

На правах рукописи

СОЛОВЬЕВА
Анна Ивановна

**МОБИЛЬНЫЕ ЭЛЕМЕНТЫ УЧАСТВУЮТ В ОБРАЗОВАНИИ
КЛОНАЛЬНОЙ ИЗМЕНЧИВОСТИ *HIMASTHLA ELONGATA* (TREMATODA,
HIMASTHLIDAE)**

03.03.04 - Клеточная биология, цитология, гистология.

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата биологических наук

Санкт-Петербург
2017

Работа выполнена в Федеральном государственном бюджетном учреждении науки
Институте цитологии Российской академии наук

Научный руководитель: доктор биологических наук, профессор
Подгорная Ольга Игоревна
ведущий научный сотрудник, зав. группой
некодирующей ДНК лаборатории морфологии клетки
Федерального государственного бюджетного
учреждения науки Института цитологии Российской
академии наук, г. Санкт-Петербург

**Официальные
оппоненты:** **Цымбаленко Надежда Васильевна**, доктор
биологических наук, профессор, ведущий научный
сотрудник отдела молекулярной генетики Федерального
государственного бюджетного научного учреждения
«Института экспериментальной медицины»
г. Санкт-Петербург
Родионов Александр Викентьевич, доктор
биологических наук, профессор, зав. лабораторией
биосистематики и цитологии Федерального
государственного бюджетного учреждения науки
Ботанический институт им. В.Л. Комарова Российской
академии наук, г. Санкт-Петербург

Ведущая организация: Федеральное государственное бюджетное учреждение
науки Институт молекулярной и клеточной биологии
Сибирского отделения Российской академии наук

Защита диссертации состоится 21 апреля 2017г. в 12 часов на заседании
Диссертационного совета Д 002.230.01 на базе Института цитологии РАН по адресу:
194064, Санкт-Петербург, Тихорецкий проспект, д.4.

Адрес электронной почты института: cellbio@mail.cytspb.rssi.ru

Сайт института: <http://www.cytspb.rssi.ru>

Факс: 8 (812) 297-35-41

С диссертацией можно ознакомиться в библиотеке Института цитологии РАН и на
сайте института по адресу <http://www.cytspb.rssi.ru>.

Автореферат разослан «__» _____ 2017 г.

Учёный секретарь диссертационного совета,
кандидат биологических наук

Е.В. Каминская

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы

В современной биологии накопление данных существенно опережает их обработку и осмысление. На сегодняшний день количество прочитанных геномов эукариот приближается к десяти тысячам, а проаннотировано чуть более 300 (National Center for Biotechnology Information, NCBI). Эукариотические геномы значительно различаются по размерам и структуре, но одним из общих и характерных их свойств является многообразие и обилие повторяющихся последовательностей (Tollis and Boissinot, 2012). Повторы традиционно принято делить на *диспергированные* и *тандемные* в зависимости от их структуры и организации в геноме. К диспергированным повторам относят мобильные элементы, или транспозоны (Transposable elements, TE) — это фрагменты ДНК, способные к перемещению и увеличению числа копий в геноме. На сегодняшний день нет устоявшейся классификации мобильных элементов, и в настоящей работе использована одна из существующих (Wicker et al., 2007). На основании структуры и механизмов перемещения мобильные элементы делят на два класса, которые состоят из отрядов, суперсемейств и семейств. К классу I относят ретроэлементы. Они используют механизм перемещения «копировать - вставить» и увеличивают число копий в геноме за счет обратной транскрипции посредством РНК-интермедиата. Ретротранспозоны можно условно разделить на элементы, содержащие длинные концевые повторы (Long Terminal Repeats, LTR), и элементы, не содержащие длинные концевые повторы — non-LTR элементы. Наиболее известными отрядами non-LTR элементов являются LINEs (Long Interspersed Nuclear Elements) и SINEs (Short Interspersed Nuclear Elements).

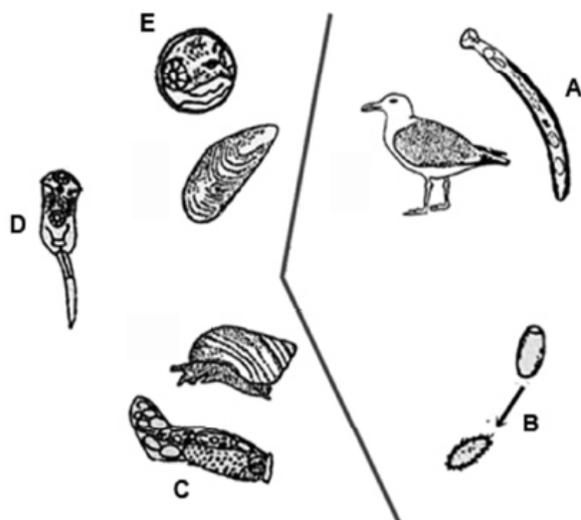


Рис.1. Жизненный цикл *H. elongata* по Werding, 1969. А. Марита, В. Мирацидий, С. Редия, Д. Церкария, Е. Метациркария.

Составляя от 3 до 80 % генома у разных видов (Brown, 2002), транспозоны не могут не участвовать в клеточных процессах. Подавляющее большинство мобильных элементов неактивны, но их копии представляют собой гомологичные последовательности ДНК, которые могут участвовать в нерегулярной или «неразрешенной» рекомбинации, называемой эктопической. Транспозоны также влияют на экспрессию и структуру генов, являются мишенями для малых РНК, эпигенетических модификаций, участвуют в организации структуры хроматина и ответственны за синтении (Abrusan and Krambeck, 2006; Slotkin and Martienssen, 2007). Таким образом,

мобильные элементы часто рассматриваются как один из основных факторов нестабильности генома. Перемещения TE выявлены как в половых, так и соматических клетках (Kazazian, 2011). Мозаичность распределения ретропозона LINE-1 в геномах соматических клеток человека (Muotri et al., 2005) и транспозона

Tc1 в клетках *Caenorhabditis elegans* дают основание говорить о том, что TE способны обеспечивать генетическую изменчивость, не сцепленную с половым процессом. Подходящей моделью для таких исследований являются представители класса Trematoda — одной из наиболее успешных и широко распространенных групп паразитических плоских червей. Трематоды обладают сложным жизненным циклом, который протекает с чередованием гермафродитного и партеногенетических поколений в нескольких животных-хозяевах. Объектом настоящей работы является *Himasthla elongata* (Mehlis, 1831, рис. 1) из семейства Himasthliidae (Tkach et al., 2016). Долгое время считали, что особи партеногенетического поколения трематод — редии или дочерние спороцисты, происходящие от одной материнской особи и обитающие в первом промежуточном хозяине, моллюске, равно как и производимые ими расселительные личинки, церкарии, представляют собой генетически идентичные клоны, так как являются продуктом диплоидного (апомиктического) партеногенеза. Тем не менее, на ряде видов трематод с помощью молекулярных методов показали, что у особей партеногенетического поколения имеет место генетический полиморфизм — клональная изменчивость (Grevelding, 1999; Халтурин и др., 2000; Семенова и др. 2005). Остается неясным, какие последовательности генома отвечают за возникновение этой изменчивости, но, согласно некоторым гипотезам, ее источником могут быть мобильные элементы (Семенова и др., 2005; Корсуненко и др., 2009).

Цель и задачи исследования

Цель — выявить изменчивость в геноме партенит *H. elongata* методом S-SAP и охарактеризовать изменяющиеся последовательности.

Задачи:

1. Провести генотипирование партенит *H. elongata* методом S-SAP и охарактеризовать последовательности из полиморфных и консервативных зон
2. Определить распределение клонированных фрагментов в паттернах генотипирования методом дот-блота.
3. Определить наличие или отсутствие последовательностей клонированных фрагментов мобильных элементов в транскриптом *H. elongata*.
4. Изучить распределение найденных клонированных фрагментов в геноме *H. elongata* методом FISH.

Основные положения, выносимые на защиту:

1. Клональная изменчивость (КИ) показана методами S-SAP и дот-гибридизации на одиночных церкариях трематоды *H. elongata*. В консервативных зонах располагаются LINE-подобные ретроэлементы, в полиморфных зонах присутствуют спейсерные участки LINEs, а также фрагменты LTR-ретротранспозонов.
2. Обе группы основанных на LINE фрагментов активно транскрибируются и, следовательно, вносят вклад в КИ.
3. Кариотип *H. elongata* содержит 12 пар хромосом ($2n=24$), среди которых 3 пары несут ядрышковые организаторы.
4. Флуоресцентная *in situ* гибридизация (FISH) обнаружила дисперсное распределение тех фрагментов LINE, которые основаны на открытой рамке считывания обратной транскриптазы (RT) и обогащение конститутивного гетерохроматина тем фрагментом, который основан на спейсерных участках LINE.

Научная новизна работы

Впервые получены данные о кариотипе *H.elongata*, выявлен полиморфизм в морфологии хромосом и в их окраске дифференциальными красителями. До настоящего момента кариологические данные какого-либо из видов рода *Himasthla* отсутствовали. Доказана пригодность метода S-SAP для генотипирования одной личинки, имеющей микроскопические размеры. Подобная техника может применяться и для других объектов с малым количеством материала. Проведена характеристика последовательностей из консервативных и полиморфных зон в полученных методом S-SAP паттернах генотипирования церкарий. Впервые в геноме *H.elongata* выявлены фрагменты ретроэлементов семейств CR1, RTE, BEL, copia и, таким образом, определены потенциальные носители клональной изменчивости. Проведен анализ последовательностей, составляющих картины генотипирования партеногенетических личинок. Впервые показано, что фрагменты кодирующей части CR1 ретроэлемента (non-LTR) и некодирующей части элемента представлены в транскриптом в значительных количествах, но в разных транскриптах. Локализация кодирующей и некодирующей частей CR1 (non-LTR) впервые показана в разных частях хромосом — факультативном или конститутивном гетерохроматине. Конкретные данные настоящей работы являются вкладом в выявление связи повторяющихся последовательностей и некодирующей РНК, определение их вклада в упаковку генома, его нестабильность и эволюцию.

Теоретическое и практическое значение работы

Работа имеет фундаментальную направленность и расширяет представления о роли мобильных элементов в геноме. Результаты работы важны для понимания механизмов появления полиморфизма при отсутствии полового процесса. Объект исследования принадлежит к классу Trematoda, среди представителей которого много возбудителей опасных заболеваний, а клональная изменчивость трематод служит препятствием для разработки противопаразитарных средств. Вид *H. elongata* безопасен для человека и является удобной моделью для изучения механизмов клональной изменчивости. Материалы диссертации используются в курсах лекций для бакалавров и магистров Биологического факультета Санкт-Петербургского государственного университета и могут быть использованы в общих и специальных курсах лекций биологических факультетов других университетов.

Апробация работы

По теме диссертации опубликовано 12 работ, из них 3 — статьи. Основные положения представлены и обсуждены на XII сессии Морской биологической станции СПбГУ (2011), III и V конференциях молодых ученых Института цитологии РАН (2012, 2016 гг); 16-й и 19-й Пущинской Международной школе-конференции молодых ученых (2012, 2015 гг); Международной научной конференции студентов, аспирантов и молодых учёных «Ломоносов-2013» (2013); Международной конференции «Хромосома 2015» (2015); 49th European Marine Biology Symposium (2014); EMBL Symposium: The Mobile genome: Genetic and Physiological impacts of transposable elements (2015).

