

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

На правах рукописи

Помазной Михаил Юрьевич

**Транскриптомный анализ
трематоды *Opisthorchis felineus***

Генетика – 03.02.07

Диссертация на соискание
ученой степени кандидата
биологических наук

Научный руководитель:
д.б.н. Мордвинов В.А.

Новосибирск – 2015

Оглавление

1 Введение.....	4
1.1 Актуальность темы исследования и степень ее разработанности.....	4
1.2 Цели и задачи.....	5
1.3 Научная новизна работы.....	5
1.4 Теоретическая и практическая значимость работы.....	6
1.5 Положения выносимые на защиту.....	6
1.6 Степень достоверности и апробация результатов.....	7
2 Обзор литературы.....	8
2.1 Краткая характеристика <i>O. felineus</i> и вызываемого им заболевания.....	8
2.1.1 Систематическое положение.....	8
2.1.2 Жизненный цикл.....	10
2.1.3 Ареал распространения.....	12
2.1.4 Описторхоз как заболевание.....	13
2.2 Функциональная геномика и транскриптомика трематод.....	15
2.2.1 Методология исследования транскриптомов.....	15
2.2.2 Метаболизм трематод.....	16
2.2.3 Покровные ткани.....	18
2.2.4 Строение тела и двигательная система.....	18
2.2.5 Пути трансдукции сигналов.....	20
2.2.6 Стволовые клетки трематод.....	21
2.2.7 Гельминтозы и онкологические заболевания.....	22
2.2.8 Взаимодействие паразита и хозяина на гуморальном уровне.....	23
2.2.9 Ответ на внешние факторы.....	25
2.2.10 Транскриптом разных стадий жизненного цикла.....	27
2.2.11 Особенности молекулярно-генетической организации трематод.....	28
2.2.12 Потенциальные мишени для диагностики, лечения и вакцинирования.....	29
2.3 Заключение.....	30
3 Материалы и методы.....	32
3.1 Бактериальные штаммы:.....	32
3.2 Реактивы и смеси.....	32
3.3 Экспериментальные методы.....	33
3.3.1 Сбор биологического материала.....	33
3.3.2 Создание кДНК библиотеки и ее анализ методом секвенирования по Сэнгеру.....	33

3.3.3 Создание и секвенирование библиотеки с помощью технологии Illumina Solexa.....	36
3.4 Биоинформатический анализ последовательностей.....	37
3.4.1 Сборка транскриптома.....	37
3.4.2 Анализ экспрессии.....	37
3.4.3 Аннотация транскриптов.....	38
3.4.4 Филогенетический анализ.....	39
4 Результаты и обсуждение.....	40
4.1 Общая характеристика полученных данных.....	40
4.2 Состав доминантных компонент транскриптома.....	43
4.2.1 HDM белки.....	43
4.2.2 Миоглобин.....	45
4.2.3 Вителлин (белок оболочки яйца).....	47
4.2.4 Цистеиновые протеазы.....	48
4.2.5 28 кДа глутатион трансфераза.....	49
4.2.6 Транскрипты кДНК библиотеки, кодирующие экскреторно-секреторные белки.....	51
4.3 Транскриптом <i>O. felineus</i> и консервативные гены эукариот.....	52
4.4 Классификация предсказанных белков <i>O. felineus</i> терминах Генной онтологии.....	55
4.5 Анализ метаболических путей с помощью KEGG.....	58
4.6 Транскриптом метацеркарии <i>O. felineus</i> – сравнение со взрослой стадией.....	61
4.7 Рибосомальные белки и филогения описторхид.....	66
4.8 Потенциальная канцерогенность <i>O. felineus</i>	69
5 Заключение.....	72
6 Выводы.....	75
7 Список сокращений.....	76
8 Список литературы.....	77
9 Приложения.....	91

1 Введение

1.1 Актуальность темы исследования и степень ее разработанности

Описторхиды представляют собой семейство паразитических плоских червей (класса Трематод), включающее более 30 родов [1], среди которых есть представители с большой медицинской значимостью, такие как *Opisthorchis felineus*, *O. viverrini* и *Clonorchis sinensis* [2]. Эти виды вызывают опасные хронические заболевания – описторхоз и клонорхоз, приводящие к повреждениям гепатобилиарной системы [3].

O. viverrini и *C. sinensis* распространены в таких странах как Китай, Тайланд, Лаос, Камбоджия и некоторые другие, и поэтому называются «азиатскими» печёночными сосальщиками. Случаи же заражения *O. felineus* в основном наблюдаются в России, Белоруссии, Казахстане, особенно, в бассейне сибирской реки Обь, где встречаемость заболевания может достигать 30% локальной популяции [4]. Менее часто встречаются случаи заражения им людей и домашних животных в западной Европе [5,6].

Описторхоз может протекать как асимптоматически, так и в тяжёлой форме. При отсутствии своевременного лечения инфекция чревата множественными осложнениями, которые включают повреждения печени и поджелудочной железы, холангит и образование камней в желчевыводящих путях [7]. Кроме того, *C. sinensis* и *O. viverrini* классифицированы как канцерогены класса 1: показано, что *O. viverrini* является одной из основных причин холангиокарциномы в Тайланде [8,9]. На настоящий момент неясно, способен ли *O. felineus* вызывать рак, но филогенетическое родство с двумя другими видами и данные о его более высокой патогенности по сравнению с азиатскими печёночными сосальщиками [10] позволяют предположить и его канцерогенность.

Не смотря на высокую медицинскую значимость этих паразитов, их диагностика и лечение представляют трудную медицинскую задачу. Для диагностики применяются такие методы, как копрологическое исследование и дуоденальное зондирование, которые, однако, низко чувствительны. Лечение же проводится в основном празиквантелом [11]. Массовое применение последнего может понизить эффективность лечения [12] и потенциально, привести к формированию лекарственной устойчивости, что делает разработку новых методик в лечении этих паразитозов приоритетной проблемой для некоторых регионов.

Создание новых подходов в лечении описторхоза тормозится низким уровнем его

изученности. Описторхиды являются немодельными организмами и мало что известно об их биологии на молекулярном уровне. Более того, отсутствие полногеномных или полнотранскриптомных сведений дополнительно затрудняют молекулярно-биологические исследования описторхид. В связи с этим в последние 5 лет были сделаны усилия по расшифровке геномов и транскриптомов азиатских печеночных сосальщиков *O. viverrini* и *C. sinensis* [13–15], но соответствующие данные для *O. felineus* отсутствуют. В связи с этим мы решили исследовать транскриптом этого вида методами РНК секвенирования на двух стадиях его развития: взрослой стадии, обитающей в желчевыводящих протоках окончательного хозяина и инвазивной стадии (метацеркарии), обитающей в рыбе – промежуточном хозяине.

1.2 Цели и задачи

В связи с актуальностью проблемы описторхоза мы поставили перед собой следующую цель исследования:

Полнотранскриптомный анализ мариты и метацеркарии *O. felineus*.

Для осуществления поставленной цели были сформулированы задачи:

1. Определить последовательности белок-кодирующих мРНК транскриптомов мариты и метацеркарии *O. felineus*.
2. Провести биоинформатический анализ полученных транскриптов.
3. Провести сравнительный анализ транскриптомных профилей мариты и метацеркарии *O. felineus*.

1.3 Научная новизна работы

В данной работе впервые отсеквенирован и аннотирован полный транскриптом трематоды *O. felineus* на стадиях мариты и метацеркарии. Обнаружены особенности генной организации этого паразита в свете его образа жизни и специфики взаимодействия паразит-хозяин. В частности оказалось, что у *O. felineus* редуцированы такие молекулярные системы, как синтез полиаминов и метаболизм метионина, а также полностью отсутствует молекулярный аппарат формирования пероксисом. С другой стороны, выявлено, что катепсины, кальмодулины и MD-2 домен содержащие белки представлены большим числом копий, чем характерно для других эукариот. Обнаружены новые гранулиноподобные белки, не имеющие ортологов у родственных описторхид.

Полнотранскриптомные данные для инвазивной стадии азиатских печеночных сосальщиков отсутствуют, таким образом транскриптом метацеркарии описторхид

отсеквенирован впервые. В результате сравнения двух исследованных стадий развития были установлены существенные различия в их транскриптомных профилях: помимо генов экспрессирующихся на обеих стадиях (около 11 тыс.) развития имеются 895 и 632 гена, экспрессирующихся только в марите или метацеркарии соответственно.

Установлены филогенетические взаимоотношения внутри семейства описторхид с использованием рибосомальных белков. Оказалось, что *C. sinensis* и *O. felineus* более близки друг к другу, чем каждый из них к *O. viverrini*.

1.4 Теоретическая и практическая значимость работы

Транскриптом организма, исследуемого в данной работе отсеквенирован впервые. Также, метацеркарии описторхид не были ранее исследованы с помощью методов полнотранскриптомного секвенирования. Полученные EST последовательности находятся в базе данных NCBI под номерами JK624271–JK626790, JK006511–JK006547, JK649790–JK649792, а белок-кодирующие транскрипты, полученные на основании данных массового параллельного секвенирования находятся в TSA архиве базы данных NCBI под номером GBJA01000000000. Находящиеся в общем доступе нуклеотидные последовательности образуют важную основу для дальнейших молекулярно-биологических исследований *O. felineus*. Кроме того, обнаруженные особенности генетической организации являются важным источником знаний об особенностях явления паразитизма на примере этого печёночного сосальщика.

1.5 Положения выносимые на защиту

1. У *O. felineus* редуцированы молекулярные системы синтеза полиаминов и метаболизма метионина, а также отсутствует молекулярный аппарат формирования пероксисом.
2. У мариты *O. felineus* среди наиболее экспрессируемых идентифицированы гены катепсина F, миоглобина, белка оболочки яйца, глутатион трансферазы, HDM белка, а у метацеркарии – гены, кодирующие белки домашнего хозяйства, такие как рибосомальные белки и убиквитин.
3. Филогенетически *O. felineus* и *C. sinensis* ближе друг к другу, чем каждый из них к *O. viverrini*.

1.6 Степень достоверности и апробация результатов

Достоверность полученных последовательностей мРНК является ключевой для всех последующих рассуждений данной работы. Так как были использованы две методики

конструирования библиотек и секвенирования, данные подвергались взаимной проверке. Методика секвенирования по Сэнгеру считается более надежной, поэтому этап сборки транскриптов из прочтений массового параллельного секвенирования обрабатывался по имеющимся EST последовательностям: нами были опробованы несколько ассемблеров и протоколов сборки, пока не были получены результаты, согласующиеся с данными секвенирования по Сэнгеру.

Основные результаты и выводы данной работы, построенные на основании качественной и количественной информации о транскриптоме *O. felineus* были представлены и обсуждались на следующих российских и международных конференциях:

1. Pomaznoy M., Tatkov S., Brusentsov I., Sivkov A., Nayakshin A., Guselnikov S., Brenner E., Vasilev G., Katokhin A., Mordvinov V. Search of potential targets for diagnostics of parasite *Opisthorchis felineus* // *Molecular Diagnostics 2010.* , 2010. P. 120–1.

2. Tatkov S., Brusentsov I., Pomaznoy M., Nosareva O., Nayakshin A.M., Guselnikov S. V, Brenner E. V, Vasiliev G. V, Katokhin A. V, Mordvinov V.A. Investigation of *Opisthorchis felineus* transcription profile by direct sequencing cDNA library's clones // *7th International conference on bioinformatics of genome regulation and structure\system biology (BGRS/SB'2010).* , 2010.

3. Помазной М., Логачева М., Катохин А., Мордвинов В. Полнотранскриптомный анализ возбудителя описторхоза *Opisthorchis felineus* методом RNA-seq // *Open Bio. Koltsovo:* , 2014. P. 131–2.

2 Обзор литературы

2.1 Краткая характеристика *O. felineus* и вызываемого им заболевания

2.1.1 Систематическое положение

Общее систематическое положение *O. felineus* описывается следующим образом:

Надцарство	Eukaryota
Царство	Metazoa (животные)
Тип	Platyhelminthes (плоские черви)
Класс	Trematoda (сосальщикообразные)
Подкласс	Digenea (двуустки)
Отряд	Opisthorchiida
Подотряд	Opisthorchiata
Семейство	Opisthorchiidae
Род	<i>Opisthorchis</i>
Вид	<i>felineus</i>

Помимо положения в классической систематике довольно важным является положение в «новой филогении животных», построенной на основании нуклеотидных и белковых последовательностей соответствующих видов [16]. По этой классификации все билатерально-симметричные животные разделились на три крупные монофилетические группы Лофотрохозоа, Экдизозоа и Вторичноротые. Данная гипотеза за вот уже полтора десятилетия с момента опубликования нашла множество подтверждений. Согласно этой классификации тип плоских червей принадлежит к Лофотрохозоа (Рисунок 1) [17]. Подобное филогенетическое расположение Platyhelminthes подтверждается не только при анализе рибосомальных генов, но и других генетических семейств, например, аннексинов [18]. Особенностью Лофотрохозоа является отсутствие модельного организма в отличие от других крупных систематических групп. Это затрудняет прогресс молекулярно-биологических исследований, направленных на искоренение заболеваний, вызываемых описторхидами.

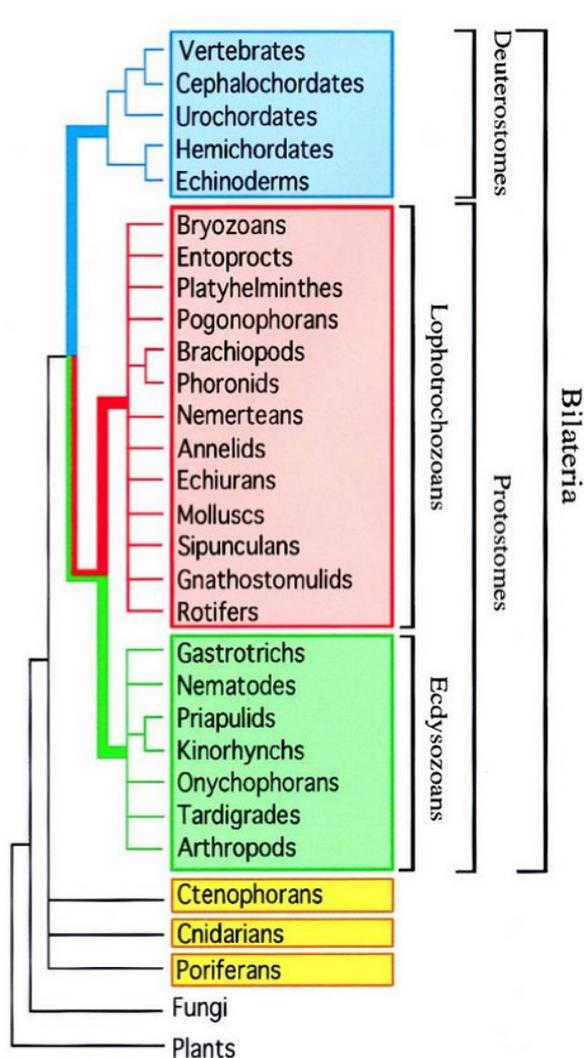


Рисунок 1: Филогенетические взаимоотношения животных согласно «новой филогении животных», построенной на основании анализа нуклеотидных и белковых последовательностей. Следует обратить внимание на достаточно позднее отделение плоских червей от общего с нематодами ствола, что говорит о нецелесообразности рассмотрения плоских червей как базового типа для всех билатеральных животных (взято из [Adoutte и др., 2000]).

Класс трематода состоит целиком из паразитических видов. В данной работе я буду часто обращаться к семействам шистосоматид (кровяные сосальщики) и описторхид (печёночные сосальщики). Первые примечательны тем, что в них входят шистосомы – важнейшие паразиты человека. Вызываемый ими паразитоз по количеству жертв в мире уступает только малярии. Их относительная филогенетическая близость к описторхидам (по сравнению с классическими объектами молекулярной биологии) и относительно высокая степень изученности позволяет во многом использовать известные о них данные и для исследования печеночных сосальщиков.

Из 33 родов [1], входящих в семейство описторхид следует особенно отметить три вида, преобладающие на сегодняшний день как паразиты человека и животных: это *O. felinus*,

O. viverrini и *C. sinensis*, первый из которых является объектом настоящего исследования. Что касается взаимоотношений этих трёх видов внутри семейства описторхид, то известно, что они очень близкородственны, однако, на их взаимоотношения внутри клады существуют различные точки зрения: согласно первой, *O. felineus* и *C. sinensis* ближе друг к другу, чем к *O. viverrini*; согласно второй, *O. viverrini* и *C. sinensis* ближе друг к другу, чем к *O. felineus*. Обе точки зрения находят подтверждение как и по морфологическим признакам, так и в результатах молекулярно-генетических исследований с применением митохондриальных и ядерных маркеров [19,20]. Ранее, при исследовании единичного ядерного маркера Pm-int9, включающего 9-й интрон гена парамиозина с фланкирующими экзонными последовательностями [20], были получены свидетельства в пользу второй точки зрения.

2.1.2 Жизненный цикл

Описторхидам требуется 3 различных вида животных для завершения паразитического жизненного цикла: моллюск, рыба и позвоночное животное, являющееся окончательным хозяином, в котором происходит половое размножение (Рисунок 2) Для большинства описторхид характерна полигостальность, то есть широкий спектр возможных окончательных хозяев, который может варьировать в пределах представителей как различных семейств, так и различных классов.

Взрослая стадия паразита (марита), паразитирует в желчных ходах печени и поджелудочной железе окончательного хозяина. Мариты в большом количестве производят яйца. Для продолжения жизненного цикла яйца паразита должны попасть в водоём, а там в пищеварительный тракт определённого вида моллюсков. Как правило, первым промежуточным хозяином являются представители семейства Vithyniidae [21]. Мирацидий – жизненная форма паразита, формирующаяся в яйце – выходит из яйца и проникает через стенку кишечника в полость тела моллюска, где образует спороцисту – следующую жизненную форму паразита. Из спороцист развиваются или новые спороцисты, или редии, активно питающиеся тканями моллюска. Внутри редий образуются либо новые поколения редий, либо церкарии – свободно живущие жизненные формы паразита. Церкарии выходят из моллюска в водную среду, откуда они должны попасть во второго промежуточного хозяина – рыбу.

При проникновении сквозь чешую рыбы паразит инцистируется образуя метацеркарию. Последняя находится в ткани рыбы, пока вместе с ней не будет поглощена окончательным хозяином. В качестве второго промежуточного хозяина как правило выступают рыбы семейства карповых, однако, есть данные что *C. sinensis* может переноситься и ракообразными [22]. Если заражённость первого промежуточного хозяина в природных очагах описторхоза невелика

(доли процента), то карповые рыбы, находящиеся в близлежащих водоёмах инфицированы как правило на 75% и более [21].

После употребления инфицированной рыбы окончательным хозяином, метацеркарии эксцистируются в двенадцатиперстной кишке, молодые мариты мигрируют в общий желчный проток через ампулу Фатера и проникают в желчные протоки. Мариты достигают половой зрелости в течение 3-4 недель, после чего начинают откладывать яйца.

Окончательным хозяином описторхид помимо человека могут быть рыбацкие животные как естественного, так и антропогенного происхождения. Так на территории Европы естественным резервуаром *O. felinus* являются лисы [23,24] и представители семейства куньих [25,26], на территории России и в восточной Европе антропогенный фактор начинает сильнее преобладать и зараженность кошек, собак и людей становится более выраженной [27]. Для *O. viverrini* и *C. sinensis* антропогенный фактор является определяющим [21].

Таким образом, в ходе жизненного цикла описторхиды претерпевают множество крупных морфологических перестроек (или метаморфозов), сопровождающихся глубокими изменениями на органном, тканевом, клеточном и молекулярном уровнях. Ясно, что все эти изменения отражены в профилях транскрипции соответствующих стадий и являются непосредственной причиной изменения белкового состава всех клеток. Полнотранскриптомное исследование различных жизненных форм могло бы дать существенную информацию об этих процессах, однако, подобных исследований описторхид еще не проводилось.

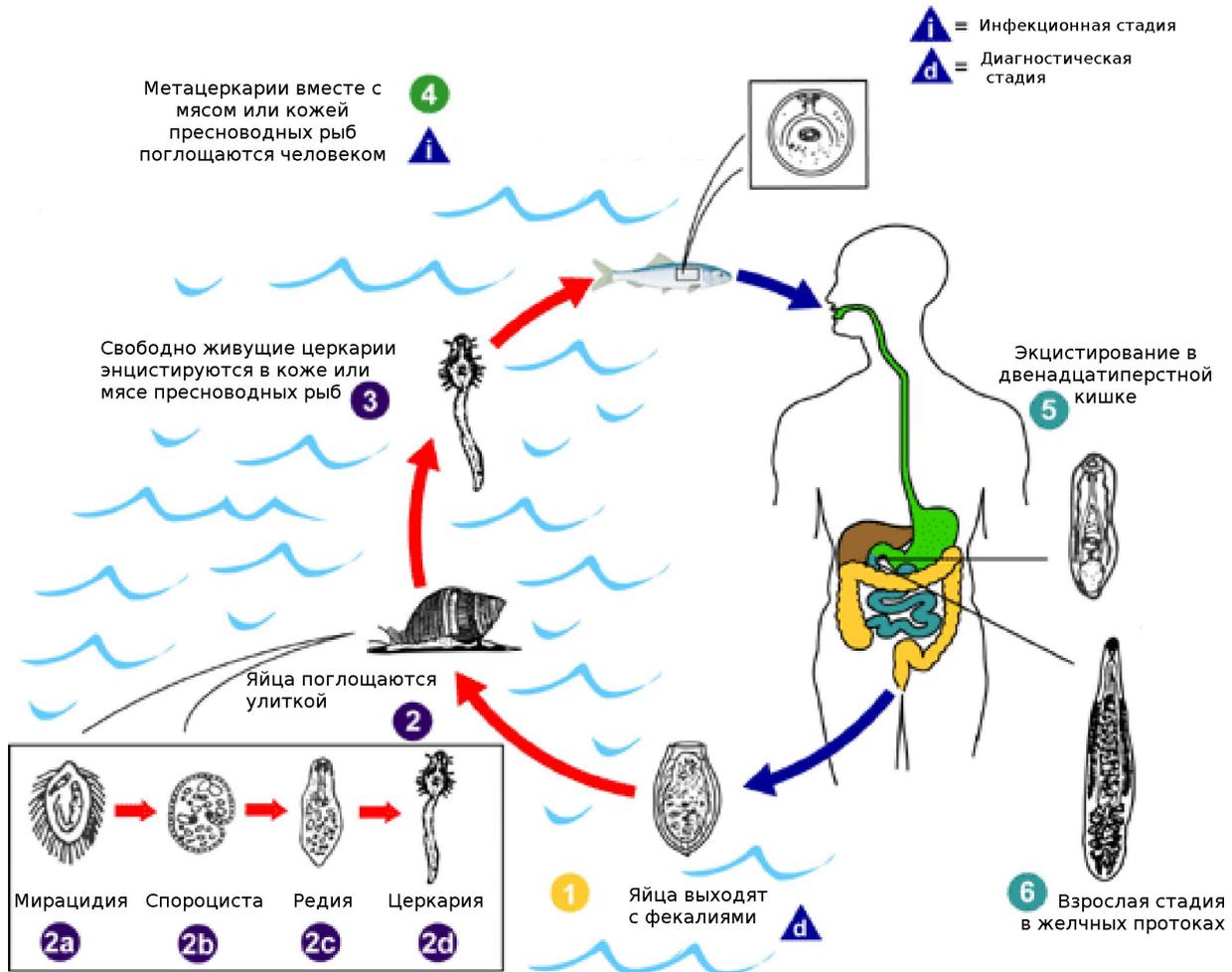


Рисунок 2: Жизненный цикл описторхид на примере *O. felineus* (переработано с сайта [183])

2.1.3 Ареал распространения

Под риском заражения печеночными сосальщикам родов *Clonorchis*, *Fasciola* (также представители подкласса дигеней) и *Opisthorchis* находятся 601 миллион, 91 миллион и 80 миллионов людей соответственно [28]. Описторхиды в основном распространены по евразийскому континенту. В частности, *C. sinensis* распространен в регионах Китая, Кореи и Северного Вьетнама и Японии [29], а *O. viverrini* во Вьетнаме, Лаосе, Тайланде и Камбоджи [7,8,30]. Паразитом *O. felineus* заражены люди в основном в республиках бывшего СССР и некоторых странах Восточной Европы [4], хотя имеются данные о случаях заражения человека в западной Европе [6,31] (Рисунок 3).

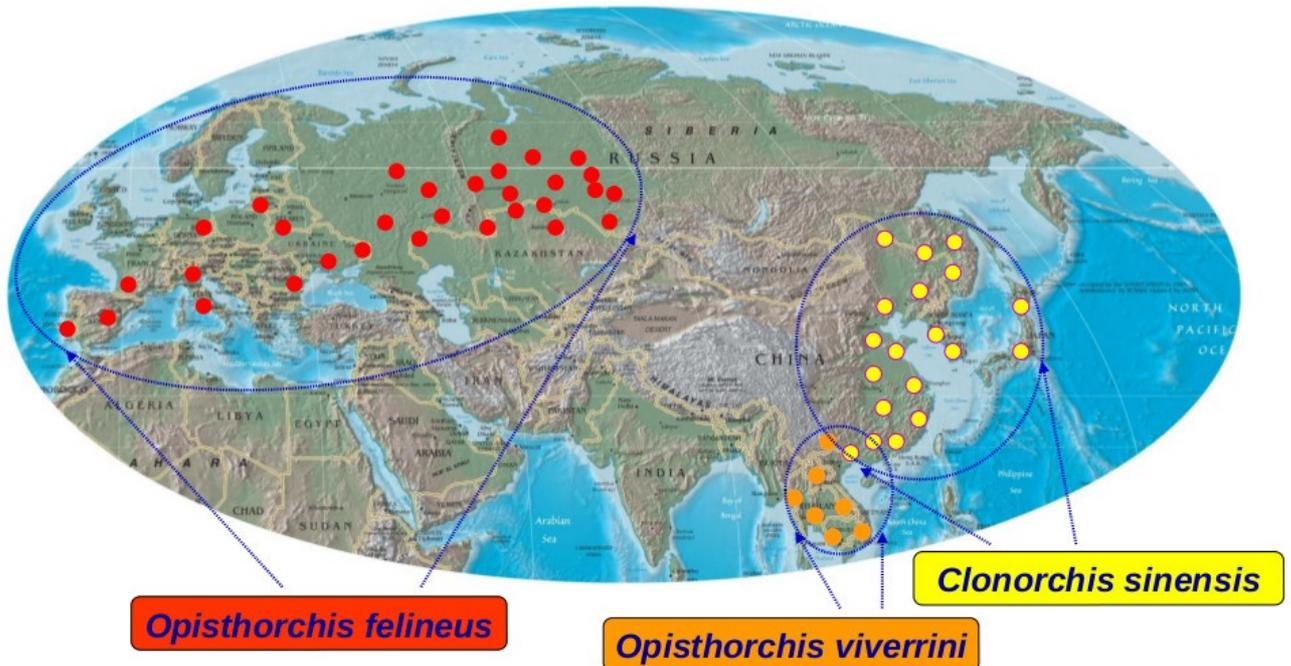


Рисунок 3: Ареалы обитания патогенных для человека видов описторхид

2.1.4 Описторхоз как заболевание

При заражении метацеркариями *O. felineus* острая фаза наступает в среднем через одну неделю. Симптомами острой фазы являются жар, боль в животе, астения, артралгия, диарея и тошнота [32]. Иногда острая фаза может проходить асимптоматически [5]. Выраженность проявлений заболевания зависят от индивидуальных особенностей и степени заражения. При отсутствии своевременного лечения (к примеру, при асимптоматическом течении инфекции) наступает хроническая фаза, в которой внешние проявления и лабораторные показатели могут прийти в норму [5]. Одним из наиболее опасных и удаленных по времени последствий заражения описторхозом, вызванным *O. viverrini*, является холангиокарцинома. Канцерогенность этого вида подтверждена в экспериментах на животной модели, а также объясняет частую встречаемость холангиокарциномы в районах Тайланда, в которых степень заражения описторхозом высока (см., например, [9] и ссылки там). Канцерогенность *O. felineus* пока не была экспериментально доказана, хотя есть данные о его большей патогенности по сравнению с *O. viverrini* и *C. sinensis* [10].

Диагностика заболеваний, вызываемых описторхидами, является медицински трудной задачей. В случае же позитивного диагноза часто трудно определить каким именно представителем описторхид произошло заражение, т. к. морфология яиц очень схожа [33]. Основным методом, применяемым на практике является копрологическое исследование,

имеющее свои недостатки, связанные со сложностями стандартизации и низкой чувствительностью [34]. Также применяется дуоденальное зондирование, при котором визуально исследуется содержимое желчи на наличие яиц гельминтов. Общим недостатком этих методов является то, что их результаты плохо коррелируют с самой важной количественной характеристикой заражения – паразитарной нагрузкой.

