реже двум-трем хромосомам курицы. Гомология к человеку использовалась в качестве дополнительной проверки.

Анализ состава повторённой ДНК проводили непосредственно на ридах с помощью инструмента RepeatExplorer. Данная программа производит

кластеризацию прочтений по перекрытию, аннотацию по базе данных повторов RepeatMasker для позвоночных, а также сборку консенсусных последовательностей кластеров. Автоматическую аннотацию перепроверяли вручную с помощью инструмента RepeatMasker Sensor, в дальнейший анализ отбирались только кластеры с сигналами чешуйчатых, остальные отбрасывались как контаминационные. Сравнение состава повторов аутосом каролинского анолиса и полного генома с базы данных RepeatMasker показало значительное сходство между количеством и частотой основных семейств повторённых элементов.

Улучшение геномной сборки каролинского анолиса

Референсная сборка *Anolis carolinensis* anoCar2.0 содержит собранные шесть макрохромосом и оставшиеся 12 микрохромосом в виде скэффолдов и групп сцепления различного размера. Используя разработанную нами методологию, удалось определить состав всех микрохромосом каролинского анолиса (Таблица 1). Всего на 12 микрохромосом было распределено 419 скэффолдов общим размером в 189 млн. п.н. На основе этих данных была построена все последующие эксперименты, так как появилась основа для сравнения данных уникальных районов микрохромосом чешуйчатых.

Половые хромосомы каролинского анолиса гомоморфны и являются микрохромосомами. С помощью цитогенетических экспериментов не удавалось обнаружить XY хромосомы. В ходе проекта по секвенированию и сборке генома каролинского анолиса с помощью ВАС-клонов было показано, что группа сцепления LGb — часть X хромосомы этого вида. После секвенирования микрохромосом сигнал на группу сцепления LGb отчётливо давали риды с хромосомы 13. Таким образом мы впервые установили, что X хромосомой у *Anolis carolinensis* по размеру является хромосома 13.

Таблица 1. Результат дополнения референсной сборки каролинского анолиса *Anolis carolinensis*. Гомология целевых скэффолдов курице и человеку использовались как дополнительный параметр для разрешения пограничных случаев. Группа сцепления — расположение групп сцепления из референсной сборки каролинского анолиса

Хромосома	Гомология курице (GGA)	Гомология человеку (HSA)	Группа сцепления	Общий размер уникальных районов, млн. п. н.
7	GGA22+GGA24/6	HSA11+HSA2/8	LGa	31.6-31.9
8	GGA11	HSA16/19	LGc, LGd	27.0-27.8
9	GGA23+GGA1/3/5/6	HSA1+HSA19q	LGf	24.6-24.9

10	GGA10	HSA15	-	22.7-23.2
11	GGA4	HSAX	-	22.5-22.9
12	GGA14	HSA16/17/7	-	16.8-18.1
13(X)	GGA15	HSA12/22	LGb	9.7-10.6
14	GGA3/5+GGA25	HSA11+HSA1	-	9.4-10.9
15	GGA19	HSA17/7	-	8.2-9.5
16	GGA17	HSA9	LGg	7.7-9.7
17	GGA21	HSA1	LGh	6.8-8.2
18	GGA28	HSA19	-	2.1-3.0

Анализ состава половых хромосом чешуйчатых

На следующем этапе были секвенированы хромосомспецифичные библиотеки микрохромосом коричневого анолиса *Anolis sagrei* и плоскохвостого домового геккона *Hemidactylus platyurus*. С помощью дополненного референсного генома каролинского анолиса нам удалось оценить состав половых хромосом исследуемых видов, а также их гомологию относительно *Anolis carolinensis*. Ранее, цитогенетическим анализом было показано, что половые хромосомы коричневого анолиса состоят из блоков синтенных X хромосоме и аутосомам 9 и 12 каролинского анолиса. Анализ данных секвенирования показал, что на X хромосоме коричневого анолиса также присутствует, незамеченный ранее, блок синтенный аутосоме 18 каролинского анолиса. После транслокации аутосомных блоков на половые хромосомы Y хромосома коричневого анолиса потеряла как минимум 25% материала в ходе дегенерации.

