Федеральное государственное бюджетное учреждение «Федеральный научно-клинический центр физико-химической медицины Федерального медико-биологического агентства»

На правах рукописи

Графская Екатерина Николаевна

АНТИМИКРОБНЫЕ ПЕПТИДЫ СЕКРЕТА СЛЮННЫХ КЛЕТОК МЕДИЦИНСКОЙ ПИЯВКИ *HIRUDO MEDICINALIS*

03.01.04 — Биохимия

Диссертация на соискание ученой степени кандидата биологических наук

> Научный руководитель: доктор биологических наук, доцент Лазарев Василий Николаевич

ОГЛАВЛЕНИЕ

ВВЕДЕ	ЕНИВ	3	4
ГЛАВА	A 1. C	ЭБЗОР ЛИТЕРАТУРЫ	9
1.1.	Ан	гимикробные пептиды	9
1.2.	Стр	уктурная организация антимикробных пептидов	15
1.3.	Me	ханизмы действия антимикробных пептидов	17
1.4.	Ког	мпьютерные подходы для поиска и анализа антимикробных пептидов	20
1.4	l.1.	Компьютерное моделирование лекарственных средств	21
1.4	.2.	Предсказательные алгоритмы	23
1.4	.3.	Базы данных антимикробных пептидов	25
1.5.	Me	дицинская пиявка <i>H. medicinalis</i>	27
1.5	5.1.	Секрет слюнных клеток <i>H. medicinalis</i>	28
1.5	5.2.	Антимикробные пептиды медицинских пиявок	33
ГЛАВА	A 2. N	ИАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ	38
2.1.	Ma	териалы	38
2.2.	Me	тоды	40
2.2	2.1.	Генотипирование медицинских пиявок	40
2.2	2.2.	Выделение геномной ДНК <i>H. medicinalis</i>	40
2.2	2.3.	Очистка геномной ДНК <i>H. medicinalis</i>	41
2.2	2.4.	Сборка генома <i>H. medicinalis</i>	42
2.2	2.5.	Аннотация генома H. medicinalis	43
2.2	2.6.	Выделение секрета слюнных клеток <i>H. medicinalis</i>	43
2.2	2.7.	Пробоподготовка секрета слюнных клеток <i>H. medicinalis</i> для протеомного анал 44	иза
2.2	2.8.	Масс-спектрометрический анализ секрета слюнных клеток <i>H. medicinalis</i>	46
2.2	2.9.	Биоинформатический анализ аннотации генома <i>H. medicinalis</i>	46
2.2	2.10.	Предсказание вторичной структуры пептидов	47
2.2	2.11.	Химический синтез пептидов	47
2.2	2.12.	Оценка антимикробной активности пептидов	49
2.2	2.13.	Определение гемолитической активности пептидов	50
2.2	2.14.	Исследование цитотоксической активности пептидов	51
2.2	2.15.	Приготовление липосом разного липидного состава	52
2.2	2.16.	Измерение и оценка кинетики связывания пептидов с липосомами	52
2.2	2.17.	Электронная микроскопия E.coli и B. subtilis после инкубации с пептидами	53
2.2	2.18.	Статистический анализ	53

ГЛАВА	. 3. РЕЗУЛЬТАТЫ	54
3.1.	Определение полной нуклеотидной последовательности генома <i>H. medicinalis</i>	54
3.2.	Аннотация генома H. medicinalis	.57
3.3.	Протеомный анализ секрета слюнных клеток <i>H. medicinalis</i>	61
3.4.	Известные антимикробные пептиды <i>H. medicinalis</i>	.62
3.5. пепти	Использование биоинформатического алгоритма для поиска новых антимикробных адов <i>H. medicinalis</i>	.63
3.6. medic	Изучение биологической активности кандидатных антимикробных пептидов <i>H</i> . cinalis	68
3.7.	Определение вторичной структуры антимикробных пептидов	71
3.8. антим	Сканирующая электронная микроскопия клеток <i>E. coli</i> и <i>B. subtilis</i> после инкубации о ликробными пептидами	с 74
3.9.	Взаимодействие антимикробных пептидов с липосомами	.76
ГЛАВА	4. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ	.79
ЗАКЛЮ	ОЧЕНИЕ	. 88
вывод	цы	.90
СПИСС	Ж СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ	.92
СПИСС	ОК ЛИТЕРАТУРЫ	94
ПРИЛО	ЖЕНИЯ1	09

введение

Актуальность и степень разработанности темы исследования

Первые антибактериальные препараты для лечения и профилактики инфекций были разработаны в начале XX века. С течением времени, массовое бесконтрольное использование антибиотиков, в том числе и в сельском хозяйстве, привело к возникновению и широкому распространению резистентных к ним патогенных микроорганизмов. Инфекционные заболевания, вызванные такими бактериями, часто становятся причиной смерти пациентов, что представляет собой серьёзную проблему для здравоохранения во всем мире. Разработка новых антибиотиков является временным решением проблемы вследствие того, что устойчивость в отношении большинства антибиотиков у бактерий развивается в течение нескольких лет. Новые стратегии борьбы с резистентными микроорганизмами требуют поиска эффективных препаратов, принципиально отличающихся по механизму действия от используемых в настоящее время антибактериальных средств. Альтернативой синтетическим антибиотикам могут служить соединения природного происхождения, а именно, компоненты системы врождённого иммунитета многих организмов - антимикробные пептиды (АМП).

АМП ингибируют рост бактерий, обладают рядом преимуществ по сравнению с традиционными антибиотиками и рассматриваются как новый класс противомикробных средств. На данный момент описано несколько тысяч АМП, синтетического или природного происхождения. Тем не менее, только несколько пептидов зарегистрированы в качестве лекарственных препаратов, и почти три десятка проходят разные стадии клинических исследований. Основным препятствием для использования АМП в медицинской практике является их цитотоксичность по отношению к клеткам млекопитающих. В то же время, секреты животных могут служить источником АПМ природного происхождения, обладающих пониженной токсичностью в отношении клеток млекопитающих.

Медицинская пиявка *Hirudo medicinalis*, в течении нескольких тысячелетий применяемая для лечения различных заболеваний, является гематофагом - организмом, приспособившимся к питанию кровью животных. Секрет слюнных клеток (ССК) *H. medicinalis* является источником соединений, разнообразных по структуре и биологической активности, и имеющих большой потенциал для разработки лекарственных препаратов на их основе. Известно, что ССК *H. medicinalis*, помимо тромболитического, антикоагуляционного, сосудорасширяющего, антигипоксического и противовоспалительного действия, обладает антибактериальным эффектом. Кроме того, медицинская пиявка способна долгое время хранить кровь интактной в собственной пищеварительной системе, что говорит о наличии в её секрете противомикробных соединений с низким гемолитическим действием. Системных исследований ССК *H. medicinalis* ранее не проводилось, данные о белковом составе ССК скудны и фрагментарны. Хорошо изучены лишь белки, присутствующие в высоких концентрациях в секрете *H. medicinalis*, такие как гирудин, бделлины, калин, дестабилаза. На данный момент получены транскриптомы нескольких пиявок [1–7]. В 2016 году был проведён сравнительный анализ нуклеотидных последовательностей митогеномов (митохондриальных геномов) трёх кровососущих пиявок *Haementeria officinalis*, *Placobdella lamothei* и *Placobdella parasitica* [8]. В этом же году, нами впервые была получена нуклеотидная последовательность митохондриального генома *H. verbana* и *H. medicinalis* [9]. Определена нуклеотидная последовательность генома пресноводной пиявки *Helobdella robusta* без аннотации данных [10]. Наиболее изученной остаётся лишь микробиота медицинских пиявок *H. verbana* [11–13]. Таким образом, исчерпывающий анализ генома и белкового состава ССК *H. medicinalis* является важным и актуальным исследованием в области поиска и изучения новых антимикробных соединений белковой природы.

Цель и задачи

Цель данной работы заключалась в идентификация новых антимикробных пептидов ССК *Н. medicinalis* и изучении их структурно-функциональных свойств. В соответствии с целью исследования были сформулированы следующие задачи:

1. Определить нуклеотидную последовательность генома *H. medicinalis*, провести аннотацию генов *H. medicinalis*, проанализировать протеом ССК *H. medicinalis*.

2. Оптимизировать алгоритм идентификации антимикробных пептидов в геноме и применить его для поиска новых АМП в геноме и протеоме ССК *H. medicinalis*.

3. Химически синтезировать и определить антимикробную активность найденных АМП в отношении *Escherichia coli*, *Bacillus subtilis*, внутриклеточного патогена *Chlamydia trachomatis*, оценить их гемолитическое действие и цитотоксичность.

4. Изучить вторичную структуру активных АМП методом ЯМР-спектроскопии высокого разрешения.

5. Исследовать мембранолитическую активность АМП с использованием сканирующей электронной микроскопии и модельных липосом.

Научная новизна

В ходе исследования впервые была определена нуклеотидная последовательность генома *H. medicinalis*, была проведена его аннотация. Впервые был проанализирован протеом нативного секрета слюнных клеток *H. medicinalis*. Проведенная оптимизация биоинформатического алгоритма поиска АМП, применённого к аннотации генома и протеома ССК *H. medicinalis*, позволила найти новые антимикробные пептиды. Идентифицированные

пептиды обладают антимикробной активностью в отношении грамположительных и грамотрицательных бактерий, низкой гемолитической активностью и не проявляют цитотоксического эффекта по отношению к клеткам млекопитающих. Впервые методом ЯМРспектроскопии было продемонстрировано, что обнаруженные антимикробные пептиды принимают α-спиральную конформацию в мембранном окружении. Было показано, что пептиды *H. medicinalis* ингибируют рост бактерий путём дезинтеграции их клеточной мембраны.

Теоретическая и практическая значимость

Практическая значимость данной работы состоит в том, что аннотация генома *H. medicinalis* и анализ белкового состава ССК *H. medicinalis* представляют собой базу новых белков, которые могут послужить основой для разработки новых лекарственных препаратов. Алгоритм поиска АМП, применённый в работе, позволил идентифицировать АМП широкого спектра действия со сниженной цитотоксической активностью по отношению к клеткам млекопитающих. Новые АМП в дальнейшем могут рассматриваться в качестве потенциальных антибактериальных препаратов. Биоинформатический алгоритм для анализа геномных данных с целью поиска новых АМП может быть применён для поиска АМП в других организмах.

Методология и методы исследования

В данной работе определение нуклеотидной последовательности генома H. medicinalis и его сборка выполнены на современном оборудовании согласно стандартным методикам с использованием готовых компьютерных приложений, электронных баз данных. Все данные, полученные в ходе исследования, были депонированы в базе данных NCBI под регистрационными номерами BioProject acc. num. PRJNA257563 и PRJNA256119 и доступны другим исследователям для проверки и использования. Аннотация генома H. medicinalis и протеома ССК *H. medicinalis* проведена с использованием готового программного обеспечения и самостоятельно написанных скриптов на двух языках программирования (R, Python). Аналитические методы исследования вторичной структуры пептидов (ЯМР-спектроскопия), биохимические (определение антибактериальной, гемолитической, шитотоксической активностей) и физико-химические методы исследования пептидов (изучение механизма действия пептидов) выполнены на современном оборудовании и с использованием актуальных методик.

Положения, выносимые на защиту

1. Размер генома *H. medicinalis* составляет 187,5 Мп.н. и включает в себя 14596 белоккодирующих генов. Анализ протеома ССК *H. medicinalis* позволил идентифицировать в секрете слюнных клеток 189 белковых молекул. 2. Оптимизированный биоинформатический алгоритм идентификации АМП в геноме, нацеленный на поиск последовательностей, кодирующих короткие положительно заряженные пептиды с тенденцией к образованию α-спиральных структур, позволяет находить антимикробные пептиды.

3. Восемь из двенадцати идентифицированных пептидов *H. medicinalis* обладают антимикробной активностью в отношении *E. coli*, *B. subtilis* и *C. thrachomatis*. Исследуемые пептиды не обладают цитотоксической активностью в отношении клеток эукариот и вызывают слабый или умеренный гемолиз эритроцитов при концентрации 100 мкМ.

4. Антимикробные пептиды 3967, 536_2 и 12530 принимают α-спиральную конформацию в растворе, содержащем додецилфосфохолиновые мицеллы.

5. Антимикробные пептиды 536_1 и 3967 обладают литической активностью в отношении бактериальных мембран.

Личный вклад автора

Автором был выполнен поиск и анализ научной литературы по теме исследования, проведено планирование экспериментов, оптимизирован биоинформатический алгоритм анализа нуклеотидной последовательности генома *H. medicinalis* для идентификации АМП, были проведены работы по сбору данных, создание референсных баз АМП. Автором были выполнены работы с использованием микробиологических, биохимических (определение антибактериальной, гемолитической активностей) и физико-химических методов исследования активности пептидов (изучение механизма действия пептидов), был проведен анализ и интерпретация результатов, подготовка публикаций. Работа была выполнена в лаборатории генной инженерии ФНКЦ ФХМ ФМБА России в период с 2016 по 2020 год.

Степень достоверности и апробация работы

Основные результаты исследования были представлены и обсуждены на российских и международных конференциях: 20-я Международная Пущинская школа-конференция молодых ученых «Биология – наука XXI века» (Пущино, 2016); VII Международная школа молодых ученых по молекулярной генетике «Геномика и биология живых систем» (Звенигород, 2016); The 8th International Young Scientist School «Systems Biology And Bioinformatics - SBB-2016 (Новосибирск, 2016); 42nd Congress of the Federation-of-European-Biochemical-Societies (FEBS) on From Molecules to Cells and Back (Израиль, 2017); Hayчная конференция молодых ученых по медицинской биологии Φ ГБУ Φ HKЦ Φ XM Φ MБА (Москва, 2017); «Наука будущего - наука молодых» (Нижний Новгород, 2017); (МССМВ 2017) 8-th International Moscow Conference (Москва, 2017); Systems Biology and Bioinformatics The Ninth International Young Scientists School SBB-2017 (Ялта, 2017); 43rd FEBS Congress, Biochemistry Forever (Прага, 2018); V

Международная конференция «Постгеном'2018» (Казань, 2018); Научная конференция молодых ученых по медицинской биологии ФГБУ ФНКЦ ФХМ ФМБА (Москва, 2018); 7-ая Ежегодная Конференция по Медицинской химии (Ноттингем, 2019), IV Всероссийская научная конференция молодых ученых «Медико-биологические аспекты химической безопасности» (Санкт-Петербург, 2020).

Сведения о публикациях по теме диссертации

Результаты исследований опубликованы в 26 печатных работах, из них 5 статей в журналах из перечня рецензируемых научных журналов ВАК Минобрнауки РФ, 21 тезис в материалах российских и международных конференций.

Структура и объем диссертации

Диссертационная работа изложена на 119 страницах печатного текста и состоит из введения, обзора литературы, материалов и методов исследования, результатов, обсуждения результатов, заключения, выводов, списка сокращений и условных обозначений, списка литературы и приложения. Диссертационная работа содержит 16 таблиц и 24 рисунка. Библиографический указатель включает 293 литературных источника.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Антимикробные пептиды

Антимикробные пептиды (АМП) являются ключевым звеном врождённого иммунитета многоклеточных организмов и обеспечивают полноценную защиту от патогенов совместно с другими эндогенными факторами. [14]. У многоклеточных организмов АМП локализуются в иммунных клетках, жидкостях организма и в клетках эпителиальной ткани [15]. Последняя является основным барьером между внутренней и внешней средой организма и чаще взаимодействует с патогенными микроорганизмами. Для беспозвоночных животных показано, что при введении им антигенов микробного происхождения повышается синтез АМП и выброс их в гемолимфу [15]. АМП также были обнаружены в секретах ядовитых животных [16–18].

Многочисленные исследования последних двух десятилетий показывают, что помимо бактерицидной активности, АМП могут выполнять и другие защитные функции [19, 20]. Например, АМП участвуют во множестве процессов, связанных с иммунным ответом, от воспаления и хемоаттракции фагоцитов до заживления ран (Рисунок 1) [21–28]. АМП играют важную роль во взаимодействии с врожденным и адаптивным иммунитетом [29]. Также некоторые АМП обладают противовирусным и противораковым действием [30, 31].

АМП по сравнению с традиционными антибиотиками обладают рядом преимуществ. Пептиды имеют широкий спектр антибактериального действия, обладают активностью при низких концентрациях [32,33]. Кроме того, уникальный механизм действия большинства АМП заключается в нарушении целостности клеточной мембраны, что в дальнейшем приводит к гибели бактерии. Патогену для нейтрализации активности пептида необходимо производить сложные структурные и электрофизиологические изменения клеточной мембраны. Таким образом, ещё одним преимуществом АМП перед антибиотиками является то, что развитие устойчивости патогенов к ним крайне ограничено. Все эти свойства делают АМП перспективными соединениями, на основе которых могут быть разработаны лекарственные препараты следующего поколения, обладающие многообещающим антимикробным потенциалом.