Существуют иммунологические методы выявления инфекции, использующие тотальный лизат паразита. Такая иммунодиагностика клонорхоза широко применяется в Корее [35]. Выделенный экскреторно-секреторный продукт является более эффективным, чем тотальный лизат [36], однако, его значительно труднее наработать в достаточных количествах. В целом главной проблемой методов детекции антител в крови пациентов является длительное время присутствия антител. Даже спустя месяцы после излечения и отсутствии паразитарной нагрузки иммунологические методы могут дать ложно позитивные результаты [34].

Лечение описторхоза в основном проводится празиквантелом [37], и хотя были обнаружены и некоторые другие вещества с выраженным антигельминтным эффектом (например, плюбангин [38]), первый является единственным применяемым в клинике средством. В нашей лаборатории ранее было показано наличие популяционной гетерогенности *O. felineus* по ответу на воздействие празиквантелом [39]. В долгосрочной перспективе наибольшую опасность использования одного и того же средства представляет возможность возникновения лекарственно устойчивой популяции.

Таким образом, создание новых препаратов для лечения трематодозов является важной задачей, однако, плохая изученность этих видов и отсутствие модельного организма для Лофотрохозоа затрудняют процесс молекулярно-биологических исследований, направленных на поиск новых лекарств. Заложить основу таких исследований могут помочь транскриптомный и геномный подходы, ставшие легко доступными с развитием технологий массового параллельного секвенирования. Имеющиеся на сегодняшний день результаты в области функциональной геномики и транскриптомики трематод освещены в следующей части обзора литературы.

2.2 Функциональная геномика и транскриптомика трематод

2.2.1 Методология исследования транскриптомов

Технологии секвенирования ДНК прочно вошли в методологическую базу современной биологии. За последние тридцать лет в этой области произошло множество прорывных изменений, которые ознаменовали начало так называемой геномной эры в биологии. Первой методикой секвенирования ДНК был метод, предложенный Фредериком Сэнгером в 1977г. [40]. Этот метод вскоре стал основным в лабораторной практике, а с некоторыми изменениями он дошел и до сегодняшнего дня. Основными модификациями, позволившими осуществлять секвенирование по Сэнгеру в коммерческом масштабе были использование флюоресцентных аналогов дидезоксинуклеотидов и применение капиллярного электрофореза. В таком виде метод используется широко и по сегодняшний день как надежный и дешевый способ рутинного лабораторного секвенирования, *post factum* называемый методом «первого поколения». Использование сэнгеровского секвенирования для исследования транскриптомов сводится к синтезу молекул кДНК, их клонированию в прокариотическую систему с последующим секвенированием клонов полученной библиотеки.

Развитие технологий привело к появлению методик NGS секвенирования (или методик «следующего поколения»). Их основным отличием является массовый параллельный анализ множества (на данный момент миллиардов) молекул ДНК. Массовость получения прочтений на NGS платформах и относительная дешевизна (точнее, цена, отнесенная к количеству прочитанных нуклеотидов) произвели качественный прорыв в молекулярной биологии. Стало возможным относительно легко получить полногеномные и полнотранскриптомные данные. На сегодня NGS методики реализованы в нескольких коммерческих платформах, каждая из которых имеет свои достоинства и недостатки. За деталями сравнения этих платформ можно обратиться к одному из множества существующих в данной области обзоров (которые, однако, очень быстро устаревают в связи с бурным развитием этой отрасли) [41,42].

По мере развития NGS технологий их стали активно применять и при исследовании немодельных организмов, в частности паразитов. Первые усилия в области транскриптомики и геномики трематод были брошены на исследование видов *Schistosoma mansoni* и *S. japonicum* в связи с наибольшей социо-экономической значимостью данных видов по сравнению с другими представителями класса. Статус исследований геномов наиболее значимых трематод на момент

написания этого текста с соответствующими ссылками представлены в сводной таблице 1.

Таблица 1: Сводная таблица, отражающая наличие геномных данных для основных паразитов класса *Trematoda* (на 2015 г.)

Вид	Геномные данные	Ссылки
<i>S. mansoni</i>	9'516 контигов (N50 76'721); 885 скаффолдов (N50 32'115'376)	[43–50]
<i>S. japonicum</i>	95'269 контигов (N50 6'121); 25'048 скаффолдов (N50 176'869)	[51]
<i>Fasciola hepatica</i>	195'709 контигов (N50 12'161); 20'158 скаффолдов (N50 204'014)	[52]
<i>C. sinensis</i>	6'190 контигов (N50 233'037); 4'348 скаффолдов (N50 417'486)	[14,15,53–56]
<i>O. viverrini</i>	50'095 контигов (N50 19'046); 158'119 скаффолдов (N50 191'538)	[13,15,57,58]

В своей работе мы использовали секвенирование на платформе Illumina HiSeq 2000, т. к. традиционно она превосходит конкурентов по таким параметрам, как производительность (количество полученных за запуск прибора данных) и цене, отнесенной к количеству прочитанных нуклеотидов.

2.2.2 Метаболизм трематод

Белки, связанные с метаболизмом составляют основную часть всех транскрипционно активных генов трематод. В связи с переходом к паразитическому образу жизни многие другие системы органов (к примеру нервная система и органы чувств) подвергаются значительному упрощению, однако метаболические потребности остаются высокими, что необходимо для производства большого количества потомства, поэтому экспрессия метаболических ферментов в целом достаточно высока.

Внутри группы ферментов метаболизма существенную фракцию составляют протеазы. Все их многообразие на основании механизма гидролиза подразделяется на сериновые, треониновые, аспаратные, цистеиновые и металло-протеазы. В относительно недавнее эволюционное время у паразитических трематод произошла геномная экспансия семейства цистеиновых протеаз, причем большинство из них относятся к катепсинам – их более узкому классу. Прежде всего протеазы трематод являются ферментами пищеварения: катепсин В *S. mansoni* способен очень эффективно расщеплять человеческий гемоглобин [59], а для катепсина В *F. hepatica* показана высокая пептидазная активность против IgG и альбуминов крови [60]. Образующиеся аминокислоты могут быть захвачены имеющимися транспортерами,

обнаруживаемыми в геномах описторхид [13]. Катепсин L *F. hepatica* способен активно расщеплять коллаген, в связи с чем рассматривается его потенциальная роль в инвазии и проникновении через ткани хозяина [61]. Имеются данные и о способности этих ферментов участвовать в модуляции иммунного ответа [62–65]. О важности этих ферментов для трематод говорят эксперименты, в которых целенаправленное ингибирование катепсинов В на фасциоле (как с использованием биохимического ингибирования, так и с помощью РНК-интерференции) приводит к понижению двигательной функции, неспособности проникать в ткань носителя, и к смерти через 14 часов [66].

Анализ EST *C. sinensis* показывает, что катепсины составляют значительную часть всех активно транскрибируемых генов, при этом основная масса протеаз этого паразита гомологичны семейству катепсинов F [53,54]. Кроме того, в недавнем исследовании на *O. viverrini* показано присутствие катепсинов в клетках желчных путей заражённых животных [65], в связи с чем предполагается что эти белки играют важную роль в воспалении, а значит могут участвовать в канцерогенезе. Основными секретируемыми протеазами *O. viverrini* являются катепсины В и F, совместное действие которых расщепляет человеческий гемоглобин эффективнее, чем каждый из них по отдельности [67]. Эти белки синтезируются в виде менее активных прозимогенов и активируются как автокаталитически, так и под действием транс-активирующих протеаз. Считается что все вместе они образуют сеть взаимодействующих ферментов, действующих на ткани хозяина. Большая важность различных семейств протеаз в развитии и жизнедеятельности паразитических трематод делает их важными мишенями для новых вакцин и терапевтических агентов [63,68,69].

Помимо белков, в качестве источников углерода и энергии могут выступать и другие соединения, например жирные кислоты и липопротеины, что более характерно для описторхид в отличие от кровяных сосальщиков, т.к. первые обитают в желчных протоках. В желчевыводящей системе млекопитающих много желчных и жирных кислот [70], которые могут являться важным нутриентом для описторхид. В подтверждение этого в геноме клонорха были обнаружены все гены, участвующие в β -окислении жирных кислот, в то время как у шистосомы – только 4 гена [71]. При этом система синтеза жирных кислот вторично редуцирована и у печеночных и кровяных сосальщиков: у шистосом есть только три соответствующих гена [14]. Так как цикл трикарбоновых кислот имеется в полноценном виде и у тех и у других, то можно утверждать, что трематоды обладают всей системой извлечения энергии из углеводородного остова жирных кислот. Однако катаболизм этих соединений требует больших количеств кислорода. Возможно это является причиной больших количеств миоглобина, синтезируемых печеночными сосальщиками и его большой аффинности к

кислороду [72]. Высокая транскрипция миоглобина характерна для всех описторхид (см. раздел 4.2.2).

2.2.3 Покровные ткани

Эпителиальные ткани трематод представлены поверхностным тегументом, гастродермисом, покрывающим изнутри пищеварительные органы и эпителием, выстилающим протонефридиальные каналы. Наружный тегумент представляет собой клеточную плазматическую мембрану, секретирующую мембранокаликс.

Питание паразита во многом осуществляется внешними механизмами и в этом особенно способствуют процессы эндоцитоза и транспорта высокомолекулярных соединений. У шистосомы обнаруживается более десятка белков транспортеров различных сахаров, а также липидов, аминокислот и нуклеотидов. Обнаруживаются все гены, составляющие молекулярную основу процесса эндоцитоза [47].

Тегумент соприкасается с внешней для паразита средой (т.е. внутренней средой организма хозяина), и поэтому активно взаимодействует с иммунной системой, что в целом определяет именно белки тегумента как потенциальные мишени для вакцин против шистосом и трематод вообще [69]. При исследовании *S. mansoni* были обнаружены транскрипты белков тетраспанинов, функция которых у трематод остаётся плохо изученной. Однако, иммуногистохимически была подтверждена их поверхностная локализация. Далее было показано, что сыворотки людей, перенесших шистосомоз проявляют иммуногенность к этим белкам. Наконец, эксперименты на животных моделях подтвердили эффективность предварительной иммунизации рекомбинантными тетраспанинами [73].

Другой пример использования белков тегумента для иммунизации отражен в работе [74]. Авторы использовали рекомбинантный аналог, сшитый с белком спор *Bacillus subtilis*. Белок показал значительную эффективность и перспективность для дальнейшей разработки. Есть и другие работы по использованию белков тегумента для иммунизации в виде рекомбинантов, а также ДНК вакцин [75].

2.2.4 Строение тела и двигательная система

В работе [43] было проведено сравнение генома шистосомы с морской губкой *Nematostella vectensis*, что, как считают авторы, позволяет выделить изменения генома, произошедшие при переходе к билатеральной симметрии, как общего плана строения тела. Выяснилось, что у трематоды увеличено количество генов, которые вовлечены в клеточную адгезию (кадгерины).

Также, семейство белков тетраспанинов, упоминавшихся выше, представлено большим числом копий в геноме шистосомы.

На уровне генома выявлено, что шистосомы имеют молекулярный аппарат, необходимый для нейрогенеза и контроля роста аксонов и миграции нервных клеток. Последнее наблюдение, с учетом большого филогенетического расстояния между шистосомами и высшими позвоночными, подтверждает древнее происхождение механизмов нервной системы.

Мышечная система трематод характеризуется преобладанием гладкой разновидности. У *S. mansoni* обнаружены гены нескольких типов миозинов и актинов, тропомиозин, парамиозин, тропонины С, I и Т, транскрипты белков, участвующих в прикреплении актиновых филаментов – альфа-актинин, винкулин и титин. Транскрипты, соответствующие белкам, специфичным к поперечно-полосатой мускулатуре не найдены [43].

При развитии систем органов и плана тела ключевыми молекулярными регуляторами являются гомеодомен содержащие и НОХ гены (или гомейозисные гены), что характерно для всех изученных эукариотических организмов [76]. Так, у *S. mansoni* обнаруживается 9 НОХ генов [77], что меньше, чем количество паралогических НОХ групп у вторичноротых. Другой особенностью организации НОХ генов трематод является то, что эти гены не образуют уникального (в смысле единственности по геному) упорядоченного кластера [78], четко характерного для млекопитающих. У последних они представлены в виде четырех кластеров, образовавшихся в результате дупликации, которые вместе составляют 39 генов тринадцати различных типов (например, [79]).

У *S. japonicum* обнаруживается 8 НОХ генов, а именно Нох1, Нох2, Нох3, Нох4, Лох5, Лох4 и Post-2. Что характерно для всех плоских червей, Нох5, Лох2, Post-1 и Нох7 гены отсутствуют, хотя по имеющимся данным характерны для предка всех Лофотрохозоа (Рисунок 4) [80].

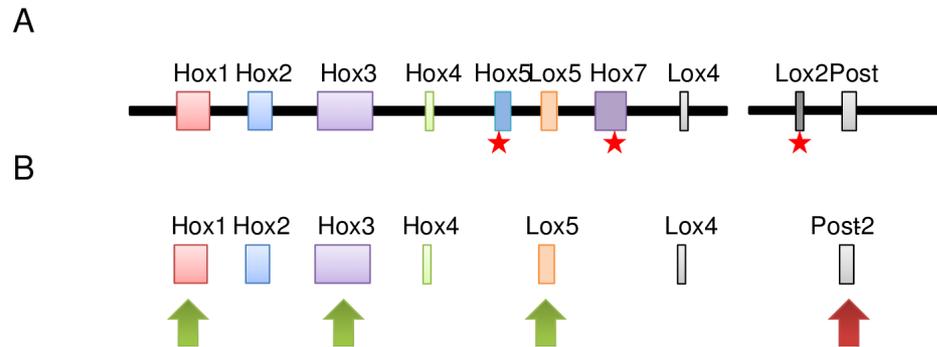


Рисунок 4: Предковый кластер НОХ генов у ближайшего общего предка всех Лофотрохозоа. Звездочками показаны гены, потерянные в ходе эволюции у плоских червей. Б. Набор НОХ генов, характерный для плоских червей. Стрелками показаны гены, которые имеются в количестве 2х в геноме *S. japonicum* и некоторых других плоских червей. (по [80])

2.2.5 Пути трансдукции сигналов

Наши знания о нервно-гуморальной регуляции функций трематод очень фрагментарны, и геномика является пока единственным источником знаний о генах трематод, участвующих в передаче сигнала. Изучение нервных и гуморальных путей регуляции является крайне интересной задачей с точки зрения разработки новых терапевтических средств, т.к. их блокада как правило сопровождается существенными нарушениями общего гомеостаза. Очень многие фармацевтические препараты направлены на регуляторные системы передачи сигнала [81].

В транскриптоме шистосомы обнаружены лиганд-зависимые каналы (никотиновые рецепторы ацетилхолина, мускариновые рецепторы), холин О-ацетилтрансфераза и ацетилхолин эстераза, что говорит о полной представленности ацетилхолинэргической составляющей нервной системы [47]. Также найдены рецепторы таких распространенных медиаторов, как глутаминовая и гамма-аминомасляная кислоты. Обнаруживаются GPCR рецепторы глутамата и возбуждающего трансмиттера серотонина (там же). Хотя установлено, что *S. mansoni* может отвечать на воздействие катехоламинами, транскриптов соответствующих рецепторов не найдено.

Некоторые молекулы передачи нейромедиаторного сигнала также были обнаружены у трематод: у шистосомы найден предшественник аллатостатина, а также белок гормон пропетид конвертаза 2, ее регуляторный белок и фермент глицин-пептидил- α -амид монооксигеназа, участвующая в синтезе пептидных гормонов. Обнаруживаются транскрипты ядерных рецепторов (retinoid-X и fushi tarazu factor 1), рецепторы ретиноевой кислоты, рецепторы тиреоидных гормонов и др [47].

Любопытно, что у *S. mansoni* найден ортолог тиронид-пероксидазы, необходимый для синтеза тиреоидных гормонов, но тироглобулин – ее типичный субстрат у высших эукариот – отсутствует в транскриптоме. Возможно, шистосомой используется другой тирозин-богатый субстрат, хотя пока подтверждений этому нет.

Наличие нескольких цитохромов P450, тестостерон 6- β -гидроксилазы и 17 β -гидроксистероид дегидрогеназы говорят о способности шистосомы осуществлять собственный синтез стероидных гормонов из холестерина. Обнаружены и белковые молекулы, способные передавать сигналы стероидных гормонов (в том числе возможно и экзогенных, т.е. принадлежащих хозяину) [47]. Последний факт, вместе с наличием рецепторов инсулина и фактора роста фибробластов и в то же время отсутствием мРНК их лигандов подтверждает концепцию воздействия экзогенных гормонов на системы паразита [45]. Таким образом трематоды способны реагировать на гормональный статус хозяина и, вероятно, каким-то образом адаптироваться к его системно-гормональным изменениям. В экспериментах по воздействию TGF на мариту шистосомы в системе *in vitro* и полнотранскриптомном сравнении с контрольными паразитами было показано значимое изменение экспрессии 386 генов [82].

2.2.6 Стволовые клетки трематод

Регенеративный потенциал плоских червей удивительно высок. Известно, что планарии способны восстанавливать целый организм лишь из некоторой его части [83], что сделало таких представителей Platyhelminthes как *Schmidtea mediterranea* и *Dugesia japonica* классическими объектами для изучения регенерации и роли стволовых клеток в ней [84]. Есть данные и о регенеративных способностях паразитических плоских червей: шистосомы могут восстанавливать поврежденные органы после воздействия сублетальных доз празиквантела [85]. Стволовые клетки шистосом были обнаружены в работе [86]. Авторы продемонстрировали наличие стволовых клеток и их схожесть по морфологии и некоторым элементам молекулярно-генетического статуса с необластами планарии. Они же показали необходимость FGF сигнального пути в поддержании статуса плюрипотентности обнаруженных клеток. Роль FGF/Erk сигналинга для регуляции плюрипотентности была продемонстрирована ранее на клетках млекопитающих [87].

С помощью методов геномики были обнаружены и явные отличия трематод от млекопитающих. Так у шистосомы отсутствуют такие важные гены поддержания плюрипотентного статуса клеток как *vasa* [88] (хотя обнаруживаются 3 *vasa*-подобных гена), *piwi* и *tudor* [86]. В то же время имеется гомолог *nanos*. В работе [89] показали отсутствие этих генов и в геноме ленточных червей. Возможно, у плоских червей есть и свои, особые

механизмы поддержания плюрипотентного статуса, либо очень древние, либо появившиеся только у этого таксона.

2.2.7 Гельминтозы и онкологические заболевания

Азиатские печеночные сосальщики классифицированы Агентством по исследованию рака как канцерогены 1ой группы [90]. Известно, что длительная инфекция *O. viverrini* и *C. sinensis* может привести к развитию холангиокарциномы [8,91]. Молекулярные механизмы, ответственные за возникновение этого смертельного заболевания, на сегодняшний день считаются мультифакторными [9,92] и включают: 1) механическое повреждение эпителия желчный протоков, 2) иммунопатология, связанная с паразит индуцированным воспалением и 3) токсические эффекты экскреторно-секреторных молекул.

Поиск компонентов экскреторно-секреторного продукта описторхид, позволил обнаружить [93] у *O. viverrini* гомологи гранулинов – ростовых факторов клеток млекопитающих [94]. Их функция в организме гельминта пока остаётся неизвестной, однако, удалось показать его способность даже в небольших дозах увеличивать митогенный потенциал эукариотических клеток [95]. Аналогичные эффекты обнаружены на северине *C. sinensis* – другом компоненте экскреторно-секреторного продукта (гомологе эукариотических гелизолинов – белков, ответственных за перестройки цитоскелета). Показана его способность ингибировать апоптоз в клетках человеческой гепатокарциномы [96].

Взаимосвязанными с онкогенезом считаются также калликреино-подобные сериновые протеазы, суперпродукция которых ассоциирована со злокачественной трансформацией клеток [97]. Такие протеазы также обнаруживаются у трематод.

Также здесь стоит упомянуть любопытную аналогию между паразитами и раковой опухолью, подмеченную авторами в работе [89]. И те и другие демонстрируют неконтролируемую пролиферацию, инвазию и метастазис. Как паразитирующий организм, так и опухоль преследуют свои интересы, а не организма носителя, и при этом, на них трудно воздействовать, не задевая окружающие ткани. В этом контексте особенно примечательно, что некоторые антигельминтные препараты (никлозамид, мебендазол и альбендазол) уже продемонстрировали противораковые эффекты [98].

2.2.8 Взаимодействие паразита и хозяина на гуморальном уровне

За долгое время совместной эволюции с хозяином паразиты установили с ним не только

тесные трофические связи, но и связи на уровне гуморальной регуляции. Показано, что паразиты способны реагировать на гормоны и факторы, циркулирующие в крови хозяина [82,99,100] (также см. раздел 2.2.5). Подобные взаимодействия носят двусторонний характер.

Интересный пример гуморального воздействия паразита на своего хозяина обнаружен в экспериментах на трематоды *F. hepatica*. При исследовании белковых фракций экскреторно-секреторного продукта этого паразита авторы натолкнулись на небольшой белок размером 8 кДа, обладавший ярко выраженными иммуномодулирующими свойствами. Основой для поиска функции вновь обнаруженного белка стало его сходство с кателицидинами млекопитающих на уровне вторичной структуры (Рисунок 5). Кателицидины (подкласс дефензинов) являются белками врождённого иммунитета человека, основными функциями которых является иммунологическая регуляция [101] и защита от некоторых патогенов, в основном посредством их лизиса [102]. Оказалось, что этот небольшой белок фасциолы, названный авторами HDM (от англ. Helminth Defense Molecule по аналогии с Host Defense Molecule у человека), проявляет как схожие с дефензинами функциональные характеристики, так и собственные, уникальные биологические свойства [103].

Имуномодулирующие эффекты HDM белка *F. hepatica* затрагивают как врождённый, так и приобретённый иммунитет млекопитающих и включают:

- связывание бактериальных липополисахаридов (LPS), предотвращая их взаимодействие с TLR системой [104], и тем самым, проявляя противовоспалительный эффект;
- ингибирование секреции TNF макрофагами [105];
- ослабление антигенного представления на поверхности макрофагов за счёт ингибирования лизосомальной АТФазы [106];
- наконец, молекулы HDM могут воздействовать и на В клетки, увеличивая продукцию IgG1 в IL-4 индуцированных лимфоцитах, но при этом супрессируя синтез IgG2a в IFN гамма [105].

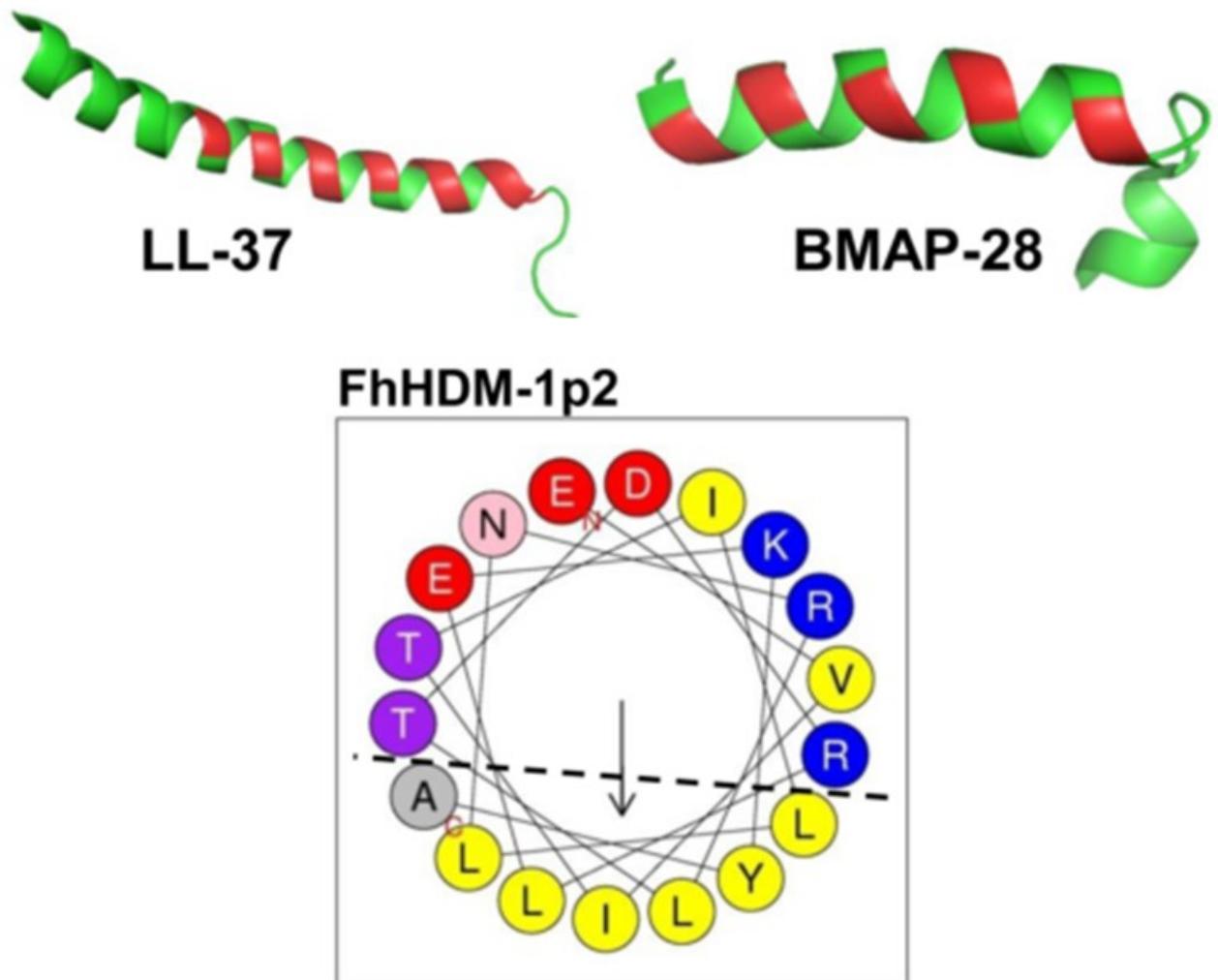


Рисунок 5: Сходство вторичной структуры кателицидинов человека (LL-37 и BMAP-28) с HDM белком фасциолы (FhHDM-1p2). Для человеческих белков показана трехмерная структура, на которой гидрофильные аминокислоты изображены красным. Для пептида *F. hepatica* изображен вид сверху соответствующей альфа-спирали. Обратите внимание на аналогичный кластер гидрофильных аминокислот.

Сходство вторичной, но не первичной структуры между кателицидинами и HDM белками породили идею, что последние мимикрируют под эффекторные молекулы хозяина, перенимая некоторые их функции выгодные для паразита. Во время инвазии фасциола должна проникнуть сквозь стенку пищеварительного тракта, что чревато попаданием содержащихся в ЖКТ микроорганизмов кровь. Действительно, микроорганизмы в значительных количествах наблюдаются в крови заражённых шистосомами животных [107]. Попадающие при этом в кровь эндотоксины могут сильно навредить носителю и даже привести к состоянию септического шока. Последний может иметь летальный исход, что невыгодно в том числе и самому паразиту. В связи с этим авторы выдвигают гипотезу, что подобная «защита» хозяина является

необходимой и самой фасциоле.

Другие элементы экскреторно-секреторного продукта фасциолы, такие как пероксиредоксин и катепсин L, также взаимодействуют с иммунитетом. Первый способен активировать макрофаги по альтернативному пути (M2 фенотип) [108]. Второй деградирует эндосомальный TLR-3 и поэтому инактивирует MyD88-независимый TRIF-зависимый TLR сигнальный путь [109]. Катепсин L также способен предотвратить септический шок (там же).

Обнаруживаемые в экскреторно-секреторном продукте *C. sinensis* белки серпины (распространенные ингибиторы протеаз, имеющиеся во всех таксонах живых организмов) также обладают гуморальными эффектами. Их транскрипция максимальна в метацеркариях (в 9 раз больше чем во взрослом паразите). *In vivo* показана способность серпинов клонорха индуцировать смешанный Th1/Th2 иммунный ответ на фоне повышенного уровня IFN- γ , при этом не действуя на Th1-обусловленный ответ [110]. Также существует гипотеза о возможном воздействии серпинов на систему коагуляции крови, что было показано у кровососущих насекомых [111].