Ген регулирующий половой каскад у разных таксонов позвоночных, как правило, сам является частью каскада. Мы провели поиск кандидатов на роль ключевого гена в каскаде анолисовых, опираясь на набор известных генов из каскада позвоночных (Таблица 2). К сожалению очевидного кандидата обнаружить не удалось. Стоит отметить, что гены NR0B1, SOX8 и RSPO1 оказались на половых хромосомах коричневого анолиса, но они вряд ли являются ключевыми в направлении каскада, так как такой ген, скорее всего должен быть консервативен у двух близкородственных видов с идентичной системой хромосом.

Таблица 2 Расположение генов из каскада позвоночных на хромосомах анолисовых. Un — неизвестная хромосома

Ген	Положение в референсной сборке каролинского анолиса <i>Anolis carolinensis</i>	Хромосома <i>A. carolinensis</i>	Хромосома <i>A. sagrei</i>
DHH	chr1, chr6, chrUn_GL343198	1, 6, Un	6, Un
DMRT1	chr2	2	Un
FGF9	chr3	3	Un
ZFPM2	chr4	4	Un
GATA4	chr4, chrUn_GL343704	4, Un	4, Un
Map3k1	chr2	2	Un
Map3k4	chr1	1	Un
NR0B1	chr3, chrUn_GL343488	3, 9	X, Y, Un
SOX3	chr3, chrUn_GL343310	3, 11	11, Un
SOX8	chrUn_GL343263	12	X, Y
SOX9	chrUn_GL343258	Un	Un
SOX10	chrUn_GL343460	Un	Un
CTNNB1	chrUn_GL343224	Un	Un
FOXL2	chr3	3	Un
FST	chr2	2	Un
RSPO1	chrUn_GL343426	9	X, Y
WNT4	chr2, chr5, chr6,	2, 5, 6	6, Un
CBX2	chr2	2	Un

Секвенирование хромосомспецифичных библиотек плоскохвостого домового геккона *Hemidactylus platyurus* показало, что Z хромосома этого вида состоит из блока синтенного аутосоме 9 каролинского анолиса. Примечательно, что этот

блок также содержит гены NR0B1 и RSP01, что может говорить о важности связывания этих генов с полоспецифичным хромосомным блоком.

Сравнение относительного состава повторённых последовательностей хромосом исследуемых видов показало, что процентное соотношение основных семейств повторов сравнимо с таковым в геномной базе данных Repeat Masker. На половых хромосомах каролинского анолиса снижается относительное количество LINE (с 30% до 20%) и SINE (с 15% до 10%) элементов и значительно повышается количество LTR GYPSY элементов (с 10% до 19%). Схожая картина наблюдается и для коричневого анолиса, только повышается количество LTR DIRS и LTR ERV1 элементов. У плоскохвостого домового геккона на Z хромосоме отмечается увеличение LTR GYPSY и LTR COPIA элементов.

Из анализа данных из этой работы, из других совместных проектов и литературных данных мы заметили, что в различных таксонах чешуйчатых в формировании половых хромосом независимо вовлекаются одни и те же синтенные блоки. Группы сцепления гомологичные аутосомам каролинского 9, 16 и 18 участвуют в формировании половых хромосом в четырех семействах чешуйчатых. Так X хромосома коричневого анолиса *Anolis sagrei* содержит блоки 9, 12 и 18, Z хромосома плоскохвостого домового геккона *Hemidactylus platyurus* содержит блок 9, Z хромосома бородатой агамы *Pogona vitticeps* содержит блоки 9 и 16, а Z хромосома комодского варана *Varanus komodoensis* содержит блоки 16 и 18. Сравнение с другими представителями чешуйчатых, а также более тщательное изучение самих синтенных блоков может объяснить важность этих групп сцепления в эволюции половых хромосом отряда.

Выводы

1. Ранее разработанный метод биоинформатического анализа данных высокопроизводительного секвенирования отдельных хромосом адаптирован под неполные сборки референсного генома.

2. Применение данного метода к данным Illumina MiSeq секвенирования библиотек сортированных хромосом каролинского анолиса (*Anolis carolinensis*) позволило дополнить сборку референсного генома, определив состав всех микрохромосом этого вида. Всего по 12 микрохромосомам было распределено 419 скэффолдов общим размером в 189 млн. п.н.