На данный момент несколько АМП применяются в клинической практике. Например, циклический липопептид полимиксин В, продукт ферментации бактерий *Bacillus polymyxa*, эффективен против резистентных грамотрицательных штаммов и широко используется для лечения инфекций мочевыводящих путей, менингита, мукоцита, отита, пародонтита, инфекций легких, ушей, глаз и раневых инфекций [35,36]. Однако, было показано, что полимиксин В в высоких концентрациях обладает нейротоксичным и нефротоксичным действиями [37]. В исследованиях по повышению биологической доступности (способности препарата усваиваться) полимиксина В и снижению побочных эффектов пептид помещали в липидные наночастицы или липосомы [38]. Данный подход позволил снизить значения минимальных ингибирующих концентраций (МИК) инкапсулированной формы пептида по сравнению с его свободной формой, тем самым усилить антимикробный эффект полимиксина В.



Рисунок 1 - Многообразие функциональных активностей АМП

АМП привлекают антигенпрезентирующие клетки путём индукции хемокинов, активируют нейтрофильные внеклеточные ловушки, изменяют эндотоксин-опосредованные сигнальные пути, подавляют провоспалительные цитокины, усиливают фагоцитоз и провоспалительные реакций на нуклеиновые кислоты, индуцируют противовоспалительные цитокины, влияют на дифференцировку дендритных клеток и поляризацию Т-клеток и заживление ран. ЛПС, липополисахарид. Адаптировано из [34].

Другой пептид, низин, продуцируемый молочнокислой бактерией Lactococcus lactis, является бактериоцином с модифицированными аминокислотами, и относится к классу лантабиотиков [39]. Низин активен в отношении грамположительных бактерий, не токсичен для клеток млекопитающих при концентрациях равных МИК и не вызывает гемолиза эритроцитов [40]. Низин используется как консервант в пищевой промышленности и применяется в качестве противогрибкового терапевтического средства при лечении мастита, кандидоза и других гинекологических инфекций у крупного рогатого скота [41].

Ряд АМП в настоящее время проходят доклинические и клинические исследования (Таблица 1) [42–45]. Например, пептид АА-139, модифицированный аналог антимикробного

пептида ареницина-3, выделенного из червя-пескожила *Arenicola marina*, активен в отношении микробов, обладающих множественной лекарственной устойчивостью [46]. Пептид AA-139 кроме того осуществляет противомикробную активность за счет нарушения целостности клеточной стенки и мембраны, а так же способен ингибировать синтез белка [47]. В настоящее время пептид AA-139 проходит доклинические исследования эффективности при инфекциях мочевыводящих путей, внутрибольничной и респиратор-ассоциированной пневмоний.

Синтетический катионный пептид пексиганан является аналогом известного антимикробного пептида магаинина, выделенного из лягушки *Xenopus laevis* [48,49] Пексиганан обладает антибактериальной активностью *in vitro* в отношении грамположительных и грамотрицательных аэробных бактерий. В настоящее время пексиганан тестируется в двух клинических исследованиях для лечения бактериальных инфекций, связанных с язвами, обусловленными синдромом диабетической стопы [50].

синтетический пептид омиганан, модифицированный АМП Другой вариант индолицидина, тридекапептида, выделенного ИЗ цитоплазматических гранул бычьих нейтрофилов, проходит клинические исследования для лечения инфекций, связанных с катетеризацией, и кожных заболеваний, таких как розацеа, атопический дерматит, генитальные бородавки и интраэпителиальная неоплазия вульвы [51,52]. По результатам II фазы исследований терапия данным пептидом оказалась эффективной.

Одним из антимикробных пептидов с большим терапевтическим потенциалом является синтетический пептид LTX-109. Этот пептид является литическим для ряда метициллинрезистентных S. aureus, включая штаммы, устойчивые к ванкомицину, и штаммы с пониженной чувствительностью к даптомицину и линезолиду [53,54]. Исследования показали, что бактерии не развивают устойчивость к пептиду LTX-109, а сам пептид обладает фунгицидной активностью. В 2014 году была успешно завершена II фаза исследований по лечению импетиго этим пептидом. В 2011 году подошли к концу клинические исследования по использованию лечения неосложнённых инфекций, этого пептида для кожных вызванных грамположительными бактериями, И назальных инфекций, вызванных метициллинрезистентными S. aureus [54].

11

12

Название препарата (коммерческое название)	Описание препарата	Механизм действия	ИД DrugBank ¹	Разработчик	Оказания для применения	Клинические исследования	Ссылка
AA-139	Аналог ареницина	Разрушает клеточную мембрану	-	Adenium Biotech	Инфекции мочеполовой системы, пневмония	Преклинические исследованияия	[46]
Боцепревир (Boceprevir, Victrelis)	Синтетический циклический пептид	Ингибирует репликацию вируса гепатита С	DB08873	Merck Sharp & Dohme Corp.	Вирус гепатита С, ВИЧ	NCT01353911 (II) NCT01335529 (II) NCT01425203 (III) NCT01591460 (IV)	[59]
Далбаванцин (Dalvance, Zeven, Xydalba)	Липогликопептидный антибиотик, разработан на основе ванкомицина и тейкопланина	Встраивается в в клеточную стенку и нарушает синтез пептидогликана	DB06219	Pfizer Inc./ Durata Therapeutics Inc./ Allergan	Острые бактериальные инфекции кожи и кожных структур	NCT00678106 (I) NCT02814916 (III) NCT02127970 (III) NCT02961764 (IV) NCT03372941 (IV)	[60]
Даптомицин (Daptomycin, Cubicin, Surotomycin)	Природный циклический липопептидный антибиотик	Разрушает клеточную мембрану и ингибирует синтез белков, ДНК и РНК	DB00080	Cubist Pharmaceuticals Inc./ Merck Sharp & Dohme Corp.	Острые бактериальные инфекции кожи и кожных структур, клостридиальная диарея	NCT01019395 (I) NCT00428844(II) NCT01922011 (III) NCT01597505 (III) NCT01419184 (IV) NCT01728376 (IV) NCT02835105 (I)	[61]
Литиксар (Lytixar, LTX-109)	Синтетический пептид	Разрушает клеточную мембрану	DB12711	Lytix Biopharma AS	Атопический дерматит, легкая экзема/дерматоз, бактериальные кожные инфекции	NCT01158235 (II) NCT01223222 (II)	[53]

Таблица 1 - Лекарственные препараты на основе АМП, тестируемые в доклинических и клинических исследованиях

Продолжение Таблицы 1

Мурепавадин (Murepavadin, POL-7080)	Циклический пептид	Ингибирует синтез клеточной стенки	DB14777	Polyphor Ltd.	Внутрибольничная пневмония, респиратор- ассоциированная пневмония	NCT03582007 (III) NCT03409679 (III)	[62]
Омиганан (Omiganan, Omigard, CLS001)	Синтетический аналог индолицидина	Вызывает деполяризацию цитоплазматичес кой клеточной мембраны	DB06610	Maruho Co., Ltd.	Себорейная экзема, атопический дерматит, пустулезная розацеа, интраэпителиальная неоплазия вульвы	NCT02028286 (I) NCT03688971 (II) NCT03091426 (II) NCT01784133 (II) NCT02596074 (II) NCT03091426 (II) NCT02576847 (III)	[63]
Оритаванцин (Oritavancin, LY333328)	Синтетический гликопептид	Разрушает клеточную мембрану	DB04911	The Medicines Company Melinta Therapeutics, Inc.	Острые бактериальные инфекции кожи и кожных структур	NCT02134301 (I) NCT03873987 (I) NCT02925416 (IV) NCT03761953 (IV) NCT02452918 (IV)	[64]
Пексиганан (Pexiganan, Locilex, MSI-78)	Аналог магаинина 2	Разрушает клеточную стенку	-	Dipexium Pharmaceuticals Inc./ Genaera Corporation	Инфицированные язвы, обусловленные синдромом диабетической стопы	NCT01594762 (III) NCT01590758 (III) NCT00563433 (III) NCT00563394 (III)	[55]

Продолжение Таблицы 1

Телаванцин (Telavancin, Vibativ)	Синтетический липогликопептид, аналог ванкомицина	Встраивается в в клеточную стенку и нарушает синтез пептидогликана	DB06402	Cumberland Pharmaceuticals	Осложнённые инфекции кожи и кожных структур, вызванных грамположительными бактериями и метициллин- резистентными <i>S.</i> <i>аиreus</i> , кистозный фиброз, внебольничная пневмония	NCT00061633 (II) NCT00062647 (II) NCT00077675 (II) NCT00091819 (III) NCT00124020 (III) NCT00107952 (III) NCT03172793 (IV)	[56,57]
Энфувиртид (Enfuvirtide, Fuzeon, T-20)	Синтетический липопептид	Препятствует проникновению ВИЧ-1 в клетки путём связывания с субъединицей gp41	DB00109	Genentech, Inc./ Roche Laboratories Inc.	Вирус иммунодефицита человека, прогрессирующая мультифокальная лейкоэнцефалопатия	NCT00615134 (II) NCT00002228 (II) NCT00454337 (III) NCT02733419 (III) NCT00657761 (IV)	[58]

¹ - идентификационный номер в базе данных лекарственных веществ с химической, фармакологической и фармацевтической информацией DrugBank

1.2. Структурная организация антимикробных пептидов

Семейство АМП достаточно обширно и разнообразно, в зависимости от пространственной структуры АМП можно разделить на пять групп (Рисунок 2) [14,65,66]: 1) αспиральные АМП; 2) АМП, содержащие β-лист, часто с двумя или более дисульфидными связями; 3) АМП с конформацией β-шпильки или петли, стабилизированной присутствием одинарной дисульфидной связи и/или циклизацией пептидной цепи; 4) АМП смешанного типа, содержащие домены разной структуры; 5) АМП с неупорядоченной конформацией.



Рисунок 2 - Структурная организация антимикробных пептидов Представлены аминокислотная и пространственная структуры (A) α-спиральных пептидов (магаинин-2, PDB ID: 2MAG); (Б) пептидов, содержащих β-листы с дисульфидными связями (кроличий дефензин-1, PDB ID: 1EWS); (В) β-шпилечных пептидов (θ-дефензин, PDB ID: 2LZI); (Г, Д) АМП смешанного типа (танатин, PDB ID: 8TFV; человеческий β-дефензин-2, PDB ID: 1FQQ) и (Е) АМП с неупорядоченной структурой (индолицидин, PDB ID: 1G89). Участки первичной структуры, принимающие α-спиральную конформацию, обозначены чёрным прямоугольником, β-складчатую структуру — серыми стрелками, остатки цистеинов, образующие дисульфидные связи, выделены сплошными линиями. На пространственных структурах обозначены дисульфидные связи (жёлтым), α-спирали (красным) и β-структуры (синим). Иллюстрация разработана автором с использованием графической программы РуМОL (версия 1.2.2.3) [67].

Наиболее распространёнными АМП являются α-спиральные пептиды, такие как мелиттин, магаинины, цекропины и кателицидины [68,69]. Большинство этих пептидов были выделены из секретов насекомых, лягушек, млекопитающих и других позвоночных [70,71].

Подобные АМП как правило неструктурированны в водном растворе, но при взаимодействии с липидами бактериальной мембраны формируют α-спираль [72,73]. Пептиды, принимающие конформацию β-листа или β-шпильки, имеют определённое количество β-тяжей, организованных по амфипатической схеме, с выраженными гидрофобными и гидрофильными областями. Примерами таких пептидов являются α- и β-дефензины, танатин [74,75].

Большинство АМП - амфифильные положительно заряженные молекулы, но встречаются и анионные АМП [76]. Длина последовательности АМП варьируется от 5 до 150 а.о. [69,77-79]. В основном АМП синтезируются с помощью рибосомального синтеза и нечасто подвергаются посттрансляционным модификациям [80–82]. Известно, что у антимикробных пептидов морского ежа, центроцинов, происходит бромирование остатков триптофана [80]. У антимикробного пептида дрозоцина, выделенного из *Drosophila melanogaster*, было обнаружено О-гликозилирование остатка треонина [81]. Кроме того, было продемонстрировано, что отсутствие посттрансляционной модификации ведет к снижению биологической активности пептида.

Подавляющая часть АМП образуется из молекул-предшественников после специфического протеолиза (Рисунок 3). Иногда зрелый пептид формируется путём объединения фрагментов разных полипептидных цепей через дисульфидные связи [80]. Так как АМП чаще всего являются секретируемыми молекулами, то самый простой предшественник АМП, помимо зрелой цепи, содержит сигнальный пептид. По указанному механизму синтезируется большинство растительных дефензинов и абаецин из пчелы *Apis mellifera* [68,82].



Рисунок 3 - Структурная организация пептидов-предшественников АМП

(A) - «простые», (Б) - «простые» с пропоследовательностью, (В) - «сложные» двухдоменные, (Г) - крупный функциональный белок с АМП в его составе, (Д) - «сложные» конкатемерные. Адаптировано из [87].

Другие АМП имеют дополнительные просегменты, которые могут располагаться как на N-, так и на С-конце белка. Такие просегменты обычно богаты кислыми и амфипатическими аминокислотными остатками. Препропептидный тип прекурсоров предназначен для эффективного подавления литической активности зрелых АМП. Такой тип предшественников АМП широко распространён и характерен для большинства дефензинов, пчелинного мелиттина, дермасептинов земноводных, растительных тионинов, кателицидинов млекопитающих и т.д [83–86].

Некоторые АМП образуются в результате протеолитической деградации больших белков. Например, буфорин азиатской жабы Bufo gargarizans формируется из N-концевой части гистона Н2А, а лактоферрицин возникает путём протеолиза белка лактоферрина млекопитающих [88,89]. Более сложные типы архитектуры белков-предшественников АМП предназначены для синтеза нескольких копий одного пептида или разных пептидов одновременно. Сложные предшественники кодируются конкатемерной ДНК. гле повторяющиеся последовательности зрелых пептидов разделены спейсерами, содержащими специфические сайты расщепления. Магаинины из лягушки X. laevis, пчелиные апидаецины и АМП растений обладают описанным сложным типом организации белка-предшественника [90,91].

1.3. Механизмы действия антимикробных пептидов

Механизмы антибактериального действия АМП также разнообразны как и их структуры. Тем не менее, можно выделить мембранолитические и нелитические способы действия АМП (Рисунок 4). При мембранолитическом действии АМП вызывает дезинтеграцию цитоплазматической мембраны патогена, что приводит к гибели клетки. При нелитическом механизме действия пептиды нарушают ключевые внутриклеточные процессы, такие как синтез нуклеиновых кислот, белков, биосинтез клеточной стенки и т.д. [20,32].

Мембранолитический механизм действия АМП осуществляется следующим образом. Сначала катионные АМП адсорбируются на поверхности клетки-мишени за счёт электростатических и гидрофобных взаимодействий с отрицательно заряженными консервативными липидными и полисахаридными компонентами мембраны клетки [92]. А именно, молекулы АМП конкурентно замещают стабилизирующие мембрану двухвалентные катионы (Mg²⁺ и Ca²⁺). Далее происходит встраивание пептида в поверхностный слой дестабилизированной мембраны. Гидрофильная часть молекулы АМП направлена наружу, гидрофобная часть погружена в слой углеводородных радикалов жирных кислот мембраны клетки. В результате пептид деформирует фосфолипидную мембрану и снижает её

термодинамическую стабильность, что ведёт к локальному нарушению целостности мембраны и последующей интеграции пептида в толщу липидного бислоя [93,94]. Пороговая концентрация, необходимая для проникновения АМП в гидрофобный слой мембраны, зависит от физико-химических характеристик пептида, биохимического состава и электрофизиологических свойств мембраны [95].



Рисунок 4 - Механизмы действия антимикробных пептидов: мембранолитические (A) и внутриклеточные (Б)

Агрегированные на поверхности мембраны пептиды достигают пороговой концентрации и дестабилизируют её. Пептиды, воздействующие на внутриклеточные мишени патогена, способны ингибировать синтез нуклеиновых кислот, белков, препятствовать правильному сворачиванию белков.

Выделяют три основных модели мембранолитических механизмов действия АМП. В действительности, подвидов и модификаций литических механизмов гораздо больше. Согласно первой «модели цилиндрической поры» (англ. *barrel stave model*) молекулы АМП встраиваются в мембрану и образуют олигомерные поры, внутренняя поверхность которых образована гидрофильными, положительно заряженными аминокислотными остатками пептида [92]. Для формирования пор пептид должен обладать определённой пространственной структурой, такой как α-спираль и/или β-лист [20]. Данная модель описывает действие анион-селективных, слабо заряженных пептидов, таких как писцидина, АМП выделенного из рыбы *Oplegnathus fasciatus* [96].

Вторая «модель тороидальной поры» (англ. *toroidal pore model*) предполагает, что олигомерные поры-каналов образованы не только АМП, но и липидами бактериальной мембраны [97,98]. Таким образом, стабильность комплекса повышается за счёт электростатических взаимодействий катионного пептида и анионных голов фосфолипидов.