Вклад автора

Автор провел: анализ литературы по теме работы, планирование экспериментов, получение основной части результатов, написание статей и

подготовку докладов на конференциях, сбор животных, выделение ДНК, проведение всех этапов S-SAP, дот-гибридизации, анализ клонированных последовательностей, подготовку проб для измерения генома и кариотипирования. Анализ кариотипа выполнен совместно с Деминым С.Ю. Анализ транскриптов и оценка размера генома выполнены совместно с Галактионовым Н. К. Эксперименты по проточной цитофлуорометрии проведены в лаборатории внутриклеточной сигнализации под руководством Аксенова Н. Д. Материалы, вошедшие в диссертацию, обсуждались и публиковались совместно с соавторами и научным руководителем.

Финансовая поддержка работы

Работа выполнена при поддержке РФФИ (гранты № 05-04-49156-а, 11-04-01700, 15-04-01857, 16-34-00603 мол_а, 17-04-02161), РНФ (гранты 15-15-20026, 14-14-00621) гранта Президиума академии наук по программе «Клеточная и молекулярная биология» (рук-ль Боголюбов Д.С.), гранта Министерства образования и науки РФ № 6.1278.2014/К (2014-2016 гг).

Объем и структура диссертации

Диссертация содержит разделы «Введение», «Обзор литературы», «Материалы и методы», «Результаты», «Обсуждение результатов», «Выводы», «Список цитируемой литературы», содержащий 141 наименование, «Приложение 1», «Приложение 2», изложена на 116 страницах и проиллюстрирована 65 рисунками (включая приложения) и 5 таблицами.

МАТЕРИАЛЫ И МЕТОДЫ

Сбор животных и выделение ДНК

Сбор и первичную обработку материала проводили на Беломорской Биологической Станции Зоологического института РАН «Картеш» в июле-августе 2013-2015 гг. Источником ДНК паразитов служили церкарии *Himasthla elongata* (Mehlis, 1831) (Trematoda; Himasthlidae). Церкарий получали путем эмиссии из первого промежуточного хозяина — моллюсков *Littorina littorea*. Личинок,

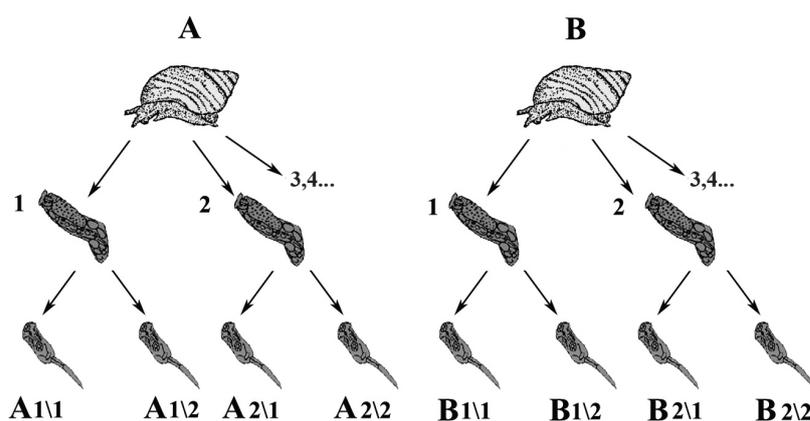


Рис.2. Номенклатура церкарий из разных клональных популяций (А, В).

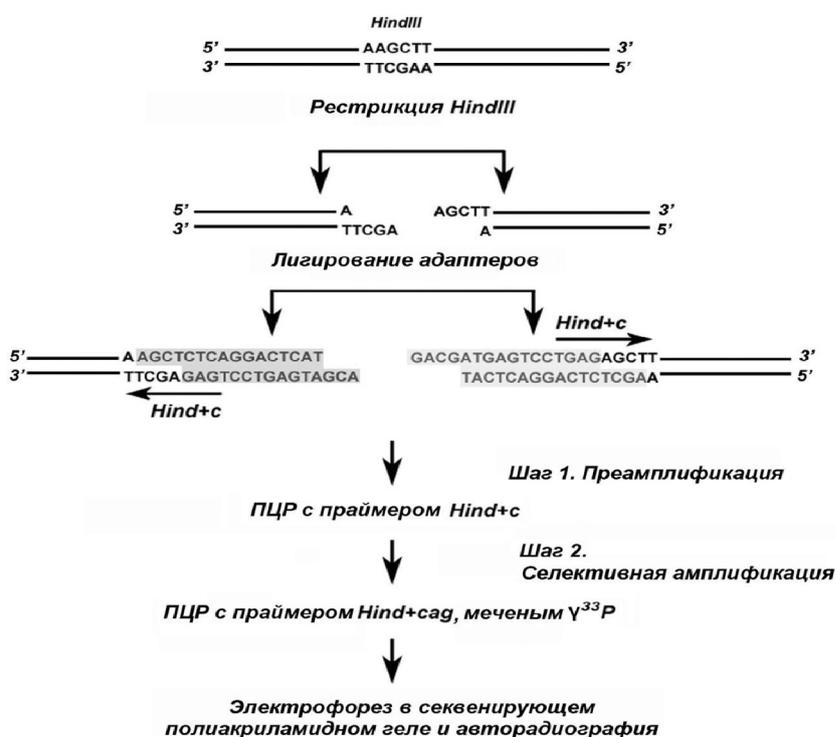
выделенных из одного моллюска, считали потомками одного мирацидия и рассматривали как принадлежащих одной клональной популяции или пула. ДНК отдельных церкарий выделяли с применением детергента СТАВ (Winneperenninckx et al., 1993). Образцы для S-SAP пронумерованы согласно схеме на рис. 2.

Определение размера генома *H. elongata*

Определение размера генома *H. elongata* провели методом проточной цитофлуориметрии. Источником диплоидных ядер трематоды стали редии, в качестве стандарта использовали спленоциты *M. musculus*, так как геном мыши заведомо больше генома червя. Подготовку проб проводили согласно опубликованной методике (Hare and Johnston, 2011) с изменениями (Galaktionov, Solovyeva et al., 2014; Solovyeva et al., 2016).

Метод S-SAP

Порядок проведения S-SAP показан на рис.3. За основу взяты опубликованные методы (Vos et al., 1995, Waugh et al., 1997, Drew and Brindley, 1995). Подробное описание обработки образцов описано в статье (Соловьева и др., 2013). Список использованных олигонуклеотидов - адаптеров для лигирования и праймеров для ПЦР приведен в таблице 1. Адаптеры необходимы для повышения специфичности отжига праймеров.



Преамплификацию проводили с праймером *HindIII+c*. Селективную амплификацию проводили с праймером *HindIII+cag*, помеченным по 5'-концу $\gamma^{33}P$ (ОАО «ИРМ») по протоколу «Силекс-М» (Соловьева и др., 2013). Использовали программу ПЦР с градиентным понижением температуры отжига праймеров (touch-down PCR).

Рис.3. Схема проведения S-SAP.

Таблица 1. Список использованных олигонуклеотидов.

№	Название	Нуклеотидная последовательность
1	Адаптер 1	5'-AGCTCTCAGGACTCAT-3'
2	Адаптер 2	5'-AGCTCTCAGGACTCAT-3'
3	<i>HindIII+c</i>	5'-GAGTCCTGAGAGCTTC-3'
4	<i>HindIII+cag</i>	5'-GAGTCCTGAGAGCTTCAG-3'
5	7_5OuF	5'-ATTAGGCAAAGTCGAACCAAC-3'
6	7_5OuR	5'-AAAGGTTAGCCCTACGAAAGT-3'

Электрофорез в секвенирующем акриламидном и агарозном гелях

Для разделения продуктов ПЦР использовали 5 %-ный денатурирующий акриламидный секвенирующий гель. Электрофорез проводили в течение 3 ч с напряжением 10 В/см, по завершении электрофореза гель высушивали и экспонировали 7 сут при - 80 °С (Sambrook et al., 1989). Визуализацию первого шага S-SAP, разделение продуктов реамплификации и клонирования проводили методом электрофореза в 1 %-ном агарозном геле в 1× TAE, pH 7,6 (Sambrook et al., 1989).

Клонирование консервативных и переменных фрагментов и выявление фланкирующих последовательностей для 7_5

Фрагменты B1, B1', A1 и A1' продуктов первого шага S-SAP вырезаны из геля после разделения методом агарозного электрофореза, реамплифицированы с праймером HindIII (рис. 3. шаг 1) и клонированы в вектор pTZ57r/t (Termofisher scientific) по стандартному протоколу (Sambrook et al., 1989). Наборы консервативных и полиморфных фрагментов (1-15) вырезали из сухого геля после полного S-SAP по экспонированной пленке (рис. 7), реамплифицировали в соответствующих условиях с праймером HindIII+cag и также клонировали в векторы pTZ57r/t и pJet2.1 (Termofisher scientific) по стандартному протоколу (Sambrook et al., 1989). Обозначения плазмид в клонотеке проводили следующим образом: для фрагментов из зон 1-15 первая цифра означает зону, которую вырезали, а вторая — № плазмиды. Амплификацию с целью выявления фланкирующих последовательностей для фрагмента 7_5 проводили с праймерами 7_5OuF и 7_5OuR в течение 25 циклов по следующей программе: денатурация 45 с при 95 °С, отжиг 45 с при 53 °С, синтез 1 мин при 72 °С. В качестве отрицательного контроля ПЦР использовали реакционную смесь без добавления матрицы. Зоны 1000-1200 п. н. и зоны 300-500 п. н. вырезаны и клонированы в вектор pTZ57r/t (Termofisher scientific) по стандартному протоколу (Sambrook et al., 1989). Вставки плазмид, получившихся при клонировании, амплифицировали и секвенировали в фирме ЗАО «Евроген» (Москва). Всего секвенировано 36 клонированных продуктов S-SAP (таблица 2) и два фланкирующих 7_5 фрагмента.

Дот-гибридизация

Для дот-гибридизации 50 нг денатурированных в 50 %-ном формамиде ПЦР-продуктов с соответствующих плазмид (таблица 2) и зонда M9 — фрагмент Nema1 (JX470542.1) — наносили на нейлоновые мембраны Hybond N+ (Amersham) согласно схеме (рис. 4). Кроме последовательностей, заведомо содержащих участки мобильных элементов, на мембраны наносили также последовательности с неидентифицированными вставками — 7_5, 7_14, 2_1, а также последовательности 18S рДНК (18S) и гена гельзолина (2_3). Отрицательным контролем служила геномная ДНК *E.coli*. Зонды для дот-блотов готовили следующим образом (рис. 4): ПЦР-продукты S-SAP (в варианте с HindIII и с добавлением EcoRI) разделяли в 5 %-ном ПААГ, выявляли путем окрашивания геля бромистым этидием, а затем вырезали наборы полос в разных диапазонах длин (800-1000 п. н., 500-700 п. н., 300-500 п. н.). Фрагменты ДНК выделяли из геля и метили biotin-11-dUTP (ДНК-Синтез) с помощью ПЦР с соответствующими праймерами. Гибридизацию проводили по стандартному протоколу (Wahl, 1979) с изменениями. Детекцию сигналов проводили с использованием диаминобензидина (Sigma) по протоколу производителя (Leary et al., 1983). В результате можно оценить приблизительное количество отдельных

фрагментов среди консервативных или вариабельных зон в паттернах S-SAP (рис. 7). Всего сделали 17 дот-блотов с зондами из продуктов S-SAP разных личинок.