Обнаруженные у гельминтов белки-иммуномодуляторы составляют молекулярную основу явления взаимодействий в системе паразит-хозяин. Т. к. различные гельминтные инфекции были распространены в человеческих популяциях на протяжении всей истории антропогенеза, за этот период времени произошла взаимная «настройка» позволяющая паразиту активней размножаться, а хозяину с минимальным для себя вредом переносить инфекцию. Лишь в недавнее время, в некоторых регионах мира, возросшие стандарты гигиены позволили резко снизить паразитарную нагрузку. Именно с этим связывают возросшее количество аутоиммунных заболеваний в развитых странах мира так называемая «гигиеническая теория» [112]. Иммуномодулирующие же свойства обнаруженных белков наводят на идею их использования в качестве средств лечения аутоиммунных заболеваний. Подобные препараты будут лишены побочных эффектов, сопряженных с паразитарной инфекцией. Так например показано, что секретируемые белки фасциолы оказывают положительный терапевтический эффект при аутоиммунном диабете у мышей NOD [113].

2.2.9 Ответ на внешние факторы

Известными ферментами защиты от ксенобиотиков у многоклеточных организмов являются глутатион-трансферазы (GST). Все они подразделяются на ряд классов: α , μ , π , θ , ζ , σ , и ω . Гомологи GST обнаруживаются у всех трематод, однако, важно заметить, что эта классификация первоначально появилась при исследовании млекопитающих и ее значение для

трематод не так однозначно.

У *C. sinensis* обнаруживаются четыре формы глутатион-трансфераз. Три из них гомологичны 28 kDa GST, и еще одна гомологична 26 kDa GST. Ферменты массой 28 kDa авторами были обозначены как принадлежащие к сигма классу, а 26 kDa – как принадлежащие к мю классу. На сегодняшний день известны как минимум три формы, каждая из которых по-разному отвечает на имеющиеся в желчи вещества [114]. Эти ферменты обнаруживаются в тегументе, мезенхимальных тканях и в яйцах в половых путях. Последнее говорит об их возможной роли в репродукции паразита. Кроме того, иммунологические эксперименты показывают, что при протекании заболевания образуются специфические антитела к этим двум ферментам паразита, что делает возможным их применение как диагностических маркеров [115]. На шистосоме показано, что иммунизация модельных животных вакциной на основе 28 kDa GST позволяет впоследствии снизить паразитарную нагрузку и плодовитость самок шистосомы [116].

Находясь под атакой клеток иммунной системы, взрослый паразит сталкивается с потоком эндогенных и экзогенных окислительных соединений. О существовании систем защиты от этих повреждающих агентов говорит наличие в транскриптом генов глутатион-пероксидаз, тиоредоксин-пероксидаз, пероксид-дисмутаза. Также наличествуют несколько пероксиредоксинов, различающиеся по паттерну экспрессии и каталитическим свойствам [117], и тиоредоксин [118]. Ген каталазы отсутствует у таких трематод, как *F. hepatica* [119], а также, к примеру, у малярии [120]. Последний факт говорит о том, что возможно редукция каталазы важна при переходе к паразитическому образу жизни, т.к. генетическая потеря этого фермента произошла в двух филогенетически отдаленных линиях паразитов. Любопытно биохимической особенностью трематод является восстановление глутатиона и тиоредоксина одним ферментом, хотя обычно у эукариот имеются две отдельные системы, осуществляющие эти процессы [121]. Система цитохромов редуцирована до единственного гена как у печеночных сосальщиков [122], так и ленточных червей [89].

Кроме защиты от молекулярных агентов, у трематод должны присутствовать хотя бы примитивные системы защиты и от патогенов. Обнаружены белки РНК-интерференции, что говорит о возможности трематод противостоять некоторым вирусам. Наличие этой молекулярной машины, в норме деградирующей двухцепочечную РНК, позволяет говорить о возможности проведения экспериментов, основанных на РНК-интерференции. Уже известен удачный эксперимент по подавлению транскрипции катепсина В у *O. viverrini* [123].

2.2.10 Транскриптом разных стадий жизненного цикла

Сложный жизненный цикл, включающий крупные изменения в морфологии, клеточном и молекулярном составе – отличительная черта трематод. Обычно наибольшее внимание уделяется стадии, паразитирующей в окончательном хозяине, т. е. в том числе и в человеке; другие стадии, обитающие в промежуточных хозяевах и природных средах, изучены значительно хуже. Тем не менее имеется ряд работ, посвященных сравнительной транскриптомике различных стадий жизненного цикла. Так, Гоберт с соавторами использовал гибридизацию на микрочипах, разработанных специально для исследования *S. japonicum*, для сравнения сразу нескольких стадий (легочная шистосомула, 4х недельная незрелая особь, половозрелая особь, мирацидий, спороциста и церкария) и у обоих полов [124]. В экспериментах были найдены множества дифференциально экспрессирующихся генов. Одной из примечательных находок являются существенные различия в путях передачи сигнала с помощью ионов кальция. Авторы считают, что эти пути задействованы в запуске перепрограммирования развития, т. е. в смене стадии жизненного цикла.

В другом исследовании шистосоматид акцент был сделан на молекулярные изменения, происходящие при преобразовании церкарии в шистосомулу (события через 3 часа после трансформации; авторы имитировали события, происходящие при пересечении кожного барьера) и шистосомулы во взрослую особь (события через 5 дней; авторы имитировали события, происходящие при попадании шистосомул в легкие). Выяснилось, что при переходе из церкарии в шистосомулу происходило понижение экспрессии генов, ответственных за наработку энергии и увеличивалась экспрессия генов, кодирующих структурные белки [125]; также происходит увеличение экспрессии белков, ответственных за стресс, что авторы связывают с резкой сменой внешних факторов среды при попадании в окончательного хозяина. При созревании шистосомулы во взрослую стадию происходит скачок экспрессии таких поверхностных белков как аннексинов и тетраспанинов. Последние являются перспективными мишенями для разработки вакцин [73] (см. также раздел 2.2.3). Также авторы подмечают существенную разницу в некоторых белках метаболизма, таких как липопротеиновые рецепторы и протеазы.

Исследователями описторхид предпринимались попытки сравнить транскриптомными методами взрослую стадию и метацеркарию, не имеющего прямого аналога у шистосоматид. Так, в одной из первых работ по этой тематике [54] были получены 419 EST последовательностей метацеркарии *C. sinensis*. Авторами был обнаружен дифференциально (различно от взрослой стадии) экспрессирующийся фактор транскрипции EBP1, который как они предполагают участвует в процессе метаморфоза метацеркарии в мариту. Также обнаружен

фермент бета-глюкозидаза желчных кислот, способная их расщеплять. В желчных путях, где собственно и находится взрослая особь, гликозилированные желчные кислоты имеются в большом количестве, чем и объясняется высокая экспрессия глюкозидазы.

Работа [56] представляет другую попытку сравнения транскриптомных профилей стадий развития, но в несколько большем масштабе. Этой группой были отсекуены 20 тыс. EST из метацеркарии и около 10 тыс. EST из яиц. Авторы пришли к выводу, что дифференциально экспрессирующиеся группы генов хорошо согласуются с образом жизни соответствующей стадии. Генами, высоко экспрессируемыми во взрослой стадии по сравнению с метацеркариями и яйцами, в основном являются структурные белки и белки, связанные с репродукцией (β -тубулин, ферритин), белки детоксификации (глутатион S-трансферазы), белки транспорта (клонорпорин, натрий/глюкоза транспортер), белки продукции энергии (GAPDH, митохондриальная малат дегидрогеназа) и некоторые ферменты (цистеиновые протеазы, PNHGPx изоформа 1). На стадии яйца высокой экспрессией отличается ацил-КоА синтетаза (длинная цепь).

Замечу, что полнотранскриптомного исследования метацеркарии трематод проведено не было.

2.2.11 Особенности молекулярно-генетической организации трематод.

Уникальной особенностью обнаруженной у шистосоматид является наличие генов с микроэкзонной структурой [126]. Авторы считают, что это механизм белковой вариации, осуществляемый посредством альтернативного сплайсинга коротких симметричных тандемно организованных экзонов. Анализ экспрессии этих генов показал, что они включаются при проникновении через кожу млекопитающего. Посредством протеомного анализа секрета шистосомулы удалось получить прямое свидетельство вариации этих белков за счет использования различных экзонов. Иммунологическое исследование подтвердило поверхностную локализацию этих белков. Все это является свидетельством использования данных белков паразитом для противодействия адаптивному иммунному ответу хозяина.

С помощью методик секвенирования РНК на шистосоме было показано явление транссплайсинга для 11% генов с участием лидерной последовательности SL [49]. В этой же работе авторы обнаружили полицистронные РНК, которые в процессе сплайсинга образуют отдельные гены. Для этого сплайсинга у шистосом похоже используется та же самая лидерная последовательность SL (для сравнения у *C. elegans* в аналогичных процессах используются

различные лидерные последовательности).

У представителей ленточных червей также обнаружены полицистронные РНК [89]. Любопытно, что 40% генов, находящихся в полицистронах обнаруживают высокую степень консервативности в сравнении с шистосомами, т. е. гены в полицистронных РНК подвержены стабилизирующему отбору.

Обнаруживается любопытная особенность интрон-экзонной структуры генов шистосомы: длина интронов меньше в 5' районе гена, где они могут составлять 26 нуклеотидов в длину, и больше в 3' районе, где они могут достигать килобазы в длину (самый большой 33.8 кб) [43]. Авторы предполагают, что это говорит о существовании особенных механизмов транскрипционного контроля. т. к. у млекопитающих наблюдается обратная тенденция – уменьшение длины интронов к 3' концу гена [43].

Любопытно, что обнаруживается значительный пул молекул с большей гомологией к молекулам позвоночных, чем к нематодам [127]. Это можно объяснить с одной стороны мимикрией к белкам окончательного хозяина, а с другой стороны чисто эволюционными соображениями, связанными с более ранней дивергенцией этих видов от общего ствола, чем нематод [47].

2.2.12 Потенциальные мишени для диагностики, лечения и вакцинирования

Вакцинирование против бактерий и вирусов является основной мерой борьбы с ними, вот уже более двух столетий спасающей миллионы жизней. Однако, до сих пор нет ни одной успешно реализованной и применяемой на человеке вакцины против гельминтов [128]. Одним из возможных объяснений этого факта может служить то, что в организме носителя гельминты используют тонкие механизмы иммунной модуляции, которые нивелируют любые приобретённые иммунные эффекты. Тем не менее многие лаборатории предпринимали попытки создания рекомбинантных аналогов белков паразитов с целью получить средства иммунизации. Так, катепсин В2 фасциолы является ключевой мишенью для вакцинирования этого гельминта. Комбинированные мультивалентные вакцины из различным катепсинов показывают высокий уровень протекции (показана высокая иммуногенность данного белка на крысах и овцах). Вакцина на основе FhCB2 была высоко иммуногенной и снижала количество червей на 60%, а их размеры – на 63%. Эти результаты показывают что вакцинация ювенильными протеазами и снижает паразитарную нагрузку, и ингибирует развитие выживших паразитов. Также, эта вакцина значительно снижала повреждения печени (на 61%) [129].

В ходе работ по разработке иммунодиагностики на описторхид значительные результаты показали цистеиновые протеазы. Их рекомбинантные аналоги показали уровень чувствительности 96%, а специфичности 96,2%. [130].

Другой разрабатываемой мишенью для создания иммунодиагностики был парамиозин [131]. Данный белок экспрессируется во всех стадиях, особенно во взрослой, где локализуется поверхностно, а значит способен взаимодействовать с иммунной системой.

2.3 Заключение

Гельминтозы по прежнему составляют существенную проблему здравоохранения в некоторых регионах мира. Описторхоз и клонорхоз вызываются гельминтами, относящимися к семейству описторхид. Вызываемое ими заболевание поражает гепатобилиарную систему и, при отсутствии своевременного лечения, чревато множеством осложнений. Описторхиды представлены печеночными сосальщиками азиатского региона – *C. sinensis* и *O. viverrini*, а также включают объект настоящего исследования – *O. felineus*. В отличие от азиатских видов, последний распространен на территории северной Евразии и особенно важен для сибирского региона.

Хотя описторхиды составляют существенную медицинскую проблему, разработка новых лекарственных препаратов и средств диагностики затруднена, что связано в том числе и с малой изученностью этих видов на молекулярном уровне. Тем не менее, уже имеющиеся молекулярно-биологические исследования трематод выявили целый ряд особенностей их строения и физиологии, которые отражены на геномном и транскриптомном уровнях. С одной стороны, для этих видов животных характерно некоторое упрощение организации, затрагивающее в частности нервную систему и план строения тела. На геномном уровне это проявляется в отсутствии некоторых характерных для других эукариот молекул рецепторов и лигандов нейротрансмиссии, в меньшем количестве НОХ генов, отсутствии пероксисомального аппарата и пр. С другой стороны, переход к паразитизму привел к сильному эволюционному отбору, направленному на специализацию и гипертрофию других систем. На геномном уровне это проявляется в экспансии некоторых генных семейств (таких как катепсины, MD2-домен содержащие белки), а на транскриптомном уровне наблюдается доминантных характер экспрессии некоторых белков, составляющих в сумме десятки процентов всей белок-кодирующей РНК (миоглобин, катепсин, белки оболочки яиц). Особенно интересными являются некоторые данные о взаимодействии паразит-хозяин на молекулярном уровне, представляющих уникальное для паразитизма эволюционное приобретение. Примерами обнаруженных молекулярных посредников таких взаимодействий являются HDM белки,

некоторые катепсины и серпины.

Для получения этих данных и дальнейших исследований важно наличие геномной и транскриптомной информации, которая стала значительно более доступной в связи с недавним развитием методов секвенирования ДНК. Эти методы были использованы для исследования упомянутых здесь и выше трематод, но не для *O. felineus*, остающегося среди них наименее изученным. В данном исследовании нами было осуществлено секвенирование его полного транскриптома, точнее, транскриптома его взрослой (мариты) и инвазивной стадий (метацеркарии). Исследование инвазивной стадии несет дополнительную новизну, т. к. полный транскриптом метацеркарии описторхид ранее изучен не был, хотя и предпринимались некоторые попытки его расшифровки с помощью методов Сэнгера. Полученные данные могут быть использованы для сравнения *O. felineus* как с другими описторхидами, так и с более отдаленными таксонами, пролить свет на молекулярные основы развития разных стадий жизненного цикла, а также составить основу для дальнейших молекулярно-биологических исследований этого вида.

Урацил ДНК-гликозилаза (10 е.а./мкл, SibEnzyme, Россия)

3.3 Экспериментальные методы

3.3.1 Сбор биологического материала

Метацеркарии собирали из зараженной рыбы семейства карповых (в основном вида *Leuciscus leuciscus baicalensis*), пойманной в реке Обь, в Новосибирске и в реке Томь близ г. Томска. Для этого, фарш из мышечной ткани и подкожной клетчатки зараженной рыбы обрабатывали искусственным желудочным соком в течение ночи при 37° С. Опустившиеся на дно метацеркарии промывали физиологическим раствором, контролируя под микроскопом их чистоту.

Золотистых (сирийских) хомячков *Mesocricetus auratus* перорально заражали метацеркариями в количестве около 50 штук на животное. Через как минимум три недели хомячков забивали и из их печени извлекали взрослых особей *O. felineus*. Животных содержали и использовали согласно протоколам этического комитета Института цитологии и генетики СО РАН. Взрослых особей *O. felineus* тщательно промывали стерильным физиологическим раствором и либо сразу использовали для выделения РНК, либо хранили в реагенте RNAlater (Ambion).

3.3.2 Создание кДНК библиотеки и ее анализ методом секвенирования по Сэнгеру.

3.3.2.1 Клонирование библиотеки

Тотальную РНК выделяли из 120 взрослых особей *O. felineus*, используя набор ToTally RNA kit (Ambion). Полиаденилированную фракцию выделяли из тотальной РНК с помощью Micropoly(A) Purist kit (Ambion). poly(A)+ РНК использовали для создания ненормализованной кДНК библиотеки. Первую цепь кДНК синтезировали используя 0,5 мкг poly(A)+ РНК и набора Mint-universal kit (Evrogen). Затем, используя набор того же производителя, 1/30 часть полученной первой цепи кДНК была амплифицирована ПЦР реакцией (20 циклов). При амплификации в ДНК были введены ассимметричные SfiI сайты узнавания по концам.

Полученную таким образом кДНК (в количестве 2,5 мкг) очищали от неиспользованных олигонуклеотидов и нуклеозидтрифосфатов используя MiniElute Reaction Cleanup Kit (Qiagen) и гидролизовали 100 единицами активности фермента SfiI в течение 2х часов при 50°С. Эндонуклеазу рестрикции инактивировали в течение 20 минут при 65°С, 1/10 часть кДНК

использовали для лигирования с 200 нг вектора pBluescript II SK (с некоторыми модификациями полилинкерного района), предварительно гидролизованного SfiI. Продукты лигирования переосаждали этанолом в присутствии гликогена и использовали для электропорации компетентных клеток *E. coli* штамма XL1-Blue MRF.

3.3.2.2 Культивирование и консервирование клонов

Клоны из кДНК библиотеки, выросшие на чашках Петри с агаризованной питательной средой, перекалывали для культивирования в глубокие 96-луночные планшеты с 150 мкл 2X LB питательной среды. Нарращивание производилось 1 сут. при температуре 37° С при интенсивном перемешивании на качалке с частотой 250 об/мин. После этого к каждой лунке с выращенными бактериями добавляли 150 мкл раствора глицерина в 2X LB (50% v/v). Полученные законсервированные клоны поместили в низкотемпературный холодильник на -70°С для длительного хранения.

3.3.2.3 Амплификация вставок кДНК

Сначала готовили ПЦР-смесь (состав см. ниже), которую с помощью роботизированной станции Тесап раскапывали по 15 мкл в лунку по 384-луночному планшету. В качестве матрицы в соответствующую лунку переносили следовые количества бактериальной биомассы. Готовую ПЦР смесь с внесённой матрицей помещали в амплификатор, где проводили реакцию при следующем температурно-временном режиме:

1. деградация урацил-содержащей ДНК ферментом UNG при 37°С, 10мин
 2. предварительная денатурация крупной плазмидной ДНК при 95°С, 2мин
 3. денатурация при 95 °С, 10 с
 4. отжиг праймеров при 52° С, 20 с
 5. элонгация при 72 °С, 1 мин
 6. завершающий синтез при 72° С, 5 мин
- } 25 циклов

Состав реакционной смеси: ПЦР буфер до концентрации 1x, dNTP 0,25 ммоль/л каждого, Т3 праймер 500 мкмоль/л, Т7 праймер 500 мкмоль/л, Таq ДНК-полимераза 0,25 е.а./реак., UNG 0,25 е.а./реак., H₂O до объема. Объем каждой реакции – 15 мкл.

3.3.2.4 Электрофорез ПЦР фрагментов

Проводили в 1,5% агарозном геле, содержащем 0,1% (v/v) бромистого этидия. Образцы

наносили с использованием 6х буфера для нанесения (бромфеноловый синий, ксиленцианол, 30% (v/v) глицерин). Условия электрофореза: 40 мин при напряженности поля 4 В/см.

3.3.2.5 Очистка ПЦР продукта

Очистку производили набором Millipore PCR Clean96 Up Kit на роботизированной станции Tecan согласно инструкции производителя.

3.3.2.6 Выделение плазмидной ДНК

Некоторые клоны анализировали отдельно и с этой целью производили выделение их плазмидной ДНК, содержащей вставку. Для этого использовали набор QIaprep Spin Miniprep Kit (QIAGEN, Германия). Процесс был автоматизирован с помощью станции QiaCube (QIAGEN, Германия).

3.3.2.7 Секвенирование плазмидной ДНК

Секвенирование проводилось методом Сэнгера. Для этого в микропробирке приготавливали реакционную смесь объемом 25 мкл, имеющую следующий состав:

Big-Dye terminators 1 мкл/реакцию

5х реакционный буфер 5мкл

секвенирующий праймер 10 пмоль

раствор плазмидной ДНК (С ~50 мкг/мл) 5мкл

Далее продукты реакции Сэнгера нарабатывались на амплификаторе при следующем температурном режиме:

- | | | |
|--|---|-----------|
| 1. предварительная денатурация при 97° С, 9 с. | } | 25 циклов |
| 2. денатурация при 97° С, 11с | | |
| 3. отжиг праймера при 50° С, 5 с | | |
| 4. синтез продуктов элонгации при 60° С, 4мин | | |
| 5. завершающий синтез при 92° С, 3 мин | | |

Продукты амплификации далее очищали методом гель-фильтрации через Sephadex G50 на спин-колонках (см. далее).

3.3.2.8 Очистка продуктов элонгации реакции Сэнгера

Сухой сефадекс G50 заливали 20 объёмами бидистиллированной воды и смесь выдерживали на термо-миксере при 60°C в течение 1ч. Затем в пробирку вставляли спин-колонку со стеклянным фильтром и в нее добавляли 750 мкл сефадекса. Смеси давали отстояться, после чего убирали излишки воды сверху. Затем колонку центрифугировали при 800g 2 мин. Добавляли 10 мкл бидистиллированной воды и вновь центрифугировали 2 мин при 800g. Далее на сефадекс наносили продукты реакции Сэнгера, колонку переставляли в новую пробирку и после очередного центрифугирования в течение 2 мин при 800g в пробирке оказывался очищенный продукт. Последний упаривали под вакуумом при 60°C в течение 30 мин и анализировали в межинститутском центре секвенирования ЦКП «Геномика» СО РАН.

3.3.3 Создание и секвенирование библиотеки с помощью технологии Illumina Solexa.

Образцы примерно 100 взрослых особей и 2000 метацеркарий *O. felinus* гомогенизировали в жидком азоте и из них выделяли тотальную РНК с помощью набора RNEasy Mini Kit (Qiagen) согласно инструкции производителя. Для каждого образца измеряли количество РНК с помощью флуориметра Qubit (Invitrogen), а качество образцов проверяли с помощью капиллярного электрофореза на приборе Bioanalyzer 2100 (Agilent). Библиотеки кДНК изготавливали из 250-500 нг тотальной РНК. Коротко, мРНК очищали с помощью магнитных частиц, покрытых олиго-dT олигонуклеотидом (Invitrogen). Далее синтезировали дцДНК с помощью обратной транскриптазы SuperscriptII (Invitrogen) и случайного гексамерного праймера и очищали с помощью набора Ampure XP (Invitrogen). К кДНК лигировали адапторы Illumina, с помощью которых проводили реакцию амплификации по протоколу

денатурация при 98 °C, 10 с	}	15 циклов
отжиг праймеров при 60 °C, 30 с		
элонгация при 72 °C, 30 с		

Полученный ПЦР-продукт очищали на магнитных частицах (Invitrogen), проводили электрофорез через 2% агарозный гель и вырезали фракцию 300-500 нуклеотидов. Библиотеки денатурировали в 0,1 М NaOH и растворяли в HT1 буфере (Illumina) до концентрации 10 пМ. Далее из библиотек генерировали кластеры на приборе cBot (Illumina) с помощью набора TruSeq PE Cluster Kit v3-cBot-HS и проводили секвенирование по парному протоколу (2x101bp) с помощью химии HiSeq200 и TruSeq SBS chemistry. Демультимплексирование проводили с

помощью CASAVA-1.8.2 (Illumina).

3.4 Биоинформатический анализ последовательностей

3.4.1 Сборка транскриптома

Из EST последовательностей удаляли участки вектора и проводили сборку программой MIRA [132].

Из парных прочтений, полученных на платформе Illumina удаляли нуклеотиды адапторных последовательностей, а также нуклеотидные тракты, имеющие низкие значения Phred. После этого удаляли слишком короткие прочтения (длиной менее 70 нуклеотидов). Эти этапы проводили одновременно программой Trimmomatic [133] с параметрами ILLUMINACLIP:../TruSeq3-PE-2.fa:2:30:10 LEADING:10 TRAILING:10 SLIDINGWINDOW:5:15 MINLEN:70 HEADCROP:10. Далее проводилась очистка от потенциальных контаминантов. Для этого удалялись все прочтения, которые картировались на доступные в Genbank RefSeq последовательности *Mesocricetus auratus* с помощью Bowtie2 [134]. Из оставшихся прочтений проводили сборку транскриптома *de novo* с помощью Trinity (используя параметры по умолчанию). [135].

3.4.2 Анализ экспрессии

Оценкой экспрессии транскрипта в данных секвенирования по Сэнгеру являлось количество прочтений, из которых состоит контиг. Для этого выходные данные сборки программы MIRA в формате ace анализировали модулем Bio.AlignIO из библиотеки BioPerl (<http://www.bioperl.org>).

Чтобы оценить уровень экспрессии генов в данных массового параллельного секвенирования, те же прочтения, что использовали для *de novo* сборки, выравнивали на полученные транскрипты с помощью алгоритма RSEM (интегрирован в Trinity). Далее, статистическую значимость дифференциальной экспрессии между стадиями устанавливали с помощью пакета edgeR [136].

Для определения подмножеств генов (или путей KEGG), обогащенных дифференциально экспрессирующимися генами, использовали точный тест Фишера. Чтобы получить более достоверные результаты, значения P в тесте на дифференциальную экспрессию нормировали с помощью поправки Бонферрони, умножив его на количество генов в исследуемом подмножестве. Гены, для которых скорректированное P было меньше 10^{-2} считались

дифференциально экспрессирующимися. Далее, вычислялось значение P в точном тесте Фишера с помощью пакета `python-scipy` (<http://www.scipy.org/>), анализируя таблицу сопряженности вида

Количество дифференциально экспрессирующихся генов в подмножестве	Количество дифференциально экспрессирующихся генов, не относящихся к подмножеству
Количество генов подмножества, не являющихся дифференциально экспрессирующимися	Количество генов не являющихся дифференциально экспрессирующимися и не относящиеся к подмножеству

3.4.3 Аннотация транскриптов

В полученных транскриптах были предсказаны открытые рамки трансляции с помощью программы Transdecoder (<http://transdecoder.sf.net>) с использованием параметров по умолчанию.

Поиск гомологов осуществляли алгоритмами BLASTp и BLASTx [137] против баз данных NCBI Nr (<http://www.ncbi.nlm.nih.gov/>), Swissprot (<http://www.uniprot.org/>), KEGG (<http://www.genome.jp/kegg/>), база данных лаборатории Робина Гассера (Ветеринарный факультет, Университет Мельбурна, Мельбурн, Австралия) (<http://bioinfosecond.vet.unimelb.edu.au/wblast2.html>), SchistoDB (<http://schistodb.net/schisto/>); GeneDB (<http://www.genedb.org/>). При поиске гомологии значение $E\text{-value} = 10^{-5}$ использовали как пороговое.

Консервативные гены эукариот определяли с помощью tBLASTn с пороговым значением $E\text{ value} < 10^{-10}$ против доступных белков человека ([138], <http://korflab.ucdavis.edu/datasets/cegma/>). Отнесение транскриптов к ортологическим группам KEGG и биологическим путям проводили с помощью программы KOBAS [139]. В предсказанных белковых последовательностях предсказывали консервативные домены с помощью InterProScan [140]. С ее же помощью проводили аннотацию транскриптов в терминах Gene Ontology [141]. Распределение терминов Gene Ontology по различным классам делали с помощью программы CateGOrizer [142]. Обогащение определенных терминов генной онтологии определяли с помощью Cytoscape [143] и плагина BINGO [144]. Поиск гомеодомен-содержащих белков в транскриптом осуществляли с помощью локальной версии HMMER [145], используя тренировочный сет гомеобокса эукариот.

Различные выходные данные программ анализировались с помощью собственных скриптов, написанных на языках программирования Python и Perl.

3.4.4 Филогенетический анализ

Для анализа рибосомальных белков кодирующие их нуклеотидные последовательности выравнивали алгоритмом ClustalW [146]. Поиск подходящей модели эволюции и реконструкцию филогенетических взаимоотношений (алгоритмом максимальной парсимонии) проводили с помощью пакета MEGA 5 [147]. Филогенетические взаимоотношения алгоритмом максимального правдоподобия устанавливали с помощью программы PhyML [148].

Для филогенетического анализа гранулиноподобных белков, нуклеотидные последовательности, кодирующие гранулиновый домен рассматриваемых пептидов, извлекали вручную из соответствующих последовательностей кДНК трематод. Выравнивания полученных последовательностей анализировали с помощью программы PHYLIP [149].

Этапы биоинформатического анализа, требующие высоких вычислительных мощностей, производили в центре коллективного пользования «Биоинформатика» и вычислительном кластере Новосибирского Государственного Университета.

4 Результаты и обсуждение

4.1 Общая характеристика полученных данных

В работе исследовались результаты секвенирования 4 тыс. клонов кДНК библиотеки по Сэнгеру, а также около 150 млн прочтений тотальной РНК мариты и метацеркарии *O. felineus*, полученных на платформе Illumina HiSeq, из которых 110 млн относятся к взрослой стадии и 40 млн относятся к метацеркарии.