Данная модель впервые была предложена для магаинина 2, линейного АМП, выделенного из шпорцевой лягушки *X. laevis* [97]. Показано, что такой механизм действия характерен и для других магаининов, протегринов и мелиттина [20]. Существует модифицированный вариант этой модели - «модель неупорядоченной тороидальной поры» (англ. *disordered toroidal pore model*), которая показывает, что для стабилизации трансмембранного канала, сильно отличающейся от цилиндрической и классической тороидальной поры, достаточно нескольких молекул положительно заряженного пептида, причём конформация АМП может быть любой [99].

Последняя из основных мембранолитических моделей - «модель ковра» (англ. carpet model) предполагает, что лизис бактериальной клетки осуществляется не путём образования специфических пептид-пептидных взаимодействий, стабильных пор И а за счёт детергентподобного действия молекул АМП [100,101]. Из-за высокой концентрации пептида на поверхности бактерии происходит утоньшение поверхностного слоя мембраны, образование пор, дезинтеграция липидного слоя, что ведет к гибели клетки. Данная модель описывает механизм действия цекропинов и дермасептина земноводных [85,102]. Амфифильные молекулы АМП в больших концентрациях способны образовывать олигомерные агрегаты, которые встраиваются в гидрофобный слой мембраны и формируют трансмембранные поры. Такая модель взаимодействия АМП с бактериальной мембраной получила название «модели двух состояний» (англ. two state model) [103]. Несмотря на описанные мембранолитические модели, процессы, происходящие во время воздействия АМП на бактериальные клетки до сих пор остаются предметом дискуссий.

В последнее время все больше исследований посвящено изучению нелитических механизмов действия АМП, при которых АМП действуют на внутриклеточные мишени патогена и нарушают жизненно важные процессы в клетке [104,105]. Например, глицинбогатые АМП ингибируют транскрипцию генов мембранных белков, что вызывает лизис мембраны. Аттацины, выделенные из гемолимфы насекомых, связываются с ЛПС бактерий, которые вступают в качестве рецептора [106]. Затем пептиды активируют сигнальный путь и ингибируют синтез белков наружной мембраны, при этом не проникая сквозь внутреннюю мембрану клетки. В ряде публикаций продемонстрировано, что дефензины млекопитающих и беспозвоночных являются с пецифическими ингибиторами биосинтеза клеточной стенки бактерий [107]. Таким ингибитором может выступать и низин, известный бактериоцин, продуцирующийся бактериями родов *Lactococcus* и *Streptococcus* родов. Низин связывается с липидом II, основным структурным белком при биосинтезе пептидогликана [108].

Пролин-богатые АМП насекомых и животных в малых концентрациях способны

избирательно ингибировать метаболические процессы у грамотрицательных бактерий, при этом не повреждая клеточной мембраны [109]. Например, пептид бактенецин-7, выделенный из гранул бычьих нейтрофилов, связываясь с бактериальными рибосомальными белками, нарушает трансляцию белков в клетке [110]. Пептид буфорин-2 азиатской жабы *Bufo gargarizans* нарушает синтез ДНК и мРНК путём неспецифического связывания с нуклеиновыми кислотами, тогда как индолицидин, выделенный из цитоплазматических гранул бычьих нейтрофилов, и тахиплезин, выделенный из подковообразного краба, ингибируют репликацию ДНК [111–113]. Пиррхорицин, полученный из насекомых, связывается с шапероном hsp70 и препятствует правильному сворачиванию белков, что приводит к накоплению неактивных белков и последующей гибели клеток [114]. Для таких АМП, как ареницин 1 из целомоцитов *Arenicola marina*, сколопендина 1 из многоножки *S. subspinipes mutilans*, продемонстрирован ещё один механизм гибели патогенов путём принудительной индукции апоптоза [115,116].

Помимо антимикробного действия, для некоторых нелитических АМП показано, что они проявляют противовирусную и противораковую активности. Например, пептид энфувиртид структурно сходен с ингибиторами слияния ВИЧ с клеточной мембраной [30]. Для данного пептида показано, что он препятствует слиянию ВИЧ-1 с клетками CD4. Другой пептид ауреин 1.2, выделенный из лягушки *Litoria aurea*, обладает широким спектром действия в отношении бактерий, а так же проявляет высокую активность в отношении 55 раковых линий клеток *in vitro*, не проявляя при этом значительной цитотоксической активности в отношении нативных клеточных линий млекопитающих [31]. В ходе исследований человеческого α -дефензина HNP-1 продемонстрирована перспективность применения данного пептида при лечении раковых заболеваний [21].

Обобщая все вышесказанное, важно подчеркнуть, что модели антимикробного действия АМП разнообразны и многие пептиды воздействуют на бактерии путём реализации нескольких механизмов действия одновременно. Реализация того или иного механизма зависит от физикохимических свойств пептида, его концентрации и состава окружающей среды, а также от особенностей клетки-мишени.

1.4. Компьютерные подходы для поиска и анализа антимикробных пептидов

Существующие подходы для поиска корреляций между активностью АМП и их физикохимическими и структурными характеристиками не дают конкретных результатов в связи с огромным разнообразием АМП. Вычислительная биология в сочетании с высокопроизводительными экспериментальными методами предлагает новый путь развития биоинформатических алгоритмов для поиска и оптимизации АМП, которые в дальнейшем могут способствовать разработке новых эффективных противомикробных лекарств.

1.4.1. Компьютерное моделирование лекарственных средств

В качестве многообещающего подхода для поиска новых терапевтических препаратов, в числе обладающих антимикробной активностью, рассматривается компьютерное том моделирование (англ. computer-aided drug design) (Рисунок 5) [117,118]. Компьютерный молекулярный дизайн заключается в определении количественной зависимости между активностью соединения и его химической структурой [119]. Таким образом, можно определить наиболее важные характеристики молекулы, определяющие её биологические функции. На основе полученных данных, новые активные соединения могут быть разработаны либо путём модификации существующих молекул, либо путём комбинирования функциональных групп для создания новых [120].

Для оптимизации соединения необходима информация структуры 0 фармакокинетических свойствах вещества, определяющих его конкретные функции и активность [121,122]. А именно, данные об абсорбции, распределении и локализации в тканях, метаболизме, экскреции, токсичности и т.д.. Структурные И физико-химические характеристики вещества описывают с помощью специальных молекулярных параметров дескрипторов [123]. По определению, молекулярный дескриптор является конечным результатом логической и математической процедуры, которая преобразует химическую информацию, закодированную в символическом представлении молекулы, в число или результат какого-либо стандартизированного эксперимента [124]. Примеры дескрипторов: молекулярный вес, наличие ароматических колец, наличие вращающихся связей, межатомные расстояния, поляризуемость, показатели ароматичности и сольватационные свойства. Значения дескрипторов получают на основе теоретико-графовых, молекулярно-механических или квантово-механических моделей [125,126]. В зависимости от наличия или отсутствия данных о химической структуре соединения, расчёт дескрипторов молекулы может быть выполнен либо на основе физико-химических свойств вещества - лиганда (англ. ligand-based drug design), либо его структурных характеристик (англ. structure-based drug design). Метод на основе лиганда позволяет предсказывать сайты взаимодействия между молекулой-мишенью и действующим лекарственным соединением [117]. Структурно-ориентированный подход позволяет рассчитать энергию такого взаимодействия [119].

Дизайн антимикробных пептидов может быть проведен как лиганд-, так и структурноориентированными методами. В исследовании Даи с соавторами методом компьютерного моделирования на основе структуры индольного алкалоида эвдистомина U были разработаны 32 новых его производных [127]. Было показано, что шестнадцать из 3,9-дизамещенных производных эвдистомина U проявляют антибактериальное действие в отношении грамположительных (метициллин-резистентных *S. aureus*, *B. cereus*) и грамотрицательной бактерии *Ralstonia solanacearum*. Согласно результатам исследования, противомикробный эффект соединения бр оказался выше, чем у антибиотиков фосфомицина, ципрофлоксацина и пропинеба. Бактерицидный эффект соединений обусловлен одновременным повреждением бактериальной мембраны и нарушением функций ДНК-гиразы.



Рисунок 5 - Схема разработки лекарственных средств с помощью компьютерного моделирования

В другом исследовании, Ли с коллегами показали, что такие характеристики, как суммарный положительный заряд, длина, аминокислотный состав пептида и влияние растворителя, являются наиболее значимыми для проявления антимикробной активности пептида [128]. Именно эти параметры влияют на дестабилизирующее действие, оказываемое пептидом на бактериальную мембрану.

1.4.2. Предсказательные алгоритмы

Класс АМП очень разнообразен, пептиды реализуют множество биологических функций, поэтому их невозможно рассматривать как одну категорию соединений. Последние достижения в области вычислительных наук, такие как генетические алгоритмы (англ. genetic algorithm), методы машинного обучения (англ. machine learning methods) и методы глубоких нейронных сетей (англ. deep neural networks), позволили использовать полученные экспериментальные данные для рационального дизайна антимикробных белков и пептидов. Далее будут рассмотрены основные алгоритмы для поиска и оптимизации АМП.

Генетические алгоритмы

В основе генетических алгоритмов лежат базовые принципы естественного отбора: молекула «эволюционирует» в направлении желаемой биологической функции. С помощью данного подхода можно анализировать любые последовательности АМП. Для этого нужна лишь функция качества (англ. *fitness function*), построенная на основе дескрипторов активности молекул и информация, собранная из баз данных АМП. Несмотря на избыточность последовательностей, генерируемых такими алгоритмами, этот метод способен предсказывать разнообразные АМП [123]. Использование генетического алгоритма для оптимизации последовательности глицин-богатого пептида из гуавы *Psidium guajava*, Pg-AMP1, позволило разработать гуаванин 2 - пептид, проявляющий бактерицидное действие в отношении *P. aeruginosa* [129]. В исследовании Супади с коллегами с помощью генетических алгоритмов был разработан программный код, позволяющий модифицировать структуру исходного соединения и выделять наиболее активные конформеры [130].

Алгоритмы машинного обучения

Машинное обучение является одной из бурно развивающихся областей вычислительных наук, которая нашла широкое применение в биологической инженерии и синтетической биологии. Метод работает следующим образом. Сначала строится модель, которая обучается на определённом статистически достоверном массиве данных. Затем модифицированная модель тестируется на дополнительной выборке данных. Этот процесс можно повторять непрерывно, причём модель будет учиться не только на выделенных данных самостоятельно, но и на новых полученных данных. Такой метод позволяет быстро и эффективно анализировать и интерпретировать большой массив данных [131].

Йошида с соавторами разработали метод с использованием алгоритмов машинного обучения для оптимизации активности природного АМП [132]. На основе темпорина из *Amolops jingdongensis* авторы синтезировали 44 новых пептида, активность которых более чем в 160 превышала активность исходной молекулы. Кроме того, авторы изменили конформацию

идентифицированных молекул. Они преобразовали неструктурированные пептиды в αспиральные, что также повлияло на их активность. Ли с соавторами, применяя алгоритмы машинного обучения, построили классификатор для разработки АМП с заданным механизмом действия [133]. Классификатор, построенный с использованием метода опорных векторов, позволил предсказать 16 антимикробных пептидов различной конформации, активных в отношении *S. mutans, E. coli, P. aeruginosa*.

Методы глубокого обучения

Методы глубокого обучения, также известные как иерархически встроенные глубокие нейронные сети, в настоящее время считаются наиболее перспективными из доступных методов искусственного интеллекта и компьютерного моделирования для разработки новых лекарственных препаратов [134]. Алгоритмы глубокого обучения на вход получают неструктурированные и неразмеченные данные. Затем модель итеративно обучается на большом количестве «скрытых слоёв» (англ. *hidden layer*), пока не достигнет определённого уровня, заданного разработчиком. Преимущество данных алгоритмов состоит в том, что для процесса самообучения можно использовать большой массив разрозненных данных.

На данный момент только несколько групп исследователей использовали алгоритмы глубокого обучения для создания предсказательных моделей. Велтри с коллегами разработали простую глубокую нейросеть, позволяющую достаточно точно идентифицировать АМП [135]. В другом исследовании был представлен альтернативный подход для комбинаторного компьютерного моделирования АМП *de novo*, разработанный на основе рекуррентных нейронных сетей с генеративной долгой краткосрочной памятью (англ. *generative long short-term memory, LSTM*) [136]. Разработанная модель, обученная на аминокислотных последовательностях α-спиральных АМП, полученных из баз данных АМП, сгенерировала новые активные молекулы.

Гибридные модели

Предполагается, что не существует одного универсального метода дизайна АМП. Вероятно, сочетание вычислительных подходов на основе структурных и физико-химических свойств этих молекул обеспечит необходимую точность для успешной разработки новых лекарственных препаратов. Согласно недавним исследованиям, объединение описанных выше подходов с методами молекулярной динамики и молекулярного моделирования повышает предсказательную способность алгоритма. Разработанная Шнайдер и коллегами гибридная модель для создания АМП сочетает в себе вычислительную сеть и нейронную сеть глубокого обучения [137]. Данная модель, обученная на основе последовательностей α-спиральных известных АМП и α-спиральных нетрансмембранных белков, обладает большей точностью

классификации и прогностической способностью по сравнению с более простыми моделями. Среди пептидов, сгенерированных случайным образом с помощью модели, один ингибировал рост *S. aureus* при значении МИК равному 10 мкМ. Мело с коллегами разработали интегрированный, настраиваемый и простой в использовании программный пакет, который позволяет проводить высококачественный и точный анализ свойств и характеристик белков или пептидов [138]. В данном программном пакете были объединены подходы молекулярной динамики (NAMD и VMD) и квантовой химии (ORCA и MOPAC) [138–140].

Конструирование молекул, обладающих специфичным механизмом антимикробного действия в отношении определённых штаммов, является следующим шагом для моделей прогнозирования. Эти модели позволят определить сложные зависимости между дескрипторами и активностью АПМ. Вишнепольский и соавторы описали прогностическую модель для разработки линейных АМП, активных в отношении грамотрицательных бактерий [141]. Сочетание машинной модели с частичным обучением и пространственной кластеризацией позволило создать пептиды, обладающие специфической антимикробной активностью в отношении E.coli ATCC 25922 и P. aeruginosa ATCC 27853. Другая группа учёных под руководством Велтри разработала модель, которая на основе аминокислотных последовательностей АМП предсказывает не только антимикробную активность молекулы, но и в отношении какого рода бактерий пептид будет более активен [142].

Описанные вычислительные методы предлагают альтернативный подход для поиска новых АМП, устраняя необходимость в проведении трудоёмких и дорогостоящих тестирований библиотек биологически активных соединений. Тем не менее, каждый из алгоритмов имеет свои границы применимости, зависящие от исходных данных, используемых при создании модели.

1.4.3. Базы данных антимикробных пептидов

За десятилетия исследований АМП природного и синтетического происхождения был накоплен огромный массив данных о их физико-химических, структурных свойствах и активностях. Ввиду этого, были созданы базы данных, объединяющие и систематизирующие полученные знания. По мере обнаружения новых активных молекул базы данных дополнялись или же создавались новые. На сегодняшний день их насчитывается порядка 20.

Первая база данных по антимикробным пептидам APD (англ. <u>Antimicrobial Peptide</u> <u>Database</u>) была создана в 2003 году [143]. С тех пор она постоянно обновляется и на данный момент в ней содержится 3199 аминокислотных последовательностей АМП. Среди антимикробных соединений большинство составляют пептиды и белки животного происхождения, включая синтетические. Остальные — это АМП растений, бактерий, грибов, простейших и архей. В базе данных помимо информации об активности пептида и его аминокислотной последовательности, указаны данные о его структуре (при наличии).

Следующей из широко используемых является база данных CAMP_{R3} (англ. <u>Collection of</u> <u>Anti-Microbial Peptides</u>), в которой собрано 8164 аминокислотных последовательностей AMП [144]. Среди этих соединений антимикробное действие подтверждено только для 2766 последовательностей, активность остальных предсказана по гомологии с подтверждёнными AMП или же математическими методами. По сравнению с другими базами данных, CAMP_{R3} содержит подробную информацию о специфических аминокислотных мотивах, характерных для семейств эукариотических и прокариотических AMП. Экспериментальные данные использовались для создания инструментов прогнозирования с применением подходов машинного обучения (метод случайного леса (англ. *random forest*), дискриминантный анализ (англ. *discriminant analysis*), метод опорных векторов (англ. *support vector machine*)) и методы глубокого обучения (метод искусственной нейронной сети (англ. *artificial neural network*)) для идентификации и классификации новых соединений. Предикторы принимают в качестве входных данных матрицу аминокислотной последовательности пептида, преобразованную в физико-химические свойства и структурные характеристики аминокислот, а также частоты встречаемости дипептидов в последовательности.

Другая база данных антимикробных пептидов ADAM (англ. <u>A Database of Anti-Microbial</u> peptides) содержит 7007 уникальных аминокислотных последовательностей AMП и 759 структур [145]. На сервере базы данных можно проанализировать аминокислотные последовательности пептидов с помощью двух алгоритмов, построенных с использованием метода опорных векторов и скрытой марковской модели (англ. *hidden markov model*). Предсказывающие алгоритмы были обучены на аминокислотных последовательностях AMП базы данных, используя аминокислотный состав в качестве признака.