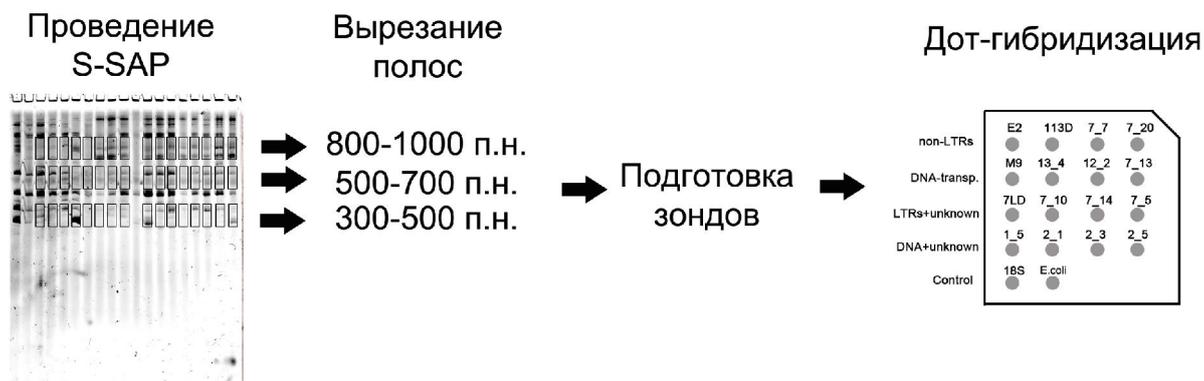


Рис.4. Схема эксперимента по проведению дот-гибридизации.

Флуоресцентная гибридизация *in situ* (FISH)

Условия для получения метафазных хромосом подбирали с учетом существующих в литературе данных (Richard and Voltz, 1987; Birstein and Mikhailova, 1989; Mutafova et al., 1991; Mutafova, 1994; Hirai, 2004). Источником метафазных хромосом служили редии *H. elongata*, изолированные при вскрытии зараженных моллюсков *L. littorea*. Препараты готовили согласно опубликованной методике (Solovyeva et al., 2016). Зонды для гибридизации *in situ* метили biotin-dUTP с помощью ПЦР в соответствующих условиях. Предобработку препаратов проводили по опубликованной методике с изменениями (Khodyuchenko et al., 2012). Гибридизацию проводили по стандартному протоколу Rouché (Nonradioactive *In Situ* Hybridization Application Manual), биотин выявляли конъюгатами streptavidin-Alexa488, streptavidin-Alexa594, ядра и хромосомы контрастировали DAPI. Препараты анализировали с помощью эпифлуоресцентных микроскопов DMI3000B Leica Microsystems GmbH в ЦКП «Хромас» и DMI6000B Leica Microsystems GmbH в ресурсном центре «Развитие молекулярных и клеточных технологий» СПбГУ. Для компьютерной обработки изображений использовали пакет программ Adobe Photoshop™ 4.0.

Секвенирование транскриптома *H. elongata*

Этот этап работы проведен А.М. Горбушиным (Лаборатория сравнительной физиологии сенсорных систем, ИЭФБ РАН). Секвенирование транскриптома редий *H. elongata* осуществили на платформе Illumina HiSeq в Ресурсном центре «Биобанк» СПбГУ с использованием набора реагентов TruSeq Stranded mRNA LT Sample Prep Kit (Illumina). На основе полученных 133 млн коротких (100 п.н.) парных фрагментов, собран референсный транскриптом, обогащенный poly-A последовательностями. Фильтрация исходных фрагментов, сборка транскриптома и аннотация кодирующих последовательностей проведена с использованием программного пакета Trinity (Grabherr et al., 2011). Оценка уровня копийности каждого транскрипта проведена *in silico* в программе Kallisto (Bray et al., 2015). Повторы из транскриптома редий *H. elongata* выявляли посредством поиска в программе BLAST (Altschul et al., 1990) по базе RepBase. Затем выявленные в ходе поиска транскрипты были переданы нашей группе для анализа.

Компьютерные методы анализа

Анализ параметров последовательностей адаптеров и праймеров осуществляли с использованием программы FastPCR 6.1 (Kalendar et al., 2009). Для сравнения фрагментов ДНК использовали программы BLAST (Altschul et al., 1990), RepeatMasker (Kohany et al., 2006), Tandem Repeat Finder (Benson, 1999) и ресурса SINE Base (Vassetzky and Kramerov, 2013). Сравнительный анализ результатов дот-блот гибридизации проводили на основании измерений интенсивности проявленных сигналов с помощью программы Image J (<http://rsb.info.nih.gov/ij/>). Поиск мобильных элементов из транскриптома, соответствующих клонированным фрагментам *H. elongata* проводили в программах BioEdit (Hall, 1999) и Usearch (Edgar, 2012) по алгоритму BLAST с ограничением E-value $1e-40$. Поиск консервативных доменов в последовательностях выполнен с помощью онлайн-сервиса NCBI Conserved Domains (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>).

РЕЗУЛЬТАТЫ

Размер генома *H. elongata*

Геном *H. elongata* не секвенирован. Мы провели оценку его размера методом флуоресцентной проточной цитометрии. У партенит *H. elongata* невозможно выделить гаметы, поэтому сравнение проводили по диплоидным ядрам, которые получали из целых редий; в качестве стандарта взяли ядра, выделенные в аналогичных условиях из спленоцитов *Mus musculus*.

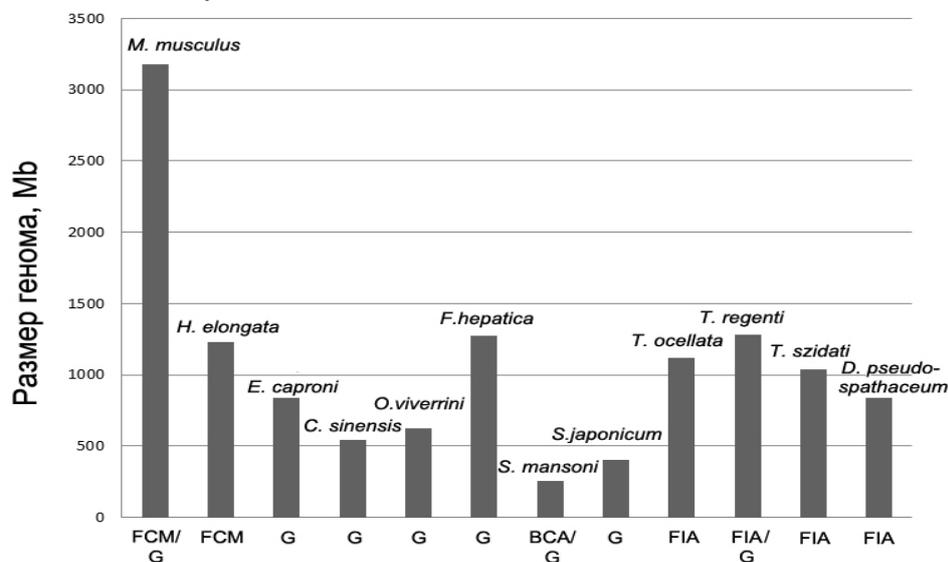


Рис.5. Соотношение размера генома *H. elongata* по сравнению с геномами мыши (PRJNA20689) и представителей класса Digenea. По горизонтали указан метод определения содержания ДНК для каждого из видов: FCM — проточная цитофлуориметрия; BCA — биохимический анализ (Simpson et al., 1982); FIA — денситометрия изображений после окраски препаратов по Фельгену (www.genomesize.com/); G — секвенирование генома (<http://parasite.wormbase.org/>).

В результате анализа ~10 тыс. ядер показали, что гаплоидный геном *H. elongata* содержит ~1.25 пг ДНК или приблизительно 1.227×10^9 п. н. (рис.5; Galaktionov, Solovyeva et al., 2014, Solovyeva et al., 2016). В сравнении с геномами других

трематод, по своим размерам геном *H. elongata* сопоставим с геномами *Fasciola hepatica* и *Trichobilharzia regenti*.

Генотипирование церкарий методом S-SAP: анализ консервативных и полиморфных фрагментов

В качестве основной эндонуклеазы рестрикции для S-SAP использовали HindIII. Такой выбор обусловлен тем, что сайты узнавания HindIII (5'- A↓AGCTT- 3') относительно редки и расположены, как правило, в АТ-богатых областях генома, обедненных генами, что позволяет ожидать формирования картин генотипирования церкарий преимущественно повторами. Известно, что именно такие районы обогащены LINEs (Sadhu and Gill, 2002; Федоров, 2008; Томилин, 2009).

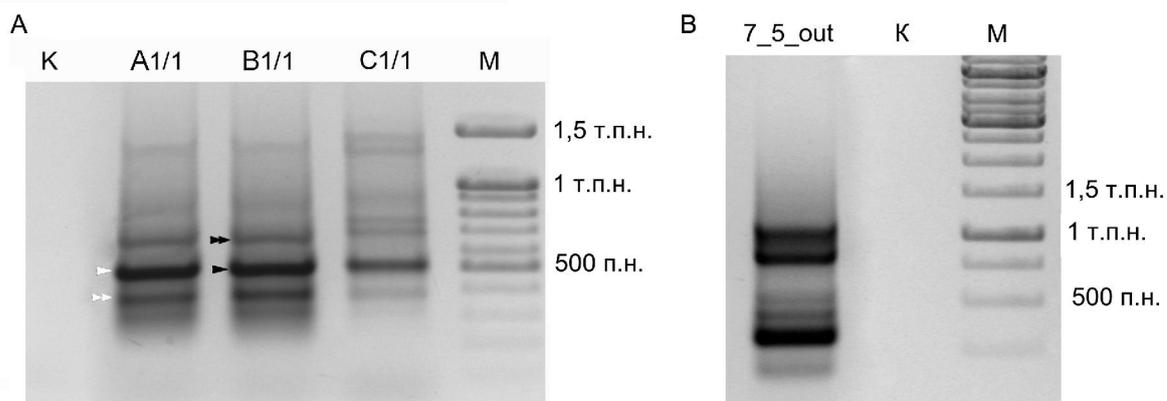


Рис.6. А. Электрофореграмма S-SAP геномной ДНК трех церкарий из клональных популяций А, В, С в 1%-ном агарозном геле. Стрелками обозначены выбранные для клонирования зоны (черная стрелка — В, двойная черная стрелка — В1', белая стрелка — А1, двойная белая стрелка — А1'). В. Результаты амплификации фланкирующих 7_5 последовательностей в 1%-ном агарозном геле. Обозначения маркеров длин фрагментов: п. н. — пара нуклеотидов, т. п. н. — тысяча пар нуклеотидов.