В кДНК библиотеке основную массу составляли вставки длиной от 400 до 1000 нуклеотидов. После препроцессинга последовательностей остались 2560 EST, которые доступны в GenBank под номерами: JK624271–JK626790, JK006511–JK006547, JK649790–JK649792. При кластеризации этих EST последовательностей было получено 267 контигов. Длина последовательностей среди контигов варьирует от 200 до 1600 пн, а основная часть их лежит в диапазоне 400–800 пн (Рисунок 6). Поиск гомологичных белков с помощью алгоритма BLASTx в базе данных NCBI Nr обнаружил для 219 (82%) последовательностей *O. felineus* статистически значимых гомологов.

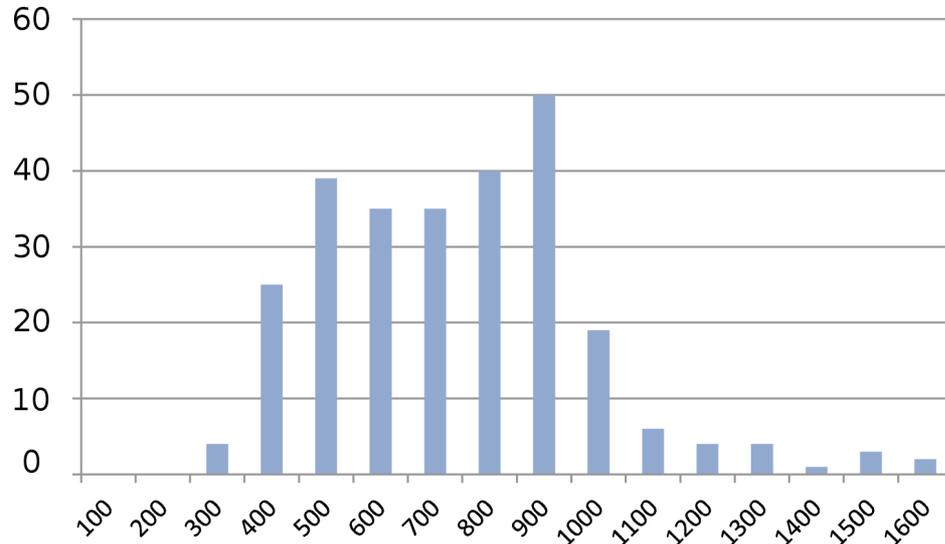


Рисунок 6: Гистограмма распределения длин контигов. На абсциссе изображен интервал длины контигов (в пн), а ордината показывает количество контигов с соответствующей длиной.

Белки, обнаружившие наиболее близкую гомологию к белкам *O. felineus* принадлежали в основном *C. sinensis*, *S. japonicum* и *S. mansoni* (все представители класса Трематода) (Рисунок 7). Существенная доля шистосоматид в этом распределении скорее всего объясняется высокой степенью представленности (изученности) их последовательностей в NCBI Nr (на 2013 г.) базе данных в сравнении с другими трематодами. Для 18% полученных последовательностей

BLASTx не обнаружил гомологов в базе Nr. Эта часть транскриптов вероятно представляет из себя новые последовательности, характерные только для *O. felineus*.

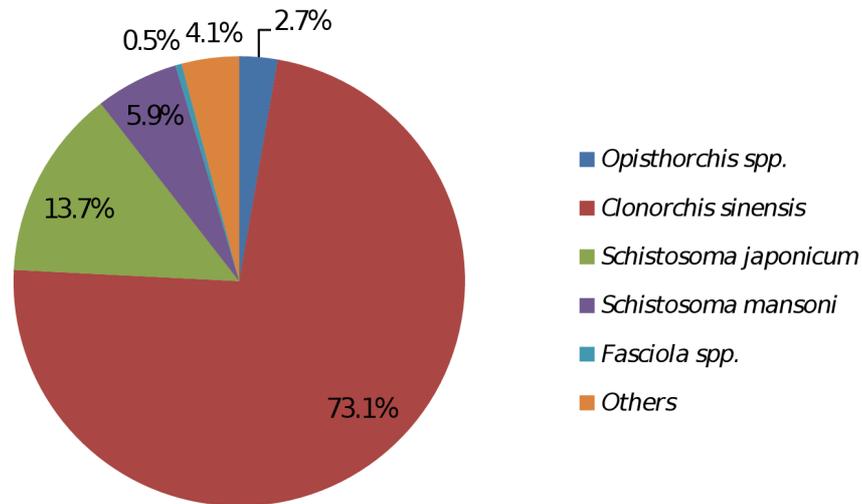


Рисунок 7: Видовое распределение организмов, обнаруживающих наиболее близкую гомологию с EST последовательностями *O. felineus*.

Из данных NGS секвенирования перед сборкой удалялись адаптеры, недостоверно прочитанные нуклеотиды и контаминации. Основным контаминантом являлись последовательности, принадлежащие золотистому хомячку (*M. auratus*), который использовался для выращивания взрослых особей *O. felineus*. Контаминации рыбой, из которой брались метацеркарии обнаружено не было. Прошедшие фильтрацию прочтения были ассемблированы в 89'727 транскриптов (средняя длина и соответствующее стандартное отклонение равны 773 нт \pm 1099 нт соответственно), среди которых большинство являются сплайс-вариантами транскриптов одних и тех же генов и артефакты сборки. Для того, чтобы сузить анализ до последовательностей с большей биологической значимостью, нами были отобраны уникальные белок-кодирующие транскрипты. Согласно предсказанию рамок трансляции с помощью Transdecoder таких транскриптов оказалось 12'665 (1952 нт \pm 1713 нт). Количество белок-кодирующих последовательностей хорошо согласуется с аналогичными данными, полученными для близкородственных трематод (14'269 для *O. viverrini* [13], 13'634 для *C. sinensis* [14]). Эти 12'265 белок-кодирующих транскриптов использовались для дальнейшего анализа (Таблица 2).

Таблица 2: Характеристики сборки транскриптома, полученной программой Trinity

Общее количество прочтений Illumina	150 млн
Общее количество полученных транскриптов	12'665
N50 полученных контигов	2781 нт
Медианная длина контигов	1591 нт
Средняя длина контига	1952 нт
Транскриптов со значимой гомологией против Nr	11'059 (87.3%)
Транскриптов со значимой гомологией против Swissprot	7'753 (61.2%)

При поиске гомологии с помощью BLASTx для 11'130 (87,8%) и 7'943 (62,7%) транскриптов были обнаружены гомологи в базах данных NCBI Nr и Swissprot соответственно. Существенное количество транскриптов (1533, 681 нт \pm 505 нт) не обнаружили статистически значимой гомологии и, возможно, представляют собой последовательности специфичные для *O. felinus*. Транскрипты полученные из NGS данных проявляли наибольшую гомологию к *C. sinensis* и *O. viverrini*, которые вместе составляли наиболее близких гомологов для 95% последовательностей (Таблица 3)

Таблица 3: Видовое распределение гомологичных последовательностей. Во второй колонке отражен процент белок-кодирующих транскриптов со статистически значимой гомологией против базы данных Nr, для которых данный вид содержит наиболее близкого гомолога.

Вид	Процент транскриптов
<i>C. sinensis</i>	50,0%
<i>O. viverrini</i>	45,0%
<i>S. mansoni</i>	1,4%
<i>S. japonicum</i>	1,2%
Другие	2,2%

Данные NGS секвенирования также являются публично доступными: в базе данных NCBI зарегистрирован соответствующих биопроект (номер PRJNA257351), первичные («сырые») данные выложены в SRA архив под номерами SRX669893 (для метацеркарии) и SRX669356 (для взрослой стадии), а собранные транскрипты находятся в NCBI TSA архиве под номером GBJA01000000.

4.2 Состав доминантных компонент транскриптома

Мы определили гены с наибольшей экспрессией в различных стадиях развития паразита (приложения 1 и 2). Среди генов с наибольшей экспрессией во взрослой стадии определяются такие гены как катепсин F, миоглобин, вителлин, белок оболочки яйца и глутатион трансфераза. Это подтверждается как по EST последовательностям, так и по NGS данным. Однако, любопытно, что транскриптом с самым большим значением FPKM оказался гомолог HDM белка. В метацеркарии наибольшую экспрессию имели рибосомальные белки (11 из 20 транскриптов с самой высокой экспрессией, см. приложение 2). В общем следует заметить, что в метацеркарии в отличие от мариты наибольшую экспрессию имеют в основном гены домашнего хозяйства, такие как уже упомянутые рибосомальные белки и убиквитин. Возможно, это отражает тот факт, что на данной стадии паразит находится в сравнительно малоактивной фазе жизненного цикла, ожидая когда он будет поглощён окончательным хозяином. Таким образом, низкие метаболические запросы приводят к низкой экспрессии генов помимо генов домашнего хозяйства. Наличие белков теплового шока среди доминантных компонент скорее подтверждает такой взгляд на физиологию метацеркарии, т. к. при попадании в млекопитающее происходит существенное увеличение температуры по сравнению с рыбой и, возможно, эти белки являются необходимым приспособлением именно к этому. Другим исключением является специфичный для метацеркарий HDM белок (см. далее).

4.2.1 HDM белки

В экспериментах на трематоды *F. hepatica* были обнаружены молекулы экскреторно-секреторного продукта, обладающие обширными иммуномодулирующими свойствами [103], названные HDM белками (см. раздел 2.2.8). Обнаружившие их авторы считают HDM белки примером молекулярной мимикрии под эффекторные молекулы врожденного иммунитета хозяина [104]. Эти белки гельминтов способны подавлять секрецию воспалительного цитокина ФНО из макрофагов [105], связывать бактериальные липополисахариды бактерий и ингибировать транспорт протонов в лизосомы макрофагов, тем самым препятствуя антигенному представлению [106].

В геноме *O. viverrini* ранее было обнаружено пять HDM белков [13]. Мы обнаружили четыре различных HDM гомолога в транскриптоме *O. felinus*. У двух из них была высокая экспрессия во взрослой стадии (один из них – белок с самым большим значением FPKM), а у двух – в метацеркарии (Таблица 4). Таким образом, эти молекулы у *O. felinus* проявляют высокую специфичность к конкретной стадии.

Таблица 4: Экспрессия HDM белков (в FPKM) в различных стадиях *O. felineus*

Транскрипт	Экспрессия во взрослой стадии	Экспрессия в метацеркарии
GBJA01004968	111083	797
GBJA01000004	403	~0
GBJA01000005	~0	13183
GBJA01000006	~0	1088

Одним из свойств HDM белков, натолкнувших их исследователей на сходство с эффекторными молекулами врожденного иммунитета человека, а именно дефензинами и кателицидинами это наличие характерной амфипатической α -спирали. Исследование первичной структуры HDM белков *O. felineus* показало, что это структурное свойство в них также имеет место (Рисунок 8). В каждом соответствующем участке α -спирали видна гидрофобная часть, состоящая в основном из лейцина, валина и изолейцина, а также гидрофильная сторона, содержащая в основном глутаминовую кислоту и некоторые другие полярные аминокислоты.

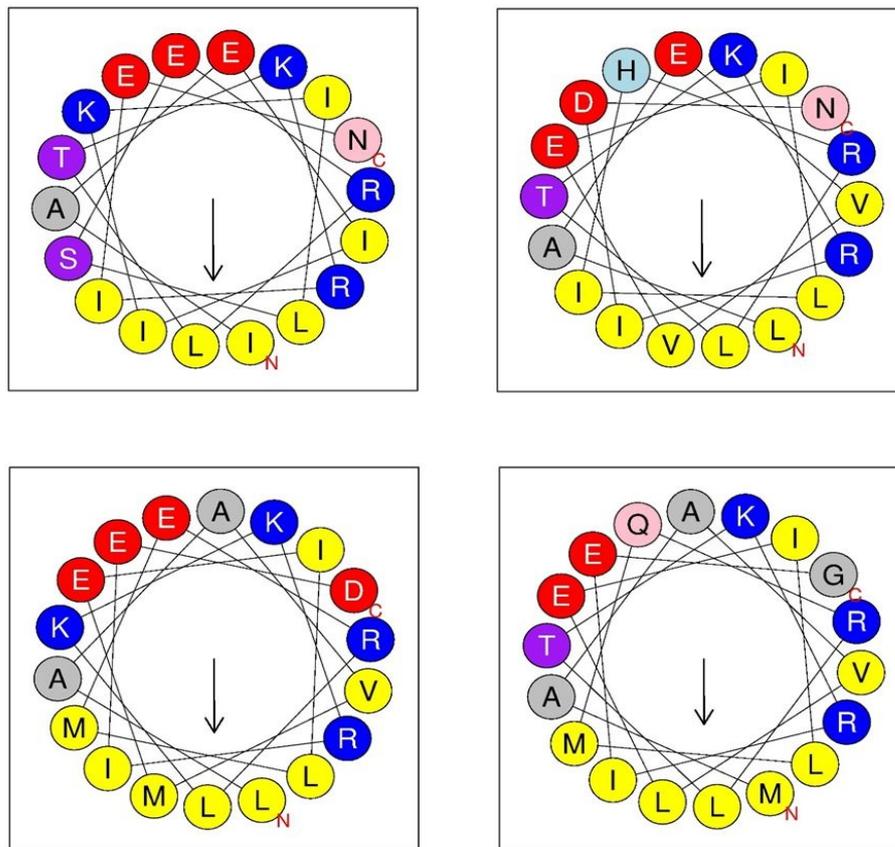


Рисунок 8: Участок амфипатической спирали четырех HDM белков, обнаруженных в транскриптом *O. felineus*.

HDM белки, экспрессируемые в метацеркарии представляются особенно любопытными. Неясно на кого должны оказывать иммуномодуляторный эффект эти белки (если таковой характерен для всех гомологов): на окончательного или промежуточного хозяина? Мы считаем, что производимые в большом количестве метацеркарные HDM белки способны произвести временный локальный эффект на окончательного хозяина в момент экцистирования, но экспериментальных данных в подтверждение этой гипотезы в настоящий момент нет. С другой стороны рыбы принадлежат к челюстным позвоночным, а значит имеют многие характеристики иммунной системы, характерные и для млекопитающих, следовательно потенциальное иммуномодуляторное воздействие на промежуточного хозяина также возможно.

4.2.2 Миоглобин

Одним из самых транскрибируемых транскриптов оказалась мРНК миоглобина, содержащая рамку трансляции длиной 450 нуклеотидов. Соответствующий контиг, полученный из данных кДНК библиотеки содержал 387 последовательностей EST, что составляет 15,1%. Транскрипция этого гена по данным высокопроизводительного секвенирования тоже высока – миоглобин является третьим по значению FPKM транскриптом во взрослой стадии (приложение 1), однако, его экспрессия в метацеркарии ниже более чем в десять раз. В транскриптомах близкородственных трематод *S. sinensis* и *O. viverrini* миоглобин также оказался одним из самых представленных транскриптов [53,58,150]. Столь активная транскрипция гена, кодирующего этот гем-содержащий белок по всей видимости является адаптацией к окружающей среде паразита. *O. felineus* проводит большую часть своего жизненного цикла в условиях гипоксии и повышенная экспрессия миоглобина, способного запасать кислород, похоже, является необходимой для его выживания функцией [151]. Важной особенностью гемоглобинов паразитических видов является их высокая аффинность к кислороду многократно превышающая таковую у млекопитающих [152].

На рисунке 9 изображено выравнивание аминокислотной последовательности миоглобина *O. felineus*, предсказанной из наших данных и миоглобина близкородственных видов. Консенсус гем-связывающего сайта (консервативный домен cd01040) эукариот также показан на рисунке. Следует заметить, что в гем-связывающем сайте семь из десяти позиций, участвующих в связывании гема консервативны среди всех перечисленных видов. Имеются данные показывающие значение двух из этих семи позиций в повышенной аффинности паразитических миоглобинов к кислороду. Это позиции Tyr(34)B10 (расположена в альфа спирали, обозначенной “В”), и Tyr(68)E7 (расположена в альфа спирали, обозначенной “Е”): в них у трематод расположены тирозины в отличие от большинства остальных животных [153,154].

В ходе построения выравнивания был обнаружен полиморфизм *O. viverrini* в позиции (34)B10, с примерно равной представленностью тирозина и фенилаланина. Значимость фенилаланина в этой позиции неясна, т. к. никакие другие трематоды не имеют в этой позиции такую аминокислоту. Также имеются позиции, отличающие *O. felineus* (и описторхид вообще) от шистосом – это аминокислотные остатки Leu (47), Glu (99), and Ala (113). В позициях 47 и 99 у *F. hepatica* обнаруживается тирозин и аспарагиновая кислота, что отличает ее от всех остальных представленных здесь видов.

```

BBBBBBBBBBBBBBBB
      1   5   10
myoglobin_C_sinensis      MAPLSKDEVDALFEELNPLVSTTEQRTEFGKAVYMALFSAYPEYIQLFTK 50
myoglobin_O_viverrini     MAPLTKDEVDAILEELNPSVSTGEQRTEFGKAIYMALFTAHPPEYIGLFTK 50
myoglobin_O_felineus      MAPLTKDEVDVLLAELNPLVSTSEQKIEFGKAIYMALFTAYPEYIGLFSK 50
hemoglobin_F2_F_hepatica  MAVLTQTQIDSILADLAHHTDTEHITEMGVSIIYKTLFAAHPEYISYFSK 50
Globin-3_S_japonicum      MAAVTQSQVDHLITELEPHVDTEAHKLELGLKVYECFLKDRPEYICKFSR 50
myoglobin_S_mansoni      MTAVTQSQVDHLITELEPQVDTEAHQEELGLKVYECFLKARPEYICKFSR 50
*: ::: **: :* ..* : ** : : : **** *::
# # #
cd01040                    SAEKKLVKASWAKL---KADREIIGLEFYERLFKAHPETRALFSR

EEEEEEEEEEEEEEEEEEEE
      1   7
myoglobin_C_sinensis      MQGLTKDNVEASEGIKYGQIFADSILEMLQCASDDGKLEAVLEKSGKEH 100
myoglobin_O_viverrini     MEGLTKDNVEASEGIKYGRIFADSILEMLQASDDGGELEAVLEKSGKEH 100
myoglobin_O_felineus      MQGLTKDNVEASEGIKYGRIFLADSILEILQASDDGELDALLEKNGKEH 100
hemoglobin_F2_F_hepatica  LQGLTKDNVQSEGIRIYGRIFLGEELIRLLKAASNPSVLEERIVQAKDH 100
Globin-3_S_japonicum      LQGLDASNVAQSEGIKYARIFVAAFVPMIQAAANKCELDKLCLEEAILEH 100
myoglobin_S_mansoni      LQGLDVSNFAQSEGLKYARIFVSALLPMIKAAANKCELDKLCLEAVVLEH 100
:*** ..* ***:***:***: : : : : : * : . *
# # #
cd01040                    FGGLS-AALKGS PKFKAHGKRVLNALDEAIKNLDDLEALKALLAKLGRKEH

myoglobin_C_sinensis      ITRNVTKQQFLSAEEVFIKHFSGVLTk---EENKQSMERFLKHIVPKVAGFLG 150
myoglobin_O_viverrini     TTRNVTKQQFLSAEPVFIHCKGVLTk---EANKQSMKFFKHIIPKIANHLG 150
myoglobin_O_felineus      VTRSVTKQQFLSCEAVFIKHFNSVLTk---EENKQSMKFIKHIMPKIASQIG 150
hemoglobin_F2_F_hepatica  KARPVTKDQFTGAAPIFIKFFQGLLKK---QEDKDAIEKFLLVHMQAIAAKM- 149
Globin-3_S_japonicum      RTRPVDEKIFQDSLPIFIKIFNLIKN---QQNKETMSKILTYTFTMIGSQI- 149
myoglobin_S_mansoni      RNRQVNEQIFLDSLPIFIDFNNYIND---QQNKETMSKILTYVFKTIGSQI- 149
* * :. * .. :***. .. : : : : : : : : : : : : :
# #
cd01040                    AKRGVDPEHFKLFGALLEVLAELVLDGdfPEVKAAWDKLLDVIADAL

```

Рисунок 9: Выравнивание аминокислотных последовательностей миоглобина *O. felineus* и его ортологов из *C. sinensis* GenBank:AAM18464.1, *O. viverrini* (предсказан из данных доступных в NCBI EST), *F. hepatica* GenBank: ABW96608.1, *S. mansoni* GenBank:XP_002578632.1, и *S. japonicum* GenBank:CAX75468.1. Также изображен консервативный домен cd01040 из базы данных CDD. Цвет отражает различия между последовательностями. Важные альфа-спирали, обсуждаемые в тексте маркированы как “..EE..” and “..BB..”.

4.2.3 Вителлин (белок оболочки яйца)

Другой транскрипт с высокой экспрессией кодирует предшественник вителлина (GBJA01003468). Длина его транскрипта составляет около 817 нуклеотидов. Белки вителлины необходимы для формирования оболочки яиц трематод и иммунологически обнаруживаются только в вителлиновых гранулах [155]. Таким образом они играют важную роль в репродуктивной функции паразита. Столь высокая экспрессия согласуется с тем, что взрослые особи продуцируют тысячи яиц каждый день. Для *C. sinensis* и *O. viverrini* также показана высокая степень экспрессии этого гена [15,53,58].

Анализ выравнивания (Рисунок 10) данного белка *O. felineus* с организмами *O. viverrini*, *C. sinensis*, *S. mansoni* и *S. japonicum*. показывает высокую степень консервативности TES (trematode eggshell synthesis, синтез оболочки яйца трематод, pfam08034) домена среди всех семи трематод. Другие же части этого белка сильно варьируют, что возможно отражает глубокие различия в жизненном цикле этих видов.

```

vitelline_O_felineus      --MKLVAILLVLLPAVLGDGYGPRSRGHGRFYRGAYGDHTGRQTAASE 48
vitelline_C_sinensis     ---ARARAVLVLLPAVLGDGYGPRSRGHGRFYRGAYGDRTGRQTAASE 47
vitelline_O_viverrini    --MKLVAILLVALLPAVLGDGYGPHSRGHGRFYRGAYGDHTGGQTAASE 48
vitelline_F_hepatica     MKMKFTLVLLLAIVPLTLARHPHGKFNH--HAS YDDREKHRGRYKENDYLN 49
eggshell_prot_chorion_S_mansoni -----MFKSTETSKYGLVLAKPKRVIADPDLHIYRNKYLTQIKSPG 42
eggshell_precursor_S_japonicum -----MKR-----ISRKKN----- 9
                               .*

vitelline_O_felineus     YDLRGRLSQRGESSYSGDWDTHAREDDKDFYEIGGRYAVYGKAADDTDYG 98
vitelline_C_sinensis     YDLRGRLSQRGESSYSGDWDTHATEDDKDFYEVGGRYAVYGKASDDTDYG 97
vitelline_O_viverrini    YDLRGRLSQRGESSYSGEWDHAREDDKDFYEMGGRYAVYGKASDDTDYG 98
vitelline_F_hepatica     YDLKKGKFAGRGKAYLHGSFDDKYGNENERGRYDDQKYLKAGKSAHDGKYG 99
eggshell_prot_chorion_S_mansoni VEID-----VNEKTNNYKDYGSSDISPTSEYGISSIGTGVISGSKYS- 84
eggshell_precursor_S_japonicum -----THQSNFVPPGGS----IDSSMPSFTTGTISGSKHS- 42
                               ..: . . : * : .

vitelline_O_felineus     LNAYLKAKGKFYGHGNEEQGSKFEQFTKFRGGGHDSYGKKKHNDYDTY 148
vitelline_C_sinensis     LNAYLKAKGKFYGHGNEEQGSKFEQFTKFRGGGHDSYGKKKHNDYDTY 147
vitelline_O_viverrini    LNAYLKAKGRFYGHGNEEQGSKFEQFTKFRGGGHDSYGKKKHNDYDYSY 148
vitelline_F_hepatica     MYGNMYAKGDFKAYGNEDEGAKFEVVTFRGGGYDSYGKKKSYDDYDTK 149
eggshell_prot_chorion_S_mansoni -HSRLNVHGKMTGVGHQDYGVHYLSLTKFRKGGGYDIYGKKRQFLDYDTQ 133
eggshell_precursor_S_japonicum -HGRLDVRGKMAGVGHQDYGVNYLSLTKFRKGGGYDIYGKKRQFLDYDTQ 91
                               . : .*: : . *:: * : : ..*.*:*:*:* *::: : *:::
pfam08034                LKAKGKFRSYGNEDKGSKYSETTKFRKGGGYDKYGKKKHYSYDPTY

vitelline_O_felineus     GQMKKYGNRDVSSKFEMYGKLLKAK-----GKFDAYGKSDVASE 186
vitelline_C_sinensis     GQMKKYGNRDVSSKFEMYGKLLKAK-----GKFDAYGKSDVASE 185
vitelline_O_viverrini    GQMKKYGNRDVSSKFELYGKLLKAK-----GKFDAYGKSDVASE 186
vitelline_F_hepatica     GHLKKFANKGRQSKFDMYGNVKADGQAI SNGNMNAYGMFDSYGKYDQYGK 199
eggshell_prot_chorion_S_mansoni GHVKKYGNQHMKAKEVVGNLQGY-----ST 159
eggshell_precursor_S_japonicum GHVKKYGNQHMRTKFEVVGNLNGY-----ST 117
                               *: *: *: *: : *::: *::: .
pfam08034                GQVKKYGNNHVKNKFDLYGKL

vitelline_O_felineus     FDYKGYGQSGSSKDYADRDVYGKLSGYG-----KYKSYG--KLSG--- 225
vitelline_C_sinensis     FDYKGYGQSGSSKDYTD RDVYGKLG YG-----KYKAYG--KLSG--- 224
vitelline_O_viverrini    FDYKGYGQSGSSKDYTD RDSYGKLSGYG-----KYKAYG--KLSG--- 225
vitelline_F_hepatica     MNDQGYEAEAGKYNAHGNL DLYGHLRGYGGSSAASKSENYGNARESGRYE 249
eggshell_prot_chorion_S_mansoni FKDKSSFKPSGST-----YGNYGNG----- 180
eggshell_precursor_S_japonicum YKQKSSFTPPS-----GGYSSK----- 134
                               .. .: . . * .

vitelline_O_felineus     -YGSQDDYS-KHGRHADYDELGY 246
vitelline_C_sinensis     -YGSQDDYS-KYGRHADYDELGY 245
vitelline_O_viverrini    -YGSKDDYS-KYGRHADYDELGY 246
vitelline_F_hepatica     PYGRYEKYEDDYARETPYDKYSY 272
eggshell_prot_chorion_S_mansoni --NGGQSY----- 186
eggshell_precursor_S_japonicum --RGRRSY----- 140
                               .*

```

Рисунок 10: Выравнивание аминокислотных последовательностей предшественника вителлина из *O. felineus* и его ортологов из *C. sinensis* GenBank:AAS92977.1, *O. viverrini* GenBank:AAL23712.1, *F. gigantica* GenBank:ABL14193.1, *F. hepatica* GenBank:AAA29143.1, *S. mansoni* GenBank:XP_002571507.1, и *S. japonicum* GenBank:AAN39278.1. Также показан консервативный домен pfam08034 из базы данных CDD. Цвет букв консенсуса отражает разницу между последовательностью *O. felineus* и последовательностью консервативного домена.

4.2.4 Цистеиновые протеазы

O. felineus активно синтезирует различные протеазы. Эти белки играют многозначную роль в физиологии паразитов. Помимо участия в пищеварении эти белки играют и другие важные функции в физиологии паразита, являясь одним из основных компонентов экскреторно-секреторного продукта, что делает их перспективными с точки зрения разработки средств иммунодиагностики [156,157].

Протеазная активность позволяет гельминтам утилизировать гемоглобин хозяина [67] и коллаген [158], что возможно является ключевым агентом проникновения сквозь ткани на ранних стадиях заражения. Последнее особенно характерно для фасциол, мигрирующих к печени через внутренние органы). Также считается важной их роль в ассоциированном с паразитами канцерогенезе. Разные трематоды имеют специфичный для них репертуар экспрессируемых протеаз, отличающийся в зависимости от текущей локализации паразита и стадии развития. *F. hepatica* в основном секретирует катепсины В и L [159], причём L доминирует в экскреторно-секреторном продукте. В случае же описторхид доминирующим является класс катепсинов F [160], представляющих собой большое семейство генов, в относительно недавнее эволюционное время дуплицировавшееся до более десятка копий [57]. Также, они вовлечены во взаимодействия паразит-хозяин посредством модуляции иммунного ответа [161] (см. раздел «Обзор литературы»)

Согласно нашим данным *O. felineus* экспрессирует катепсин F наиболее активно, что согласуется с поведением других описторхид. EST последовательности, соответствующие этой мРНК составляют 3,6% всей кДНК библиотеки (92 EST), а транскрипт предшественника катепсина F, собранный из данных РНК-секвенирования (GBJA01002241) имеет значение экспрессии 55'291 FPKM.