Многие базы данных АМП используются в качестве обучающей выборки для построения предсказательных алгоритмов. Существенным недостатком большинства баз данных АМП является ограниченность представленной информации. Например, лишь немногие содержат данные об активности АМП в отношении протестированных патогенов. Как показывает практика, такая информация важна для точности предсказательного алгоритма. Базы данных являются полезным инструментом для более детального анализа данных об АМП.

В базах данных преобладают АМП млекопитающих, амфибий, насекомых и практически отсутствует информация об АМП червей, в том числе медицинских пиявок, хотя они представляют собой источник перспективных противомикробных агентов.

1.5. Медицинская пиявка H. medicinalis

В качестве объекта исследования была выбрана медицинская пиявка H. medicinalis, кровососущий хищник, который используется в медицинских целях с давних времён. Пиявка относится к типу кольчатых червей (Annelida) и является родственником дождевых червей и полихет. В медицине используются животные, принадлежащие семейству Hirudinidae, panee называвшегося Gnathobdellidae, «пиявки с челюстями». В отверстии рта червя расположены челюсти, которые содержат ряд обызвествлённых зубов, позволяющих червю прорезать поверхность тела потенциальной жертвы и отбирать кровь из раны [146]. Пиявки, способные питаться кровью млекопитающих, относятся к родам Hirudo, Macrobdella и Haemadipsa. Род Hirudo подразделён на несколько видов [147,148]. Биология видов в основном одинакова, возможно, с некоторыми различиями в температурных предпочтениях. Животные живут в прудах с пресной водой и медленно текущими ручьями [149]. Центральноевропейская медицинская пиявка, *H. medicinalis*, применяется на практике отечественными и зарубежными врачами [150–152]. В России пиявка занесена в Государственный реестр лекарственных средств (Лекарственная форма: пиявки медицинские, Фармако-терапевтическая группа: действия, Анатомо-терапевтическая антикоагулянтное средство прямого химическая классификация: Группа гепарина). В США медицинская пиявка в 2004 году была зарегистрирована как «медицинское изделие» (англ. medical device), использующееся в реконструктивной хирургии для предотвращения отторжения трансплантируемых тканей [153].

Главная особенность медицинских пиявок, которая привлекает многих исследователей, заключается в их эволюционном приспособлении к питанию кровью животных и человека. Такой способ питания создаёт определённые трудности, поскольку хищнику необходимо подавлять нормальные реакции жертвы при укусе, а именно снижать отёк, воспаление и болевые ощущения, чтобы оставаться незамеченным [146]. В ходе эволюции кровососущие пиявки научились регулировать гемостаз жертвы биологически активными веществами собственного секрета. Стоит отметить, что у челюстных пиявок, в отличии от хоботковых, клетки вырабатывающие слюнной секрет не образуют отдельный орган, они расположены диффузно в мышечной и соединительной ткани, что существенно затрудняет исследование данного типа клеток. Известно также, что в слюнном секрете пиявки содержатся вещества, способствующие поддержанию крови жертвы в интактном состоянии в течение длительного времени [146]. Кроме того, в ССК находятся вещества, необходимые для защиты организма от экзогенных патогенов, попадающих внутрь пиявки вместе с кровью во время кормления, а также для контроля микрофлоры желудочно-кишечного тракта, где обеспечивается долговременное хранение набранной крови. Мы предположили, что описанные процессы осуществляются с использованием пептидов, которые, помимо антибактериальной активности, должны быть низкотоксичными для клеток крови для предотвращения чрезмерного гемолиза.

На данный момент не было проведено масштабного системного исследования медицинской пиявки с использованием высокопроизводительного секвенирования и анализа протеома ССК. В специализированных базах данных находятся транскриптомы нескольких пиявок [1–7]. Для пресноводной пиявки *Helobdella robusta* была получена лишь нуклеотидная последовательности генома, без какой-либо аннотации данных [10]. Также, исследователи углублённо изучали микробиоту медицинских пиявок *H. verbana* [11-13]. В 2016 году был проведён сравнительный анализ нуклеотидных последовательностей митогеномов (митохондриальных геномов) трёх кровососущих пиявок Haementeria officinalis, Placobdella lamothei и Placobdella parasitica [8]. В этом же году, нами впервые была получена нуклеотидная последовательность митохондриального генома *H. verbana* и *H. medicinalis* [9]. Таким образом, медицинская пиявка *H. medicinalis* представляется перспективным объектом исследования для идентификации новых АМП, потенциальных противомикробных препаратов с низкой токсичностью.

1.5.1. Секрет слюнных клеток *H. medicinalis*

Секрет слюнных клеток (ССК) медицинской пиявки содержит сложную смесь соединений различной химической природы, которые выполняют функции, связанные исключительно со способом и типом питания медицинских пиявок [154,155]. Основные активности, проявляемые компонентами ССК: противотромботическое, антикоагуляционное, обезболивающее и противовоспалительное действия, деградация внеклеточного матрикса, И антимикробный эффект [149,156]. ССК ускорение кровотока Также оказывает рефлексогенное, противоишемическое, антигипоксическое, гипотензивное, иммуностимулирующее действия и многие другие. Исследования нативного секрета Н. medicinalis показали, что секрет не оказывает токсического действия и не вызывает гемолиз человеческих эритроцитов. Некоторые из этих биологически активных компонентов были изучены и охарактеризованы на молекулярном и биохимическом уровнях (Таблица 2).

До сих пор в секрете пиявки не было идентифицировано соединений, обладающих анальгезирующим свойством [157]. Тем не менее, показано, что укусы пиявки гораздо менее болезненны, чем раны подобного рода. Таким образом, в ССК пиявки содержатся компоненты, обеспечивающие анальгезирующее лействие. Известно, ЧТО тканевые медиаторы, сигнализирующие организму 0 воспалительных процессах, являются сильными сенсибилизаторами, которые способны вызывать болевые ощущения [158,159]. Вероятно, обезболивающий эффект является опосредованным процессом. Антистазин, предполагаемый анальгетик, ингибирует тканевые калликреины, протеиназы, расщепляющие неактивные формы кининогенов до биологически активных кининов [160]. Таким образом, предполагается, что в секрете пиявки содержатся вещества подавляющие действие кининов. Провоспалительным агентом является эглин С, ингибитор сериновых протеиназ, который нарушает хемотаксис нейтрофилов и запускает защитные реакции, которые способствуют воспалению [157,161,162]. Специфический ингибитор триптазы (англ. leech-derived LDTI) tryptase inhibitor, предположительно подавляет воспалительные реакции, вызванные тучными клетками [163,164]. Также высказывалась гипотеза, что компонент ССК, ингибитор компонента комплемента С1, обладает противовоспалительным эффектом [164,165].

Антикоагулянтные и тромболитические компоненты секрета слюнных клеток *H. medicinalis*

Антикоагулянты предотвращают процесс свёртывания крови, что способствует свободному току крови из повреждённых сосудов. ССК медицинской пиявки содержит вещества, влияющие на все основные звенья гемостаза [149]. Агрегация тромбоцитов в месте повреждения сосуда является первой стадией механизма активации системы свёртывания крови. Ингибитором адгезии тромбоцитов к стенкам повреждённых сосудов является саратин, который связывается с экспонированным фибриллярным коллагеном и конкурентным образом ингибирует активацию фактора фон Виллебранда, гликопротеина плазмы крови [167,168]. Механизм действия саратина локальный и не вызывает побочных эффектов. Другой компонент ССК, калин, также связывается с коллагеном I и предотвращает его связывание с фактором фон Виллебранда [169,170]. Агрегация тромбоцитов в месте повреждения сосуда ускоряется за счёт автоматической активации тромбоцитов, что приводит к высвобождению аденозин-5'-дифосфата (АДФ) из внутриклеточных везикул. Локальное повышение концентрации АДФ приводит к увеличению сродства к фактору фон Виллебранда, что в свою очередь стимулирует образование агрегатов тромбоцитов, связанных с коллагеном. Специфический фермент, апираза, расщепляет АДФ, тем самым ингибируя активацию тромбоцитов [157].

Запуск протеазного каскада приводит к активации тромбина, сериновой протеиназы, которая играет ключевую роль в процессе свёртывания крови. Тромбин модифицирует протеазоактивируемые рецепторы на тромбоцитах, что приводит к их активации [171]. Также тромбин расщепляет фибриноген, образуя короткие молекулы фибрина, которые впоследствии полимеризуются и образуют нерастворимую сетку. При привлечении дополнительных факторов из плазмы крови в месте повреждения сосуда образуется массивный сгусток нерастворимого фибрина с агрегированными тромбоцитами — тромб. В дальнейшем происходит прогрессирующий рост тромба и окклюзия кровеносного сосуда [172]. В ССК

29

пиявки был обнаружен белок гирудин, который является высокоспецифичным ингибитором тромбина [173]. Гирудин, благодаря высокому сродству к протеазе, блокирует все известные реакции, активатором которых выступает тромбин [174–176]. Рекомбинантный белок используется в клинической практике, так как обладает рядом преимуществ по сравнению с остальными используемыми антикоагулянтами [177].



Рисунок 6 - Первичная и пространственная структура дестабилазы

Участки первичной структуры, принимающие α-спиральную конформацию, обозначены чёрным прямоугольником, остатки цистеинов, образуюшие дисульфидные связи, выделены сплошными линиями. На пространственных структурах обозначены дисульфидные связи (жёлтым), α-спирали (красным). Иллюстрация разработана автором с использованием графической программы РуМОL (версия 1.2.2.3) [67] на основе данных из [178].

Очень интересен полифункциональный фермент, обнаруженный в ССК медицинской пиявки — дестабилаза (Рисунок 6) [179]. Помимо лизоцимной, мурамидазной и антимикробной активностей белок обладает фибринолитическим действием [179]. Растворение фибрина под действием дестабилазы осуществляется за счёт эндо-є-(γ-Glu)-Lys-изопептидазной активности, которая опосредует гидролиз изопептидных связей между γ-карбоксамидом глутаминового остатка и є-аминогруппой остатка лизина стабилизированного фибрина, что приводит к расщеплению Д-димеров (продукт распада тромба) до мономеров [180]. Также было показано, что дестабилаза обладает противотромбической активностью [181]. Показано, что фермент ингибирует спонтанную и индуцированную агрегацию тромбоцитов.

Дестабилаза в растворе формирует прочный белково-липидный комплекс совместно с

гирудином, простагландиновым компонентом и ингибитором калликреина плазмы крови [165,182]. «Дестабилазный комплекс» обеспечивает стабилизацию входящих в его состав компонентов, что сохраняет их биологическую функцию. В зависимости от растворителя комплекс способен изменять пространственную конформацию и легко проникает через мембрану клетки [182]. Комплекс проявляет противотромботическое, тромболитическое и гипотензивное действие.

Таблица 2 - Белковые компоненты ССК пиявки и их предполагаемая функция

Название	Молекулярный вес	Функция	Ссылки
Антистазин	15 кДа	Ингибирует тканевые калликреины, фактор Ха	[186]
Апираза	45 кДа	Расщепляет АДФ, ингибирует активацию тромбоцитов	[157]
Бделлины	8,1 кДа	Ингибируют трипсин, плазмин и акрозин	[187]
Гементерин	3,5 Да	Активирует плазминоген	[188]
Гементин	4,3 Да	Разрушает фибрин и фибриноген	[189]
Гиалуронидаза	27,5 кДа	Расщепляет гиалуроновую кислоту внеклеточного матрикса	[190]
Гирудин	7,1 кДа	Ингибирует тромбин	[175]
Гирустазин	5,9 Да	Ингибирует катепсин G, тканевой калликреин, трипсин и химотрипсин	[160]
Дестабилаза	12,6-12,9 кДа	Расщепляет изопептидные связи в стабилизированном фибрине, обладает лизоцимной,	[179]
Ингибитор компонента С1 системы комплемента	60–70 кДа	Ингибирует адгезию и агрегацию тромбоцитов, а также активацию фактора фон Виллебранда	[165,166]
Калин	65 кДа	Ингибирует адгезию тромбоцитов, блокирует образование богатых тромбоцитами тромбов и связывание фактора фон Виллебранда с поверхностью коллагена	[169]
Коллагеназа	~ 100 кДа	Ингибирует агрегацию тромбоцитов, разрушает коллаген внеклеточного матрикса	[157]
Саратин	12 кДа	Ингибирует адгезию и агрегацию тромбоцитов к повреждённой поверхности сосуда	[168]
Пиявочный ингибитор триптаз	4,3-4,8 кДа	Ингибирует триптазу	[163]
Пиявочный ингибитор карбоксипептидаз	7,2-7,3 кДа	Ингибирует карбоксипептидазу В, опосредует фибринолиз	[184]
Эглин С	8,1 кДа	Ингибирует нейтрофильную эластазу и катепсин G	[185]
Эглины	5-38 кДа	Ингибирует химотрипсин, субтилизин и нейтральные протеиназы гранулоцитов человека	[157]

Антимикробный потенциал секрета слюнных клеток H. medicinalis

ССК *H. medicinalis*, как эволюционно сбалансированный комплекс природных антимикробных соединений, вызывает особый интерес. Ранее была определена антимикробная активность ССК [183]. Было показано, что ССК проявляет антимикробное действие по отношению к грамотрицательным бактериям *E. coli* и *Pseudomonas aeruginosa*, а также оказывает антимикотическое действие на условно-патогенные грибы *Candida guillermondii*. В итоге было предложено использование ССК медицинской пиявки вместо традиционно используемых антибиотиков при лечении микозов у пациентов с ослабленным иммунитетом, нарушениями функционирования сердечно-сосудистой системы, печени, почек, и ряда других заболеваний [183].

1.5.2. Антимикробные пептиды медицинских пиявок

На данный момент охарактеризовано семь антимикробных пептидов, обнаруженных в разных пиявках (Таблица 3). В 2004 году из пиявки *Theromyzon tessulatum* был выделен первый АМП - Тt-теромацин [191]. Позже были обнаружены другие пиявочные представители семейства мацинов, катионных цистеин-богатых АМП, а именно гирудомацин, выделенный из *H. nipponica* Whitman, Hm-нейромацин и Hm-теромацин, выделенные из *H. medicinalis* [191–193].

Мацины представляют собой довольно сложные пептиды, длина этих АМП составляет более 60 аминокислотных остатков. Аминокислотная последовательность мацинов высоко консервативна благодаря наличию сигнального пептида (за исключением Hm-теромацина), четырёх дисульфидных мостиков и пятой внутримолекулярной дисульфидной связи (C31: C73) в теромацинах (Рисунок 7) [194,195]. Пространственная структура АМП состоит из цистеинового узла (cystine knottin), сформированного цистеиновыми мостиками, дополнительной α-спирали на N-конце и двух длинных гибких петель. Теромацины обладают более длинным C-концевым доменом, чем Hm-нейромацин. Поэтому конформации этих двух пептидов отличаются, что и определяет различия в биологических функциях [196].

Согласно проведённым исследованиям Нт-теромацин и Нт-нейромацин проявляют бактерицидную активность в отношении *Bacillus megaterium* и *Micrococcus luteus*, а также обладают низким антибактериальным эффектом в отношении *Escherichia coli* [191,196]. Для Нт-нейромацина показано, что он активен в отношении микрококка *M. nishinomiyaensis* [192]. Гирудомацин лизирует как грамположительные (*Staphylococcus aureus*, *B. subtilis*, *Listeria monocytogenes*), так и грамотрицательные бактерии (*E. coli*) [197]. Механизм действия мацинов реализуется по «модели ракушек» (англ. *barnacle model*), отличающейся от канонических механизмов действия [196]. Благодаря электростатическому и гидрофобному взаимодействию

пептид-мембрана, мацины воздействуют одновременно на две отдельные бактериальные клетки. Таким образом, каждая бактериальная клетка через пептид «прилипает» к нескольким другим, что приводит к образованию клеточных агрегатов. Ввиду своей амфипатичности, мацины повышают проницаемость мембраны грамположительных бактерий, что позволяет им самим проникают внутрь клетки и убивать патоген [194,196]. Исключительно для Hm-нейромацина было показано, что он может проявлять мембранолитическое действие за счёт формирования пор [196]. Кроме того, показано, что порообразующая активность не ингибируется высокой концентрацией солей и зависит от pH [198]. Так как солеустойчивость Hm-нейромацина зависела от pH, она, по-видимому, опосредована де-/протонированием остатков гистидина, которые отсутствуют в Hm-теромацине. Кроме того, Hm-нейромацин этого не делает. Более того, Hm-теромацин не достаточно амфипатичен, чтобы действовать по «модели ракушек».