Методом электрофореза в агарозном геле обнаружили различия в картинах разделения (паттернах) ПЦР - фрагментов у церкарий из разных моллюсков, то есть из разных клонов (рис.6, А). Консервативная зона длиной ~ 500 п. н., а также зоны длиной ~ 400 п. н. и ~ 650 п.н. видны во всех случаях (Соловьева и др., 2013). Продукты длиной более 700 п.н. имеют недостаточное разрешение в агарозе, поэтому на этом этапе S-SAP провели анализ консервативных полос. Фрагменты В1, В2, В1', А1 и А1' секвенировали, их характеристики приведены в таблице 2. Для повышения чувствительности метода на втором шаге S-SAP (селективной ПЦР) в праймер Hind+cag вводили изотоп фосфора-33, а для разделения продуктов амплификации применяли электрофорез в денатурирующем полиакриламидном геле. На рис.7 представлены результаты генотипирования редии (1) и 15 церкарий из клональных популяций А, В, С. Оказалось, что в геномах всех личинок присутствуют как консервативные, так и полиморфные фрагменты. При этом положение консервативных фрагментов неизменно у особей как внутри одного клона, так и при сравнении личинок из разных моллюсков. Наличие консервативных фрагментов и сходные паттерны генотипирования церкарий свидетельствуют о надежной воспроизводимости результатов S-SAP. Для клонирования отобрали наборы фрагментов разной длины (рамки на рис.7). Успешную реамплификацию прошли все вырезанные зоны, кроме зон 4-6. Клонированные последовательности изучали по

нескольким параметрам: 1) проводили анализ на наличие фрагментов повторов; 2) проводили поиск консервативных доменов; 3) изучали распределение в картине продуктов S-SAP; 4) изучали распределение в геноме *H. elongata*; 5) проводили поиск сходных последовательностей в транскриптоме *H. elongata*.

Тандемные повторы не выявили ни в одной из клонированных последовательностей. По результатам поиска программой RepeatMasker из 36 клонированных фрагментов 20 последовательностей имеют сходство с мобильными элементами (таблица 2); в некоторых обнаружены консервативные домены обратных транскриптаз (A1, A1', B1, B1', B2, 7_7) и интегразы (8_1, 15_1). В 15 клонированных фрагментах по результатам поиска в RepBase не выявлено каких-либо последовательностей мобильных элементов, но в отдельных (2_3, 3_5) обнаружены не относящиеся к транспозонам домены. Все неизвестные фрагменты подвергли поиску по базе SINE Base (Vassetsky and Kramerov, 2013). В результате в последовательностях 1_2, 1_5, 2_1, 2_5, 3_10, 7_14, 9_2, 14_1, 15_2 и 15_3 обнаружили фрагменты SINE элементов. Аналогичные результаты получили при анализе последовательностей 12_2, 13_1 и 13_4, в которых фрагменты SINE обнаруживаются при поиске RepeatMasker, а также последовательности 12_1. Во фракции повторов транскриптома найдены 25 клонированных последовательностей. Фрагменты non-LTR ретроэлементов обнаружены в консервативных и полиморфных зонах выше 300 п. н., а фрагменты LTR-содержащих ретроэлементов — в зонах более 500 п. н. (15_1, 7_9, 7_10 и 12_1).

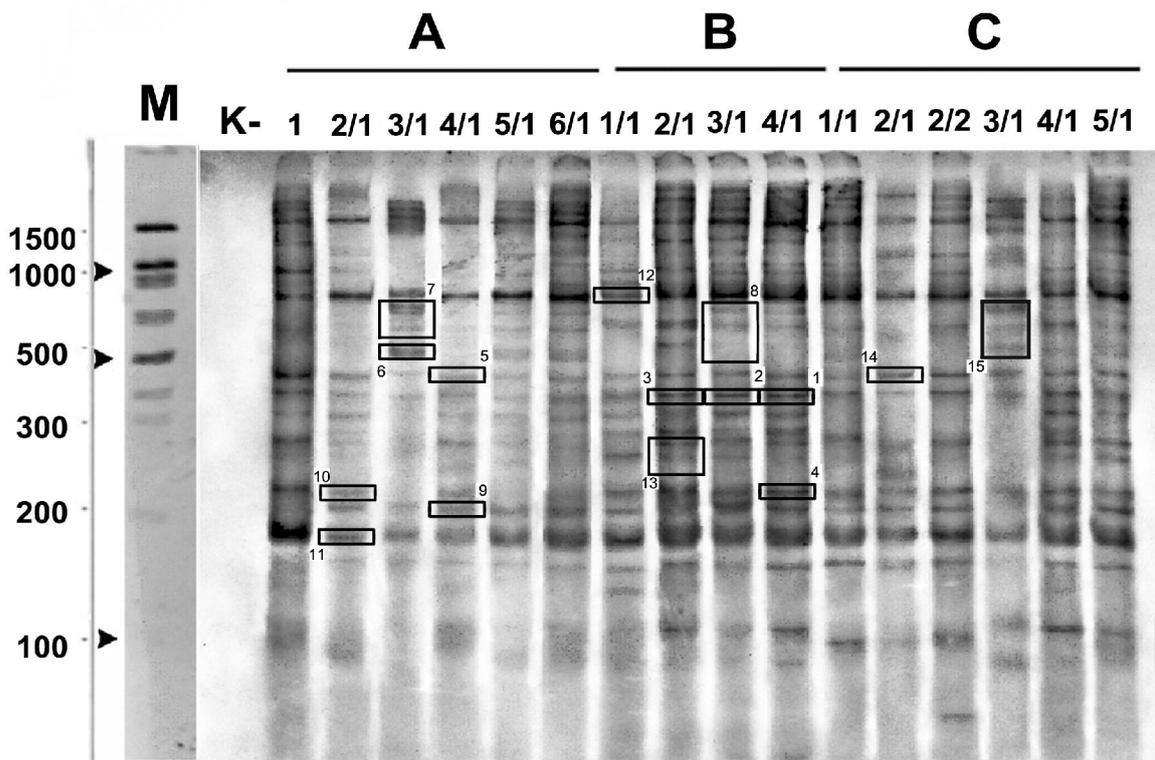


Рис.7. Авторадиограмма результатов S-SAP, поставленного с редией (1) и церкариями (1\1,2\1...) из клональных популяций А, В, С. Каждая дорожка соответствует распределению продуктов S-SAP отдельной личинки. М — маркер молекулярной длины, цифры слева — обозначения длин фрагментов в п. н.; линии справа — области наиболее варибельных фрагментов. Рамки обозначают вырезанные наборы зон; для каждой рамки указано цифровое обозначение.

Среди обнаруженных non-LTR ретроэлементов подавляющее большинство имеет гомологию с транспозонами семейства CR1.

Таблица 2. Анализ секвенированных фрагментов.

№	Фргм	Длина п. н.	RBase	Класс	Sim	Зона	СКД	T	DOT
1	A1	545	CR1-25_SP	Non-LTR /CR1	0,695	K	RT	+	-
2	A1'	477	RTEX-14_ACar	Non-LTR /RTEX	0,750	K	RT	+	+
3	B1	529	CR1-25_SP	Non-LTR /CR1	0,621	K	RT	+	+
4	B2	548	CR1-X1_Pass, CR1-6_CTe	Non-LTR /CR1	0,719, 0,671	K	RT	+	-
5	B1'	525	CR1-X1_Pass, CR1-23_HM	Non-LTR /CR1	0,724, 0,674	K	RT	+	-
6	1_1	564	-	-	-	K	-	+	-
7	1_2	530	-	-	-	K	-	+	-
8	1_5	551	MuDR-1_ROr	DNA/MuDR	0,736	K	-	+	+
9	2_1	570	-	-	-	K	-	-	+
10	2_3	567	Daphne-3_HMM	Non-LTR /Daphne	0,796	K	ADF_ gelsolin	-	+
11	2_5	549	MuDR-1_ROr	DNA/MuDR	0,717	K	-	+	+
12	3_4	348	-	-	-	K	-	+	-
13	3_5	220	-	-	-	K	MFS	-	-
14	3_8	261	-	-	-	K	-	+	-
15	3_10	569	-	-	-	K	-	-	-
16	7_1	930	RTE-2_NVi	Non-LTR /RTE	0,703	Π	-	+	-
17	7_5	471	-	-	-	Π	-	+	+
18	7_9	423	Gypsy-32_DP-LTR	LTR/Gypsy	0,766	Π	-	-	-
19	7_7	527	CR1-25_SP, CR1-X1_Pass	Non-LTR /CR1	0,617, 0,759	Π	RT	+	+
20	7_10	573	Copia-90_ALY-I	LTR/Copia	0,788	Π	-	-	+
21	7_13	831	MtPH-A6-2-Ia	DNA/Harbinger	0,710	Π	-	+	+
22	7_14	416	-	-	-	Π	-	+	+
23	7_20	498	Tad1-35_BG	Non-LTR /Tad1	0,792	Π	-	-	+
24	8_1	733	BEL-1_CSa-I	LTR/BEL	0,655	Π	RVE	+	-
25	9_2	320	-	-	-	K	-	+	-
26	10_5	338	-	-	-	K	-	-	-
27	11_1	318	-	-	-	K	-	+	-
28	11_3	327	-	-	-	K	-	+	-
29	12_1	429	Gypsy-32_DP-LTR	LTR/Gypsy	0,766	K	-	-	-
30	12_2	413	EptSINE1	Non-LTR /SINE2	0,750	K	-	+	+
31	13_1	456	EptSINE1	Non-LTR /SINE2	0,750	Π	-	+	-
32	13_4	471	EptSINE1	Non-LTR /SINE2	0,750	Π	-	+	+
33	14_1	405	SR2	Non-LTR/RTE	0,895	K	SLC5-6	+	-
34	15_1	726	BEL-1_CSa-I	LTR/BEL	0,666	Π	RVE	+	+
35	15_2	351	-	-	-	Π	-	-	-
36	15_3	468	-	-	-	Π	-	-	-

Обозначения: № — номер в таблице, Фргм — собственное название клонированного фрагмента. Длина — длина клонированного фрагмента в п. н.; RBase — совпадение с RepBase по результату поиска программой RepeatMasker, Класс — классификация элемента, класс

(отряд)/семейство, Sim — степень сходства с элементом из RepBase, З — характеристика зоны, из которой вырезан клонированный фрагмент: консервативная (К) или полиморфная (П), СКД — суперсемейство к которому относится найденный консервативный домен по результатам поиска в Conserved Domain Database: RT — reverse transcriptase, RVE — Integrase core domain, GTPase — домен ГТФазы, SLC5-6 — solute carrying families 5, 6-like, MFS — Major Facilitator Superfamily, ADF-gelsolin - Actin depolymerization factor/cofilin- and gelsolin-like domains; Т — транскрипция (фрагмент найден во фракции повторов транскриптома (+) или отсутствует (-); DOT — фрагмент использовали (+) или не использовали (-) для дот-гибридизации.

Некоторые клонированные последовательности несут короткие фрагменты ретроэлементов RTE и Tad1. В некоторых последовательностях найдены короткие участки до 100 п. н., гомологичные ДНК-транспозонам семейства MuDR (1_5 и 2_5) и Harbinger (7_13).

Далее анализировали только те последовательности, которые найдены во фракции повторов транскриптома *H. elongata*.