Также любопытно, что метацеркария активно транскрибирует катепсин семейства L, характерный только для этой стадии. Протеазы данного семейства по некоторым данным вовлечены в процесс экзистирования [162]. Наши данные, по которым имеется катепсин L, характерный для метацеркарии (GBJA01003923), являются дополнительным подтверждением этой гипотезы.

4.2.5 28 кДа глутатион трансфераза

Глутатион S-трансферазы (GST) – это белки второй фазы инактивации ксенобиотиков. Однако, их роль в случае гельминтов оказывается более многозначной и включает молекулярные механизмы взаимодействия паразит – хозяин. Так, например, в экспериментах на *F. hepatica* была показана способность этого фермента к простагландин-синтазной активности [163]. Трематоды, так же как и другие организмы экспрессируют несколько классов изоформ глутатион трансфераз. Как минимум, всегда имеются представители σ - и μ - классов. Экспрессия GST зависит от стадии развития, достигая самых высоких значений во взрослой особи [56], что возможно связано с иммунной атакой хозяина на именно эту стадию. Соотношение упомянутых двух классов GST также зависит от стадий, однако, для описторхид характерно превалирование фермента класса σ .

В кДНК библиотеке *O. felineus* содержались последовательности гомологичные 28 и 26 кДа GST. 28 кДа GST составляет существенную фракцию всех мРНК, т. к. в библиотеке было 119 (4,6 %) последовательностей EST, соответствующих этой мРНК. Значение экспрессии соответствующего транскрипта (GBJA01007477) из данных РНК-секвенирования составляет 8778 FPKM. Длина мРНК составляет около 801 нуклеотида. Высокая экспрессия данного фермента наблюдалась и у *O. viverrini* [58].

На рисунке 11 изображено выравнивание аминокислотной последовательности GST *O. felineus*, предсказанной из наших данных, с наиболее близкими молекулами родственных представителей трематод. Каждая молекула GST содержит два консервативных домена: N-концевой домен, который включает глутатион связывающий сайт (так называемый G-сайт) и C-концевой домен, который включает субстрат-связывающий карман (так называемый H-сайт). Из выравнивания видно, что G-сайт имеет все две переменные позиции. Одна из них это позиция 73, в которой белок *C. sinensis* имеет остаток глицина вместо серина во всех остальных приведенных видах. Второй позицией является позиция 53, в которой различаются описторхиды (у которых закодирован метионин) от шистосоматид (лейцин) и фасциолид (валин), хотя эти аминокислоты могут быть функционально близкими. H-сайт, который ответственен за связывание различных субстратов показывает значительно большую переменность. Из семи аминокислот, составляющих H-сайт, четыре были идентичны в описторхидах, но отличались от шистосоматид и фасциолид (позиции 112, 113 и 168 аминокислотной последовательности *O. felineus*). В добавок к разнице между семействами паразитов, в позиции 108 у *O. felineus* закодирован лейцин, в отличие от других описторхид, имеющих фенилаланин в этой позиции и шистосоматид и фасциолид, которые имеют в этой позиции остаток тирозина. Эта схема замен возможно связана с несколько различной субстратной специфичностью и специализацией фермента к жизненному циклу паразита.

потенциальной канцерогенностью [95] (см. также раздел 4.8).

4.3 Транскриптом *O. felineus* и консервативные гены эукариот

В работе [138] авторы попытались найти множество генов, отличающихся консервативностью и легкостью в аннотации и, при этом, в идеале, характерное для всех эукариотических организмов. Для этого они рассмотрели полные геномы шести модельных организмов: *Homo sapiens*, *Drosophila melanogaster*, *Arabidopsis thaliana*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae* и *Schizosaccharomyces pombe*. В белок кодирующей части этих геномов они отобрали 456 генов, имеющих у всех шести видов, а также удовлетворяющих определенным критериям консервативности. Обнаруженные гены использовались в ими же разработанном программном обеспечении, названном SEGMA (Core Eukaryotic Genes Mapping Approach). Мы решили сравнить полученный транскриптом *O. felineus* с этими 456 генами. Ясно, что так как изначальные шесть видов, являющихся источником обнаруженных генов затрагивают широчайший круг эукариотических таксонов – от растений до животных – то разумно ожидать наличия большей их части и у *O. felineus*. Такое сравнение полученного нами транскриптома может дать двоякую информацию:

1. Если считать, что у *O. felineus* количество транскрибируемых (или имеющих) генов из этого набора очень близко к 456 (а этого можно ожидать в связи с высокой консервативностью этих белков), то можно дать оценку степени полноты покрытия реального транскриптома паразита в нашем исследовании.
2. Отсутствие каких-либо генов из набора SEGMA может указывать не на случайные ошибки методики, а отражать реальное упрощение на генетическом (или транскрипционном) уровне, характерное для исследуемого вида.

Для установления ортологии с этими генами был осуществлен поиск алгоритмом tBLASTn против соответствующих человеческих белков. Оказалось, что с пороговым значением $E \leq 10^{-10}$ в имеющемся у нас транскриптоме были идентифицированы 445 из 456 генов, а 11 генов из набора SEGMA не были найдены. Номера доступа и описание соответствующих 11 человеческих белков отображены в таблице 5.

Таблица 5: Консервативные белки *H. sapiens* из CEGMA, отсутствующие в исследуемом транскриптом *O. felineus*

Номер белка <i>H. sapiens</i>	Описание белка
NP_115661	Метилтиорибозо-1-фосфат изомераза
NP_955378	Профилин-4
NP_005099	Ксилулозо киназа
NP_057736	Crooked neck-like protein
NP_009111	Трегалаза
NP_001743	Каталаза
NP_003123	Спермидин синтаза
NP_002530	Орнитин декарбоксилаза изоформа 1
NP_001625	S-аденозилметионин декарбоксилаза изоформа 1
NP_036255	МТО1 гомолог

Факт отсутствия некоторых консервативных эукариотических белков в транскриптом *O. felineus* нами был подтвержден их отсутствием и в геномах родственных трематод *O. viverrini*, *C. sinensis* и *S. mansoni*. Исходя из этого можно сделать заключение, что отсутствие этих генов не случайно, а характерно для транскриптома (и, скорее всего, генома) *O. felineus*. Следовательно, мы можем заключить, что полученные данные о транскриптоме паразита представляет собой представительную информацию о практически всех транскрибируемых им генах.

Одним из отсутствующих генов является каталаза (NP_001743). Это фермент, повсеместно встречающийся у животных, вовлечен в деградацию пероксида водорода. В исследованиях на *F. hepatica* авторы обнаружили отсутствие каталазной активности [119]. По всей видимости, этого фермента нет у всех трематод, что следует из проведенной нами проверки доступных геномов. Отсутствие каталазной активности подчеркивает важность пероксиредоксинов для дезактивации пероксида водорода и других реактивных форм кислорода [164]. Кроме того, возможными участниками этого процесса могут быть глутатион, тиоредоксин и тиоредоксин пероксидаза. Возможно, также это следствие особой специализации системы окислительно-восстановительного баланса у трематод. Кроме того, имеются данные об отсутствии каталазы и у малярийного плазмодия [120]. Так как эти виды филогенетически очень удалены, то генетическая потеря каталазы похоже является чертой, связанной с переходом к паразитическому образу жизни.

Другой интересной особенностью является отсутствие трех генов синтеза полиаминов: спермидин синтазы (NP_003123), орнитин декарбоксилазы (NP_002530) и S-аденозилметионин декарбоксилазы (NP_001625). Эти молекулы имеют множество различных функций в эукариотах: регуляция клеточного роста и генной экспрессии, поддержание стабильности

мембран, пост-трансляционная модификация фактора инициации транскрипции eIF5A и защита от окислительного стресса [165]. Анаболизм полиаминов включает помимо трех упомянутых в SEGMA ферментов еще один [166], но оказалось, что и он не транскрибируется у описторха. Способность полиаминов инактивировать активные формы кислорода поднимает вопрос о существенности редукции (или как минимум выключение на уровне транскрипции) традиционных для других эукариот средств борьбы с активными кислородными соединениями при переходе к паразитизму в условиях с низким содержанием кислорода. Этот факт возможно связан также и с отсутствием каталазы.

Другие три отсутствующих консервативных транскрипта вовлечены в метаболизм углеводов: ксилулозо киназа (NP_005099), трегалаза (NP_009111) и метилтиорибозо-1-фосфат изомеразы (NP_115661). Генетическая потеря некоторых метаболических путей – характерная черта паразитического существования, что уже обсуждалось в соответствующем разделе обзора литературы. Их редукция является общей чертой паразитов, т. к. они могут использовать молекулы синтезируемые хозяином в качестве нутриентов [89]. Это хотя бы частично объясняет отсутствие этих трех ферментов, однако, в случае метилтиорибозо-1-фосфат изомеразы причины могут быть намного глубже. Этот фермент вовлечён в метаболизм метионина – цепь реакций непосредственно взаимосвязанной с синтезом полиаминов [167]. Метилтиoadенозин, являющийся побочным продуктом синтеза полиаминов является важным промежуточным соединением при метаболизме метионина. Этот факт подтолкнул нас проверить остальные ферменты, участвующие в метаболизме метионина. Оказалось, что большинство этих ферментов отсутствует в транскриптом, за исключением метилтиoadенозин фосфорилазы и S-аденозилметионин синтетазы (то же самое имеет место в случае *S. mansoni*). Таким образом, связь потери обоих этих путей становится более рациональной. Более того, метаболизм метионина является кислородозависимым и ранее была показана его редукция у анаэробов [168]. Малые количества кислорода, при которых приходится существовать печеночным сосальщикам могут объяснить эти изменения в метаболизме.

Отсутствие гомолога crooked neck-подобного белка (NP_057736) говорит о значительных изменениях в путях развития в сравнении с дрозофилой и человеком. Ранее было показано, что этот белок вовлечен в эмбриогенез и ассоциирован со сплайсосомой и вовлечен в древнюю консервативную систему сплайсинга пре-мРНК [169].

4.4 Классификация предсказанных белков *O. felinus* терминах Генной онтологии

Генная онтология (Gene Ontology; <http://www.geneontology.org>) представляет собой контролируемый словарь биологических терминов и иерархических связей между ними. Все содержащиеся там описательные термины принадлежат одному из трех больших доменов: молекулярная функция, биологический процесс и клеточный компонент. Мы попытались классифицировать имеющиеся последовательности в терминах GO с помощью программы Interproscan.

Функциональная классификация EST последовательностей кДНК библиотеки в терминах Gene Ontology была получена для 168 (63%) из 267 последовательностей EST (рисунки 12 – 14). Соответствующее распределение внутри трех классификационных доменов показывает, что наиболее представленными группами в категории молекулярной функции были каталитическая и связывающая активности, составляющие 29,8% и 40,4% соответственно. Категория биологических процессов в основном представлена белками вовлеченными в клеточные и метаболические процессы (33,4% и 42,6%, соответственно).

Аналогичный анализ данных РНК-секвенирования позволил выявить консервативные домены в 8998 транскриптах (71 %), при этом для 6684 из них (52,7% от всех белок-кодирующих транскриптов) были получены термины Генной онтологии. В этих транскриптах наиболее представленными (таблица 6) категориями были метаболизм азот содержащих соединений (GO:0034641, эта категории в основном состоит из нуклеотидного метаболизма) и транспорт (GO:0006810; в основном представленный мембранным транспортом). Наличие большого количества генов, связанных с транспортом указывает на большое количество секретируемых в желчный проток и обратно соединений. Другими представленными категориями были биосинтетический процесс и метаболизм малых молекул (GO:0044281, включает метаболизм аминокислот). Здесь имеются многие гены трансляции указывая на большие анаболические запросы паразита. Любопытно, что 475 генов вовлечены в передачу сигнала (GO:0007165), многие из которых могут быть важными мишенями для разработки лекарственных препаратов, т.к. эти пути уязвимы для фармацевтического вмешательства. Также любопытно наличие молекул, принадлежащих к категории взаимодействия с иммунной системой (GO: 0002376) таких как гомолог фактор связывания энхансера энтерлейкина, фактор ассоциированный с TNF-рецептором и два гомолога лектина.

В домене молекулярных функций основными группами Генной онтологии были

связывание ионов (GO:0043176), связывание ДНК (GO: 0003677), киназная активность (GO:0016301), активность трансмембранных транспортеров (GO:0022857), пептидазная активность (GO:0008233), связывание РНК (GO: 0003723), оксидоредуктазная активность (GO: 0016491) и активность передачи сигнала (GO:0004871). Замечу, что аналогично с доменом биологических процессов здесь в большом количестве представлены молекулярные функции передачи сигнала.

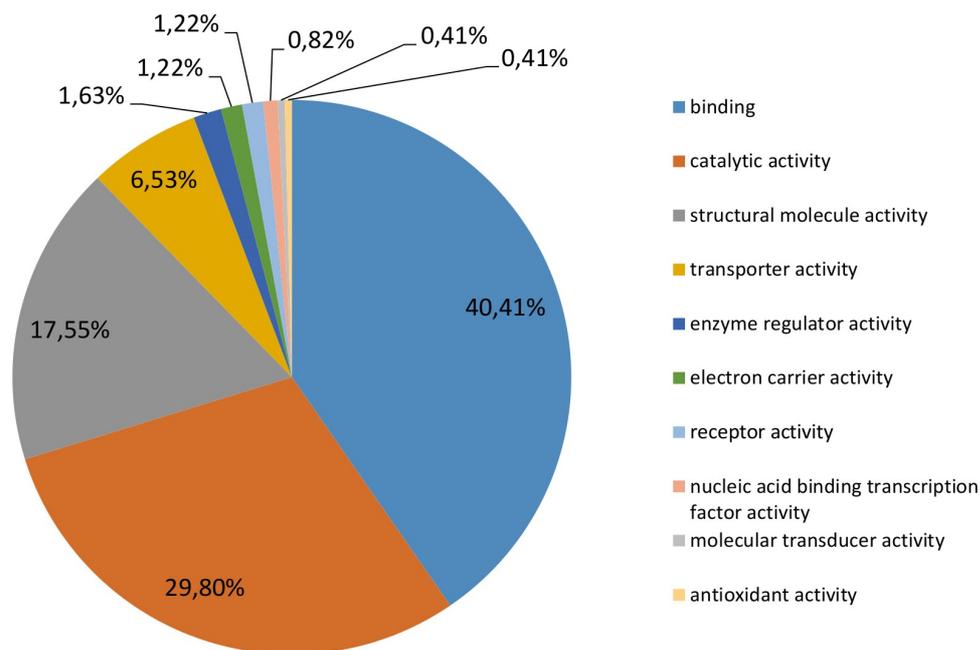


Рисунок 12: Классификация белков *O. felineus*, предсказанных из кДНК библиотеки в терминах молекулярных функций

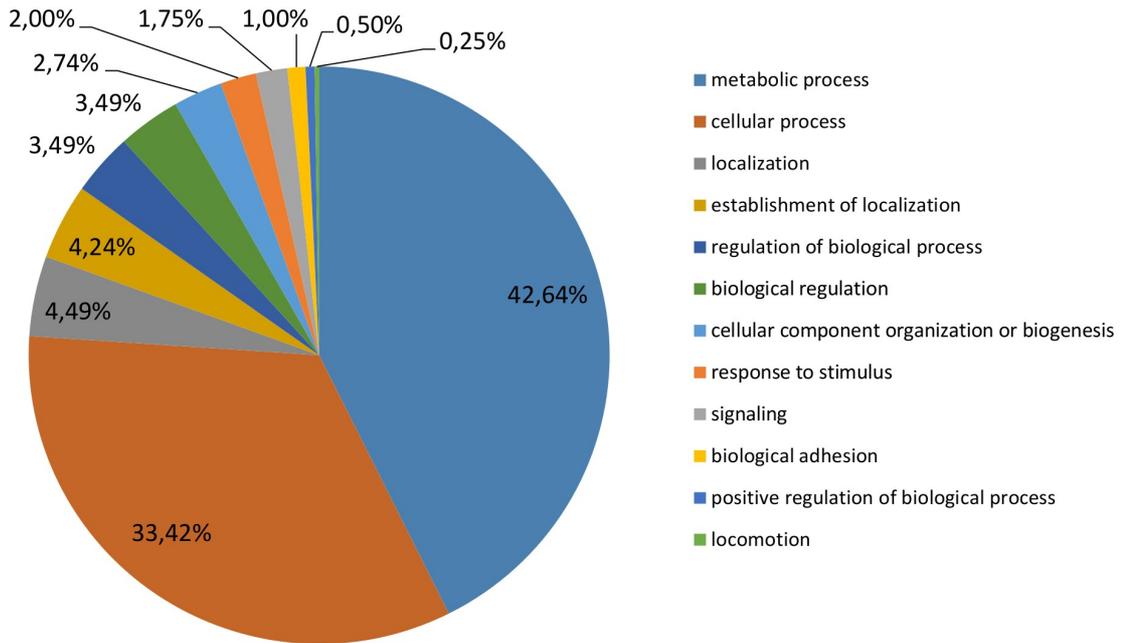


Рисунок 13: Классификация белков *O. felinus*, предсказанных из кДНК библиотеки в терминах биологических процессов

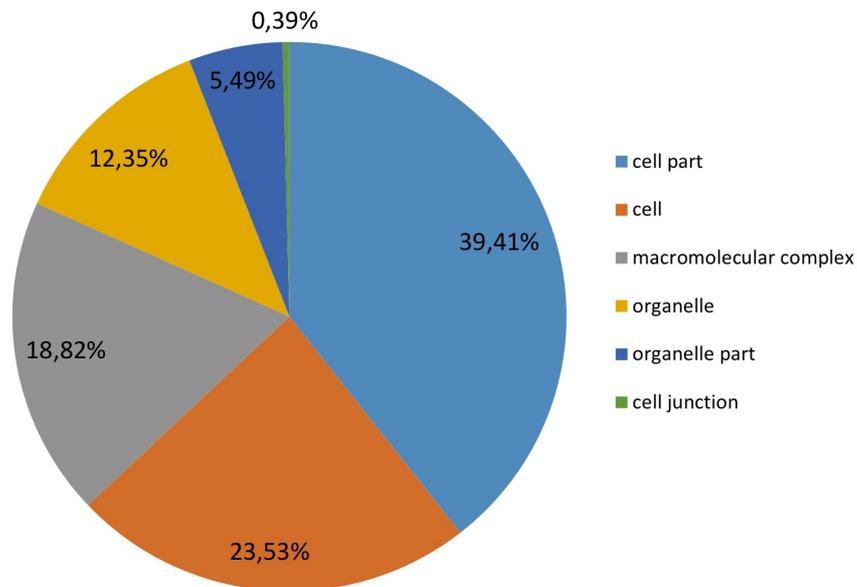


Рисунок 14: Классификация белков *O. felinus*, предсказанных из кДНК библиотеки в терминах клеточных компонентов

Таблица 6: Наиболее представленные термины Генной онтологии в транскриптах, полученных методом РНК-секвенирования

Термин Генной онтологии	Описание термина	Процент ГО-аннотированных генов, аннотированных термином
Биологические процессы		
GO:0034641	cellular nitrogen compound metabolic process	20,20%
GO:0006810	transport	17,10%
GO:0009058	biosynthetic process	13,70%
GO:0044281	small molecule metabolic process	12,30%
GO:0006464	cellular protein modification process	11,10%
GO:0007165	signal transduction	10,70%
GO:0009056	catabolic process	9,60%
GO:0055085	transmembrane transport	9,50%
GO:0034655	nucleobase-containing compound catabolic process	6,70%
GO:0006259	DNA metabolic process	5,00%
Молекулярные функции		
GO:0043167	ion binding	32,60%
GO:0003677	DNA binding	6,60%
GO:0016301	kinase activity	5,60%
GO:0022857	transmembrane transporter activity	5,50%
GO:0008233	peptidase activity	4,30%
GO:0003723	RNA binding	4,00%
GO:0016491	oxidoreductase activity	3,90%
GO:0004871	signal transducer activity	3,40%
GO:0001071	nucleic acid binding transcription factor activity	3,00%
GO:0005198	structural molecule activity	2,90%

4.5 Анализ метаболических путей с помощью KEGG

База данных KEGG (Kyoto Encyclopedia of Genes and Genomes, <http://www.kegg.jp/>) наряду с инициативой Gene Ontology представляет собой попытку формализации молекулярно-биологических данных. Однако, отличительной особенностью KEGG является то, что основной акцент сделан на метаболические пути. Также важно, что в этой базе данных присутствуют вручную проверенные данные по биохимическим путям *S. mansoni*, что облегчает работу с *O. felineus*.

Для 5816 белок-кодирующих транскриптов (46%) была установлена KEGG ортология. Из них 3559 относились хотя бы к одному KEGG пути. Помимо шистосомы в сравнение путей были добавлены модельные организмы отдаленных таксонов – *H. sapiens* и *C. elegans* (приложение 3).

Среди наиболее похожих между *O. felinus* и млекопитающими (по количеству имеющихся ортологов) путей представлены такие фундаментальные метаболические пути как цикл трикарбоновых кислот (ko00020) и пентозо-фосфатный путь (ko00030), а также некоторые другие консервативные пути: белковый экспорт (ko3060), путь контроля мРНК (ko03015), репликация ДНК (ko03030), экцизионная репарация нуклеотидов (ko03420), рибосома (ko03010) и сплайсома (ko03040). Однако в большинстве путей *O. felinus* имеет меньшее количество ортологов, чем человек. Среди наиболее редуцированных путей следует отметить такие пути как метаболизм ретинола (ko00830) и Jak-STAT сигнальный путь (ko04630). Практически полное отсутствие ортологов во втором из этих путей говорит либо о его редукции, либо о филогенетической удаленности его эффекторных молекул. В KEGG пути взаимодействия нейроактивный лигандов и их рецепторов (ko04080) 39 ортологов обнаружены у *O. felinus*, однако, у таких видов как *S. mansoni* и *C. elegans* их еще меньше – 15 и 22 соответственно.

Анализ пероксисомы (ko04146) выявил только 17 ортологов у *O. felinus* (для сравнения человек и нематода имеют 71 и 39 ортологов соответственно). Генетическая потеря пероксисомального аппарата также была ранее обнаружена на моногенетических плоских червях [170] и цестодах [89]. Более детальное рассмотрение показало отсутствие транскрипции всех PEX генов (Рисунок 15), являющихся основными регуляторами образования и поддержания пероксисом [171]. В то же время обнаруженные 17 ортологов осуществляют процессы у большинства эукариот происходящие в пероксисоме. Мы считаем, что эта находка отражает практически анаэробные условия существования печеночных сосальщиков. С другой стороны, пероксисомы обычно вовлечены в метаболизм очень длинных и разветвленных жирных кислот [172] (другие жирные кислоты могут быть метаболизированы в митохондриях [173]), которых может и не быть в желчном протоке.

Сравнение метаболических путей с шистосомой показало обратную тенденцию: как правило, у *O. felinus* большее количество ортологов, чем в том же пути *S. mansoni*. Возможно, это показывает приспособленность к сильно различающимся средам обитания, таким как кровеносный сосуд и желчевыводящие пути. К примеру, путь деградации жирных кислот более развит у описторха. Очевидно, что жирные и желчные кислоты являются важным источником углерода и энергии для печеночных сосальщиков [70], но не для кровяных сосальщиков. Важность липидного метаболизма для описторхид была показана на геномных данных *O. viverrini* в работе [13]. Янг с соавторами идентифицировал различные ферменты липолиза: Ацил-КоА:холестерин ацилтрансфераза, фосфолипазы, холинэстеразы, липаза лизосомальных липидов, галактозилцерамидаза и сфингомиелин фосфодиэстеразы. Мы обнаружили

соответствующие ортологи и у *O. felineus*, что показывает его схожую способность использовать липиды как источник энергии.

Кроме того, в той же работе [13] было обнаружено 25 белков содержащих консервативный MD-2 липид-связывающий домен. Было показано, что некоторые белки с этим доменом вовлечены в транспорт холестерина из лизосом в другие органеллы и плазматическую мембрану человеческих клеток [174]. Янг с соавторами выдвинули гипотезу, что эти белки возможно участвуют в транспорте стеролов, однако, оказалось (по нашим данным), что они активны во втором промежуточном хозяине, т. к. подавляющее большинство из них были транскрипционно активны в метацеркарии (Рисунок 16, В). Кроме того, к нашему удивлению у *O. felineus* было обнаружено 52 транскрипта, кодирующих такие белки. Для сравнения, большинство эукариот имеют только одну копию такого гена [13,175].

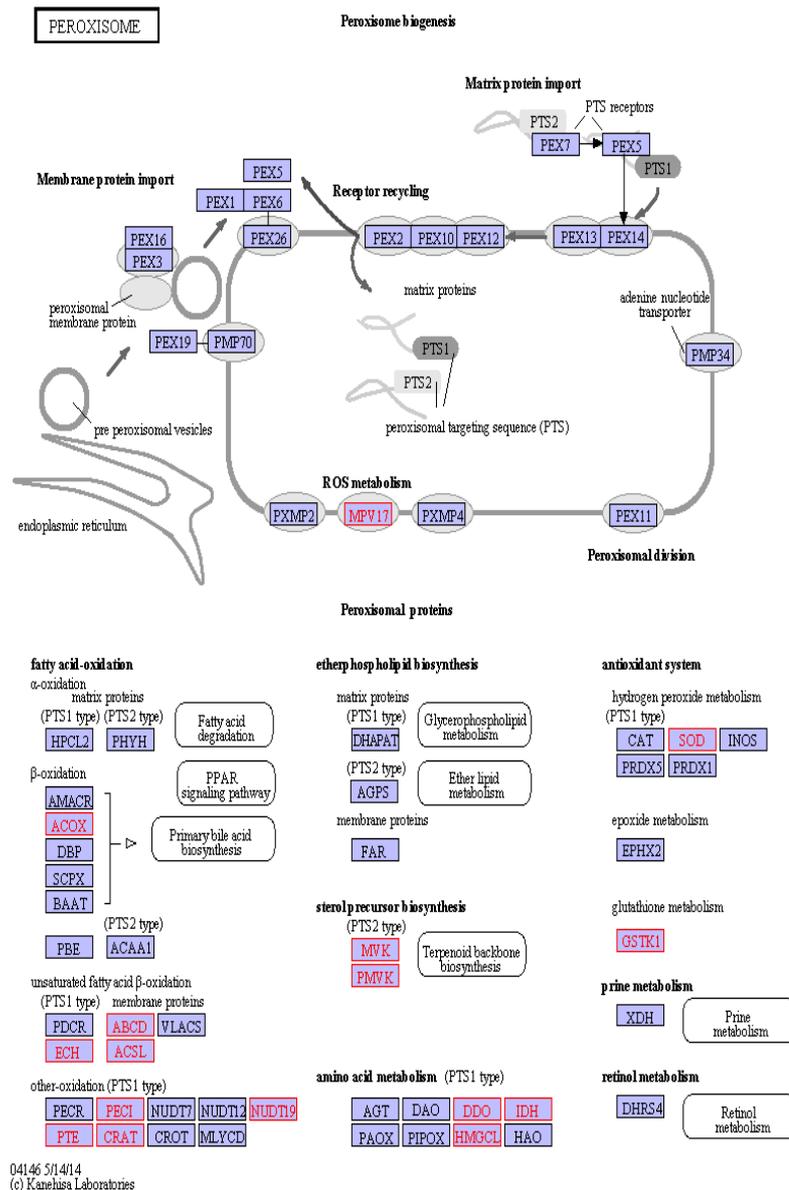


Рисунок 15: Редукция пероксисомального аппарата *O. felineus*. Красным обозначены белки, обнаруженные в транскриптом *O. felineus*.

4.6 Транскриптом метацеркарии *O. felineus* – сравнение со взрослой стадией

Основной проблемой, возникающей при молекулярно-биологических исследованиях метацеркарий описторхид – это контаминация другими паразитами. В азиатском регионе рыба заражена разными видами трематод: помимо печеночных сосальщиков обнаруживаются кишечные сосальщики семейства гетерофиид (Heterophyidae) [176], которых очень трудно отличить морфологически. Однако, особенностью *O. felineus* является то, что в ареале его обитания вероятность обнаружения метацеркарий других паразитов крайне мала: наиболее

вероятным сопутствующим паразитозом может являться *Metorchis bilis* [4], однако, в имеющихся у нас образцах его обнаружено не было, т. к. в хомячках после заражения выделенными метацеркариями выявлялись только мариты *O. felineus*.