Рисунок 7 - Консервативные участки мацинов

(А) Первичная и пространственная структуры Тt-теромацина (PDB ID: 2LN8). Участки первичной структуры, принимающие α-спиральную конформацию обозначены чёрным прямоугольником, β-складчатую структуру — светлыми стрелками. На пространственной структуре обозначены дисульфидные связи (жёлтым), α-спирали (красным) и β-тяжи (синим). (Б) Выравнивание аминокислотных последовательностей пиявочных мацинов: TtTh - Tt-теромацин, HmTh — Hm-теромацин, Hmc — гирудомацин и HmN — Hm-нейромацин. Консервативные участки выделены серым, остатки цистеинов, формирующие дисульфидные связи, выделены жёлтым. Дисульфидные связи между соответствующими остатками цистеина, характерные для всех мацинов, обозначены сплошными линиями. Пунктирной линией отмечены два остатка цистеина в теромацинах, образующих дополнительную пятую дисульфидную связь. Иллюстрация разработана автором с использованием графической программы РуМОL (версия 1.2.2.3) [67].

Для всех пиявочных мацинов было показано, что в дополнение к их бактерицидной активности, они проявляют нейротрофическое и пролиферативное действие [192,196]. Процесс восстановления нейронов не связан с регенерацией целых нейронов *de novo*. Пептиды опосредуют миграцию микроглии к месту поражения нерва и способствуют росту аксона [199]. Нт-теромацин вырабатывается циркулирующими клетками крови пиявки и затем высвобождается в плазму, тогда как Hm-нейромацин вырабатывается микроглиальными, нервными клетками и эпителиальными клетками кишечника [191,192]. Для обоих пептидов показано, что они накапливаются в раневом участке повреждённого нерва пиявки. Тт-теромацин экспрессируется в жировых клетках и высвобождается в целомическую жидкость в ответ на инфекцию или повреждение нерва [191,192]. Для гирудомацина показано, что его экспрессия повышается в кишечнике и слюнных клетках пиявки при кормлении [197].

Помимо теромацина у *T. tessulatum* был найден анионный Tt-теромизин [191]. Ttтеромизин является первым анионным АМП, обнаруженным у беспозвоночных, и проявляет бактериостатическую активность в отношении грамположительных бактерий и не влияет на жизнеспособность грамотрицательных бактерий [200]. Tt-теромизин, как и Tt-теромацин процессируется из белка-предшественника, после отщепления сигнального пептида, за которым непосредственно следует активный пептид.

H. medicinalis экспрессирует Нт-люмбрицин, пролин-богатый пептид, помимо антибактериальной активности обладает нейрорегенеративной активностью [192]. Пространственная структура пептида на данный момент не получена. Было показано, что Нт-люмбрицин, как и Нт-нейромацин продуцируется клетками микроглии и самими нейронами в ответ на повреждение центральной нервной системы, накапливается в поврежденных участках нервов. Сообщается, что аналог люмбрицина был найден у *T. tessulatum*, анионный пептид Tt-люмбрицин, но более детальная информация о пептиде отсутствует.

Как было сказано выше, антимикробной активностью обладает многофункциональный фермент — дестабилаза [201]. Помимо фибринолитической активности, белок демонстрирует антимикробное действие, сходное по активности с лизоцимом. Показано, что не только структурированный белок проявляет бактерицидное действие, но и в денатурированном состоянии он вызывает гибель бактериальных клеток [201,202]. Для дестабилазы показано, что фермент обладает противомикробным действием в отношении грамположительных *M. lysodeikticus* и грамотрицательных *E. coli* бактерий, грибов *Botrytiscinerea*, дрожжей *Schizosaccharomyces pombe* и архей *Methanosarcina barkeri*.

Таким образом, АМП играют сложную регуляторную роль в защитных процессах организма. Широкий спектр природных молекул АМП представляет собой основу для

оптимизации и разработки новых антибактериальных, антигрибковых и антивирусных инфекций терапевтических соединений лечения препаратов, для паразитарных И существует большое разнообразие онкологических заболеваний. На данный момент вычислительных методов, которые могут помочь создать лекарственные препараты на основе АМП. Важно корректно подобрать алгоритм анализа данных в соответствии с задачей и спецификой изучаемого объекта. Для более точной идентификации АМП необходимо использовать набор подходов вычислительной и синтетической биологии, что и было выполнено в данной работе.

Настоящая работа посвящена всестороннему исследованию медицинской пиявки *H. medicinalis*, с целью поиска новых антимикробных пептидов. Несмотря на то, что сейчас опубликованы транскриптомы нескольких пиявок, полученные данные разрознены. В ходе выполнения данной работы была определена нуклеотидная последовательность генома *H. medicinalis*, проведена аннотация генома *H. medicinalis*, выполнен анализ белкового состава ССК *H. medicinalis*. Полученные данные позволили идентифицировать новые антимикробные пептиды, обладающие широким спектром антимикробного действия и нетоксичные для клеток млекопитающих. Кроме того, обширный системный анализ *H. medicinalis*, выполненный в ходе данного исследования, может стать основой для более детального поиска и описания новых биологически активных соединений ССК *H. medicinalis*.
Таблица 3 - Краткая характеристика АМП пиявок

Семейство	АМП	Источник	Размер	Активность ²	Функции	Ссылки
Мацины	Тt-теромацин	T. tessulatum	97 а.а., цистеин-богатый	G+, G-	Дестабилизирует мембрану грамположительных бактерий, стимулирует регенерацию нервной ткани	[187]
	Гирудомацин	H. nipponia	82 а.а., цистеин-богатый	G+, G-	Дестабилизирует мембрану грамположительных бактерий	[191]
	Hm-теромацин	H. medicinalis	75 а.а., цистеин-богатый	G+	Дестабилизирует мембрану грамположительных бактерий, стимулирует регенерацию нервной ткани	
	Hm-нейромацин	H. medicinalis	82 а.а., цистеин-богатый	G+	Дестабилизирует мембрану грамположительных бактерий, вызывает агглютинацию бактерий и липосом, стимулирует регенерацию нервной ткани	[196]
-	Теромизин	T. tessulatum	107 а.а., линейный, анионный	G+	-	[196]
Люмбрицины	Тt-люмбрицин	T. tessulatum	57 а.а., линейный, пролин-богатый	G+, G-, F	-	[191]
	Hm-люмбрицин	H. medicinalis	55 a.a., линейный, пролин-богатый	G+, G-, F	Стимулирует регенерацию нервной ткани	[191]
Лизоцимы	Дестабилаза	H. medicinalis	136 a.a.	G+, G-, F	-	[192]

² – обозначения: G+ — грамположительные бактерии, G- — грамотрицательные бактерии, F — грибы.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

2.1. Материалы

Реактивы и расходные материалы, используемые в работе: 1,2-диолеоил-глицеро-3фосфоглицерол (Sigma-Aldrich, США), 1-пальмитоил-2-олеоил-глицеро-3-фосфо-L-серин (Sigma-Aldrich, США), 1-пальмитоил-2-олеоил-глицеро-3-фосфохолин (Sigma-Aldrich, США), 3,6-диокса-1,8-октандитиол (Sigma-Aldrich, CША), 4-метилпиперидин (Sigma-Aldrich, CША), 9флуоренилметилоксикарбонил (Fmoc)-защищенные по аминогруппам производные аминокислот (Sigma-Aldrich, CША), N,N'-диизопропилкарбодиимид (Sigma-Aldrich, США), N,N-диметилформамид (Sigma-Aldrich, США), аргинин (Sigma-Aldrich, США), ацетонитрил (Sigma-Aldrich, США), бикарбонат аммония (Sigma-Aldrich, США), бромид этидия (Sigmaгексадецилтриметиламмонийбромид Aldrich, США), (Sigma-Aldrich, США), гексаметилдисилазан (Reachem, Россия), гентамицин (Sigma-Aldrich, CША), глутаральдегид (Fisher BioReagents, США), глюкоза (VWR Life Science AMRESCO, США), дезоксихолат натрия (Sigma-Aldrich, США), дейтерированные додецилфосфатидилхолиновые мицеллы (d-38, 98%; Cambridge Isotope Laboratories США), дитиотреитол (BioRad, США), диэтиловый эфир (Sigma-Aldrich, США), додецилсульфат натрия (Panreac, США), изопропанол (Sigma-Aldrich, США), йодацетамид (BioRad, CША), карбоксифлуоресцеин (Sigma-Aldrich, США), мелиттин (Sigma-Aldrich, США), мочевина (Sigma-Aldrich, США), муравьиная кислота (Sigma-Aldrich, США), параформальдегид (Sigma-Aldrich, США), поверхностно-активного вещества RapiGest SF (Waters, CIIIA), среда DMEM (ThermoScientific, CIIIA), среда MEM (ThermoScientific, CIIIA), среда МНВ (Difco, CША), триизопропилсилан (Sigma-Aldrich, CША), TrisHCl (VWR Life Science AMRESCO, США), трифторуксусная кислота (Sigma-Aldrich, США), урографин (Sigma-Aldrich, США), фенол (Sigma-Aldrich, США), хлороформ (Sigma-Aldrich, США), ЭДТА (AppliChem, США), эмбриональная бычья сыворотка (HyClone, США), этанол (Sigma-Aldrich, США), этил (гидроксиимино) цианоацетат (Sigma-Aldrich, США), HBSS (ThermoScientific, США), L-глутамин (Panreac, CША), NaCl (VWR Life Science AMRESCO, CША), Triton X-100 (Sigma-Aldrich, CIIIA).

Антитела и ферменты: антитела, специфичные к липоолигосахаридам хламидий (ХламиСлайд—рекомбинантный, ООО «Галарт-Диагностикум», Россия), протеиназа К (Amresco, США), трипсин (Promega, США), смесь нуклеаз (GE Healthcare, США).

Бактериальные штаммы, клеточные линии и условия культивирования: бактериальные штаммы *E. coli* K-12 substr. MG1655 (ATCC[®] 700926TM), *B. subtilis* 168HT (ATCC[®] 23857TM) выращивали в среде MHB при 37°C в термостате. Клетки линии мышиных

фибробластов McCoy (ATCC[®] CRL-1696TM) растили в CO₂-инкубаторе при 37°C, 5% CO₂ в среде DMEM с добавлением 10% эмбриональной бычьей сыворотки, 1 г/л глюкозы, 2 мM L-глутамина.

Штамм *Chlamydia thrachomatis* D/UW-3/Cx (ATCC[®] VR-885TM) культивировали в линии клеток McCoy в среде Игла (MEM) с добавлением 0,5% глюкозы, 10% эмбриональной бычьей сыворотки. Заражение клеток McCoy инокулятом элементарных телец включения *C. thrachomatis* проводили центрифугированием в течение 1 ч при 900 g, затем клетки McCoy инкубировали при 37°C в течение 2 ч. Далее промывали клетки средой MEM, добавляли среду для роста и инкубировали клетки при 37°C на 48 ч. Элементарные тельца *C. trachomatis* очищали с помощью ультрацентрифугирования в градиенте плотности урографина, согласно стандартной методике [203].

Биологический материал, используемый в работе: медицинские пиявки *H. medicinalis* были предоставлены биофабрикой ООО НВФ «Гируд И.Н.» (Россия).

Компьютерные приложения и программы: сценарий выполнения встроенной программы matePairingSplitReads.py (Thermo Fisher Scientific, CША), ADAM (National Taiwan Ocean University, Taйвань), AUGUSTUS 125 (версия 3.7.1) (Institute of Mathematics and Computer Science, Германия), bcl2fastq v2.17.1.14 Conversion Software (Illumina, CША), BLAT (версия 34x13) (National Institutes of Health Library, CША), ImageJ (версия 1.48; RSB) (National Institutes of Health, CША), MEGAN6 (Center for Bioinformatics, Германия), NxTrim (Illumina, CША), SPAdes 3.6.09 (https://cab.spbu.ru/software/spades/), TOPSPIN (Bruker Biospin, CША), Trinity (версия r20131110) (Broad Institute of MIT and Harvard, CША; CSIRO Ecosystem Sciences, Black Mountain Labs, ABстралия), BEDtools (University of Virginia, CША), Bowtie 2 (University of Maryland, CША), EMBOSS (MRC UK HGMP Resource Centre, Великобритания), Peptides (https://github.com/dosorio/Peptides/), protr (Central South University, Китай), seqinr (http://seqinr.r-forge.r-project.org/), CARA (Institute of Molecular Biology & Biophysics, Германия), CYANA (Ueno Taito Tokyo, Япония), PyMOL (SBGrid Consortium), TALOS-N (National Institutes of Health, CША), UCSF Chimera (University of California, CША).

Онлайн-сервисы и электронные базы данных: APD3 (National Institute of Allergy and Infectious Diseases, CША), CAMPR3 (Biomedical Informatics Centre, NIRRH, Индия), DADP (<u>http://split4.pmfst.hr/dadp/</u>), UniProtKB, ADAM (National Taiwan Ocean University, Taйвань), AGGRESCAN (Universitat Autónoma de Barcelona, Испания), AMPA (Centre for Genomic Regulation (CRG) of Barcelona, Испания), I-TASSER-MR (Huazhong University of Science and Technology, Китай; University of Michigan, CША), PEP-FOLD3 (Université Paris Diderot, Франция), TANGO (Heidelberg, Германия; Vrije Universiteit Brussel, Бельгия).

2.2. Методы

2.2.1. Генотипирование медицинских пиявок

Медицинские пиявки *Н. medicinalis* (ГИРУД И.Н., г. Балаково, Саратовская область, Россия) выдерживали при 12-14°С в течение 2 месяцев до начала работ. Подтверждение видовой принадлежности пиявок проводили путём ПЦР-амплификации и последующего секвенирования участков генов, кодирующих ядерную и митохондриальную рибосомную РНК. Для амплификации участка гена 12S-митохондриальной рРНК использовали олигонуклеотиды 12A и 12B, для амплификации внутреннего транскрибируемого участка-спейсера ITS2 ядерной рРНК - праймеры IS-A и IS-B (Таблица 4). Амплифицированные фрагменты ДНК секвенировали, используя те же олигонуклеотиды, и сравнивали с последовательностью, содержащейся в базе данных GenBank: *Н. medicinalis* - АҮ763166 (ITS2) и АҮ763156 (12S). Нуклеотидную последовательность ампликонов определяли по методу по Сэнгера с использованием коммерческого набора Big DyeTM Terminator v.3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, США) и генетического анализатора ABI Prism 3730XL (Applied Biosystem, США), согласно инструкции производителя.

1 17						
Название	Структура олигонуклеотида 5 -3					
12A	GCCAGCAGCCGCGGTTA					
12B	CCTACTTTGTTACGACTTAT					
IS-A	TGGGTCGATGAAGAGCGCAG					
IS-B	GTAATCCGGTCTGATCTCAG					

Таблица 4 - Структура олигонуклеотидов, используемых в работе

2.2.2. Выделение геномной ДНК *H. medicinalis*

Геномную ДНК выделяли из взрослой особи *H. medicinalis* или новорожденных пиявок. Предварительно расчлененное тело одной пиявки (или 15-20 новорожденных пиявок) помещали в пробирку объемом 50 мл и добавляли 10 мл экстракционного буфера, содержащего 100 мМ NaCl, 10 мМ TrisHCl, pH 8, 25 мМ ЭДТА pH 8, 0,5% додецилсульфат натрия (SDS), 0,3 мг/мл протеиназы К. Образец инкубировали при 37°С в течение 18 ч, переворачивая пробирку три или четыре раза в течение первых 2 ч. Затем добавляли 10 мл фенола и перемешивали, осторожно переворачивая пробирку в течение 10 мин. После центрифугировали в течение 15 мин при 3500 g. Водную фазу, содержащую ДНК, экстрагировали второй раз 10 мл смесью фенол:хлороформ в объёмном соотношении 1:1. Затем водную фазу переносили в новую пробирку и добавляли равный объем охлаждённого (-20°С) 98%-ого этанола. ДНК осаждали в течение 20 мин при -80°С, центрифугировали в течение 20 мин при 3500 g. Супернатант

отбирали, а осадок промывали охлаждённым 70%-ым этанолом. Этанол отбирали, осадок высушивали и затем ресуспендировали в 1 мл ТЕ-буфера (10 мМ TrisHCl, pH 8, 1 мМ ЭДТА pH 8).

2.2.3. Очистка геномной ДНК *H. medicinalis*

К раствору ДНК добавляли 1 объем прогретого до 65°С буфера, содержащего 2% гексадецилтриметиламмонийбромид (ГТАБ), 100 мМ TrisHCl, pH 8, 20 мМ ЭДТА pH 8, 1,4 М NaCl. Перемешивали раствор путём переворачивания пробирки и инкубировали при 65°С в течение 10 мин. К раствору добавляли равный объем хлороформа, перемешивали на вортексе. Для разделения фаз центрифугировали раствор в течение 5 мин при 3500 g. Отбирали супернатант и добавляли 1,1 объем прогретого до 65°С буфера, содержащего 1% ГТАБ, 50 мМ TrisHCl, pH 8, 10 мМ ЭДТА pH 8). Тщательно перемешивали раствор путём переворачивания пробирки. Центрифугировали раствор в течение 5 мин при 3500 g. Отбирали надосадок, растворяли геномную ДНК в 0,5 мл буфера, содержащего 10 мМ TrisHCl, pH 8, 0,1 мМ ЭДТА, 1 М NaCl, при 65°С в течение 30 мин. К раствору добавляли 0,6 объема изопропанола. Тщательно перемешивали раствор в течение 8 мин при 3500 g. Супернатант отбирали и осадок промывали 0,5 мл 70%-ого этанола. Геномную ДНК ресуспендировали в 200 мкл буфера ТЕ.