Вклад клонированных фрагментов в формирование паттернов S-SAP

Эксперименты по дот-гибридизации провели (рис.4, рис.8), чтобы определить вклад конкретных ТЕ в формирование консервативных и варибельных участков паттерна S-SAP. При анализе сигналов гибридизации выявили, что соотношение интенсивности точек на блотах не одинаково для разных личинок на внутриклональном и межклональном уровнях. Типичные для зондов каждого диапазона картины гибридизации показаны на рис.8, часть В. Максимально выраженные на дот-блотах фрагменты различны для разных гибридизационных проб. В случае варибельных зон средней длины (рис. 8, В, 2, 3) выражены фрагменты, несущие короткие участки ДНК-транспозона MuDR, а также не идентифицированные фрагменты и фрагмент гена гельзолина (рис.8 В, 2). Также наиболее яркими по интенсивности фрагментами, представленными на всех блотах, являются 15_1 и 7_5.

Сигналы, соответствующие другим клонированным фрагментам, представлены приблизительно равномерно на всех блотах.

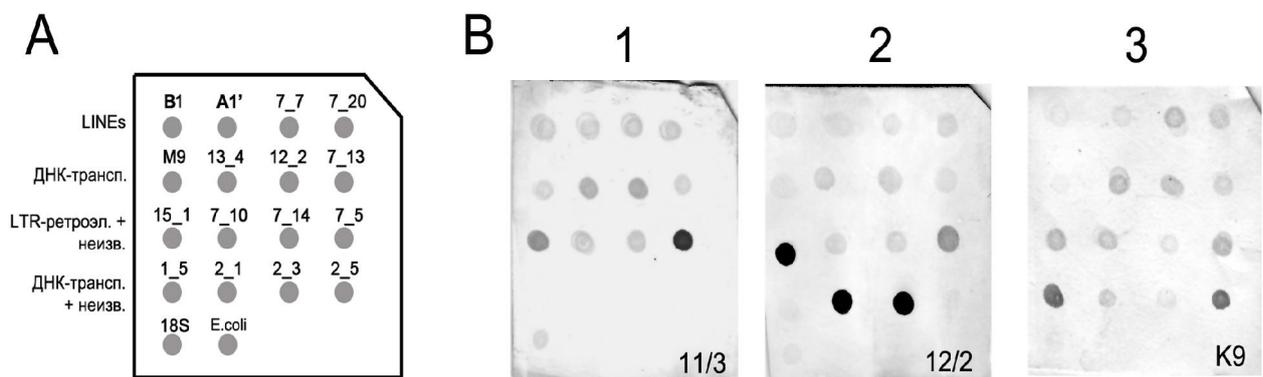


Рис.8. Дот-гибридизация. А — схема нанесения проб на мембрану с указанием рабочих названий последовательностей. В. 1 — пример гибридизации зондов длиной 800 - 1000 п. н.; 2 — пример гибридизации зондов длиной 500-700 п. н.; 3 — пример гибридизации зондов 300-500 п. н. для S-SAP с добавлением рестриктазы EcoRI. В правом нижнем углу на мембранах указаны названия использованных зондов.

Фрагмент 15_1 является частью открытой рамки считывания, кодирующей интегразу элемента BEL из класса LTR-содержащих ретроэлементов. При анализе клонированных последовательностей LTR-элементы обнаруживали только в высоких зонах фореца (более 500 п. н.), а на блотах фрагмент выявляется во всех областях. LTR-содержащие ретропозоны — длинные элементы (~5-8 т. п. н.) с несколькими потенциальными открытыми рамками считывания. Вероятно, фрагмент 15_1 представляет собой часть длинной последовательности LTR-содержащего элемента BEL, способного транскрибироваться и перемещаться целиком.

Последовательность 7_5 также присутствует в трех вариантах блотов. Этот фрагмент не несет в себе идентифицируемых доменов и не имеет совпадений с Repbase. Его решили изучить более подробно. По результатам поиска в базе European Nucleotide Archive 7_5 имеет высокую степень сходства с последовательностями генома *E. caproni* (EM_INV:LL284659, EM_INV:LL284659). При исследовании фрагментов, с которыми у 7_5 наибольшая степень сходства, обнаружены участки CR1-подобных ретротранспозонов. Амплификация фланкирующих последовательностей выявила несколько продуктов разной длины (рис.6, В). Анализ клонированных фрагментов (7_5_12, 7_5_13) выявил в них участки CR1-подобных ретроэлементов, а сами фрагменты обнаружены во фракции повторов транскриптома *H. elongata* (таблица 3). Таким образом, последовательность 7_5 принадлежит к ретротранспозонам семейства CR1 и располагается вне открытой рамки считывания (Рис. 9, С).

Использование двойной рестрикции HindIII и EcoRI привело к выравниванию интенсивности точек на блотах, а также к появлению других мажорных сигналов (рис.8, В, 2).

Анализ транскриптов, соответствующих клонированным фрагментам

Повторы в транскриптом редий *H. elongata* выявляли посредством поиска в программе BLAST (Altschul et al., 1990) по базе RepBase, и именно с этой частью базы, не содержащей кодирующие белки гены, сравнивали клонированные фрагменты из таблицы 2. В таблице 3 представлены результаты поиска в транскриптом последовательностей, соответствующих клонированным фрагментам. Самое большое число совпадений приходится на фрагменты, вырезанные из консервативных зон S-SAP и несущие домены обратных транскриптаз (рис. 9). Для всех транскриптов есть данные о числе транскриптов на миллион (TPM, transcripts per million transcripts). Эта метрика отражает изобилие или представленность определенных РНК в транскриптом, и по ней можно судить об относительном уровне экспрессии. По сравнению с TPM для гена глицеральдегид-3-фосфат-дегидрогеназы (1709,59) число транскриптов, соответствующих клонированным фрагментам в 10 – 100 раз ниже, тем не менее, они достаточно разнообразны. Среди LINE-подобных элементов найдены представители семейств RTE и CR1, для LTR-ретроэлементов найдены представители семейства BEL. Для некоторых фрагментов в соответствующих транскриптах вообще не обнаружено участков каких-либо мобильных элементов. RepeatMasker способен найти только те последовательности, которые есть в базе данных Repbase; для не секвенированного генома эта выборка не может быть репрезентативной. Сравнение с Repbase не дает возможности выявить как видоспецифичные TE, так и тандемные повторы (ТП). Видимо, к этим классам и относятся неидентифицированные фрагменты. При начальном анализе для некоторых клонированных фрагментов невозможно было определить даже класс TE.

Таблица 3. Анализ транскриптов, соответствующих клонированным фрагментам.

№	Фрм.	hits_N	TPM	RBase	Классификация	Sim	СКД
1	A1, B1, B2	60	41,59	CR1-77_HM	LINE/CR1	0,793	RT
			41,59	CR1-3_BF	LINE/CR1	0,668	RT
2	A1'	7	9,79	CR1-24_SP	LINE/CR1	0,636	RT
3	B1'	84	9,79	CR1-24_SP	LINE/CR1	0,636	RT
4	1_1	1	0,76	+/-	LTR/Copia	0,800	-
5	1_2	1	0,76	+/-	LTR/Copia	0,800	-
6	1_5, 2_5	5	35,07	CR1-6_HMM	LINE/CR1	0,803	-
7	3_4	2	27,64	+/-	DNA/MuDR	0,875	+/-
8	7_1	70	27,38	RTE_SJ	LINE/RTE	0,668	+/-
9	7_5	49	9,51	-	-	-	-
10	7_7	106	20,62	CR1-12_CTe	LINE/CR1	0,677	RT
11	7_13	33	27,38	RTE_SJ	LINE/RTE	0,668	+/-
12	8_1	25	10,24	Saci-1_I	LTR	0,652	RVE
13	9_2	1	0,41	+/-	DNA/Kolobok	0,771	-
14	11_1, 11_3	1	0,41	+/-	DNA/Kolobok	0,771	-
15	12_2	43	4,11	P-27_HM	DNA/P	0,714	-
16	13_1	64	4,11	P-27_HM	DNA/P	0,714	-
17	13_4	27	23,87	P-27_HM	DNA/P	0,714	-
18	14_1	11	2,71	SR2	LINE/RTE	0,968	-
19	15_1	25	10,24	Saci-1_I	LTR	0,652	RVE
			6,87	BEL-1_CSa-I	LTR/BEL	0,665	RVE
20	7_5_12	21	9,79	CR1-X1_Pass	LINE/CR1	0,741	+/-
21	7_5_13	156	133,38	CR1-36_HM	LINE/CR1	0,646	RT

Обозначения: № — номер в таблице, Фрм — название клонированного фрагмента, hits_N — число найденных программой BLAST последовательностей из транскриптома; TPM — метрика нормализованной экспрессии, выраженная в транскриптах на миллион и посчитанная для транскриптов, имеющих наибольшую степень сходства с клонированным фрагментом; RBase — совпадение с Rerbase по результату поиска программой RepeatMasker на участке перекрывания транскрипта и клонированного фрагмента, случай наличия совпадений вне зоны перекрывания обозначен «+/-»; Классификация — класс (отряд) элемента/семейство, Sim — степень сходства с элементом из Rerbase; СКД (суперсемейство консервативных доменов) — суперсемейство, к которому относится найденный консервативный домен по результатам поиска в Conserved Domain Database в зоне перекрывания транскрипта с клонированным фрагментом: RT — reverse transcriptase, обратная транскриптаза, RVE — Integrase core domain, коровый домен интегразы; случай совпадения вне зоны перекрывания обозначен +/-.

Так, во фрагменте 2_3 обнаружили домен гельзолина и короткий участок ретроэлемента (таблица 2). Во фракции повторов 2_3 отсутствует, поэтому мы считаем фрагмент принадлежащим скорее к «экзомной» части транскриптома. Транскрипты, соответствующие фрагменту 3_4, имеют выявляемые домены

обратной транскриптазы, а RepeatMasker находит фрагменты non-LTR ретроэлементов (Perere, CR1). Последовательность 7_5 не несет идентифицируемых доменов, однако имеет весомый вклад в формирование паттернов S-SAP. Исследование фланкирующих последовательностей позволило обнаружить фрагменты подобных CR1 ретроэлементов. В транскриптоме найдены преимущественно короткие РНК длиной до 300 п. н., соответствующие 7_5, кроме одного транскрипта длиной 1300 п. н. Этот транскрипт имеет большую степень сходства с одним из фланкирующих 7_5 фрагментов (7_5_13) и имеет участки, гомологичные ретроэлементам CR1. В некоторых транскриптах, соответствующих последовательностям 1_5 и 2_5, обнаруживаются короткие участки CR1-подобных ретроэлементов, однако поиск консервативных доменов не дал результатов (таблица 3). Во фрагменте 7_13 при первичном анализе обнаружили короткий участок ДНК-транспозона Harbinger, однако в гомологичных транскриптах в области перекрытия последовательностей находятся домены обратной транскриптазы non-LTR ретроэлементов. Транскрипты, соответствующие фрагментам 1_1, 1_2, 9_2, 11_1 и 11_3 имеют значения TPM менее единицы, это скорее говорит о том, что они являются либо частью транскрипционного шума, либо контаминацией. Фрагменты 13_1 и 13_4 несут участки SINE подобных ретроэлементов и имеют совпадения с базой данных по SINE Base, однако в некоторых гомологичных им транскриптах обнаруживаются участки ДНК-транспозонов семейства P. Консервативных доменов эти транскрипты не имеют.