Есть способы проверить контаминацию и на молекулярном уровне. С этой целью нами была проведена отдельная сборка прочтений образцов марит (которые если и загрязнены, то млекопитающим-хозяином и бактериями) и образца метацеркарий. Из полученных сборок транскриптов были выделены последовательности рРНК (в большом количестве присутствующие в любом образце тотальной РНК, даже не смотря на полиА обогащение). Их сравнение показало 100% идентичность первичной последовательности. Эти факты в совокупности позволяют говорить об относительной чистоте имеющихся образцов.

Для сравнения стадий развития мы применяли разные биоинформатические методы. В первую очередь мы попытались определить в собранных транскриптах подгруппы специфичные только для определенной стадии, т. е. применяя максимально строгий порог отсека. Для этого мы относили транскрипт к специфичным для взрослой стадии (метацеркарии), если его экспрессия (выраженная в количестве картированных фрагментов) была не равна нулю во взрослой стадии (метацеркарии) и равнялась нулю в метацеркарии (взрослой стадии). Следуя этой логике, мы идентифицировали 895 метацеркарий-специфичных транскриптов и 632 транскрипта, специфичных для взрослой стадии. Мы использовали Cytoscape чтобы исследовать термины Генной онтологии, обогащенные в каждой из подгрупп транскриптов.

В домене биологических процессов во взрослой стадии были обогащены такие группы транскриптов, как локализация липидов (GO:0010876), липидный транспорт (GO:0006869), процессы связанные с микротрубочками (GO:0007017), негативная регуляция эндопептидазной активности (GO:0010951) и регуляция протеолиза (GO:0030162). Категории обогащенные в метацеркарий-специфичных генах существенно отличались и включали адгезию клетка-клетка (GO:0098609), G-белок связанные пути трансдукции сигнала (GO:0007186), пути сигнала через поверхностные клеточные рецепторы (GO:0007166), трансмембранный транспорт ионов (GO:0034220), биологическую регуляцию (GO:0065007), транспорт ионов металлов (GO:0030001), регуляцию клеточных процессов (GO:0050794), протеолиз (GO:0006508), регуляцию биологических процессов (GO:0050789), транспорт катионов (GO:0006812), транспорт нейротрансмиттеров (GO:0006836) и трансмембранный транспорт (GO:0055085). Эти находки указывают на наличие в метацеркарии множества специфичных для этой стадии сигнальных молекул, возможно вовлеченных в процесс экцистирования и регуляции смены стадии жизненного цикла.

Для дальнейшего, более детального исследования путей трансдукции сигналов мы решили исследовать GPCR-рецепторы, а точнее гомологи уже обнаруженных 57 GPCR-рецепторов в геноме *O. viverrini* Янгом с соавторами [13]. GPCR-рецепторы трематод представляют собой большое семейство белков, относящихся к всем 5 категориям [177] системы классификации GRAFS [178]. Эти белки представляют собой особенный интерес, так как они задействованы в жизненно важных процессах, таких как регуляция генной экспрессии, подвижность, хемотаксис и представляют основную мишень (более 30%) всей фармакологической индустрии [81]. Нам удалось обнаружить гомологов 53 из 57 GPCR-рецепторов из геномных данных *O. viverrini*. Большинство из этих генов (31 из 53) являются дифференциально экспрессируемыми (Рисунок 16, А). Более того можно утверждать, что это не случайно, т.к. аналогичная процедура, проделанная для консервативных генов эукариот, показывает очень высокую степень корреляции между стадиями (Рисунок 16, Г), а применение точного теста Фишера указывает на статистически значимую обогащенность этого пути дифференциально экспрессирующимися генами ($P = 4,6 \cdot 10^{-11}$). Другой удивительной особенностью было то, что из 31 дифференциально экспрессирующихся GPCR-рецепторов 28 имели большую экспрессию на стадии метацеркарии, а значит, среди них возможно имеются регуляторы процесса экцистирования при поглощении окончательным хозяином.

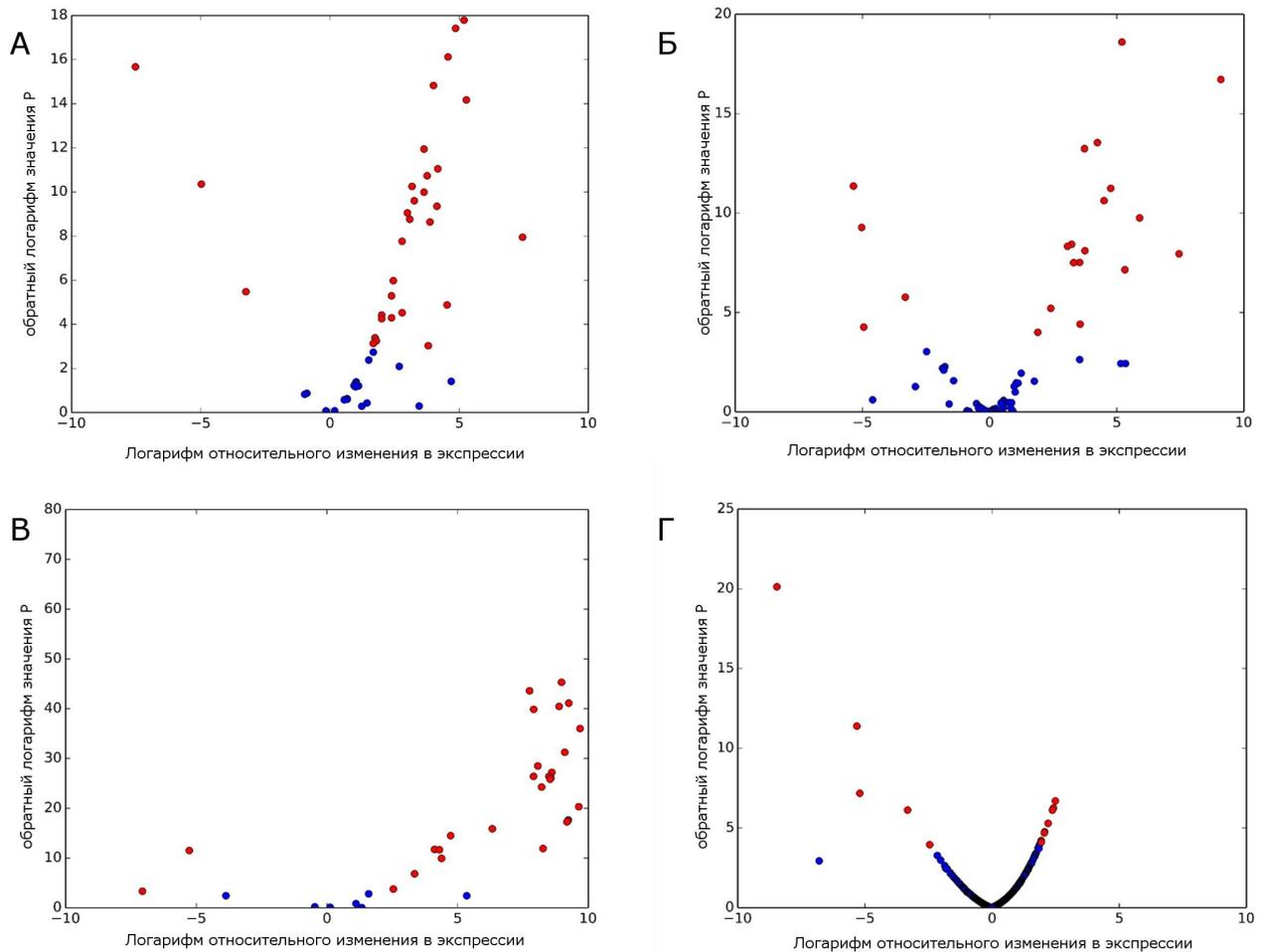


Рисунок 16: Дифференциальная экспрессия обсуждаемых в тексте подгрупп белков в разных стадиях *O. felineus*. На каждом рисунке по оси абсцисс отложен логарифм относительного изменения в экспрессии (отношение FPKM в метацеркарии к FPKM во взрослой стадии), а по оси ординат – отрицательный логарифм значения P value. Красным отмечены точки с условием $P \text{ value} < 10^{-5}/N$, где N – общее число генов в транскриптом. Остальные транскрипты изображены синим. А. GPCR рецепторы Б. Гомеодомен-содержащие белки В. MD-2 домен-содержащие белки Г. Консервативные гены эукариот (см. раздел 4.3)

Другой группой генов, которая, как мы считаем, также взаимосвязана с процессом межстадийного перехода, является суперсемейство гомеодомен-содержащих белков. Гомеодомен способен присоединяться к ДНК и обычно содержится в транскрипционных факторах, задействованных в процессе развития. НОХ гены это подгруппа гомеодомен-содержащих белков, ответственная за дифференциацию вдоль антеро-постериорной оси во многих животных [179]. Экспрессия гомеодомен-содержащих генов также показывает существенные различия между стадиями: 22 из 63 дифференциально экспрессируются (Рисунок 16, Б). Значение P в точном тесте Фишера для подгруппы гомеодомен-содержащих белков равно 10^{-3} .

Тот же статистический тест Фишера мы использовали для полнотранскриптомного поиска

биохимических путей, статистически достоверно обогащенных дифференциально экспрессирующимися между взрослой стадией и метациркуляцией транскриптами (см. раздел 3.4.2). В итоге было обнаружено семь таких путей ($P < 10^{-2}$, таблица 7). Особенно выделялся путь лизосом (map04142), включавший 94 дифференциально экспрессирующихся гена. Более подробное рассмотрение показало, что существенным источником межстадийных различий в случае лизосом являлись катепсины. Из 56 идентифицированных транскриптов катепсинов 24 дифференциально экспрессировались, что указывает на высокую специализацию пептидазного репертуара в зависимости от стадии развития. Согласно KEGG NPC2-подобные белки, рассмотренные ранее в разделе 4.5, также относятся к лизосомам и, как уже говорилось, практически все такие обнаруженные белки экспрессируются различно между стадиями. Такой паттерн экспрессии лизосомальных белков вместе с большим количеством дифференциально экспрессирующихся генов в таких путях как метаболизм аргинина и пролина (map00330) и фагосома (map04145) указывают на существенные метаболические различия между стадиями.

Также стоит упомянуть различия в системе сигналинга посредством фосфатидилинозитола (map 04070, не показано в таблице 7, значение P также достаточно низко и равно 0,013) и взаимодействие нейромедиаторных лигандов и рецепторов (map04080), так как эти пути затрагивают системы межклеточной передачи сигнала. Более подробное рассмотрение пути №04070 выявило 26 транскриптов относящихся к ортологической группе K02183, т. е. кодирующих кальмодулины – внутриклеточные регуляторы, являющиеся сенсорами ионов кальция [180]. Большинство из них экспрессировались в метациркуляции.

Итоговая иллюстрация, включающая межстадийные различия на уровне транскриптома, обсуждаемые в тексте изображена на рисунке 17

Таблица 7: KEGG пути, обогащенные дифференциально экспрессирующимися генами (ДЭГ) согласно точному тесту Фишера.

Номер KEGG пути	Описание	Значение P обогащенности ДЭГ	Количество генов <i>O. felineus</i> входящих в путь	Количество ДЭГ, входящих в путь
04142	Lysosome	3.46E-23	173	76
04080	Neuroactive ligand-receptor interaction	3.45E-08	73	29
04145	Phagosome	7.45E-06	91	29
00740	Riboflavin metabolism	3.15E-03	12	6
04512	ECM-receptor interaction	4.26E-03	29	10
00350	Tyrosine metabolism	7.84E-03	18	7
00330	Arginine and proline metabolism	8.49E-03	27	9



Рисунок 17: Обобщение различий между марицей и метацеркарией на уровне полного транскриптома.

4.7 Рибосомальные белки и филогения описторхид

Для выяснения филогенетических взаимоотношений внутри семейства описторхид мы решили использовать рибосомальные белки. Для этого мы построили набор ортологов 32 рибосомальных белков с доступными в NCBI последовательностями для *C. sinensis*, *O. viverrini*, *S. japonicum*, *S. mansoni*, *F. gigantica*, и *F. hepatica*.

Известно, что описторхиды образуют кладу близкородственных видов, но есть две точки зрения на взаимоотношения внутри семейства. Согласно первой *O. felinus* и *O. viverrini* ближе друг к другу, чем каждый из них к *C. sinensis*; согласно другой, *O. viverrini* и *C. sinensis* ближе друг к другу, чем к *O. felinus*. Обе точки зрения находят свои подтверждения в морфологических критериях и доступных молекулярно-генетических данных, полученных для митохондриальных и ядерных маркеров [19,20].

Мы выбрали рибосомальные белки в связи с универсальностью их функции среди всех

организмов и легкостью установления ортологических взаимоотношений. Полученных конкатемер аминокислотных последовательностей 32 рибосомальных белков анализировался методами максимальной парсимонии и максимального правдоподобия. Обоими методами была установлена одна и та же топология (Рисунок 16). Данная филогения позволяет сделать три заключения: во-первых, компактность клада описторхид такая же как и в случае шистосом и фасциолы, что не позволяет выделять *C. sinensis* и описторхов в разные рода; во-вторых, топология ветвления внутри описторхид поддерживает гипотезу, что *O. viverrini* и *C. sinensis* ближе друг к другу, чем каждый из них к *O. felineus*. Последнее заключение согласуется с предыдущими результатами, полученными при анализе ядерного маркера парамиозина [20].

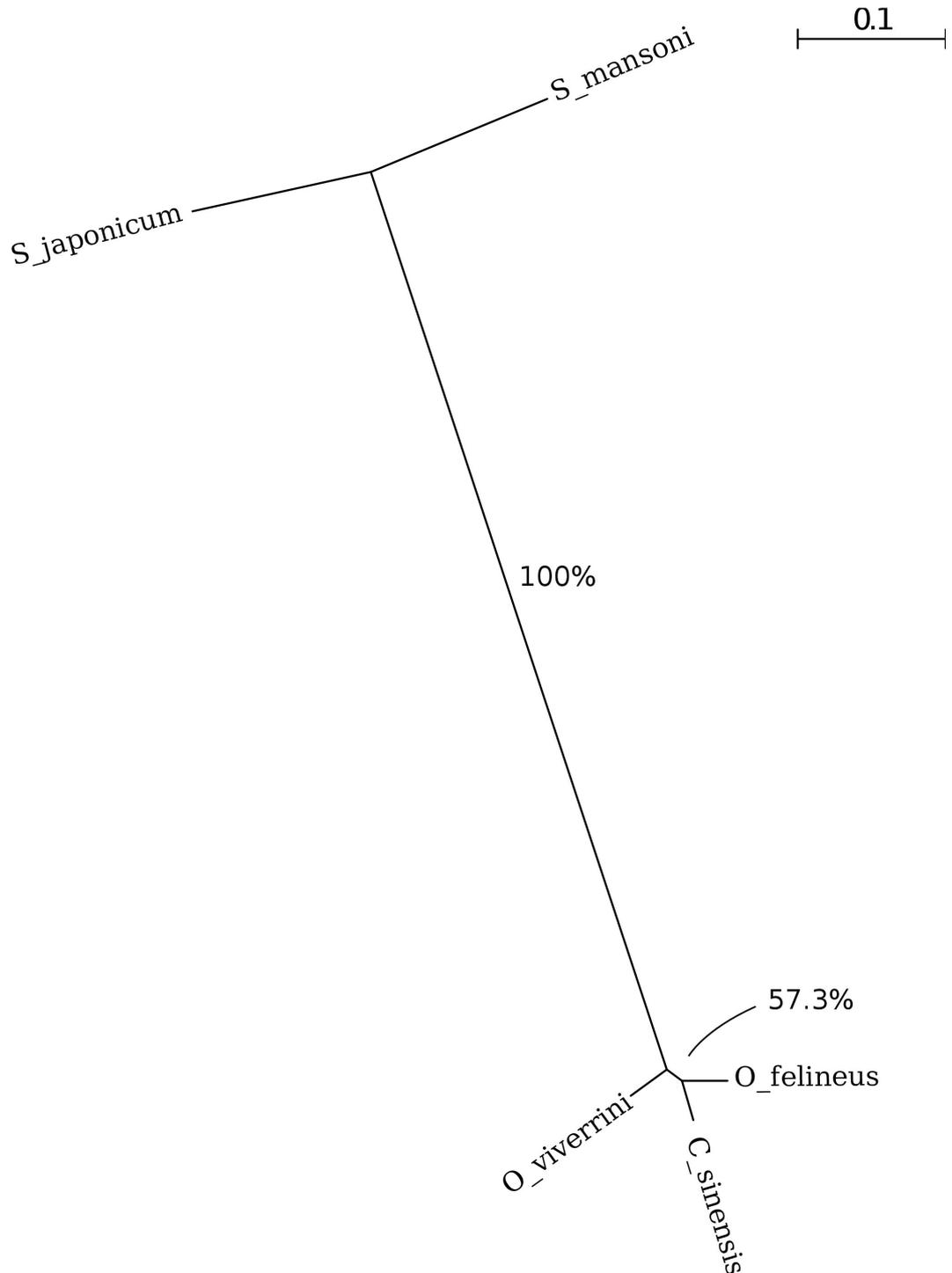


Рисунок 18: Филогенетическая взаимоотношения среди трематод. Филогенетическое дерево построено с помощью метода максимального правдоподобия по 32 рибосомальным белкам. Бутстреп поддержка всех ветвей была равна 100%.

4.8 Потенциальная канцерогенность *O. felineus*

O. viverrini и *C. sinensis* ассоциированы с повышенным риском холангиокарциномы [8,91]. Было показано, что встречаемость этого смертельного рака значительно выше в регионах Тайланда, в которых высока степень заражённости *O. viverrini* [181]. Это стало причиной классификации этих видов Агентством по исследованию рака как канцерогенов группы I («канцерогенны для человека») [90], однако, мало известно о канцерогенности *O. felineus*. Мы решили сравнить по транскриптомным данным гены, которые считают вовлечёнными в процесс паразит-индуцированного канцерогенеза.

Пути патогенеза описторхоз-ассоциированной холангиокарциномы считаются мультифакторными и включают механические повреждения, иммунопатология, вызванная воспалением и прямые эффекты ростовых факторов, секретируемых гельминтом [182]. Некоторые белки секрета *O. felineus* могут вызывать повреждения клеток и тканей. К примеру, катепсины в большом количестве обнаруживаются в транскриптоме паразита (см. раздел 4.2.4). Активные пептидазы, расщепляющие белки хозяина могут вносить существенный вклад в воспаление и злокачественное перерождение.

Секретируемые ростовые факторы (в особенности гранулины) считаются важнейшими канцерогенными факторами ассоциированными с печеночными сосальщиками [95]. В геноме *O. viverrini* закодированы по крайней мере три белка, содержащих гранулиновый домен, несколько различающихся по доменной организации [13]. У *O. viverrini* есть гомолог програнулина с восемью гранулино-подобными доменами, а также два дополнительных однодоменных гранулина. В транскриптоме *O. felineus* мы обнаружили пять гранулиноподобных транскриптов. Самый длинный из них кодирует белок с как минимум семью (GBJA01002699) гранулиновыми мотивами (при транскриптомных исследованиях трудно гарантировать, что мы располагаем полным транскриптом, который может присутствовать в реальных условиях). Это гомолог програнулина, у которого высок уровень экспрессии и во взрослой стадии и в метацеркарии (296 и 82 ФРКМ соответственно). Другие четыре транскрипта содержат по одному гранулиноподобному домену как и в случае *O. viverrini*. GRN-1 (GBJA01012525) имеет очень высокую экспрессию во взрослой стадии (5729 и 129 ФРКМ в двух стадиях развития). Примечательно, что именно его ортолог, как было показано, является сильным митогеном для клеток млекопитающих [95].

Интересно, что были обнаружены два новых гранулина, специфичных для *O. felineus* (GBJA01012526 и GBJA01006896), т. к. не удалось обнаружить их ортологов в родственных описторхидах. Эти гранулиноподобные транскрипты кодируют белки, проявляющие

консервативность в районе гранулинового домена, но сильно варьируемые на концах (Рисунок 19). Мы также провели филогенетический анализ белковых последовательностей гранулиновых доменов, кодируемых обнаруженными транскриптами *O. felineus* и имеющимися последовательностями гранулинов *O. viverrini* и *C. sinensis*. На филогенетическом дереве видна четкая кластеризация ортологов GRN-1 и GRN-2, принадлежащих разным описторхидам (Рисунок 20). Любопытно, что новые гранулиноподобные транскрипты, а особенно GRN-3, имеющий высокую экспрессию во взрослой стадии (788 FPKM), являются более филогенетически удаленными, что делает их любопытной мишенью для исследований. GRN-4 проявляет очень низкую экспрессию в двух исследованных стадиях и возможно более активен на других жизненных стадиях паразита.

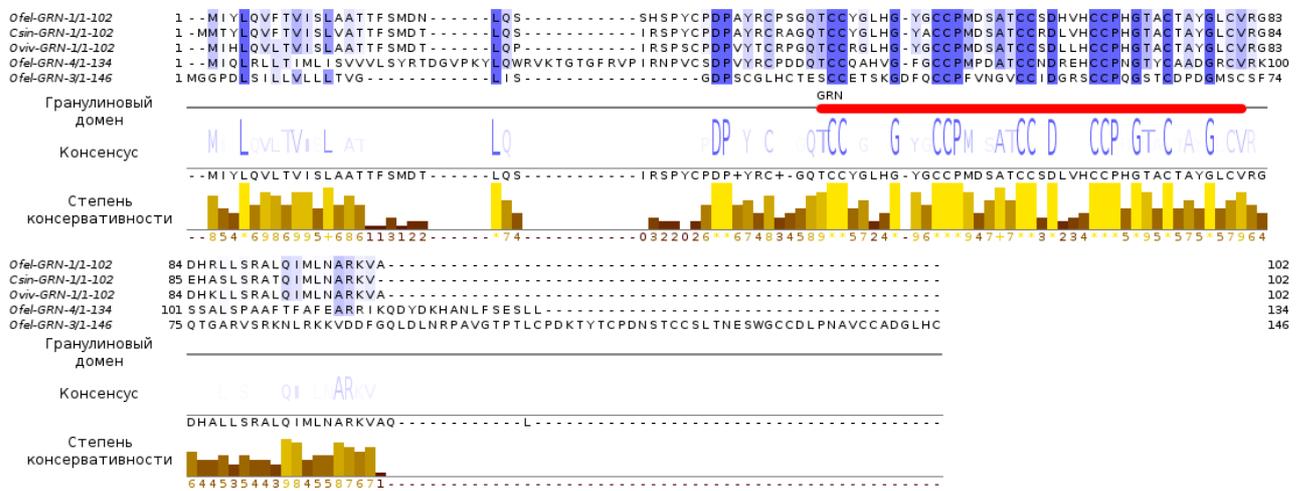


Рисунок 19: Первичная структура новых гранулиноподобных молекул, обнаруженных в транскриптом *O. felineus*. Выравнены следующие молекулы: *Ofel-GRN-1* (GBJA01012525), *Csin-GRN-1* (AT006891), *Oviv-GRN-1* (FJ436341), *Ofel-GRN-3* (GBJA01012526), *Ofel-GRN-4* (GBJA01006896). Последние две – новые (не обнаруженные у других описторхид) гранулины *O. felineus*. Также на выравнивании размечены: гранулиноподобный домен (красным), консенсусная последовательность (размер букв отражает консервативность соответствующей позиции) и степень консервативности.

Эти молекулярные данные, в сумме с повышенной в сравнении с другими описторхидами патогенностью *O. felineus* [10], являются свидетельством того, что он может быть вовлечён в развитие холангиокарциномы. Хотя текущие медицинские данные о такой ассоциации пока отсутствуют, нужно иметь ввиду потенциальные злокачественные перерождения при описторхозе, вызванном *O. felineus*.

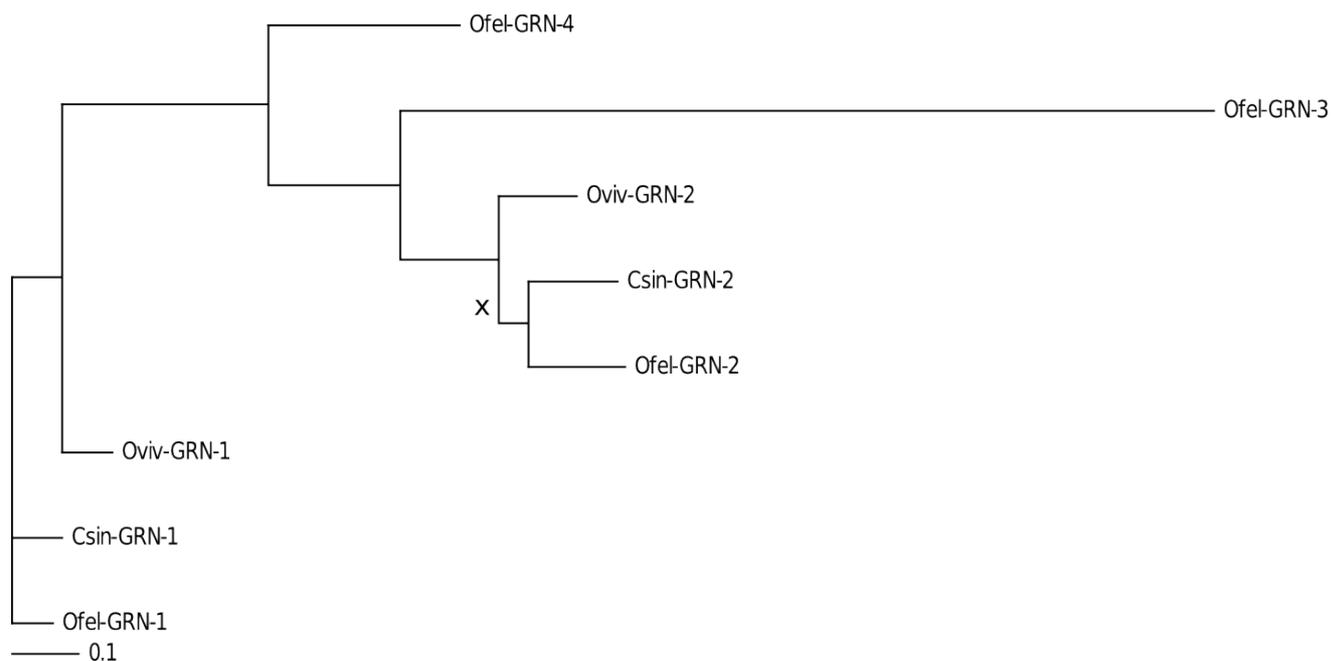


Рисунок 20. Филогенетическое дерево, построенное по аминокислотным последовательностям гранулиновых доменов, извлеченных из гранулиноподобных белков описторхид. Значение P было меньше для всех ветвей за исключением обозначенного «х» (для этой ветви значение $P < 0.14$, что тоже является достаточно низким). Обратите внимание на кластеризацию ортологов из разных видов и удаленность белка *Ofel-GRN-3*. Молекулы, использованные в выравнивании *Ofel-GRN-1* (GBJA01012525), *Csin-GRN-1* (AT006891), *Oviv-GRN-1* (FJ436341), *Ofel-GRN-2* (GBJA01001691), *Oviv-GRN-2* (T265_07410), *Csin-GRN-2* (GAA49645), *Ofel-PGRN* (GBJA01002699), *Oviv-PGRN* (T265_10096), *Csin-PGRN* (GAA54205), *Ofel-GRN-3* (GBJA01012526), *Ofel-GRN-4* (GBJA01006896).

5 Заключение

В данной работе впервые проведено транскриптомное исследование важного паразита человека и животных *O. felinus*. В работе использовались два подхода. Первый – на сегодняшний день в значительной мере устаревший – состоял в создания кДНК библиотеки с последующим ее секвенированием по Сэнгеру. Таким путем было получено около 4 тысяч EST последовательностей. Второй подход – значительно более информативный – включал массовое параллельное секвенирование на платформе Illumina HiSeq 2000. Таким путем было получено около 100 млн прочтений кДНК взрослой стадии *O. felinus* и 50 млн прочтений кДНК метацеркарии. Эти данные использовались для сборки мРНК транскриптов. В результате сборки EST последовательностей получено 267 уникальных транскриптов, а в результате сборки NGS прочтений получено 12'665 уникальных белок-кодирующих транскриптов. Последняя величина сравнима с общим количеством генов, полученным для родственных видов описторхид. Последовательности EST находятся в базе данных NCBI под номерами JK624271–JK626790, JK006511–JK006547, JK649790–JK649792, а белок-кодирующие транскрипты, полученные на основании данных массового параллельного секвенирования находятся в TSA архиве базы данных NCBI под номером GBJA01000000.