Ионное полупроводниковое секвенирование на платформе Ion Proton

Очищенную геномную ДНК (приблизительно 1000 нг) фрагментировали до 200-300 п.н. с использованием системы Covaris S220 (Covaris, США). Для лигирования вспомогательных ДНК-адаптеров к фрагментам ДНК использовали набор Ion Xpress™ Plus Fragment Library Kit (Thermo Fisher Scientific, США). Подготовку магнитных сфер осуществляли с использованием набора One Touch 2 and Template Kit v2 (Thermo Fisher Scientific, США) согласно методике производителя. Затем проводили эмульсионную ПЦР с использованием системы Ion OneTouchTM System (Thermo Fisher Scientific, США). Считывание нуклеотидной последовательности подготовленной ДНК-библиотеки выполняли с использованием набора Ion Proton[™] 200 Sequencing Kit v2 (Thermo Fisher Scientific, США) и чипа P1 Ion (Thermo Fisher Scientific, США) на полупроводниковом секвенаторе Ion Proton[™] System (Thermo Fisher Scientific, CIIIA).

Ионное полупроводниковое секвенирование на платформе Ion Torrent

Парная ДНК-библиотека (Ion TrueMate), содержащая ДНК-фрагменты размером 3-6 т.п.н., получали из очищенной геномной ДНК с использованием набора Ion TrueMate Library Reagents (Thermo Fisher Scientific, США). Для проведения эмульсионной ПЦР использовали набор Ion PGMTM template OT2 400 kit (Thermo Fisher Scientific, США). Секвенирование

выполняли на полупроводниковом секвенаторе Ion Torrent PGM (Thermo Fisher Scientific, США) с использованием чипа Ion 318 (Thermo Fisher Scientific, США) и набора Ion PGMTM Sequencing 400 Kit v2 (Thermo Fisher Scientific, США) в соответствии с рекомендациями производителя. Для разделения неспаренных прочтений от истинных парных прочтений, использовали сценарий выполнения встроенной программы matePairingSplitReads.py [204].

Секвенирование нового поколения на платформе Illumina

Парную ДНК-библиотеку (Illumina TruSeq), содержащую ДНК-фрагменты размером 8-12 т.п.н., получали из очищенной геномной ДНК с использованием наборов Nextera Mate Pair Library Prep Kit (Illumina, США) и TruSeq DNA Sample Prep Kit (Illumina, США). Секвенирование выполняли на секвенаторе MiSeq platform (Illumina, США) с использованием набора MiSeq Reagent Kit v2 (Illumina, США) в соответствии со стандартными протоколами секвенирования Illumina. Демультиплексирование ДНК-фрагментов выполняли с использованием программного обеспечения bcl2fastq v2.17.1.14 Conversion Software (Illumina, США), что позволило удалить последовательности вспомогательных ДНК-адаптеров. Для разделения одно-концевых, парно-концевых и парных прочтений использовали программное обеспечение NxTrim [205].

2.2.4. Сборка генома *H. medicinalis*

Наборы прочтений, соответствующих трём ДНК-библиотекам, объединяли для создания единой сборки генома с использованием программного обеспечения SPAdes 3.6.09 [206]. Для скаффолдинга эукариотических контигов использовали программное обеспечение Sspace с параметрами -р 1 -х 0 -1 library.txt -s Contigs.fasta -k 2 [207]. Параметры качества и длины определяли согласно сборщику SPAdes [208]. **GC-состав** контигов нуклеотидных последовательностей рассчитывали с помощью инструмента infoseq, встроенного в программный пакет EMBOSS [208]. Контиги длиной менее 500 п.н. исключали из сборки генома. Наличие тетрануклеотидов определяли с использованием сценария выполнения встроенной calc.kmerfreq.pl, который доступен программы по ссылке [https://github.com/MadsAlbertsen/miscperlscripts/blob/master/calc.kmerfreq.pl].

Прочтения каждой из ДНК-библиотек картировали на последовательности сборки генома с использованием программного пакета Bowtie 2 [209]. Глубину покрытия ДНК фрагментов для каждой ДНК библиотеки рассчитывали с использованием программного пакета BEDtools [210]. Алгоритм BLASTN и базу данных «nr» использовали для определения таксономической принадлежности контигов с применением метода «ближайшего соседа» (англ. *lowest common ancestor, LCA*), реализованного в программе MEGAN6 [211].

2.2.5. Аннотация генома H. medicinalis

Для аннотации аннотация генома *H. medicinalis* разработали три шаблона, на основании которых определяли участки нуклеотидной последовательности генома, принадлежащих экзону, интрону и т. д. В качестве первого шаблона был выбран набор последовательностей генов, кодирующих белки пиявки H. robusta (GenBank acc. Num. AMQM01000000, BioProject acc. num. PRJNA175704) [10]. Второй шаблон генерировали с использованием контигов, соответствующих de novo сборке транскриптома H. medicinalis, собранного с использованием программного обеспечения Trinity (версия r20131110) с параметрами по умолчанию [212]. Третий шаблон был создан на основе прочтений кДНК *H. medicinalis* с использованием стандартного вычислительного конвейера, полученные нами в ходе исследования [215]. Референсные последовательности для каждого шаблона картировали на последовательности скаффолдов сборки генома H. medicinalis с использованием программного обеспечения BLAT (версия 34х13) в соответствии с параметрами разработчика [213]. Затем все три шаблона объединяли и преобразовывали в алгоритмы для использования в программном обеспечении AUGUSTUS 125 (версия 3.7.1) с помощью прилагаемых сценариев выполнения встроенной программы [214]. Аннотирование генов *H. medicinalis* выполняли в соответствии с параметрами «--genemodel = complete --gff3 = on --species = fly».

Работа по типированию пиявок, выделению и очистке геномной ДНК, пробоподготовке, секвенированию и анализу ДНК библиотек медицинской пиявки осуществлялась в сотрудничестве с лабораторией геномных исследований и вычислительной биологии ФГБУ ФНКЦ ФХМ ФМБА России.

2.2.6. Выделение секрета слюнных клеток H. medicinalis

Для сбора ССК медицинской пиявки дно пробирки объёмом 15 мл отрезали, делали отверстие сбоку в дне пробирки. На вырезанном конце растягивали лабораторную ленту «Parafilm M» (Bemis Company, США). Пробирку заполняли водным раствором, содержащим 150мМ NaCl и 10 мМ аргинина. Пиявка прокусывала мембрану, всасывала раствор и выделяла секрет в пробирку. Раствор, обогащённый ССК, непрерывно перемешивали и обновляли через дополнительное отверстие, чтобы предотвратить его попадание в пиявку. Собранные образцы концентрировали на твердофазном экстракционном картирдже Sep-Pak Vac C18 (Waters, CША), используя для элюции белковой фракции в качестве буфера 0,1%-ый раствор TFA в 70%:30%-ом растворе ацетонитрил:вода. Ацетонитрил (ACN) высушивали в лиофильной сушке CoolSafe (LaboGene, Дания). Высушенный осадок хранили при -70°С.

Сбор нативного ССК *H. medicinalis* выполнялись м.н.с лаборатории генной инженерии ФГБУ ФНКЦ ФХМ ФМБА России к.б.н. А.А. Курдюмовым и с.н.с., к.б.н. В.А. Мануверой.

2.2.7. Пробоподготовка секрета слюнных клеток *H. medicinalis* для протеомного анализа

Для ферментативного гидролиза белков ССК медицинской пиявки и последующего масс-спектрометрического анализа использовали несколько методов пробоподготовки: подготовка образца с использованием центрифужных фильтров (англ. *filter-aided sample preparation, FASP*), трипсинолиз в растворе и ферментативный гидролиз в геле.

Подготовка образца с помощью фильтра

Лиофилизированный ССК медицинской пиявки ресуспендировали в 500 мкл буфера, содержащего 50 мМ TrisHCl, pH 8, 0,1% дезоксихолата натрия (DCNa). Раствор переносили в центрифужные ультрафильтрационные колонки с номинальной отсечкой по молекулярной массе 10000 Да (Millipore, США) и центрифугировали в течение 20 мин при 14000 g, 20°С. В колонки добавляли 400 мкл буфера, содержащего 50 мМ TrisHCl, pH 8, 8 М мочевину и 0,1% DCNa, ресуспендировали раствор и центрифугировали в течение 20 мин при 14000 g, 20°C. Элюаты отбирали, в колонки добавляли 400 мкл буфера, содержащего 50 мМ TrisHCl, pH 8, 0,1% DCNa, и повторяли центрифугирование. Шаг с добавлением DCNa повторяли. Цистеиновые связи восстанавливали добавлением 100 мкл буфера, содержащего 50 мМ TrisHCl, pH 8, 0,1% DCNa, 10 мМ дитиотреитола (DTT), и инкубацией в течение 30 мин при 60°С. Затем образцы центрифугировали в течение 10 мин при 14000 g, 20°C. Для алкилирования остатков цистеина в ультрафильтрационные колонки добавляли 100 мкл 55 мМ йодацетамида (IAA), инкубировали в темноте при комнатной температуре в течение 30 мин и центрифугировали в течение 10 мин при 14000 g, 20°C. Фильтры дважды промывали 200 мкл 50 мМ раствора бикарбоната аммония, содержащего 0,1% DCNa. Белки гидролизовали в 40 мкл 50 мМ раствора бикарбоната аммония, содержащего 0,1% DCNa, при 37°С в течение 16 ч, с помощью трипсина, используя соотношение фермента к белку 1:50. Полученные пептиды собирали центрифугированием в течение 15 мин при 14000 g с последующими двумя промываниями 50 мкл 0,5 М раствора NaCl. Объединенный гидролизат обессоливали с использованием колонки для твердофазной экстракции Discovery DSC-18 Tube (Supelco, Sigma-Aldrich, США) в соответствии с протоколом производителя. Пептиды элюировали 700 мкл 75%-ого раствора ACN, содержащего 0,1% трифторуксусной кислоты (TFA), затем сушили в вакуумном испарителе SpeedVac (Labconco, CША) и ресуспендировали в 3%-ом растворе ACN, содержащего 0,1% TFA до конечной концентрации белка 5 мкг/мкл.

Трипсинолиз белков в растворе

Лиофилизированный ССК медицинской пиявки растворяли в 5 мкл 10%-ого раствора поверхностно-активного вещества RapiGest SF (Waters, США), с добавлением 1 мкл смеси

нуклеаз, в течение 30 мин при 4°С. Затем смесь ресуспендировали в 45 мкл 100 мМ раствора бикарбоната аммония, перемешивали и грели при 100°С в течение 5 мин. После охлаждения до комнатной температуры раствор центрифугировали в течение 5 мин при 15000 g. Супернатант отбирали и концентрацию белка определяли методом с бицинхониновой кислотой с использованием набора BCA Assay Kit (Sigma-Aldrich, США). Цистеиновые связи восстанавливали с помощью 10 мМ DTT в течение 30 мин при 60°С и алкилировали 30 мМ IAA в темноте при комнатной температуре в течение 30 мин. Шаг с добавлением DTT повторяли. После добавления трипсина к супернатанту в объёмном соотношении фермент:белок, равном 1:50, смесь инкубировали при 37°С в течение ночи. Для инактивации трипсина и деградации кислотно-лабильного поверхностно-активного вещества исходный раствор TFA (конечная концентрация в образце 0,5%) добавляли к гидролизату, инкубировали при 37°C в течение 45 мин и центрифугировали в течение 10 мин при 15000 g для удаления сурфактанта. Гидролизат обессоливали с использованием колонки для твердофазной экстракции Discovery DSC-18 Tube в соответствии с протоколом производителя. Пептиды элюировали 700 мкл 75%-ого раствора ACN, содержащего 0,1% TFA, затем сушили в SpeedVac и ресуспендировали в 3%-ом растворе ACN, содержащего 0,1% TFA до конечной концентрации белка 5 мкг/мкл.

Ферментативный гидролиз белков в геле

Лиофилизированный ССК медицинской пиявки ресуспендировали в 500 мкл 50 мМ буфера TrisHCl, pH 6,8, раствор переносили в центрифужные ультрафильтрационные колонки с номинальной отсечкой по молекулярной массе 3000 Да (Millipore, США) и центрифугировали в течение 20 мин при 14000 g. Повторяли предыдущий шаг дважды. Концентрацию белка в элюате определяли методом Брэдфорда с использованием набора Bradford Protein Assay Kit (Bio-Rad, США). Затем к концентрированному образцу добавляли буфер Лэммли для внесения проб в отношении 1:1, нагревали в течение 5 мин при 95°С и центрифугировали в течение 10 мин при 15000 g. Супернатант отбирали, а белки разделяли с помощью электрофореза в 7,5%ом полиакриламидном геле в денатурирующих условиях. После окончания электрофореза гель окрашивали краской Кумасси G-250 согласно протоколу. Для трипсинолиза гели для каждого образца разрезали на более мелкие фрагменты (1х1 мм) и переносили в пробирку. Образцы геля очищали с помощью 50%-ого раствора ACN, содержащего 50 мМ бикарбоната аммония, при 50°С. Цистеиновые связи восстанавливали с помощью 100 мМ раствора бикарбоната аммония, содержащего 10 мМ DTT, в течение 30 мин при 56°С и алкилировали добавлением 100 мМ раствора бикарбоната аммония, содержащего 55 мМ IAA, в темноте при комнатной температуре в течение 20 мин. Шаг с добавлением DTT повторяли. При алкилировании образцы геля обезвоживали, добавляли чистый ACN. После удаления ACN к высушенным

фрагментам геля добавляли 40 мМ раствор бикарбоната аммония, содержащий 20 нг/мкл трипсина и 10% ACN так, чтобы раствор покрывал кусочки геля, и инкубировали сначала в течение 1 ч на льду, а затем в течение 16 ч при 37°С. Полученные триптические пептиды извлекали из геля последовательно следующими растворами: 5%-ым раствором муравьиной кислотой; 50%-ым раствором ACN, содержащего 5% муравьиной кислоты; 75%-ым раствором ACN, содержащего 5% муравьиной кислоты, 75%-ого растворяли в 50 мкл 3%-ого раствора ACN, содержащего 0,1% TFA. Гидролизат обессоливали с использованием колонки для твердофазной экстракции Discovery DSC-18 Tube в соответствии с протоколом производителя. Пептиды элюировали 700 мкл 75%-ого раствора ACN, содержащего 0,1% TFA, затем сушили в SpeedVac и ресуспендировали в 3%-ом растворе ACN, содержащего 0,1% TFA, до конечной концентрации белка 5 мкг/мкл.

2.2.8. Масс-спектрометрический анализ секрета слюнных клеток H. medicinalis

Масс-спектры получали для каждого приготовленного белкового образца с использованием трёх приборов: (а) масс-спектрометр TripleTOF 5600+ с источником ионов NanoSpray III (Sciex, CША), соединенным с системой нано-ВЭЖХ NanoLCUltra 2D + (Eksigent, CША), (б) масс-спектрометр Q-Exactive HF с наноспреем Flex (Thermo Scientific, Германия) и (в) масс-спектрометр maXis 3G с модернизацией ячейки HDC и источником ионов NanoElectrospray (Bruker Daltonics GmbH, Германия) в сочетании с системой ВЭЖХ Dionex Ultimate 3000 (ThermoScientific, США). Полученные исходные масс-спектры преобразовывали в наборы белковых данных, сгенерированные для отдельных образцов, на основе которых была сформирована спектральная библиотека белков ССК медицинской пиявки.

Пробоподготовку и масс-спектрометрический анализ ССК медицинской пиявки выполняли в сотрудничестве с лабораторией системной биологии ФГБНУ «Научноисследовательский институт биомедицинской химии имени В.Н. Ореховича» (ИБМХ), а также лабораторией протеомного анализа ФГБУ ФНКЦ ФХМ ФМБА России.