Провели оценку соотношения количества транскриптов, соответствующих клонированным фрагментам. В результате выявили, что наибольшее число совпадений (в сумме 93 %) приходится на фрагменты non-LTR ретроэлементов: 1) последовательности, гомологичные несущим участки обратной транскриптазы фрагментам (подгруппа LINE/RT) и 2) последовательности, которые принадлежат спейсерным участкам LINE-подобных ретроэлементов. Эти фрагменты определили в подгруппу LINE/spacer. Значительно слабее (4 %) представлены последовательности, соответствующие LTR-ретроэлементам — подгруппа LTR. Транскрипты, соответствующие фрагментам ДНК-транспозонов (подгруппа DNA), составили 1 %, и около 2 % транскриптов представлено последовательностями, в которых не обнаружено каких-либо идентифицируемых доменов — это подгруппа Unknown.

Распределение найденных мобильных элементов в геноме *H. elongata*

Некоторые фрагменты послужили зондами для определения их позиций в геноме. Определили расположение некоторых фрагментов на консенсусных представителях семейств из Repbase (рис.9). На панели А — примеры положения фрагментов по отношению к консенсусам ДНК-транспозонов. На панели В рис.9 показаны примеры расположения последовательностей, в которых обнаружены участки LTR-содержащих ретроэлементов.

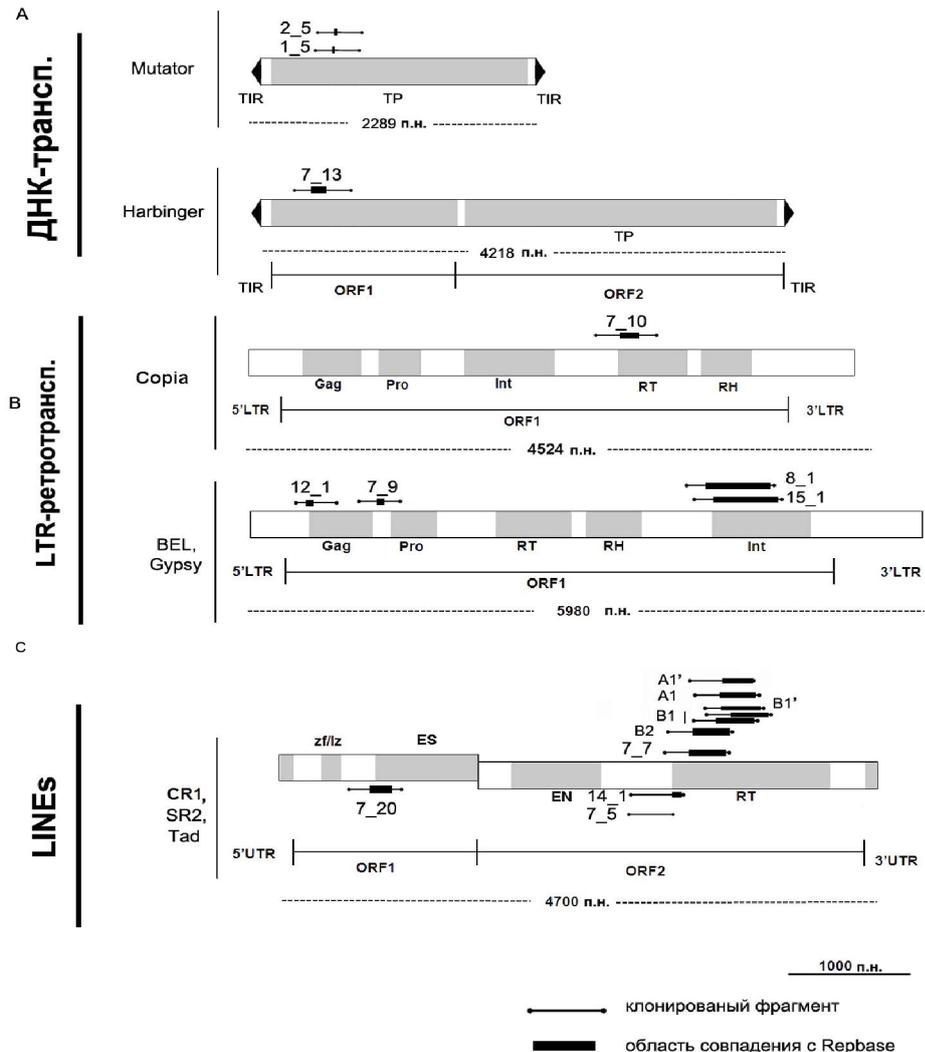


Рис.9. Схема организации полноразмерных транспозонов и положение некоторых секвенированных фрагментов. Отмечены размеры ретроэлементов, 5' и 3' нетранслируемые районы (UTR), открытые рамки считывания (ORF). Внутри ORF отмечены домены: *zf/lz* — домен «цинковых пальцев» и «лейциновой молнии», ES — esterase (эстеразы), EN — endonuclease (эндонуклеазы), RT — reverse transcriptase (обратной транскриптазы), РНКазы H (RH), интегразы (Int), ДНК-связывающего белка (Gag) и протеиназы (Pro), TP — транспозаза. TIR — терминальный инвертированный повтор. Слева от схем транспозонов указаны семейства, над схемами отмечены положения клонированных фрагментов.

Клонированные фрагменты 7_10 и 12_1, которые выравниваются на участки вне открытых рамок считывания LTR-ретроэлементов, в транскрипте не обнаружены. В то же время, участки некодирующих частей LINE, соответствующие фрагментам 7_5 и 14_1, весьма многочисленны в транскрипте. Все фрагменты над консенсусом LINE относятся к группе LINE\RT, под схемой консенсуса — к группе LINE/spase (рис.9, C). Обе группы хорошо представлены в транскрипте.



Рис.10. Кариотип *H. elongata* (Solovyeva et al., 2016). А — метафазная пластинка, окраска DAPI. В — кариотип. С — идеограмма хромосом *H. elongata*, серые прямоугольники отмечают локализацию 18S рДНК. Масштабный отрезок — 10 мкм.

Вид *H. elongata* становится все более популярным объектом молекулярно-биологических исследований, однако не проведено исследование кариотипа и секвенирование генома. Для исследования распределения найденных фрагментов мобильных элементов использовали метод FISH. Для картирования зондов исследовали кариотип *H. elongata*. Определили, что 74,4 % метафазных пластинок содержат 12 пар хромосом (1m/sm+1m+3sm/st+7st; рис10, А, В). На основании анализа лучших 64 метафазных пластинок сконструировали идеограммы (рис. 10, С). Хромосомы *H.elongata* имеют хромомерную структуру и обладают рядом особенностей. В поздней профазе наблюдали хромосомы с конъюгированными сестринскими хроматидами. Выявлен полиморфизм блоков центромерного и субтеломерного гетерохроматина, который, вероятно, может отражать популяционную или клональную изменчивость, а также дискорданс гомологов, то есть различия между гомологичными хромосомами. Дифференциальная окраска хромосом DAPI и СМА₃ выявила области хроматина, которые варьируют по содержанию АТ- и GC-пар. Картирование 18S рДНК выявило 3 ядрышковых организатора, расположенных на хромосомах 3, 6, 10 (Solovyeva et al., 2016).

Фрагмент 15_1 представляет собой участок интегразы LTR-содержащего элемента BEL. Успешная гибридизация зонда 15_1 прошла только при длительной предобработке коллагеназой (5 мин). В результате такой предобработки ядра и хромосомы приобретают зернистую структуру, и вероятно, происходит более успешное «разворачивание» гетерохроматина. Гибридизация фрагмента 15_1 обнаруживает большие конгломераты, вероятно, полноразмерных LTR-ретротраспозонов как в интерфазном ядре, так и на некоторых хромосомах (рис.11). Однако, процедура FISH потребовала жесткой обработки препаратов и мы не можем соотнести меченые хромосомы с кариотипом.

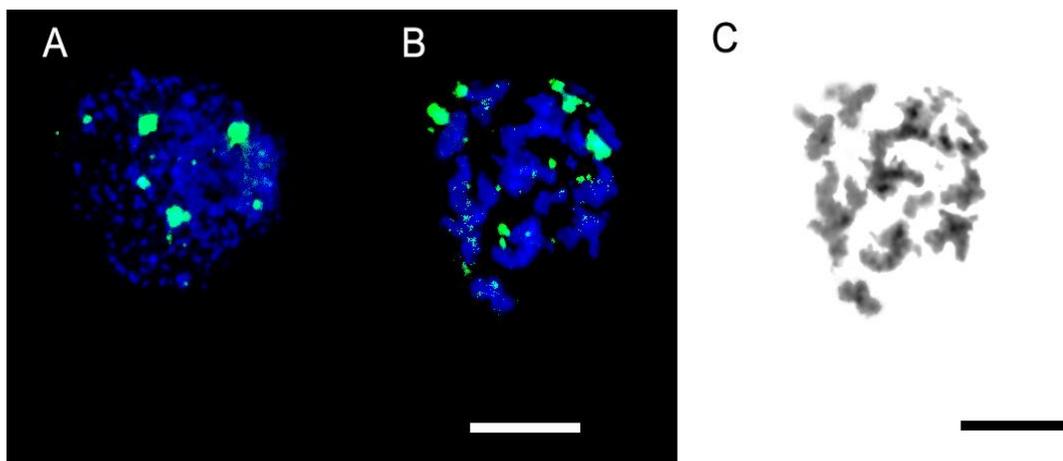


Рис.11. FISH с зондом к 15_1 (зеленый) на ядре и метафазной пластинке, контрастированных DAPI (синий) после длительной обработки коллагеназой. А. Ядро, В.Метафазная пластинка, С. Негативное черно-белое изображение метафазной пластинки (В). Масштабный отрезок — 10 мкм.

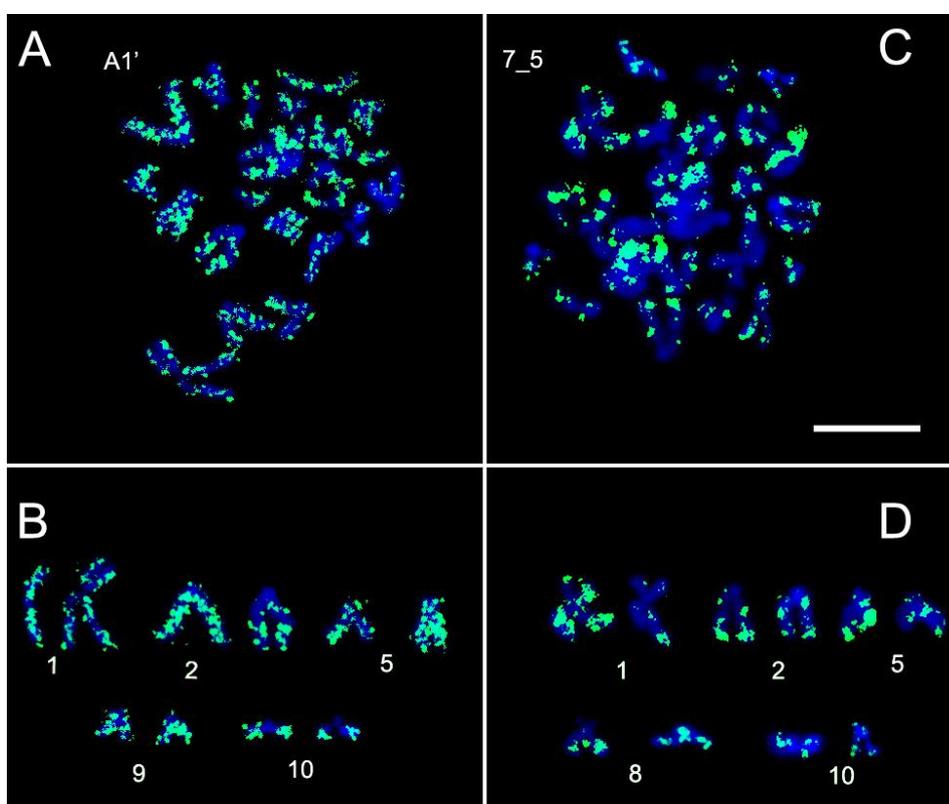


Рис.12. А, С — FISH с зондами к A1' и 7_5 (зеленый) на хромосомах *H. elongata*, контрастированных DAPI (синий). В, D — некоторые хромосомы с их номерами по кариотипу (рис. 11) с наиболее характерным расположением метки. Масштабный отрезок — 10 мкм.