3. Половой хромосомой у каролинского анолиса является хромосома 13. Половая хромосома коричневого анолиса (*Anolis sagrei*) состоит из блоков синтенных хромосомам X, 9, 12 и 18 каролинского анолиса. Z хромосома плоскохвостого домового геккона (*Hemidactylus platyurus*) гомологична хромосоме 9 каролинского анолиса.

4. Однозначного кандидата на роль ключевого гена каскада определения пола у исследуемых видов не было обнаружено. На X хромосоме коричневого

анолиса обнаружены гены NR0B1, SOX8 и RSPO1 из половго каскада. На Z хромосоме плоскохвостого домового геккона обнаружены гены NR0B1 и RSPO1.

5. На половых хромосомах чешуйчатых было замечено обогащение повторов семейства LTR. Так у каролинского анолиса повышено содержание LTR Gypsy на X хромосоме, у коричневого анолиса повышено содержание LTR DIRS и LTR ERV1 на половых хромосомах, а у плоскохвостого домового геккона повышено содержание LTR Gypsy и LTR Copia элементов на Z хромосоме.

6. В формирование половых хромосом различных семейств чешуйчатых независимо вовлекаются одни и те же синтенные блоки.

Публикации по теме научного доклада

1. **Kichigin IG**, Giovannotti M, Makunin AI, et al (2016) Evolutionary dynamics of Anolis sex chromosomes revealed by sequencing of flow sorting-derived microchromosome-specific DNA. *Mol Genet Genomics* 291:1955–1966.

2. **Kichigin IG**, Lisachov AP, Giovannotti M, et al (2018) First report on B chromosome content in a reptilian species: the case of *Anolis carolinensis*. *Mol Genet Genomics*.

3. **Kichigin IG**, Trifonov VA (2013) Genomic structure and sex determination in squamate reptiles. *Tsitologiya* 55:253–258

4. Makunin AI, **Kichigin IG**, Larkin DM, et al (2016) Contrasting origin of B chromosomes in two cervids (Siberian roe deer and grey brocket deer) unravelled by chromosome-specific DNA sequencing. *BMC Genomics* 17:618.

5. Giovannotti M, Trifonov VA, Paoletti A, **Kichigin IG**, et al (2017) New insights into sex chromosome evolution in anole lizards (Reptilia, Dactyloidae). *Chromosoma* 126:245–260.

Отчет о проверке текста
научно-квалификационной работы
на объем заимствования

Кичигин Илья Геннадьевич

«Изучение состава и эволюции половых хромосом чешуйчатых»

Оценка оригинальности: 90.59%

Заимствования: 9.41%

№ документа: 995658

Имя исходного файла: Научный_доклад_Кичигин.docx

Размер файла: 27.2 КБ

Размер текста: 16648

Слов в тексте: 2383

Число предложений: 210

Источники (доля в тексте – ссылка):

10.9% <http://www.list-org.com/company/814678>

10.9% <https://www.rusprofile.ru/id/1451999>

5.0% <http://zna4enie.ru/opredelenie/programmnyj-tip-opredelenija-pola.html>

8.6%

<https://ru.wikipedia.org/wiki/%D0%9E%D0%BF%D1%80%D0%B5%D0%B4%D0%B5%D0%BB%D0%B5%D0%BD%...>

7.3% <http://www.kinetics.nsc.ru/CMS/index.php?id=153>

21% <http://www.dslib.net/virusologia/analiz-geneticheskogo-sostava-v-hromosom-mlekoopitaju...>

5.2% <http://lib.knigi-x.ru/23biologiya/275153-1-sravnitelnyj-analiz-strukturi-sinapsisa-re...>

25.6% <http://www.mcb.nsc.ru/laboratory/lcg>

18.5% <http://reptile-database.reptarium.cz/species?genus=Aolis&species=carolinensis>

14.6% <http://www.mcb.nsc.ru/en/laboratory/457>

Оригинальность работы составляет 90.59 % , что соответствует требованиям порядка и условиям допуска научно-квалификационных работ к защите на итоговом заседании Государственной итоговой аттестации в аспирантуре ИХБФМ СО РАН.

Отчет предоставлен сервисом «Антиплагиат» на сайте www.antiplagius.ru

06.09.2018 г.

Проверку выполнила секретарь ГЭК - к.б.н. К.В. Разум