2.2.9. Биоинформатический анализ аннотации генома H. medicinalis

Для идентификации потенциальных АМП проводили *in silico* анализ белок-кодирующих последовательностей, обнаруженных в результате аннотации *de novo* собранного генома *H. medicinalis* [215]. Аминокислотные последовательности предположительных белков использовали в качестве матрицы для формирования списка всевозможных пептидов червя. Для каждой последовательности получали все пептиды длиной 50 а.о. со сдвигом на один аминокислотный остаток с использованием функции, написанной на языке программирования Рython. На начальном этапе анализа данных считали физико-химические и структурные

свойства аминокислотных последовательностей. Физико-химические свойства, такие как молекулярный вес пептида, количество аминокислотных остатков, значения изоэлектрической точки, суммарного заряда при разных значениях pH, индекса гидрофобности в соответствии со шкалами GRAVY, Кайта-Дулиттла, Эйзенберга, рассчитывали с использованием функции, написанной с использованием пакетов языка программирования R: protr, seqinr, Peptides [216-218]. Для дальнейшего отбирали аминокислотные последовательности анализа положительным суммарным зарядом и значением изоэлектрической точки в пределах от 8 до 12. Склонность пептидов к in vitro и in vivo агрегации определяли с помощью онлайн-сервисов TANGO и AGGRESCAN с использованием стандартных опции, предлагаемых разработчиками [219,220]. Отбирали пептидные последовательности с показателями AGG \leq 500, 0 \leq HELIX \leq 25, $25 \le BETA \le 100$ и $40 \le Na4vSS \le 60$ для TANGO и AGGRESCAN соответственно, где указанные параметры оценивали склонность белка к AGG - агрегации, HELIX формированию α-спиральной структуры, ВЕТА - формированию β-складчатой структуры, Na4vSS - нормализованное среднее значение склонности белка к агрегации. Значения физикохимических и структурных свойств, характерных для АМП, выбирали в соответствии с опубликованными ранее данными. Для каждой оставшейся в анализе аминокислотной последовательности получали все возможные пептидные последовательности длиной от 10 до 15 а.о. с использованием функции, написанной на языке программирования Python. После фрагментации пептидов для всех аминокислотных последовательностей оценивали антимикробный потенциал с помощью онлайн-алгоритмов АМРА (> 0,2), ADAM (> 1) и САМР_{R3} (> 0,5) [144,145,221]. Пороговые значения параметров указаны в круглых скобках для каждого предиктора соответственно. Отбирали последовательности с самым высоким антимикробным потенциалом по всем показателям. Отобранные пептидные последовательности верифицировали относительно протеомных данных секрета слюнных клеток *H. medicinalis*. Выбирали аминокислотные последовательности, которые достоверно идентифицировали в протеоме ССК пиявки.

2.2.10. Предсказание вторичной структуры пептидов

Вторичную структуру кандидатных пептидов предсказывали с использованием онлайнсервисов PEP-FOLD3 и I-TASSER-MR с опциями по умолчанию [222,223]. Для визуализации пространственного расположения боковых цепей аминокислотных остатков потенциальных АМП, анализировали последовательности в α-спиральной проекции в с помощью пакета Python modlAMP [125].

2.2.11. Химический синтез пептидов

Химический синтез пептидов осуществляли по протоколу твердофазного пептидного

синтеза на микроволновом пептидном синтезаторе Liberty Blue (CEM, CША) с использованием (Fmoc)-защищенных по аминогруппам аминокислот. Для конденсации Fmoc-аминокислот использовали этил-(гидроксиимино)-

производных

цианоацетат (ОХҮМА) в присутствии N,N'-Diisopropylcarbodiimid (DIC). В качестве растворителя в пептидном синтезе использовали N,N-диметилформамид (DMF) (ос.ч). Удаление защитных Fmoc-групп с первой аминокислоты на носителе и с растущей пептидной цепи проводили 20%-ым раствором 4-метилпиперидина (4M-pip) в DMF. Операции по наращиванию пептидной цепи проводили по протоколу для каждого синтетического цикла. Сначала снимали защиту с α-аминогруппы 3 мл 20%-ого раствора 4М-рір, инкубировали 15 сек. при 75°C, 50 сек при 90°C. Затем промывали 2 мл DMF, выполняли барботирование реакционной смеси 5 мл аргона (3 цикла промывки). Конденсировали 0,63 мл 200мМ раствора аминокислоты + 1 мл 125 мМ раствора DIC + 0,5 мл 250мМ раствора ОХҮМА, инкубировали 15 сек. при 75°С, 225 сек при 90°С (для аргинина: инкубировали 1500 сек. при 25°С, 120 сек при 75°С). Снова промывали 2 мл DMF, выполняли барботирование реакционной смеси 5 мл аргона (3 цикла промывки). По окончании наращивания пептидной цепи проводили финальное снятие защитных Fmoc-групп с α-аминогруппы пептида 3 мл 20%-ого раствора 4М-рір, инкубировали 15 сек при 75°C, 50 сек при 90°C и промывали DMF. Снятие пептида с твердой фазы и одновременное деблокирование боковых групп осуществляли в 5 мл смеси TFA:3,6-диокса-1,8октандитиол: триизопропилсилан: вода в объёмном соотношении 92,5:2,5:2,5:2,5 в течение 30 мин при 37°С. По окончании инкубации смесь отфильтровывали от полимера через фильтр реакционного сосуда в пробирки, после чего добавляли 45 мл охлаждённого до -20°С диэтилового эфира (далее – эфир). Полученную суспензию центрифугировали 10 мин. при 8000 об/мин. Надосадочную жидкость удаляли, полученный осадок повторно ресуспендировали в 25 мл охлаждённого эфира, центрифугировали, надосадочную жидкость удаляли. Осадок высушивали в вытяжном шкафу при комнатной температуре. Для очистки пептиды растворяли в 2 мл смеси (или в большем объёме, в зависимости от растворимости пептида) вода: ACN в объёмном соотношении 90:10. Степень чистоты пептидов контролировали с помощью ВЭЖХ на хроматографической системе ÄKTA Pure (GE Healthcare, США) и масс-спектрометрии с использованием Ultraflex II MALDI-TOF/TOF (Bruker, США), оснащённого УФ (Nd) лазером. концентрации определяли Молярные аминокислот с помощью количественного хроматографического анализа производных ортофталевой аминокислоты с использованием стандартов аминокислот.

Синтез, очистка и масс-спектрометрический анализ пептидов выполнялся м.н.с. лаборатории генной инженерии ФГБУ ФНКЦ ФХМ ФМБА России И.А. Лацисом.

9-флуоренилметилоксикарбонил

2.2.12. Оценка антимикробной активности пептидов

МИК пептидов в отношении штаммов E. coli и B. subtilis определяли в 96-луночных планшетах для микротитрования (Costar 3790, Corning, США) согласно стандартной методике двухкратных серийных разведений [224]. Клетки растили на питательной среде МНВ. Отдельные бактериальные клоны с агаризованных чашек культивировали в термостате в течение ночи в среде МНВ при 37°С и перемешивании со скоростью 150 об/мин. В логарифмической фазе роста культуру разбавляли средой МНВ до концентрации ~106 КОЕ/мл. Концентрацию клеток определяли путём измерения оптического поглощения при λ=570 нм с использованием микропланшетного фотометра Multiskan Ascent (ThermoScientific, CША). Ранее нами были построены кривые пересчёта оптических единиц в КОЕ/мл для обоих штаммов. В лунки планшета для микротитрования добавляли 50 мкл суспензии клеток с последующим добавлением 50 мкл разведённого средой МНВ пептида (0,5-128 мкг/мл, конечные концентрации). После инкубации при 37°С в течение 24 ч антимикробную активность пептидов в отношении тестируемых бактерий оценивали по оптической плотности раствора. За значение МИК принимали минимальную концентрацию пептида, которая полностью ингибирует рост микроорганизма. В качестве положительного контроля использовали антимикробный пептид мелиттин.

МИК пептидов для штамма C. thrachomatis определяли согласно методике, разработанной ранее в нашей лаборатории [203]. Клетки МсСоу высевали в лунки 24-луночного планшета с покровными стёклами в среде DMEM, содержащей 10% эмбриональной бычьей сывороткой и 10 мкг/мл гентамицина. Культивировали клетки в CO₂-инкубаторе при 37°C, 5% СО₂ до 90% конфлуэнтности. Элементарные тельца (ЭТ) С. trachomatis (5×10⁴ IFU) инкубировали в сахарозо-фосфат-глутаматном буфере (SPG) с пептидами (концентрация пептида составляла от 2,5 до 100 мкМ) при 25°С в течение 2 ч. Культуральную среду из лунок отбирали и к клеткам добавляли 200 мкл растворов пептидов с хламидийными ЭТ. В качестве отрицательного контроля использовали хламидийные ЭТ, инкубированные в SPG. Клеточный монослой инфицировали инокулятом C. trachomatis с последующим центрифугированием в течение 1 ч при 900 g. Затем добавляли свежую среду DMEM, содержащую 10% эмбриональной бычьей сыворотки и 10 мкг/мл гентамицина, до конечного объема 1 мл. Планшеты культивировали в CO₂-инкубаторе при 37°C, 5% CO₂ в течение 48 ч. После инкубации лунки фиксировали 4%-ым раствором параформальдегида в течение 20 мин и инкубировали с 1%-ым раствором Triton X-100 в течение 30 мин. Хламидийные включения выявляли с помощью реакции прямой иммуннофлуоресценции. Для этого клетки МсСоу обрабатывали антителами, специфичными к липоолигосахаридам хламидий, мечеными флуоресцеинизотиоционатом (FITC), в течение 1 ч. Раствор антител также содержал бромид этидия. Готовые препараты анализировали на эпифлуоресцентном микроскопе Nikon Eclipse E800 (Nikon, Япония) с объективами Plan Fluor 10× и 40×. В качестве источника света использовали ртутную лампу. Для работы с индикаторами, флуоресцирующими в зелёной и красной областях спектра, использовали соответствующими фильтрами: FITC (EX 465-495, DM 505, BA 515-555) и TRITC (EX 540/25, DM 565, BA 605/55). Обработку полученных изображений осуществляли с помощью программного обеспечения ImageJ (версия 1.48; RSB). Хламидийные включения идентифицировали как характерные включения зелёного цвета в цитоплазме клеток красного цвета. Титр, образующую включения, рассчитывали исходя из количества включений в одном поле зрения по следующей формуле:

Единица, образующая включения (количество/мл) = X * N * фактор разведения / V,

где X- количество включений в одном поле зрения, N - количество полей зрения, V – объем исследуемого образца, мл. Количество полей зрения рассчитывалось исходя из площади одного поля зрения и площади поверхности исследуемого образца. За значение МПК90 принимали минимальные концентрации пептидов, при которых уровень инфекции был меньше 90% от уровня в контрольных образцах.

Эксперимент и обработка изображений выполнялись н.с. лаборатории генной инженерии ФГБУ ФНКЦ ФХМ ФМБА России к.б.н. Н.Ф. Полиной и м.н.с. П.А. Бобровским.

2.2.13. Определение гемолитической активности пептидов

Свежевыделенные человеческие эритроциты использовали для оценки гемолитической активности пептидов. К отобранной крови здорового донора с целью избежания свёртывания добавляли цитратный буфер, центрифугированием течение 15 мин при 1000 g и трижды промывали забуференным физиологическим раствором (PBS, pH 7,4) в течение 5 мин при 2900 g. После последнего центрифугирования клетки ресуспендировали в PBS. Затем аликвоты по 100 мкл в концентрации от 6,25 до 200 мкМ пептидов добавляли в 96-луночный микропланшет, где каждая лунка содержала 100 мкл 8%-ой суспензии эритроцитов в PBS. Затем образцы инкубировали в термостате при 37°C в течение 1 ч при интенсивном перемешивании (150 об/мин). После инкубации микропланшет центрифугировали в течение 5 мин при 1000 g. Процент гемолиза определяли путём измерения оптического поглощения при λ =450 нм. 0,1%-ый раствор Triton X-100 в PBS использовали в качестве положительного контроля (K-). Процент гемолиза рассчитывали по следующей формуле:

% Гемолиза = $[(OD450 \text{ probe} - OD_{450} \text{K}^{-})/(OD_{450} \text{K}^{+} - OD_{450} \text{K}^{-})] \times 100,$

где OD₄₅₀probe - оптическая плотность раствора эритроцитов, инкубируемых с пептидом,

ОD₄₅₀К⁻ - оптическая плотность раствора эритроцитов, инкубируемых с PBS,

 $OD_{450}K^+$ - оптическая плотность раствора эритроцитов, инкубируемых с 0,1%-ым раствором Triton X-100.

2.2.14. Исследование цитотоксической активности пептидов

Оценку цитотоксического действия АМП в отношении клеток млекопитающих проводили с использованием набора LIVE/DEADTM Viability/Cytotoxicity Kit (ThermoScientific, США) в соответствии с рекомендациями производителя. Клетки МсСоу высевали в 96луночные планшеты до плотности 1×10⁵ клеток/мл и через 24 ч в лунки добавляли пептиды в концентрациях равных 1/2×, 1×, 2× и 4× значению МИК. Свежую культуральную среду без добавления пептидов использовали в качестве отрицательного контроля. После инкубации с пептидами клетки промывали сбалансированным солевым раствором Хэнкса (HBSS). Окрашивание клеток производили инкубацией в HBSS, содержащим Calcein AM и Ethidium homodimer-1, в концентрациях 2,5 мМ, инкубировали клетки при 37°С в течение 20 мин. После HBSS. окрашивания клетки дважды промывали Изображения анализировали на эпифлуоресцентном микроскопе с использованием фильтров: FITC (EX 465-495, DM 505, BA 515-555) и TRITC (EX 540/25, DM 565, BA 605/55). Обработку полученных изображений осуществляли с помощью программного обеспечения ImageJ (версия 1.48; RSB).

Эксперимент и обработка изображений выполнялись совместно с н.с. лаборатории генной инженерии ФГБУ ФНКЦ ФХМ ФМБА России к.б.н. Н.Ф. Полиной и н.с., к.б.н. О.В. Подгорным.

ЯМР-спектроскопия и расчёт пространственной структуры пептидов

Каждый пептид растворяли в 90%:10%-ом растворе H₂O:D₂O до концентрации равной 1 мМ при pH ~ 4. Регистрацию одномерных и двумерных спектров ЯМР в растворах (¹H, TOCSY (время перемешивания 60 мс), ROESY, NOESY (время перемешивания 100 мс), ¹H-¹³C HSQC, ¹H-¹⁵N HSQC) проводили на ЯМР-спектрометрах Bruker Avance III 600, 700 и 800 МГц (Bruker, США) при температуре 303 К. Все спектры ЯМР обрабатывали с помощью программного обеспечения TOPSPIN (Bruker Biospin, США), отнесение сигналов в спектрах выполняли с помощью программы CARA [225]. Вторичная структура и ограничения для двугранных углов ф и ф предсказывали с использованием программы TALOS-N [226]. С помощью программы СУАNА для каждого пептида рассчитывали 100 структур, из которых затем отбирали по 10 структур с минимальным значением энергии [227]. Для определения структуры в мембраннообразцу добавляли суспензию дейтерированных миметической среде к додецилфосфатидилхолиновых (ДФХ) мицелл (d-38, 98%; Cambridge Isotope Laboratories США) в молярном отношении пептид:детергент 1:100.

Исследование пространственной структуры пептидов проводили совместно с с.н.с. лаборатории биомолекулярной ЯМР-спектроскопии ИБХ РАН к.ф.-м.н. К.Д. Надеждиным, и студенткой И. А. Талызиной.

2.2.15. Приготовление липосом разного липидного состава

Цвиттерионные липосомы получали из 1-пальмитоил-2-олеоил-глицеро--3-фосфохолина (ПОФХ), а анионные – из смеси 1-пальмитоил-2-олеоил-глицеро-3-фосфо-L-серин: (ПОФС), 1пальмитоил-2-олеоил-глицеро—3-фосфохолин (ΠΟΦΧ) 1,2-диолеоил-глицеро-3-И фосфоглицерол (ДОФГ): ПОФХ:ДОФГ 4:1; ПОФХ:ПОФС 3:1; ПОФХ:ПОФС 1:1 и чистый ПОФС. Большие моноламеллярные везикулы (БМВ) получали методом экструзии с использованием поликарбонатных фильтров с диаметром 100 нм пор (Whatman, Великобритания) и экструдера Mini-Extruder (Avanti Polar Lipids Inc, США). Сухие липиды растворяли в хлороформе и сушили в атмосфере аргона. Оставшийся хлороформ высушивали дополнительно при вакууме (2 мбар) в течение 3 ч. Сухую липидную плёнку ресуспендировали в 100 мМ растворе карбоксифлуоресцеина (CF) (PBS, pH 7,4), осторожно встряхивая в течение 30 мин. После десяти циклов замораживания-оттаивания в жидком азоте и водяной бане при 50°С липидные суспензии обрабатывали ультразвуком и с помощью экструдера пропускали через поликарбонатный фильтр 21 раз при комнатной температуре. Неинкапсулированный CF удаляли гель-проникающей хроматографией на колонке для обессоливания HiTrap™ (Bio-Rad, США). БМВ, содержащие CF, переводили в 10 мМ TrisHCl, содержащий 150 мМ NaCl (pH 7,4).

Липосомы были приготовлены заведующим лабораторией электрофизиологии ФГБУ ФНКЦ ФХМ ФМБА России к.ф.-м.н. П.В. Башкировым.