Для успешной гибридизации остальных зондов достаточно стандартной предобработки (в случае ядер) или кратковременной инкубации в коллагеназе (1-3 мин для хромосом). Фрагмент A1' относится к наиболее представленной в транскриптом группе, то есть имеет часть открытой рамки считывания обратной транскриптазы элемента LINE (рис. 9, С) и распределен на хромосомах дисперсно (рис.12). Максимальное количество зонда наблюдали на хромосомах 1-5, 9. Сигнал покрывает все плечи хромосом без обогащения в гетерохроматиновых районах. На хромосоме 10 есть обогащение в субтеломерной зоне, что может быть связано с расположением здесь основного ядрышкового организатора. Известно обогащение

околоядрышкового хроматина элементами LINE (Ne'meth et al., 2010). Дисперсное распределение характерно для представителей класса LINE и соответствует данным литературы (Staginnus et al., 2010). Фрагмент 7_5 является частью LINE, но он выравнивается со спейсерным участком, не принадлежащим даже потенциальным открытым рамкам считывания (рис.9, С). Фрагмент 7_5 локализуется, как правило, в субтеломерной области всех хромосом набора (рис 12). У многих хромосом выражен и сигнал в перицентромерных областях (рис. 12, 7_5, С, D). Несмотря на то, что есть некоторое количество сигнала по длине плеч хромосом 1-4 и 6, фрагмент 7_5 очевидно тяготеет к гетерохроматиновым областям генома. Таким образом, LTR-ретроэлементы кластеризуются в геноме; картины распределения фрагментов группы LINE-RT (A' и др.) и группы LINE\space (7_5) различаются.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В настоящей работе с помощью методов клеточной, молекулярной и компьютерной биологии впервые проведена характеристика последовательностей из консервативных и полиморфных зон в полученных методом S-SAP картинах генотипирования церкарий, а также получены данные о кариотипе *H. elongata*. Генетическая вариабельность, необходимая для поддержания разнообразия в популяции, ранее считалась присущей лишь организмам, размножающимся половым путем. Тем не менее, существование клональной изменчивости показано у растений (Pelsy et al., 2015), грибов и ряда Metazoa. У трематод увеличение генетической вариабельности клональной изменчивостью повышает шанс на успешную инвазию хозяина. Несмотря на очевидные доказательства существования клональной изменчивости, ее механизмы возникновения остаются неизвестны. Для объекта настоящей работы, *H. elongata*, показан полиморфизм церкарий в клональных популяциях на уровне поведенческих реакций и выживаемости внутри второго промежуточного хозяина, что предполагает и генетическую вариабельность партенит (Прокофьев и др., 2011; Levakin et al., 2012, Левакин и др., 2013).

Клональная изменчивость найдена у других трематод. У метацеркарий рода *Microphallus* (Халтурин и др., 2000, К. Галактионов и др., 2008), церкарий *Trichobilharzia* (Семенова и др., 2005) и *Bucephalus polymorphus* (Корсуненко и др., 2009) она показана с помощью метода RAPD (Random Amplified Polymorphic DNA). Аналогичным способом выявлена клональная изменчивость *Trichobilharzia szidati* и *Bilharziella polonica*. С помощью клонирования и анализа вариабельных зон RAPD выяснили, что в них хорошо представлены фрагменты ретроэлементов. На основе идентифицированных ТЕ сконструировали вторичные праймеры, которые выявляли фланкирующие фрагменты, необходимые для более точной классификации. Клонирование и анализ зон показали, что фрагменты представлены в основном последовательностями non-LTR ретропозонов — 76% с CR1-подобными элементами (15% из них), LTR - ретроэлементами — 16%, а также Penelope-подобными элементами (PLE) — 8% (Корсуненко, 2010). Таким образом, мы могли ожидать появления подобных ТЕ среди клонированных фрагментов. Действительно, фрагменты non-LTR и LTR-содержащих ретроэлементов обнаружили в разных зонах S-SAP. Представители non-LTR ретроэлементов широко представлены в геномах известных трематод. Дот-гибридизация позволила получить информацию о наполнении определенных диапазонов зон клонированными фрагментами. Этот метод не является достаточно точным, но с его помощью подтвердили наличие полиморфизма в картинах генотипирования церкарий. Также выявили мажорные

фрагменты, характерные для определенного диапазона зон S-SAP. Поиск соответствующих последовательностей в транскриптоме и их анализ показали, что фрагменты non-LTR и LTR-содержащих ретроэлементов встречаются в транскриптах и имеют ненулевой уровень экспрессии, число их транскриптов на миллион (TPM) варьирует от 6,87 до 133,38.

В транскриптах, соответствующих фрагменту 7_13, обнаружен участок LINE-подобного ретроэлемента. По-видимому, в области выравнивания последовательностей 7_13 и гомологичных транскриптов находятся короткие участки, общие для ретроэлементов и ДНК-транспозонов. Фрагменты SINE найдены в последовательностях 12_2, 13_1 и 13_4 (таблица 2) по результатам поиска в двух базах данных — RepBase и SINE Base. Однако в соответствующих транскриптах выявляются последовательности ДНК-транспозонов семейства P. Элементы SINE очень видоспецифичны, и для их подробного описания и классификации необходимо отдельное исследование.

У трематод картированы на хромосомах лишь некоторые полноразмерные ретроэлементы из генома *S. mansoni*. SINE элемент Smá распространен дисперсно на всех хромосомах. LTR-содержащий ретроэлемент Boudicca не имеет дисперсного распределения и локализуется в эухроматине хромосомы 2 и женской половой хромосомы W, non-LTR ретроэлемент Perere 03 из семейства RTE также расположен в эухроматине субтеломерной зоны хромосомы 2 (Valentin et al., 2008).

На ядрах и хромосомах *H. elongata* FISH проводили с клонированными фрагментами. Фрагмент BEL скорее всего выявляет полноразмерные копии транспозонов. Есть примеры их расположения в гетерохроматине; тогда они могут быть источником перестроек в этих районах, хотя мы и не нашли их выраженной транскрипции. Известно, что LTR-содержащий BEL-подобный ретропозон Achilles из генома мушки *Vastrocera oleae* локализуется предпочтительно в гетерохроматиновых районах Y-хромосомы (Tsoumani et al., 2015). Gypsy-подобные ретроэлементы из генома кофе также локализуются преимущественно в виде кластеров и ассоциируются с блоками гетерохроматина (Yuyama et al., 2012). Мы также наблюдаем кластеризацию несущего участок интегразы зонда 15_1, вероятно, в гетерохроматиновых областях (рис. 11, B).

По условиям постановки опыта и анализа клонов перестройки tandemных повторов (ТП) не могут быть выявлены. Однако при изучении кариотипа *H. elongata* наблюдали варьирование по размерам блоков центрального и субтеломерного гетерохроматина. Источник такого полиморфизма может лежать в варьировании числа копий ТП, главных компонентов гетерохроматина, но обнаруженные фрагменты TE, а именно non-LTR элементы, также могут обеспечить такой эффект.

Все те фрагменты, которые содержат домены обратной транскриптазы, задействованы в активной транскрипции и, следовательно, возможно их перемещение, в основном, в эухроматиновых районах. Факультативный гетерохроматин обогащен non-LTR ретроэлементами; известна склонность TE встраиваться рядом с гомологичной последовательностью. Возможно, что закономерные перестройки (видимые как консервативные зоны) расположены в факультативном гетерохроматине.

В соответствии с литературными данными (Корсуненко, 2010), мы нашли обогащение консервативных зон элементами LINE. Однако есть различия по фрагментам, входящим в консервативные (К) и полиморфные (П) зоны. Если К зоны

содержат фрагмент обратной транскриптазы LINE (RT), то П зоны содержат преимущественно спейсерные области (рис.9, таблица 2).

Известно, что элементы LINE1 могут тандемно повторяться, ранее это было показано для полноразмерных элементов (Ahmed and Liang, 2012). В процессе классификации больших мышинных ТП методами биоинформатики в Группе Некодирующей ДНК обнаружили ТП, образованные из фрагментов L1 (Komissarov et al., 2011). Родственные элементу L1 мономеры ТП образованы фрагментом второй открытой рамки считывания (ORF2) и 3' нетранслируемого района (3'UTR). В лаборатории проводили работу по экспериментальному клонированию фрагментов гетерохроматина. Клонированные фрагменты оказались частью L1 и Lx ретроэлементов мыши. Они покрывают фрагмент ~2 т. п. н. в конце ORF2 и 3'UTR L1. Тот же фрагмент обнаружен и в гетерохроматиновых районах человеческого генома (Kuznetsova et al., 2016). Экспериментально полученные клоны короткие (100 - 200 п. н.), а яркость их FISH-сигналов на хромосомах и в хромоцентрах сигнала сравнима с сигналом от классических ТП — сателлитах мыши. Это является свидетельством в пользу тандемной организации фрагмента внутри гетерохроматинового района.

Впервые клональная изменчивость трематод выявлена именно на тандемно повторенных последовательностях сателлите W1 в геномах партеногенетических поколений *S. mansoni* (Bayne, Grevelding, 2003). Элемент W1 существует в 500 – 1000 копиях в геноме как ТП с мономером 476 п. н., одним сайтом рестрикции для EcoRI, соответственно, относится к сателлитной ДНК (Grevelding, 1999) и локализуется преимущественно в гетерохроматине половых хромосом (Lepesant et al., 2012). На праймерах к W1 выявили генетическую гетерогенность внутри популяций *S. mansoni* и интерпретировали это как нестабильность генома. После культивирования дочерних спороцист доказали, что внутри одного клона есть различия по содержанию W1. Геномную гетерогенность обнаружили среди клонов церкарий, произошедших от одного мирацидия. Авторы высказали предположение, что причиной клональной изменчивости стала митотическая рекомбинация при партеногенетическом размножении или так называемая эктопическая или неразрешенная рекомбинация (Bayne, Grevelding, 2003). Но результаты настоящей работы свидетельствуют о том, что в клональной изменчивости задействованы скорее фрагменты TE, которые также присутствуют в гетерохроматине (Kuznetsova et al., 2016). Фрагмент 7_5 не организован тандемно, что подтверждается анализом *in silico* (TRF; Wenson, 1999). Однако морфологические данные (рис. 12) позволяют думать, что фрагмент тяготеет к конститутивному гетерохроматину. У мыши фрагменты аналогичного района LINE выделены из хромоцентров, то есть из конститутивного гетерохроматина (Kuznetsova et al., 2016). Точная локализация фрагмента 7_5 нуждается в уточнении дополнительными методами (например, двойной fiberFISH с зондом к ТП с учетом видоспецифичности), однако этот фрагмент, безусловно, представлен в транскриптом (таблица 3).