Полнота имеющего транскриптома была исследована путем сравнения с консервативными генами эукариот. В результате было установлено, что полученный транскриптом близок к полному, а исходя из отсутствующих консервативных генов, была обнаружена редукция таких систем как синтез полиаминов и метаболизм метионина. Далее, белок-кодирующие транскрипты были проаннотированы по гомологии с транскриптами и белками других организмов. В результате этого статистически значимые гомологи были обнаружены для 87,8% транскриптов. Путем поиска консервативных доменов 57,1% транскриптов были проаннотированы в терминах Генной онтологии, а для 46% транскриптов была установлена ортология по базе данных KEGG.

Помимо качественной информации о нуклеотидных последовательностях методы NGS секвенирования позволяют дать и количественную оценку их уровня экспрессии. Используя эту информацию, нами были установлены транскрипты с наибольшей экспрессией в каждой из стадий жизненного цикла. Во взрослой стадии среди 20 транскриптов с наибольшим значением ФРКМ следует отметить транскрипты катепсина F, миоглобина, белка оболочки яйца и глутатион трансферазы. Особенно любопытно, что наибольшим значением ФРКМ обладал транскрипт, кодирующий гомолог HDM белка *F. hepatica*, для которого показано множество иммуномодуляторных эффектов. Другое имеет место для метацеркарии: наибольшее значение

ФРКМ имеют транскрипты, кодирующие белки домашнего хозяйства, такие как рибосомальные белки (11 из 20 транскриптов с наибольшей экспрессией), а также убиквитин, что показывает метаболическую специализацию метацеркарии, отличную от мариты. Это, по всей видимости, связано с тем, что метацеркарии не размножаются, а лишь ждут поглощения окончательным хозяином.

Анализ биохимических путей позволил выявить особенность, ранее обнаруженную на родственных видах: *O. felineus* не экспрессирует ни один ген формирования пероксисом (т. н. PEX гены), что возможно говорит об отсутствии этой органеллы как таковой. В сумме с редукцией синтеза полиаминов и метаболизма метионина, а также с отсутствием гена каталазы, данный факт позволяет предположить черты факультативно анаэробного метаболизма у *O. felineus*.

Сравнение профилей экспрессии двух исследованных жизненных стадий было проведено с разными порогами строгости. Один из подходов включал строгое разделение все транскриптов на подмножество тех, которые экспрессируются в одной стадии, но при этом не детектируемы (ФРКМ равно нулю) в другой. Следуя этой логике, было обнаружено 11'114 генов, экспрессирующихся в обеих стадиях и 903 и 648 гена, экспрессирующихся только в марите или метацеркарии соответственно. Транскрипты взрослой стадии были обогащены группами, задействованными в метаболизме липидов и протеолизе. Транскрипты метацеркарии были обогащены группами, задействованными в путях передачи сигнала. Другой, менее строгий подход сравнения стадий, включал поиск биохимических путей KEGG, обогащенных дифференциально экспрессирующимися генами, т. е. генами, экспрессия которых может иметь и ненулевое значение в обеих из стадий. Тест Фишера позволил выявить 7 биохимических путей обогащенных ДЭГ. Среди этих путей обнаруживается лизосома, метаболические пути и пути передачи сигнала.

Также у *O. felineus* были обнаружены гранулиноподобные транскрипты, не обнаруживающие ортологов у родственных описторхид. У гранулина *O. viverrini* ранее был показан митогенный эффект на клетках млекопитающих, что делает его одной из главных молекулярных детерминант канцерогенности этого вида.

Используя рибосомальные белки, были проанализированы филогенетические взаимоотношения исследуемого вида с родственными трематодами. Оказалось, что *O. felineus* и *C. sinensis* филогенетически ближе друг к другу, чем каждый из них к *O. viverrini*.

Таким образом, полученный массив транскриптомных данных позволил выявить некоторые особенности генетической организации *O. felineus* и впервые сравнить

полнотранскриптомные профили разных стадий жизненного цикла описторхид. Также мы надеемся, что депонированные в базы данных последовательности заложат важную основу для последующих исследований *O. felineus*, его взаимоотношений с хозяином и явления паразитизма вообще.

6 Выводы

1. Среди наиболее экспрессируемых генов во взрослой стадии *O. felineus* идентифицированы гены катепсина F, миоглобина, белка оболочки яйца, глутатион трансферазы, а также HDM белка. В метацеркарии же, среди генов с наибольшей экспрессией представлены те, которые кодируют белки домашнего хозяйства: рибосомальные белки и убиквитин.
2. *O. felineus* не экспрессирует гены синтеза полиаминов, большинство ферментов метаболизма метионина и PEX белки, участвующие в формировании пероксисом.
3. Обнаружены новые гранулиноподобные белки, не имеющие ортологов у родственных описторхид.
4. При анализе экспрессии полученных транскриптов были идентифицированы 11'114 генов, экспрессирующихся в обеих стадиях и 903 и 648 гена, экспрессирующихся только в марите или метацеркарии соответственно
5. В результате анализа рибосомальных белков выявлено, что *O. felineus* и *C. sinensis* филогенетически ближе друг к другу, чем каждый из них к *O. viverrini*.

7 Список сокращений

пн – пар нуклеотидов

ЭС белки – экскреторно-секреторные белки

ДЭГ – дифференциально экспрессирующийся ген

EST – от англ. «Expressed sequence tag». Прочтения, получаемые при секвенировании методом Сэнгера плазмид клонов кДНК библиотеки.

FPKM (от англ. «Fragments Per Kilobase of exon per Million fragments mapped») – количество фрагментов прочтений на килобазу экзона на количество фрагментов, использованных в маппировании в миллионах штук.

GO – Gene Ontology

GPCR – от англ. «G-protein Coupled Receptors». Класс мембранных рецепторов, связанных с так называемыми G-белками.

GRAFS – система классификации GPCR белков. Название образовано по первым буквам английских названий семейств рецепторов: Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2, and Secretin

HDM – от англ. «Helminth Defense Molecules». Класс белков гельминтов, мимикрирующих под кателицидины млекопитающих

НОХ-гены – от англ. «homeobox». Класс гомеодомен-содержащих белков, ответственных за строение тела эмбриона вдоль антерио-постериорной оси.

NGS – от англ. «Next Generation Sequencing». Методы секвенирования «следующего поколения»

TSA – от англ. «Transcript Shotgun assembly». Раздел базы данных NCBI, включающий транскрипты, полученные путем сборки данных массового параллельного секвенирования.

8 Список литературы

1. King S., Scholz T. Trematodes of the family Opisthorchiidae: a minireview. // *Korean J. Parasitol.* 2001. T. 39. № 3. P. 209–21.
2. Hung N.M., Madsen H., Fried B. Global status of fish-borne zoonotic trematodiasis in humans. // *Acta Parasitol.* 2013. T. 58. № 3. P. 231–58.
3. Hong S.-T., Fang Y. Clonorchis sinensis and clonorchiasis, an update. // *Parasitol. Int.* 2012. T. 61. № 1. P. 17–24.
4. Mordvinov V. a et al. Opisthorchis felinus and Metorchis bilis are the main agents of liver fluke infection of humans in Russia. // *Parasitol. Int.* 2012. T. 61. № 1. P. 25–31.
5. Armignacco O. et al. Cryptic and asymptomatic Opisthorchis felinus infections. // *Am. J. Trop. Med. Hyg.* 2013. T. 88. № 2. P. 364–6.
6. Pozio E. et al. Opisthorchis felinus, an emerging infection in Italy and its implication for the European Union. // *Acta Trop.* 2013. T. 126. № 1. P. 54–62.
7. Kaewpitoon N. et al. Opisthorchis viverrini: the carcinogenic human liver fluke // *World J Gastroenterol.* 2008. T. 14. № 5. P. 666–674.
8. Sripa B. et al. Liver fluke induces cholangiocarcinoma. // *PLoS Med.* 2007. T. 4. № 7. P. e201.
9. Sripa B. et al. The tumorigenic liver fluke Opisthorchis viverrini - multiple pathways to cancer. // *Trends Parasitol.* 2012. T. 28. № 10. P. 395–407.
10. Lvova M.N. et al. Comparative histopathology of Opisthorchis felinus and Opisthorchis viverrini in a hamster model: An implication of high pathogenicity of the European liver fluke. // *Parasitol. Int.* 2011. P. 6–11.
11. Chai J.-Y. Praziquantel treatment in trematode and cestode infections: an update. // *Infect. Chemother.* 2013. T. 45. № 1. P. 32–43.
12. Soukhathammavong P. et al. Efficacy and safety of mefloquine, artesunate, mefloquine-artesunate, tribendimidine, and praziquantel in patients with Opisthorchis viverrini: a randomised, exploratory, open-label, phase 2 trial. // *Lancet Infect. Dis.* 2011. T. 11. № 2. P. 110–8.
13. Young N.D. et al. The Opisthorchis viverrini genome provides insights into life in the bile duct. // *Nat. Commun.* 2014. T. 5. P. 4378.
14. Wang X. et al. The draft genome of the carcinogenic human liver fluke Clonorchis sinensis. //

Genome Biol. 2011. T. 12. № 10. P. R107.

15. Young N.D. et al. Unlocking the transcriptomes of two carcinogenic parasites, *Clonorchis sinensis* and *Opisthorchis viverrini*. // *PLoS Negl. Trop. Dis.* 2010. T. 4. № 6. P. e719.
16. Adoutte A. et al. The new animal phylogeny: reliability and implications. // *Proc. Natl. Acad. Sci. U. S. A.* 2000. T. 97. № 9. P. 4453–6.
17. Olson P.D., Tkach V. V. Advances and trends in the molecular systematics of the parasitic Platyhelminthes. // *Adv. Parasitol.* 2005. T. 60. P. 165–243.
18. Cantacessi C. et al. A genome-wide analysis of annexins from parasitic organisms and their vectors. // *Sci. Rep.* 2013. T. 3. P. 2893.
19. Cai X.Q. et al. Sequences and gene organization of the mitochondrial genomes of the liver flukes *Opisthorchis viverrini* and *Clonorchis sinensis* (Trematoda). // *Parasitol. Res.* 2012. T. 110. № 1. P. 235–43.
20. Shekhovtsov S. V et al. A novel nuclear marker, Pm-int9, for phylogenetic studies of *Opisthorchis felinus*, *Opisthorchis viverrini*, and *Clonorchis sinensis* (Opisthorchiidae, Trematoda). // *Parasitol. Res.* 2009. T. 106. № 1. P. 293–7.
21. Petney T.N. et al. The zoonotic, fish-borne liver flukes *Clonorchis sinensis*, *Opisthorchis felinus* and *Opisthorchis viverrini*. // *Int. J. Parasitol.* 2013. T. 43. № 12-13. P. 1031–46.
22. Chen D. et al. Epidemiological investigation of *Clonorchis sinensis* infection in freshwater fishes in the Pearl River Delta. // *Parasitol. Res.* 2010. T. 107. № 4. P. 835–9.
23. Schuster R. et al. A sero-epidemiological survey on the occurrence of opisthorchiid liver flukes in red foxes (*Vulpes vulpes*) in Berlin, Germany. // *Parasitol. Res.* 2003. T. 90. № 5. P. 400–4.
24. Schuster R. et al. Liver fluke (Opisthorchiidae) findings in red foxes (*Vulpes vulpes*) in the eastern part of the Federal State Brandenburg, Germany--a contribution to the epidemiology of opisthorchiidosis. // *Parasitol. Res.* 1999. T. 85. № 2. P. 142–6.
25. Shimalov V., Shimalov V. Helminth fauna of the stoat (*Mustela erminea* Linnaeus, 1758) and the weasel (*M. nivalis* Linnaeus, 1758) in Belorussian Polesie // *Parasitol. Res.* 2001. T. 87. № 8. P. 680–681.
26. Shimalov V. V, Shimalov V.T., Shimalov A. V. Helminth fauna of otter (*Lutra lutra* Linnaeus, 1758) in Belorussian Polesie. // *Parasitol. Res.* 2000. T. 86. № 6. P. 528.
27. Mordvinov V.A., Furman D.P. The Digenea parasite *Opisthorchis felinus*: a target for the

discovery and development of novel drugs. // *Infect. Disord. Drug Targets*. 2010. T. 10. № 5. P. 385–401.

28. Keiser J., Utzinger J. Emerging foodborne trematodiasis. // *Emerg. Infect. Dis.* 2005. T. 11. № 10. P. 1507–14.

29. Lun Z.-R. et al. Clonorchiasis: a key foodborne zoonosis in China. // *Lancet Infect. Dis.* 2005. T. 5. № 1. P. 31–41.

30. Sithithaworn P., Haswell-Elkins M. Epidemiology of *Opisthorchis viverrini*. // *Acta Trop.* 2003. T. 88. № 3. P. 187–94.

31. Armignacco O. et al. Human illnesses caused by *Opisthorchis felinus* flukes, Italy. // *Emerg. Infect. Dis.* 2008. T. 14. № 12. P. 1902–5.

32. Traverso a. et al. A large outbreak of *Opisthorchis felinus* in Italy suggests that opisthorchiasis develops as a febrile eosinophilic syndrome with cholestasis rather than a hepatitis-like syndrome // *Eur. J. Clin. Microbiol. Infect. Dis.* 2012. T. 31. № 6. P. 1089–1093.

33. McCarthy J.S. et al. A research agenda for helminth diseases of humans: diagnostics for control and elimination programmes. // *PLoS Negl. Trop. Dis.* 2012. T. 6. № 4. P. e1601.

34. Johansen M.V. et al. Towards improved diagnosis of zoonotic trematode infections in Southeast Asia. // *Adv. Parasitol.* 2010. T. 73. № 10. P. 171–95.

35. Kim Y.J. et al. Performance of an enzyme-linked immunosorbent assay for detection of *Clonorchis sinensis* infestation in high- and low-risk groups. // *J. Clin. Microbiol.* 2010. T. 48. № 7. P. 2365–7.

36. Choi M.-H. et al. Excretory-secretory antigen is better than crude antigen for the serodiagnosis of clonorchiasis by ELISA. // *Korean J. Parasitol.* 2003. T. 41. № 1. P. 35–9.

37. Hotez P.J. et al. Control of neglected tropical diseases. // *N. Engl. J. Med.* 2007. T. 357. № 10. P. 1018–27.

38. Lorsuwannarat N. et al. The anthelmintic effect of plumbagin on *Schistosoma mansoni*. // *Exp. Parasitol.* 2012. T. 133. № 1. P. 18–27.

39. Pakharukova M.Y. et al. The first comprehensive study of praziquantel effects in vivo and in vitro on European liver fluke *Opisthorchis felinus* (Trematoda) // *Int. J. Antimicrob. Agents.* 2015.

40. Sanger F., Nicklen S., Coulson A.R. DNA sequencing with chain-terminating inhibitors. // *Proc. Natl. Acad. Sci. U. S. A.* 1977. T. 74. № 12. P. 5463–7.

41. Liu L. et al. Comparison of next-generation sequencing systems. // *J. Biomed. Biotechnol.* 2012. T. 2012. P. 251364.
42. Loman N.J. et al. Performance comparison of benchtop high-throughput sequencing platforms. // *Nat. Biotechnol.* 2012. T. 30. № 5. P. 434–9.
43. Berriman M. et al. The genome of the blood fluke *Schistosoma mansoni* // *Nature.* 2009. T. 460. № 7253. P. 352–358.
44. Kelley K.W. The *Schistosoma mansoni* transcriptome: an update // *Brain. Behav. Immun.* 2008. T. 22. № 5. P. 629.
45. Oliveira G. The *Schistosoma mansoni* transcriptome: an update // *Exp Parasitol.* 2007. T. 117. № 3. P. 229–235.
46. Nawaratna S.S.K. et al. Gene Atlasing of digestive and reproductive tissues in *Schistosoma mansoni*. // *PLoS Negl. Trop. Dis.* 2011. T. 5. № 4. P. e1043.
47. Verjovski-Almeida S. et al. Transcriptome analysis of the acoelomate human parasite *Schistosoma mansoni* // *Nat Genet.* 2003. T. 35. № 2. P. 148–157.
48. Merrick J.M. et al. The *Schistosoma mansoni* gene index: gene discovery and biology by reconstruction and analysis of expressed gene sequences // *J Parasitol.* 2003. T. 89. № 2. P. 261–269.
49. Protasio A. V et al. A systematically improved high quality genome and transcriptome of the human blood fluke *Schistosoma mansoni*. // *PLoS Negl. Trop. Dis.* 2012. T. 6. № 1. P. e1455.
50. Almeida G.T. et al. Exploring the *Schistosoma mansoni* adult male transcriptome using RNA-seq. // *Exp. Parasitol.* 2012. T. 132. № 1. P. 22–31.
51. Zhou Y. et al. The *Schistosoma japonicum* genome reveals features of host-parasite interplay. // *Nature.* 2009. T. 460. № 7253. P. 345–51.
52. Cwiklinski K. et al. The *Fasciola hepatica* genome: gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution // *Genome Biol.* 2015. T. 16. № 1. P. 1–13.
53. Cho P.Y. et al. Expressed sequence tag analysis of adult *Clonorchis sinensis*, the Chinese liver fluke // *Parasitol Res.* 2006. T. 99. № 5. P. 602–608.
54. Cho P.Y. et al. Gene expression profile of *Clonorchis sinensis* metacercariae // *Parasitol Res.* 2008. T. 102. № 2. P. 277–282.
55. Lee J.-S. et al. Analysis of the genes expressed in *Clonorchis sinensis* adults using the

expressed sequence tag approach. // *Parasitol. Res.* 2003. T. 91. № 4. P. 283–9.

56. Yoo W.G. et al. Developmental Transcriptomic Features of the Carcinogenic Liver Fluke, *Clonorchis sinensis* // *PLoS Negl. Trop. Dis.* 2011. T. 5. № 6. P. e1208.

57. Jex A.R. et al. Molecular changes in *Opisthorchis viverrini* (Southeast Asian liver fluke) during the transition from the juvenile to the adult stage. // *PLoS Negl. Trop. Dis.* 2012. T. 6. № 11. P. e1916.

58. Laha T. et al. Gene discovery for the carcinogenic human liver fluke, *Opisthorchis viverrini*. // *BMC Genomics.* 2007. T. 8. P. 189.

59. Sajid M. et al. Functional expression and characterization of *Schistosoma mansoni* cathepsin B and its trans-activation by an endogenous asparaginyl endopeptidase. // *Mol. Biochem. Parasitol.* 2003. T. 131. № 1. P. 65–75.

60. Wilson L.R. et al. *Fasciola hepatica*: characterization and cloning of the major cathepsin B protease secreted by newly excysted juvenile liver fluke. // *Exp. Parasitol.* 1998. T. 88. № 2. P. 85–94.

61. Corvo I. et al. The major cathepsin L secreted by the invasive juvenile *Fasciola hepatica* prefers proline in the S2 subsite and can cleave collagen. // *Mol. Biochem. Parasitol.* 2009. T. 167. № 1. P. 41–7.

62. Dvorak J. et al. Differential use of protease families for invasion by schistosome cercariae // *Biochimie.* 2008. T. 90. № 2. P. 345–358.

63. Dalton J.P. et al. *Fasciola hepatica* cathepsin L-like proteases: biology, function, and potential in the development of first generation liver fluke vaccines // *Int J Parasitol.* 2003. T. 33. № 11. P. 1173–1181.

64. Delcroix M. et al. A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite // *J Biol Chem.* 2006. T. 281. № 51. P. 39316–39329.

65. Pinlaor P. et al. Cathepsin F cysteine protease of the human liver fluke, *Opisthorchis viverrini*. // *PLoS Negl. Trop. Dis.* 2009. T. 3. № 3. P. e398.

66. McGonigle L. et al. The silencing of cysteine proteases in *Fasciola hepatica* newly excysted juveniles using RNA interference reduces gut penetration. // *Int. J. Parasitol.* 2008. T. 38. № 2. P. 149–55.

67. Sripa J. et al. Secreted cysteine proteases of the carcinogenic liver fluke, *Opisthorchis viverrini*: regulation of cathepsin F activation by autocatalysis and trans-processing by cathepsin B. // *Cell. Microbiol.* 2010. T. 12. № 6. P. 781–95.

68. Abdulla M.H. et al. Schistosomiasis mansoni: novel chemotherapy using a cysteine protease inhibitor // *PLoS Med.* 2007. T. 4. № 1. P. e14.
69. Loukas A., Tran M., Pearson M.S. Schistosome membrane proteins as vaccines // *Int J Parasitol.* 2007. T. 37. № 3-4. P. 257–263.
70. Halpern Z. et al. Bile and plasma lipid composition in non-obese normolipidemic subjects with and without cholesterol gallstones. // *Liver.* 1993. T. 13. № 5. P. 246–52.
71. Huang Y. et al. The carcinogenic liver fluke, *Clonorchis sinensis*: new assembly, reannotation and analysis of the genome and characterization of tissue transcriptomes. // *PLoS One.* 2013. T. 8. № 1. P. e54732.
72. Kiger L. et al. Trematode hemoglobins show exceptionally high oxygen affinity. // *Biophys. J.* 1998. T. 75. № 2. P. 990–8.
73. Tran M.H. et al. Tetraspanins on the surface of *Schistosoma mansoni* are protective antigens against schistosomiasis. // *Nat. Med.* 2006. T. 12. № 7. P. 835–40.
74. Zhou Z. et al. Immunogenicity of recombinant *Bacillus subtilis* spores expressing *Clonorchis sinensis* tegumental protein // *Parasitol Res.* 2008. T. 102. № 2. P. 293–297.
75. Fonseca C.T. et al. *Schistosoma* tegument proteins in vaccine and diagnosis development: an update. // *J. Parasitol. Res.* 2012. T. 2012. P. 541268.
76. Holland P.W.H. Evolution of homeobox genes. // *Wiley Interdiscip. Rev. Dev. Biol.* 2013. T. 2. № 1. P. 31–45.
77. Koziol U., Lalanne A.I., Castillo E. Hox genes in the parasitic plathyhelminthes *Mesocestoides corti*, *Echinococcus multilocularis*, and *Schistosoma mansoni*: evidence for a reduced Hox complement. // *Biochem. Genet.* 2009. T. 47. № 1-2. P. 100–16.
78. Pierce R.J. et al. Evidence for a dispersed Hox gene cluster in the plathyhelminth parasite *Schistosoma mansoni*. // *Mol. Biol. Evol.* 2005. T. 22. № 12. P. 2491–503.
79. Garcia-Fernàndez J. Hox, ParaHox, ProtoHox: facts and guesses. // *Heredity (Edinb).* 2005. T. 94. № 2. P. 145–52.
80. Gu J.-L. et al. Hox genes from the parasitic flatworm *Schistosoma japonicum*. // *Genomics.* 2012. T. 99. № 1. P. 59–65.
81. Overington J.P., Al-Lazikani B., Hopkins A.L. How many drug targets are there? // *Nat. Rev. Drug Discov.* 2006. T. 5. № 12. P. 993–6.

82. Oliveira K.C. et al. Effect of human TGF- β on the gene expression profile of *Schistosoma mansoni* adult worms. // *Mol. Biochem. Parasitol.* 2012. T. 183. № 2. P. 132–9.
83. Roberts-Galbraith R.H., Newmark P. a. On the organ trail: insights into organ regeneration in the planarian // *Curr. Opin. Genet. Dev.* 2015. T. 32. P. 37–46.
84. Rink J.C. Stem cell systems and regeneration in planaria // *Dev. Genes Evol.* 2013. T. 223. № 1-2. P. 67–84.
85. Shaw M.K., Erasmus D.A. *Schistosoma mansoni*: structural damage and tegumental repair after in vivo treatment with praziquantel. // *Parasitology.* 1987. T. 94 (Pt 2). P. 243–54.
86. Collins J.J. et al. Adult somatic stem cells in the human parasite *Schistosoma mansoni*. // *Nature.* 2013. T. 494. № 7438. P. 476–9.
87. Lanner F., Rossant J. The role of FGF/Erk signaling in pluripotent cells. // *Development.* 2010. T. 137. № 20. P. 3351–60.
88. Skinner D.E. et al. Vasa-Like DEAD-Box RNA Helicases of *Schistosoma mansoni*. // *PLoS Negl. Trop. Dis.* 2012. T. 6. № 6. P. e1686.
89. Tsai I.J. et al. The genomes of four tapeworm species reveal adaptations to parasitism. // *Nature.* 2013. T. 496. № 7443. P. 57–63.
90. Bouvard V. et al. A review of human carcinogens—Part B: biological agents // *Lancet Oncol.* 2009. T. 10. № 4. P. 321–322.
91. Choi D. et al. Cholangiocarcinoma and *Clonorchis sinensis* infection: a case-control study in Korea. // *J. Hepatol.* 2006. T. 44. № 6. P. 1066–73.
92. Sirica A.E. Cholangiocarcinoma: molecular targeting strategies for chemoprevention and therapy // *Hepatology.* 2005. T. 41. № 1. P. 5–15.
93. Mulvenna J. et al. The secreted and surface proteomes of the adult stage of the carcinogenic human liver fluke *Opisthorchis viverrini*. // *Proteomics.* 2010. T. 10. № 5. P. 1063–78.
94. He Z., Bateman A. Progranulin (granulin-epithelin precursor, PC-cell-derived growth factor, acrogranin) mediates tissue repair and tumorigenesis // *J Mol Med.* 2003. T. 81. № 10. P. 600–612.
95. Smout M.J. et al. A granulin-like growth factor secreted by the carcinogenic liver fluke, *Opisthorchis viverrini*, promotes proliferation of host cells // *PLoS Pathog.* 2009. T. 5. № 10. P. e1000611.
96. Chen X. et al. Molecular Characterization of Severin from *Clonorchis sinensis*

Excretory/Secretory Products and Its Potential Anti-apoptotic Role in Hepatocarcinoma PLC Cells. // PLoS Negl. Trop. Dis. 2013. T. 7. № 12. P. e2606.

97. Whitbread A.K. et al. The role of kallikrein-related peptidases in prostate cancer: potential involvement in an epithelial to mesenchymal transition // Biol Chem. 2006. T. 387. № 6. P. 707–714.

98. Doudican N. et al. Mebendazole induces apoptosis via Bcl-2 inactivation in chemoresistant melanoma cells. // Mol. Cancer Res. 2008. T. 6. № 8. P. 1308–15.

99. Osman A. et al. Schistosoma mansoni TGF-beta receptor II: role in host ligand-induced regulation of a schistosome target gene // PLoS Pathog. 2006. T. 2. № 6. P. e54.

100. Beall M.J., Pearce E.J. Human transforming growth factor-beta activates a receptor serine/threonine kinase from the intravascular parasite Schistosoma mansoni // J Biol Chem. 2001. T. 276. № 34. P. 31613–31619.

101. Kahlenberg J.M., Kaplan M.J. Little peptide, big effects: the role of LL-37 in inflammation and autoimmune disease. // J. Immunol. 2013. T. 191. P. 4895–901.

102. Zanetti M. Cathelicidins, multifunctional peptides of the innate immunity. // J. Leukoc. Biol. 2004. T. 75. № 1. P. 39–48.

103. Robinson M.W., Donnelly S., Dalton J.P. Helminth defence molecules-immunomodulators designed by parasites! // Front. Microbiol. 2013. T. 4. № October. P. 296.

104. Robinson M.W. et al. A family of helminth molecules that modulate innate cell responses via molecular mimicry of host antimicrobial peptides. // PLoS Pathog. 2011. T. 7. № 5. P. e1002042.

105. Thivierge K. et al. Cathelicidin-like helminth defence molecules (HDMS): absence of cytotoxic, anti-microbial and anti-protozoan activities imply a specific adaptation to immune modulation. // PLoS Negl. Trop. Dis. 2013. T. 7. № 7. P. e2307.

106. Robinson M.W. et al. A helminth cathelicidin-like protein suppresses antigen processing and presentation in macrophages via inhibition of lysosomal vATPase. // FASEB J. 2012. T. 26. № 11. P. 4614–27.

107. Onguru D. et al. Human schistosomiasis is associated with endotoxemia and Toll-like receptor 2- and 4-bearing B cells. // Am. J. Trop. Med. Hyg. 2011. T. 84. № 2. P. 321–4.

108. Donnelly S. et al. Helminth 2-Cys peroxiredoxin drives Th2 responses through a mechanism involving alternatively activated macrophages. // FASEB J. 2008. T. 22. № 11. P. 4022–4032.

109. Donnelly S. et al. Helminth cysteine proteases inhibit TRIF-dependent activation of

macrophages via degradation of TLR3 // *J. Biol. Chem.* 2010. T. 285. № 5. P. 3383–3392.

110. Lei H. et al. The biochemical and immunological characterization of two serpins from *Clonorchis sinensis*. // *Mol. Biol. Rep.* 2013. T. 40. № 6. P. 3977–85.

111. Gulley M.M., Zhang X., Michel K. The roles of serpins in mosquito immunology and physiology. // *J. Insect Physiol.* 2012.