Подтверждение функциональной важности 3'-фрагмента LINE, к этой области относится и 7_5, для эукариотических геномов находится в работе, посвященной COT1 РНК человека (Hall et al., 2014). Недавние исследования обнаружили богатое разнообразие не кодирующих РНК, в основном, с неизвестными функциями. Но работ, посвященных изучению диспергированных повторов, составляющих существенную (до 80% у некоторых видов) часть генома, очень мало. С помощью *in situ* гибридизации зондов COT-1 ДНК с тотальной РНК клеток обнаружили, что

обилие связанной с эухроматином РНК представлено преимущественно повторами (C0T-1 РНК), включая транскрипты ретроэлементов LINE-1. C0T-1 РНК располагается между хромосомными территориями и остается стабильно связанной с ними после ингибирования транскрипции. C0T-1 РНК очень прочно связана с хроматином и ядерным матриксом. Согласно мнению авторов, РНК повторов не является транскрипционным шумом, а, напротив, увеличивает возможности регулирования и настройки транскрипции; более того, ее потеря приводит к нарушению распределения хроматина и его конденсации. Именно регуляторной ролью объясняется обилие и повсеместное распространение ТЕ. Но более того, транскрипты ORF2 LINE и 3' конца LINE доминируют во фракции C0T-1 РНК (fig.6, Hall et al., 2014), что полностью согласуется и с результатами настоящей работы.

Результаты настоящей работы подтверждают, что повторяющуюся «мусорную» часть генома стоит изучать не только как ДНК, но и как основную часть «темной материи» транскриптома. Всего лишь 5 % промоторов лежит в пределах канонических генов (Venters and Pugh, 2013). Как происходит экспрессия и регуляция остальной их, генов, части, остается непонятным. Настоящая работа является вкладом в то, что может привести к новой революции в исследованиях генома: выявлению связи РНК и повторяющихся последовательностей, их вклад в упаковку генома, его нестабильность и эволюцию.

ВЫВОДЫ

1. Метод S-SAP позволяет выявить клональную изменчивость (КИ) у одиночных церкарий трематоды *H. elongata*. В консервативных зонах располагаются LINE-подобные ретроэлементы, в полиморфных зонах присутствуют спейсерные участки LINEs, а также фрагменты LTR-ретротранспозонов.

2. Методом дот-гибридизации выявлено преобладание последовательностей ретроэлементов в разных зонах паттернов генотипирования церкарий.

3. Кариотип *H. elongata* содержит 12 пар хромосом ($2n=24$), среди которых 3 пары несут ядрышковые организаторы.

4. Методом флуоресцентной *in situ* гибридизации (FISH) обнаружено дисперсное распределение тех фрагментов LINE, которые основаны на открытой рамке считывания обратной транскриптазы и обогащение конститутивного гетерохроматина тем фрагментом, который основан на спейсерных участках LINE.

5. Обе группы основанных на LINE фрагментов активно транскрибируются и, следовательно, вносят вклад в КИ.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ:

1. **Соловьева А.И.**, Галактионов Н.К., Подгорная О.И. 2013. Ретротранспозон класса LINE является компонентом паттерна полиморфных фрагментов партенит трематоды *Himasthla elongata*. Цитология. 55 (7): 492–500.

2. Galaktionov N., **Solovyeva A.**, Fedorov A., Podgornaya O. 2014. Trematode *Himasthla elongata* mariner element (Hemar): structure and applications. JEZ Part B: Molecular and Developmental Evolution. 322(3):142-55.

3. **Solovyeva A.I.**, Stefanova V.N. Podgornaya O.I., Demin S.Iu. 2016. Karyotype fetatures of trematode *Himasthla elongata*. Molecular Cytogenetics. 9: 34

Тезисы:

1. **Соловьева А.И.**, Подгорная О.И., Галактионов Н.К. 2011. Метод Транспозон Дисплея как инструмент выявления генетического полиморфизма. Морской биологический симпозиум СПб. Сборник тезисов докладов, стр. 81

2. **Соловьева А.И.**, Демин С.Ю., Галактионов Н.К., Подгорная О.И. «ДНК-транспозон *mariner* как инструмент выявления генетического полиморфизма». Тезисы докладов и сообщений, представленных на III конференцию молодых ученых Института цитологии РАН (Санкт-Петербург, 15–16 мая 2012 г). Цитология 54 (4): 355

3. **Соловьева А.И.**, Галактионов Н.К., Подгорная О.И. Последовательность транспозона *Nemar1* варьибельна в клональных популяциях *Himasthla elongata*. 16-я Пущинская Международная школа-конференция молодых ученых. Сборник тезисов докладов, стр.147. Пущино, 2012 г.

4. **Соловьева А.И.** Мобильные элементы *Himasthla elongata* (Trematoda, Echinostomatidae). Международная научная конференция студентов, аспирантов и молодых учёных «Ломоносов-2013». Сборник тезисов докладов, стр. 231. Москва, 2013 г.

5. **Solovyeva A.**, Galaktionov N., Podgornaya O. Retrotransposons of different families participate in the trematode *Himasthla elongata* parthenitae clonal diversity. 49th European Marine Biology Symposium. Сборник тезисов докладов, стр. 98. Санкт-Петербург, 2014 г.

6. **Соловьева А.И.**, Галактионов Н.К., Подгорная О.И. Полиморфизм распределения последовательностей мобильных элементов в геномах партенит *Himasthla elongata*. 19-я Пущинская Международная школа-конференция молодых ученых. Сборник тезисов докладов, стр. 273. Пущино, 2015 г.

7. **Solovyeva A. I.**, Galaktionov N. K., Podgornaya O. I. Polymorphisms and mobile elements in trematode *Himasthla elongata* parthenitae. EMBL Symposium: The Mobile genome: Genetic and Physiological impacts of transposable elements. Сборник тезисов докладов, стр. 160. Гейдельберг, 2015 г.

8. **Соловьева А.И.**, Стефанова В.Н., Подгорная О.И., Демин С.Ю. Особенности кариотипа трематоды *Himasthla elongata*. Конференция Хромосома 2015. Сборник тезисов докладов, стр. 160. Новосибирск, 2015 г.

9. **Соловьева А. И.**, Галактионов Н. К., Горбушин А. М., Подгорная О. И.. Полиморфизм распределения последовательностей мобильных элементов в геномах партенит *Himasthla elongata*. V молодежная конференция по молекулярной и клеточной биологии Института цитологии РАН. Сборник тезисов докладов, стр. 58. Санкт-Петербург, 2016 г.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА.

Tollis and Boissinot, 2012, *Genome Dyn.* 7: 68 – 91; Wicker et al., 2007, *Nature Reviews Genetics*. 8(12):973 – 982; Brown, 2002, Oxford, Garland Science, 608p; Abrusan and Krambeck, 2006, *Theoretical Population Biology*. 70(3): 364 – 375; Slotkin and Martienssen, 2007, *Nat Rev Genet*. 8(4):272 – 285.; Kazazian, 2011, *BMC Biol*. 9:62; Muotri et al., 2005, *Nature*. 435: 903 – 910; Werding, 1969; Tkach et al., 2016, *Int J Parasitol*. 46(3):171—85; Grevelding, 1999, *Parasitol.*, 101: 207 – 216; Халтурин и др., 2000, *Паразитология*. 34 (6): 486 – 500; Семенова и др. 2005, *Генетика* 41(1): 17 – 22; Корсуненко и др., 2009, *Генетика*, 45(1): 73 – 80; Winperenninckx et al., 1993, *Trends Genet*. 9: 407; Hare and Johnston, 2011, *Methods Mol Biol.*, 772: 3 – 12; Vos et al., 1995, *Nucleic Acids Res*. 23: 4407 – 4414; Waugh et al., 1997, *Mol. and General Genetics*, 253: 687 – 694; Drew and Brindley, 1995, *Molecular and Biochemical Parasitology*. 71(2): 173 – 181; Sambrook et al., 1989, Cold Spring Harbor Laboratory Press; Wahl, 1979, *Proc Natl Acad Sci USA*. 76(8):3683 – 7; Leary et al., 1983, *Proc Natl Acad Sci USA*, 80:4045 – 49; Richard and Voltz, 1987, *Systematic Parasitology*. 9: 169 – 172.; Birstein and Mikhailova 1989, *Genetica*. 80: 159 – 165; Mutafova et al., 1991, *J. of Helminthology*. 65: 255 – 258; Mutafova, 1994, *Systematic Parasitology*. 28: 229 – 238 ; Hirai and Hirai, 2004, *Methods Mol. Biol*. 270: 379 – 394.; Khodyuchenko et al., 2012, *Histochem Cell Biol*.138(1):57 – 73; Grabherr et al., 2011, *Nature biotechnology*. 29:644 – 652; Bray et al., 2015, *Nature Biotechnology* 34,525 – 527; Altschul et al., 1990, *J. Mol. Biol*. 5: 403 – 410; Kalendar et al., 2009, *Genes, Genomes and Genomics*, 3(1): 1 – 14; Kohany et al., 2006, *BMC Bioinformatics*. 7: 474; Vassetzky and Kramerov, 2013, *Nucleic Acids Res*. 41(Database issue):D83 – 9; Benson, 1999, *Nucl.Acid.Res*. 27(2): 573 – 580; Hall, 1999, *Nucleic Acids Symposium Series*. 41: 95 – 98; Edgar, 2012, *Bioinformatics* 26(19), 2460 – 2461; Simpson et al., 1982, *Molecular and Biochemical Parasitology*. 6: 125 – 137; Sadhu and Gill, 2002, *Plant Physiol*. 128:803 – 11; Федоров, 2008, *Цитология*. 50(12): 1011 – 1022; Томилин, 2009, *Цитология*. 51 (4): 291 – 296; Pelsy et al., 2015; Прокофьев и др., 2011, *Паразитология* 45(5): 345 – 357; Levakin et al., 2012, Левакин и др., 2013, *Паразитология*. 47(5): 353 – 360; К. Галактионов и др., 2008, *Материалы IV Всероссийского Съезда Паразитологического общества при Российской академии наук*. 154; Корсуненко, 2010, Москва, 26 стр; Valentim et al., 2008, *J Parasitol*. 94(4): 993 – 5; Tsoumani et al., 2015, *PLoS One*.10(9): e0137050; Yuyama et al., 2012, *Genome*. 55(12): 825 – 33; Ahmed and Liang, 2012, *Comp Funct Genomics*. 2012: 947089; Komissarov et al., 2011, *BMC Genomics*. 12:531; Kuznetsova et al., 2016, *Chromosome Res*. 24(3): 309 – 23; Bayne, Grevelding, 2003, *J. Parasitol.*, 89:1056—1060; Lepesant et al., 2012, *Genome Biol*. 13(2): R14; Hall et al., 2014, *Cell*, 156(5): 907 – 918; Venters and Pugh, 2013, *Nature*. 513(7518): 444.