112. Allen J.E., Maizels R.M. Diversity and dialogue in immunity to helminths. // *Nat. Rev. Immunol.* 2011. T. 11. № 6. P. 375–88.

113. Lund M.E. et al. Secreted Proteins from the Helminth *Fasciola hepatica* Inhibit the Initiation of Autoreactive T Cell Responses and Prevent Diabetes in the NOD Mouse. // *PLoS One.* 2014. T. 9. № 1. P. e86289.

114. Bae Y.-A. et al. Differential activation of diverse glutathione transferases of *Clonorchis sinensis* in response to the host bile and oxidative stressors. // *PLoS Negl. Trop. Dis.* 2013. T. 7. № 5. P. e2211.

115. Hong S.-J. et al. *Clonorchis sinensis*: glutathione S-transferase as a serodiagnostic antigen for detecting IgG and IgE antibodies. // *Exp. Parasitol.* 2002. T. 101. № 4. P. 231–3.

116. Sun J.B. et al. Intranasal administration of a *Schistosoma mansoni* glutathione S-transferase-cholera toxoid conjugate vaccine evokes antiparasitic and antipathological immunity in mice. // *J. Immunol.* 1999. T. 163. № 2. P. 1045–52.

117. Bae Y. et al. Identification and biochemical characterization of two novel peroxiredoxins in a liver fluke, *Clonorchis sinensis*. // *Parasitology.* 2011. T. 138. № 9. P. 1143–53.

118. Boumis G. et al. Structural and functional characterization of *Schistosoma mansoni* thioredoxin. // *Protein Sci.* 2011. P. 1–21.

119. McGonigle S., Curley G.P., Dalton J.P. Cloning of peroxiredoxin, a novel antioxidant enzyme, from the helminth parasite *Fasciola hepatica*. // *Parasitology.* 1997. T. 115 (Pt 1. P. 101–4.

120. Jortzik E., Becker K. Thioredoxin and glutathione systems in *Plasmodium falciparum*. // *Int. J. Med. Microbiol.* 2012. T. 302. № 4-5. P. 187–94.

121. Otero L. et al. Thioredoxin and glutathione systems differ in parasitic and free-living platyhelminths. // *BMC Genomics.* 2010. T. 11. P. 237.

122. Pakharukova M.Y. et al. Cytochrome P450 in fluke *Opisthorchis felinus*: identification and characterization. // *Mol. Biochem. Parasitol.* 2012. T. 181. № 2. P. 190–4.

123. Sripa J. et al. RNA interference targeting cathepsin B of the carcinogenic liver fluke, *Opisthorchis viverrini*. // *Parasitol. Int.* 2011. T. 60. № 3. P. 283–8.
124. Gobert G.N. et al. Developmental gene expression profiles of the human pathogen *Schistosoma japonicum*. // *BMC Genomics.* 2009. T. 10. P. 128.
125. Gobert G.N. et al. Transcriptional changes in *Schistosoma mansoni* during early schistosomula development and in the presence of erythrocytes. // *PLoS Negl. Trop. Dis.* 2010. T. 4. № 2. P. e600.
126. DeMarco R. et al. Protein variation in blood-dwelling schistosome worms generated by differential splicing of micro-exon gene transcripts. // *Genome Res.* 2010. T. 20. № 8. P. 1112–21.
127. Young N.D. et al. Progress on the transcriptomics of carcinogenic liver flukes of humans--unique biological and biotechnological prospects. // *Biotechnol. Adv.* 2010. T. 28. № 6. P. 859–70.
128. Dalton J.P. et al. Immunomodulatory molecules of *Fasciola hepatica*: candidates for both vaccine and immunotherapeutic development. // *Vet. Parasitol.* 2013. T. 195. № 3-4. P. 272–85.
129. Jayaraj R. et al. Vaccination against fasciolosis by a multivalent vaccine of stage-specific antigens. // *Vet. Parasitol.* 2009. T. 160. № 3-4. P. 230–6.
130. Nagano I. et al. Molecular expression of a cysteine proteinase of *Clonorchis sinensis* and its application to an enzyme-linked immunosorbent assay for immunodiagnosis of clonorchiasis. // *Clin. Diagn. Lab. Immunol.* 2004. T. 11. № 2. P. 411–6.
131. Wang X. et al. Identification and characterization of paramyosin from cyst wall of metacercariae implicated protective efficacy against *Clonorchis sinensis* infection. // *PLoS One.* 2012. T. 7. № 3. P. e33703.
132. Chevreaux B. et al. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. // *Genome Res.* 2004. T. 14. № 6. P. 1147–59.
133. Bolger A.M., Lohse M., Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. // *Bioinformatics.* 2014.
134. Langmead B., Salzberg S.L. Fast gapped-read alignment with Bowtie 2. // *Nat. Methods.* 2012. T. 9. № 4. P. 357–9.
135. Haas B.J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. // *Nat. Protoc.* 2013. T. 8. № 8. P. 1494–512.

136. Robinson M.D., McCarthy D.J., Smyth G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. // *Bioinformatics*. 2010. T. 26. № 1. P. 139–40.
137. Altschul S.F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. // *Nucleic Acids Res.* 1997. T. 25. № 17. P. 3389–402.
138. Parra G., Bradnam K., Korf I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. // *Bioinformatics*. 2007. T. 23. № 9. P. 1061–7.
139. Wu J. et al. KOBAS server: a web-based platform for automated annotation and pathway identification // *Nucleic Acids Res.* 2006. T. 34. № Web Server issue. P. W720–4.
140. Zdobnov E.M., Apweiler R. InterProScan--an integration platform for the signature-recognition methods in InterPro. // *Bioinformatics*. 2001. T. 17. № 9. P. 847–8.
141. Ashburner M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. // *Nat. Genet.* 2000. T. 25. № 1. P. 25–9.
142. Hu Zhi-Liang, Bao J R.J. CateGORizer: A Web-Based Program to Batch Analyze Gene Ontology Classification Categories // *Online J. Bioinforma.* 2008. T. 9. P. 108–112.
143. Cline M.S. et al. Integration of biological networks and gene expression data using Cytoscape. // *Nat. Protoc.* 2007. T. 2. № 10. P. 2366–82.
144. Maere S., Heymans K., Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. // *Bioinformatics*. 2005. T. 21. № 16. P. 3448–9.
145. Finn R.D., Clements J., Eddy S.R. HMMER web server: interactive sequence similarity searching. // *Nucleic Acids Res.* 2011. T. 39. № Web Server issue. P. W29–37.
146. Larkin M.A. et al. Clustal W and Clustal X version 2.0 // *Bioinformatics*. 2007. T. 23. № 21. P. 2947–2948.
147. Tamura K. et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. // *Mol. Biol. Evol.* 2011. T. 28. № 10. P. 2731–9.
148. Guindon S., Gascuel O. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood // *Syst. Biol.* 2003. T. 52. № 5. P. 696–704.
149. Felsenstein J. PHYLIP - Phylogeny Inference Package (Version 3.2) // *Cladistics*. 1989. T. 5. № 2. P. 163–166.
150. Kim T.I., Na B.K., Hong S.J. Functional genes and proteins of *Clonorchis sinensis* // *Korean*

J Parasitol. 2009. T. 47 Suppl. P. S59–68.

151. Chung Y.-B. et al. Molecular cloning and immunolocalization of the 17 kDa myoglobin of *Clonorchis sinensis*. // Parasitol. Res. 2003. T. 90. № 5. P. 365–8.

152. Rashid a K., Weber R.E. Functional differentiation in trematode hemoglobin isoforms. // Eur. J. Biochem. 1999. T. 260. № 3. P. 717–25.

153. Dewilde S. et al. The hemoglobins of the trematodes *Fasciola hepatica* and *Paramphistomum epiclitum*: a molecular biological, physico-chemical, kinetic, and vaccination study. // Protein Sci. 2008. T. 17. № 10. P. 1653–62.

154. Pesce A. et al. Very high resolution structure of a trematode hemoglobin displaying a TyrB10-TyrE7 heme distal residue pair and high oxygen affinity. // J. Mol. Biol. 2001. T. 309. № 5. P. 1153–64.

155. Tang Y. et al. Molecular cloning and characterization of vitelline precursor protein B1 from *Clonorchis sinensis*. // J. Parasitol. 2005. T. 91. № 6. P. 1374–8.

156. Chen W. et al. Molecular characterization of cathepsin B from *Clonorchis sinensis* excretory/secretory products and assessment of its potential for serodiagnosis of clonorchiasis. // Parasit. Vectors. 2011. T. 4. P. 149.

157. Sripa J. et al. Evaluation of liver fluke recombinant cathepsin B-1 protease as a serodiagnostic antigen for human opisthorchiasis. // Parasitol. Int. 2012. T. 61. № 1. P. 191–5.

158. Stack C.M. et al. Structural and functional relationships in the virulence-associated cathepsin L proteases of the parasitic liver fluke, *Fasciola hepatica*. // J. Biol. Chem. 2008. T. 283. № 15. P. 9896–908.

159. McVeigh P. et al. *Fasciola hepatica* virulence-associated cysteine peptidases: a systems biology perspective. // Microbes Infect. 2012. T. 14. № 4. P. 301–10.

160. Kang J.-M. et al. A family of cathepsin F cysteine proteases of *Clonorchis sinensis* is the major secreted proteins that are expressed in the intestine of the parasite. // Mol. Biochem. Parasitol. 2010. T. 170. № 1. P. 7–16.

161. Dowling D.J. et al. Major secretory antigens of the helminth *Fasciola hepatica* activate a suppressive dendritic cell phenotype that attenuates Th17 cells but fails to activate Th2 immune responses. // Infect. Immun. 2010. T. 78. № 2. P. 793–801.

162. Chung Y.B. et al. Excystment of *Paragonimus westermani* metacercariae by endogenous cysteine protease. // J. Parasitol. 1995. T. 81. № 2. P. 137–42.

163. LaCourse E.J. et al. The Sigma class glutathione transferase from the liver fluke *Fasciola hepatica*. // *PLoS Negl. Trop. Dis.* 2012. T. 6. № 5. P. e1666.
164. McGonigle S., Dalton J.P., James E.R. Peroxidoxins: a new antioxidant family. // *Parasitol. Today.* 1998. T. 14. № 4. P. 139–45.
165. Pegg A.E. The function of spermine. // *IUBMB Life.* 2014. T. 66. № 1. P. 8–18.
166. Casero R. a, Marton L.J. Targeting polyamine metabolism and function in cancer and other hyperproliferative diseases. // *Nat. Rev. Drug Discov.* 2007. T. 6. № 5. P. 373–90.
167. Albers E. Metabolic characteristics and importance of the universal methionine salvage pathway recycling methionine from 5'-methylthioadenosine. // *IUBMB Life.* 2009. T. 61. № 12. P. 1132–42.
168. Sekowska A. et al. Bacterial variations on the methionine salvage pathway. // *BMC Microbiol.* 2004. T. 4. P. 9.
169. Chung S. et al. Crooked neck is a component of the human spliceosome and implicated in the splicing process. // *Biochim. Biophys. Acta.* 2002. T. 1576. № 3. P. 287–97.
170. Hahn C., Fromm B., Bachmann L. Comparative genomics of flatworms (platyhelminthes) reveals shared genomic features of ecto- and endoparasitic neodermata. // *Genome Biol. Evol.* 2014. T. 6. № 5. P. 1105–17.
171. Hettema E.H. et al. Evolving models for peroxisome biogenesis. // *Curr. Opin. Cell Biol.* 2014. T. 29C. P. 25–30.
172. Wanders R.J., Grunsvan E.G. van, Jansen G. a. Lipid metabolism in peroxisomes: enzymology, functions and dysfunctions of the fatty acid alpha- and beta-oxidation systems in humans. // *Biochem. Soc. Trans.* 2000. T. 28. № 2. P. 141–9.
173. Nelson D.L., Cox M.M. Fatty Acid Catabolism // *Lehninger principles of Biochemistry.* , 2004. P. 631–54.
174. Infante R.E. et al. NPC2 facilitates bidirectional transfer of cholesterol between NPC1 and lipid bilayers, a step in cholesterol egress from lysosomes. // *Proc. Natl. Acad. Sci. U. S. A.* 2008. T. 105. № 40. P. 15287–92.
175. Inohara N., Nuñez G. ML -- a conserved domain involved in innate immunity and lipid metabolism. // *Trends Biochem. Sci.* 2002. T. 27. № 5. P. 219–21.
176. Rim H. et al. Fishborne trematode metacercariae in Luang Prabang, Khammouane, and

Saravane Province, Lao PDR. // Korean J. Parasitol. 2013. T. 51. № 1. P. 107–14.

177. Campos T.D.L. et al. Identification of G protein-coupled receptors in *Schistosoma haematobium* and *S. mansoni* by comparative genomics. // Parasit. Vectors. 2014. T. 7. № 1. P. 242.

178. Schiöth H.B., Fredriksson R. The GRAFS classification system of G-protein coupled receptors in comparative perspective. // Gen. Comp. Endocrinol. 2005. T. 142. № 1-2. P. 94–101.

179. Burke a C. et al. Hox genes and the evolution of vertebrate axial morphology. // Development. 1995. T. 121. № 2. P. 333–46.

180. Hoeflich K.P., Ikura M. Calmodulin in action: Diversity in target recognition and activation mechanisms // Cell. 2002. T. 108. № 6. P. 739–742.

181. Sriamporn S. et al. Prevalence of *Opisthorchis viverrini* infection and incidence of cholangiocarcinoma in Khon Kaen, Northeast Thailand. // Trop. Med. Int. Health. 2004. T. 9. № 5. P. 588–94.

182. Smout M.J. et al. Infection with the carcinogenic human liver fluke, *Opisthorchis viverrini*. // Mol. Biosyst. 2011. T. 7. № 5. P. 1367–75.

183. DPDx. Parasites - *Opisthorchis* Infection [Электронный ресурс]. URL: <http://www.cdc.gov/parasites/opisthorchis/biology.html>.

9 Приложения

Приложение 1: Транскрипты с наибольшей экспрессией во взрослой стадии.

GenBank #	Ближайший гомолог по Blastx (в круглых скобках в некоторых случаях дана собственная аннотация)	Значение E	GenBank # ближайшего гомолога	ФРКМ в марите	ФРКМ в метацеркарии
GBJA01004968	unknown [Clonorchis sinensis] (HDM белок)	2e-47	AAM55183	111083.025	796.927
GBJA01002241	cathepsin F precursor [Clonorchis sinensis] (CTSF; cathepsin F [EC:3.4.22.41])	0.0	ABK91811	55291.406	183.524
GBJA01003860	gb AAN28366.1 17 kDa myoglobin [Clonorchis sinensis]	5e-82	AAM18464	32903.787	2008.604
GBJA01003468	putative eggshell protein [Clonorchis sinensis] (Вителлин)	2e-148	GAA33751	24510.389	1.515
GBJA01006740	egg protein [Clonorchis sinensis] (Белок оболочки яйца)	2e-134	AAN64160	18014.323	0.804
GBJA01004450	unknown [Paragonimus westermani]	1e-05	AAK35218	9848.134	41.337
GBJA01007810	repetin [Clonorchis sinensis]	2e-48	GAA49231	9638.584	0.531
GBJA01012095	gb KER29080.1 hypothetical protein T265_04264 [Opisthorchis viverrini] (GST, глутатион S-трансфераза [EC:2.5.1.18])	5e-157	XP_009167217	9563.114	2483.381
GBJA01007477	prostaglandin-H2 D-isomerase [Clonorchis sinensis] (E5.3.99.2, PTGDS; простагландин-H2 D-изомераза [EC:5.3.99.2])	3e-135	GAA33791	8777.73	825.848
GBJA01007961	hypothetical protein CLF_102719 [Clonorchis sinensis]	1e-101	GAA49232	7901.27	0.471
GBJA01006940	Нет гомолога			6674.371	0.0
GBJA01003640	gb KER33138.1 hypothetical protein T265_00841, partial [Opisthorchis viverrini]	0.002	XP_009163000	6383.407	0.163
GBJA01012525	Нет гомолога			5729.185	129.223
GBJA01006939	Нет гомолога			5377.125	0.265
GBJA01011268	hypothetical protein CLF_102719 [Clonorchis sinensis]	4e-19	GAA49232	5018.483	0.291
GBJA01000303	putative eggshell protein [Clonorchis sinensis]	2e-126	GAA42793	4541.562	0.103
GBJA01003535	unnamed protein product [Oikopleura dioica] (UBB; убиквитин B)	9e-57	CBY12791	4312.73	3988.972
GBJA01002131	gb KER25708.1 hypothetical protein T265_06881 [Opisthorchis viverrini]	7e-38	XP_009170533	4212.435	18241.515
GBJA01003079	large subunit ribosomal protein LP1 [Clonorchis sinensis] (RP-LP1, RPLP1; белок большой субъединицы рибосомы LP1)	7e-42	GAA30172	4156.214	10363.564
GBJA01004791	hypothetical protein CLF_112758 [Clonorchis sinensis]	2e-43	GAA57461	4016.73	1.052

Приложение 2: Транскрипты с наибольшей экспрессией в метацеркарии.

GenBank #	Ближайший гомолог по Blastx (в круглых скобках в некоторых случаях дана собственная аннотация)	Значение E	GenBank # ближайшего гомолога	ФРКМ в марите	ФРКМ в метацеркарии
GBJA01002131	hypothetical protein T265_06881 [Opisthorchis viverrini]	7e-38	XP_009170533	4212.435	18241.515
GBJA01000005	hypothetical protein T265_03708 [Opisthorchis viverrini] (HDM белок)	2e-27	XP_009166519	3.469	13183.308
GBJA01003079	large subunit ribosomal protein LP1 [Clonorchis sinensis] (RP-LP1, RPLP1; белок большой субъединицы рибосомы LP1)	7e-42	GAA30172	4156.214	10363.564
GBJA01002422	hypothetical protein CLF_104361, partial [Clonorchis sinensis]	1e-09	GAA50311	1.351	9034.557
GBJA01001270	unnamed protein product [Oikopleura dioica] (UBB; убиквитин B)	3e-110	CBY16339	2313.104	7550.854
GBJA01006186	ribosomal protein S24 [Clonorchis sinensis] (RP-S24e, RPS24; белок малой субъединицы рибосомы S24e)	1e-63	AAN04092	1961.249	6305.044
GBJA01009829	hypothetical protein POPTR_0005s04540g, partial [Populus trichocarpa]	0.009	XP_006382696	742.661	5826.374
GBJA01012532	hypothetical protein T265_14302, partial [Opisthorchis viverrini] (RP-L37Ae, RPL37A; белок большой субъединицы рибосомы L37Ae)	1e-41	XP_009171293	1557.614	5645.914
GBJA01012557	hypothetical protein T265_07192 [Opisthorchis viverrini] (RP-L31e, RPL31; белок большой субъединицы рибосомы L31e)	8e-82	XP_009170912	1525.537	4650.172
GBJA01006390	hypothetical protein T265_04276 [Opisthorchis viverrini] (RP-S17e, RPS17; белок малой субъединицы рибосомы S17e)	1e-86	XP_009167239	2018.021	4602.811
GBJA01010744	hypothetical protein T265_05623 [Opisthorchis viverrini] (RP-L27e, RPL27; белок большой субъединицы рибосомы L27e)	1e-72	XP_009168946	1663.962	4600.646
GBJA01002231	hypothetical protein T265_11418 [Opisthorchis viverrini] (RP-S14e, RPS14; белок малой субъединицы рибосомы S14e)	2e-75	XP_009176333	1394.224	4525.005
GBJA01012646	hypothetical protein T265_07425 [Opisthorchis viverrini] (RP-S11e, RPS11; белок малой субъединицы рибосомы S11e)	4e-100	XP_009171207	1369.604	4366.613
GBJA01008066	heat shock 70kDa protein 1/8 [Clonorchis sinensis] (HSPA1_8; белок теплового шока 70kDa 1/8 форма 1)	9e-59	GAA52733	1353.76	4336.87
GBJA01000686	hypothetical protein T265_12096 [Opisthorchis viverrini] (RP-S20e, RPS20; белок малой субъединицы рибосомы S20e)	4e-79	XP_009177327	1348.032	4304.842
GBJA01006052	ribosomal protein L30 [Schistosoma japonicum] (RP-L30e, RPL30; белок большой субъединицы рибосомы L30e)	6e-44	CAX72575	1620.126	4137.567
GBJA01003535	unnamed protein product [Oikopleura dioica] (UBB; ubiquitin B)	9e-57	CBY12791	4312.73	3988.972
GBJA01012539	small subunit ribosomal protein S27e [Schistosoma japonicum] (RP-S27e, RPS27; белок малой субъединицы рибосомы S27e)	1e-36	CAX71546	1489.623	3955.669
GBJA01012067	heat shock 70kDa protein 1/8 [Clonorchis sinensis] (HSPA1_8; белок теплового шока 70kDa 1/8 форма 2)	7e-76	GAA35873	139.711	3928.878
GBJA01010586	hypothetical protein [Pseudomonas chlororaphis]	8e-05	WP_025808025	5.523	3863.856

Приложение 3: KEGG пути и количество ортологов имеющихся в них у *O. felineus*, *S. mansoni*, *H. sapiens* и *C. elegans*.

Путь KEGG	Description	# of orthologs in KEGG pathway	# of orthologs in <i>O felineus</i>	# of orthologs in <i>S mansoni</i>	# of orthologs in <i>H sapiens</i>	# of orthologs in <i>C elegans</i>
00830	Retinol metabolism	48	4	3	41	6
04630	Jak-STAT signaling pathway	123	13	11	122	10
00980	Metabolism of xenobiotics by cytochrome P450	32	4	3	29	7
00531	Glycosaminoglycan degradation	14	2	2	14	4
00982	Drug metabolism - cytochrome P450	23	3	2	21	7
04080	Neuroactive ligand-receptor interaction	250	39	15	242	22
00603	Glycosphingolipid biosynthesis - globo series	14	2	2	12	1
00100	Steroid biosynthesis	30	3	2	18	4
00532	Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate	18	3	2	17	6
00514	Other types of O-glycan biosynthesis	31	5	3	25	6
00604	Glycosphingolipid biosynthesis - ganglio series	14	3	3	14	3
04146	Peroxisome	72	17	8	71	39
04512	ECM-receptor interaction	58	13	6	48	3
00590	Arachidonic acid metabolism	43	11	8	37	7
01220	Degradation of aromatic compounds	196	1	1	3	2
00040	Pentose and glucuronate interconversions	59	5	5	14	10
00512	Mucin type O-Glycan biosynthesis	11	4	2	11	3
01040	Biosynthesis of unsaturated fatty acids	30	6	4	16	6
02010	ABC transporters	433	17	6	44	7
00350	Tyrosine metabolism	68	11	9	28	14
00561	Glycerolipid metabolism	71	14	11	33	17
00053	Ascorbate and aldarate metabolism	38	3	2	7	6
00260	Glycine, serine and threonine metabolism	89	14	13	32	21
00330	Arginine and proline metabolism	134	20	16	45	20
00270	Cysteine and methionine metabolism	81	12	13	27	19
04350	TGF-beta signaling pathway	61	27	15	60	14
00534	Glycosaminoglycan biosynthesis - heparan sulfate / heparin	22	10	8	22	9
00360	Phenylalanine metabolism	75	6	4	13	10
00670	One carbon pool by folate	27	7	7	15	12

04340	Hedgehog signaling pathway	40	17	8	36	10
04068	FoxO signaling pathway	100	48	36	99	36
00600	Sphingolipid metabolism	41	16	12	32	15
00750	Vitamin B6 metabolism	15	3	2	6	1
00380	Tryptophan metabolism	65	16	8	32	18
00430	Taurine and hypotaurine metabolism	21	3	3	6	3
00983	Drug metabolism - other enzymes	22	11	7	22	17
00511	Other glycan degradation	18	7	8	14	10
00592	alpha-Linolenic acid metabolism	21	4	2	8	3
00340	Histidine metabolism	40	9	3	18	8
04310	Wnt signaling pathway	98	48	33	95	31
00970	Aminoacyl-tRNA biosynthesis	64	25	38	48	42
04145	Phagosome	95	46	32	88	37
00565	Ether lipid metabolism	26	10	8	19	10
04142	Lysosome	99	53	43	97	47
00250	Alanine, aspartate and glutamate metabolism	65	16	14	29	20
04144	Endocytosis	147	80	62	145	66
00740	Riboflavin metabolism	32	5	5	9	5
01100	Metabolic pathways	2526	519	424	883	518
00500	Starch and sucrose metabolism	97	15	13	25	16
00061	Fatty acid biosynthesis	30	3	3	5	3
00400	Phenylalanine, tyrosine and tryptophan biosynthesis	72	3	2	5	4
03460	Fanconi anemia pathway	54	31	25	51	22
04120	Ubiquitin mediated proteolysis	124	72	63	118	54
04140	Regulation of autophagy	24	11	11	18	14
00230	Purine metabolism	259	76	68	124	82
03450	Non-homologous end-joining	19	8	5	13	5
00630	Glyoxylate and dicarboxylate metabolism	74	13	10	21	16
00310	Lysine degradation	68	26	14	41	21
00860	Porphyrin and chlorophyll metabolism	104	14	14	22	8
04130	SNARE interactions in vesicular transport	34	18	14	28	17
00780	Biotin metabolism	20	2	2	3	1
00071	Fatty acid degradation	50	20	3	30	20
04330	Notch signaling pathway	28	17	13	25	14
00052	Galactose metabolism	73	15	13	22	12

00062	Fatty acid elongation	22	13	6	19	9
01212	Fatty acid metabolism	69	24	10	35	22
00562	Inositol phosphate metabolism	54	24	22	35	20
00760	Nicotinate and nicotinamide metabolism	47	11	8	16	6
00130	Ubiquinone and other terpenoid-quinone biosynthesis	48	7	5	10	8
03410	Base excision repair	41	21	17	30	11
00564	Glycerophospholipid metabolism	99	35	29	50	33
04150	mTOR signaling pathway	37	26	19	37	23
00010	Glycolysis / Gluconeogenesis	90	26	23	36	28
04070	Phosphatidylinositol signaling system	39	26	25	36	22
00770	Pantothenate and CoA biosynthesis	34	8	5	11	10
00480	Glutathione metabolism	39	16	11	22	18
00520	Amino sugar and nucleotide sugar metabolism	138	25	23	34	27
00190	Oxidative phosphorylation	214	81	76	110	93
03440	Homologous recombination	59	20	16	27	14
00920	Sulfur metabolism	91	6	6	8	9
00650	Butanoate metabolism	76	12	4	16	13
00730	Thiamine metabolism	24	3	2	4	2
04141	Protein processing in endoplasmic reticulum	140	98	85	130	91
00563	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	26	19	17	25	11
00510	N-Glycan biosynthesis	44	32	29	42	26
04320	Dorso-ventral axis formation	24	10	7	13	11
00051	Fructose and mannose metabolism	91	14	13	18	15
00620	Pyruvate metabolism	87	21	18	27	21
00640	Propanoate metabolism	77	18	9	23	20
03050	Proteasome	48	34	32	43	34
03022	Basal transcription factors	35	27	24	34	30
00072	Synthesis and degradation of ketone bodies	8	4	2	5	4
03430	Mismatch repair	45	17	15	21	14
00240	Pyrimidine metabolism	168	61	53	75	63
00280	Valine, leucine and isoleucine degradation	61	31	8	38	30
00450	Selenocompound metabolism	28	9	7	11	8
01210	2-Oxocarboxylic acid metabolism	74	9	7	11	9
01230	Biosynthesis of amino acids	225	38	33	46	41

03008	Ribosome biogenesis in eukaryotes	82	58	59	69	60
03013	RNA transport	136	105	98	124	94
03018	RNA degradation	75	48	36	56	41
01200	Carbon metabolism	308	65	55	75	69
04122	Sulfur relay system	21	8	7	9	8
03020	RNA polymerase	51	25	23	28	24
00790	Folate biosynthesis	35	9	7	10	8
03030	DNA replication	51	31	26	34	27
03420	Nucleotide excision repair	48	37	29	40	30
03010	Ribosome	143	114	109	123	117
03040	Spliceosome	120	104	99	112	97
00900	Terpenoid backbone biosynthesis	50	17	16	18	13
00030	Pentose phosphate pathway	73	17	17	18	16
03060	Protein export	39	20	19	21	21
03015	mRNA surveillance pathway	59	53	50	55	50
00910	Nitrogen metabolism	56	6	5	6	5
00785	Lipoic acid metabolism	4	3	3	3	2
00471	D-Glutamine and D-glutamate metabolism	6	2	2	2	2
00020	Citrate cycle (TCA cycle)	50	23	22	22	22
00460	Cyanoamino acid metabolism	31	4	2	3	4