2.2.16. Измерение и оценка кинетики связывания пептидов с липосомами

Анализ взаимодействия пептидов с липосомами различного состава оценивали по измерению флуоресценции, вытекшего красителя из липосом. Для оценки кинетики пептиды в концентрации 2 мкМ добавляли к суспензии БМВ, содержащих СF, и инкубировали в 96-луночном планшете при 25°C. Для измерения эффекта взаимодействия пептиды в разной концентрации (0,125–10 мкМ) добавляли к суспензии БМВ, содержащих CF, и инкубировали в 96-луночном планшете при 25°C. Интенсивность флуоресценции (λ ex = 485 нм, λ em = 538 нм) измеряли с использованием устройства для считывания микропланшетов Fluoroskan Ascent (ThermoScientific, CША) через 5 мин после смешивания растворов. Процент высвобождения CF оценивали по следующей формуле:

% Высвобождение CF = $[(F_{obs} - F_0) / (F_0 - F_{100})] \times 100$,

где F_{ob} - интенсивность флуоресценции контрольного образца с пептидом,

F₀ - интенсивность флуоресценции контрольного образца с буфером,

F₁₀₀ - интенсивность флуоресценции контрольного образца после добавления мелиттина в концентрации 1 мкМ.

2.2.17. Электронная микроскопия E.coli и B. subtilis после инкубации с пептидами

Бактериальные культуры *E.coli* и *B. subtilis* в логарифмической фазе роста обрабатывали наиболее активными пептидами 3967 и 536_1 в конечной концентрации $1/2 \times$ МИК и $1 \times$ МИК и инкубировали анаэробно в термостате при 37°С в течение 8 ч и 24 ч. Клетки отбирали, центрифугировали в течение 5 мин при 2000 g, дважды промывали и ресуспендировали в PBS. Бактериальную суспензию осаждали на чистом стерильном предметном стекле и сушили при 37°С. Образцы фиксировали с использованием 2,5%-ого раствора глутаральдегида в течение ночи при 4°С и дегидратировали батареей спиртов (10%, 20%, 30%, 50%, 70%, 80%, 90%, $2 \times 96\%$ в течение 15 мин для каждого раствора). Затем образцы инкубировали в смеси гексаметилдисилазан (HMDS):этанол в объёмном соотношении 1:1 в течение 10 мин и в 100%ом HMDS в течение ночи до полного испарения HMDS [228]. После химической сушки образцы покрывали 10 нм сплавом золота и палладия с использованием устройства для ионного напыления Sputter Coater Q150T (Quorum Technologies, Великобритания). Образцы охарактеризовывали с помощью микроскопа Zeiss Merlin, оснащенного GEMINI II Electron Optics (Zeiss, Германия) при ускоряющем напряжении 1 кВ и токе зонда 110 пА.

Анализ образцов с использованием СЭМ проводили в сотрудничестве с м.н.с. лаборатории медицинских нанотехнологий ФГБУ ФНКЦ ФХМ ФМБА России Е.Р. Павловой.

2.2.18. Статистический анализ

Эксперименты выполняли не менее чем в трёх независимых повторах для каждой экспериментальной группы. Из-за негауссовского распределения средних значений использовали непараметрический критерий Крускала-Уоллиса с последующим сравнительным анализом Данна для определения значимости между контрольными и тестовыми значениями. Результаты считали статистически значимыми при значении *p*< 0,05.

ГЛАВА З. РЕЗУЛЬТАТЫ

3.1. Определение полной нуклеотидной последовательности генома H. medicinalis

Расшифровка полной информации о геноме *H. medicinalis* важна не только для понимания структуры генома, но и является основой для исследования богатого репертуара белков медицинской пиявки, в том числе белков ССК. На первом этапе работы было проведено генотипирование отобранных особей пиявок. Участки генома, кодирующие транскрибируемый участок-спейсер ITS2 ядерной И 12S-митохондриальную рибосомную РНК были отсеквенированы и проанализированы. Выравненные нуклеотидные последовательности участков ITS2 гена ядерной рРНК и генов 12S-митохондриальной рРНК трёх видов медицинских пиявок представлены на Рисунках 8, 9. Мы подтвердили, что выбранные особи пиявок относились к виду *H. medicinalis*.



Рисунок 8 - Выравнивание нуклеотидных последовательностей участка ITS2 гена ядерной рРНК трёх видов медицинских пиявок

H. medicinalis - AY763166.1, H. orientalis - AY763170.1, H. verbana — AY763167.1.



Рисунок 9 - Выравнивание нуклеотидных последовательностей участка гена 12Sмитохондриальной рРНК трёх видов медицинских пиявок *H. medicinalis* - AY763156.1, *H. orientalis* - AY763163.1, *H. verbana* – AY763160.1.

Для расшифровки нуклеотидной последовательности генома H. medicinalis были получены ДНК-библиотека коротких прочтений (Ion Proton) и две парные ДНК-библиотеки с длиной прочтений 3-6 т.п.н. (Ion TrueMate) и 8-12 т.п.н. (Illumina MiSeq). Основные характеристики ДНК-библиотек представлены в Таблице 5. Далее был применён алгоритм SPAdes, который использует парные прочтения в качестве основы геномной сборки и заполняет пробелы с помощью высокоточных коротких прочтений [206]. Суммарно было получено 168624 контига - набора перекрывающихся последовательностей фрагментов ДНК (прочтений), с параметром N50 равному 12,9 т.п.н. (Таблица А.1 (Приложение А)). N50 - это мера оценки качества сборки генома, которая фрагментирована на контиги разной длины. Значение N50 необходимого равно длине самого короткого контига, для покрытия половины последовательности генома. С целью идентификации и разделения прокариотических и эукариотических контигов, была проведена процедура биннинга — метод предобработки данных с учётом характеристик последовательностей (GC-состав, частота встречаемости тетрануклеотидов, покрытие прочтений).

Длина прочтения, п.н.	Тип прочтения	Общее количество прочтений	Общая длина, 10 ⁹ п.н.	Покрытие			
Ion Proton, #SRR6926478 ³							
200	одно-концевые	47502374	11292	x50			
Ion Torrent, True Mate, размер вставки 5 т.п.н., #SRR6926477 ³							
400	парные	1416496	512	x2,5			
Illumina MiSeq, Nextera mate pair, размер вставки 10 т.п.н., #SRR6926476 ³							
150	одно-концевые	150286	328	x1,5			
300	парно-	6161612	1955	x9			
300	парные	6953316	2208	x10			

Таблица 5 - Результаты секвенирования ДНК-библиотек, полученных на основе геномной ДНК *H. medicinalis*

³ – идентификационный номер нуклеотидной последовательности в Архиве последовательностей прочтений (Sequence Read Archive, SRA)



Рисунок 10 Двумерная диаграмма распределения контигов, полученных в ходе полногеномного секвенирования H. medicinalis обозначены кругами, окрашенными В соответствии Контиги с таксономической принадлежностью, определенной по базе данных NCBI (зелёный - бактерии, синий - эукариоты, чёрный — не определено). Контиги представлены в осях GC-состава (ось абсцисс) и покрытия

Таксономическая классификация контигов была выполнена с использованием программного обеспечения MEGAN6 на основе результатов BLAST-анализа [211]. Таким

контигов (ось ординат).

образом, были объединены два основных подхода для разделения и кластеризации метагеномных последовательностей [229]. Статистика и визуализация биннинга представлена на Рисунке 10 и в Таблице А.2 (Приложение А). В результате, 29764 контига принадлежали последовательности генома *H. medicinalis*, тогда как 4094 контига имели прокариотическое происхождение. В целом, контиги отличаются параметром N50, для контаминационной ДНК он равен 31787 п.н., что намного больше этого значения для *H. medicinalis* (13786 п.н.). Отобранные эукариотические контиги с помощью программного обеспечения Sspace подверглись процедуре скаффолдинга — упорядочивании контигов, разделённых пробелами в покрытии [207]. Таким образом сборка генома *H. medicinalis* состояла из 14042 скаффолдов с параметром качества сборки N50 равному 98 т.п.н. (Таблица А.2 (Приложение А)). Всего было предсказано 14596 белок-кодирующих генов. Размер собранного генома *H. medicinalis* составил 85% 187.5 Мп.н., что составляет от размера теоретически предсказанного (http://www.genomesize.com/result_species.php?id=1061). Bce необработанные прочтения, полученные в ходе данного исследования, были депонированы в базе данных NCBI под регистрационными номерами BioProject acc. num. PRJNA257563 и PRJNA256119.

3.2. Аннотация генома *H. medicinalis*

Перед аннотацией генома H. medicinalis скаффолды были проверены на наличие мобильных генетических элементов с помощью программы RepeatMasker и с использованием базы данных ДНК-повторов беспозвоночных Repbase-GIRI (Рисунок 11) [230,231]. Всего было замаскировано 12% нуклеотидной последовательности генома *H. medicinalis*. Мобильные элементы двух основных классов, ретротранспозоны и ДНК-транспозоны, составляют 4,8% генома практически в равной пропорции. Доля ДНК-транспозонов (Class II) составляет 2,4% всего генома *H. medicinalis* и включает в себя самые распространенные подсемейства EnSpm, hAT и Crypton. Важно отметить, что в сборке генома H. medicinalis отсутствуют ДНКтранспозоны прокариот, что подтверждает качество выполненного биннинга. Ретротранспозоны с длинными концевыми повторами (англ. long terminal repeat, LTR) составляют 1,3% генома *H. medicinalis*, большинство из них относится к суперсемейству Gypsy. Длинные и короткие диспергированные ядерные элементы (англ. long interspersed nuclear element, LINE и short interspersed nuclear element, SINE), относящиеся к ретроэлементам, составляют около 1% генома, основная часть которых принадлежит суперсемействам CR1, L2, Daphne. Мало что известно о мобильных элементах в геномах аннелид [10]. Однако, полученные данные согласуются с результатами исследований геномов известных представителей Lophotrochozoa, например, кольчатых червей H. robusta и Capitella teleta и моллюсков Lottia gigantea и Octopus bimaculoides.



Рисунок 11 - Мобильные генетические элементы в геноме *H. medicinalis* Внутренняя диаграмма показывает процентное распределение трёх основных классов мобильных элементов (LTR, non-LTR и Class_II), а внешняя диаграмма показывает распределение суперсемейств в каждом из классов. Мобильные элементы сгруппированы согласно классификации Викера [232]. LTR - длинные концевые повторы, non-LTR - длинные и короткие диспергированные ядерные элементы, Class_II - ДНК-транспозоны.

Для картирования участков нуклеотидной последовательности генома были выбраны три шаблона: 1) набор последовательностей генов, кодирующих белки пиявки *H. robusta* [10]; 2) контиги, соответствующие *de novo* сборке транскриптома *H. medicinalis* [212]; 3) последовательности кДНК *H. medicinalis*, полученные нами в ходе исследования трех медицинских пиявок [215]. Референсные последовательности для каждого шаблона картировали на последовательности скаффолдов сборки генома *H. medicinalis* с использованием программного обеспечения BLAT (версия 34х13) [213]. Затем все три шаблона объединяли и с помощью программного обеспечения AUGUSTUS 125 (версия 3.7.1) выполняли аннотацию генома *H. medicinalis* [214]. Статистика аннотации генома *H. medicinalis* приведена в Таблице А.3 (Приложение А).

Далее было выполнено сравнение генома *H. medicinalis* с геномами гематофагов (пиявок *Helobdella robusta* и *Capitella teleta*, комаров *Anopheles gambiae* и *Culex quinquefasciatus*, постельного клопа *Cimex lectularius*, клеща *Ixodes scapularis* и летучей мыши *Myotis lucifugus*), и других спиральнодробящихся животных (моллюска *Aplysia californica*, устрицы *Crassostrea gigas*, осьминога *Octopus bimaculoides*, лингулы *Lingula anatina*) (Рисунок 12; Таблица Б.1 (Приложение Б)).



Рисунок 12 Статистика аннотации геномов *H. medicinalis*, представителей гематофагов и *Lophotrochozoa*

Гистограмма иллюстрирует (А) размеры генома (чёрный цвет), количество аннотированных генов (пунктирная линия) и среднюю длину генов (сплошная линия); (Б) количество и среднюю длину аннотированных экзонов (светло-серый столбец, сплошная линия) и интронов (тёмносерый столбец, пунктирная линия) геномов *Н. medicinalis*, 7 представителей гематофагов и 4 представителей *Lophotrochozoa*.

Среди всех организмов сравнения геном *H. medicinalis* меньше по размеру и количеству генов. Средняя длина гена составила 7002 п.н., что почти в два раза больше этого значения у других представителей класса *Annelida*, и на несколько тысяч больше средней длины гена *C. quinquefasciatus*. Тем не менее, это значение у *H. medicinalis* меньше, чем у остальных гематофагов и спиральнодробящихся организмов. Если сравнивать экзон-интронную структуру,

то все три пиявки, *H. medicinalis, H. robusta* и *C. teleta* в целом схожи между собой. Стоит отметить, что размер генома рассмотренных организмов полностью коррелирует со средней длиной интронов в геноме, что согласуется с литературными данными [233,234].

С помощью программы OrthoFinder были определены группы ортологичных функциональных генов и проведён их кластерный анализ между представителями типа Annelida (*H. medicinalis, H. robusta, C. teleta*) и клады Lophotrochozoa (L. anatina, G. gigas, O. bimaculoides, A. californica) [235]. В качестве внешней группы были выбраны позвоночные Danio rerio (BioProject acc. num. PRJNA11776) и Homo sapiens (BioProject acc. num. PRJNA31257) (Рисунок 13).



Рисунок 13 - Диаграмма Венна кластеризации ортологичных функциональных генов *H. medicinalis*, представителей гематофагов и *Lophotrochozoa* Видовые составы обозначены цветами: *Annelida* — оранжевым, *Brachiopoda* — красным, *Mollusca* — синим, *Vertebrata* — зелёным.

Всего было рассмотрено 15928 групп ортологичных функциональных генов, среди которых 7520 генов, кодирующих белки, участвующие в общих биологических процессах, были общими для всех организмов. По количеству уникальных ортологичных генов пиявки уступают лишь позвоночным. Кластерный анализ подтвердил близость таксонов Annelida и Brachipoda, и показал насколько отличаются эти таксоны от позвоночных. При более детальном белок-кодирующих последовательностей *H. medicinalis* рассмотрении помимо генов, кодирующих конститутивные белки, были идентифицированы новые гены-гомологи, кодирующие известные антикоагулянты и белки, связанные с перевариванием и хранением крови, а именно бделлины, антистазины, гирудин, дестабилаза, саратин и др. (Таблица Б.2 (Приложение Б). Большинство генов кодируют белки с неизвестной функцией.

3.3. Протеомный анализ секрета слюнных клеток H. medicinalis

ССК *Н. medicinalis* выполняет важные функции по обеспечению и поддержанию жизнедеятельности пиявок. Компоненты секрета отвечают за подавление защитных механизмов жертвы, влияют на систему свёртывания крови, регулируют хранение и переваривание крови в кропе медицинской пиявки и обеспечивают антимикробную защиту червя [149,156,183]. На данный момент в ССК *Н. medicinalis* идентифицировано лишь несколько десятков белков, поэтому исследование состава секрета все ещё остаётся важной задачей [154].

Нами был проведён комплексный анализ белкового состава ССК *H. medicinalis*. Выделение ССК *H. medicinalis* было выполнено по разработанной нами ранее методике. Ввиду сложного состава ССК было использовано три протокола пробоподготовки: подготовка образца с использованием центрифужных фильтров, трипсинолиз в растворе и геле. Образцы анализировали на трёх ВЭЖХ масс-спектрометрах: TripleTOF 5600+, Q-Exactive HF и maXis 3G. Анализ масс-спектров проводили с использованием программы ProteinPilot 4.5, ревизия 1656 [236]. идентификаций Для контроля качества применялась оценка частоты ложноположительных идентификаций (FDR) с помощью утилиты PSPEP версия 4.5.0.0, ревизия 1654. В качестве достоверных отобраны белки, для которых было обнаружено не менее 2 пептидов с баллом достоверности 95 и выше, с оценкой локальной FDR менее 1%, а также спектры с оценкой локальной FDR менее 5%. Выходные данные результатов протеомного профилирования образцов ССК *H. medicinalis* были объединены в спектральную библиотеку идентифицированных белков. В результате масс-спектрометрического анализа в ССК Н. medicinalis было обнаружено 189 белков, среди которых четверть молекул составили белки с неизвестной функцией (Рисунок 14, Таблица В.1 (Приложение В). Все детектированные белки были разделены на семь функциональных классов: 1) белки-регуляторы системы гемостаза, 2) ингибиторы протеаз, 3) цистеин-богатые секреторные белки, 4) белки, участвующие в переваривании крови, 5) протеазы 6) секретируемые белки с неизвестными функциями и 7) прочие.



Рисунок 14 — Белки, идентифицированные в ССК *H. medicinalis* Круговая диаграмма иллюстрирует процентное распределение отдельных групп белков, обнаруженных в ССК *H. medicinalis*.

3.4. Известные антимикробные пептиды H. medicinalis

На следующем этапе анализа *H. medicinalis* был выполнен целенаправленный поиск АМП H. medicinalis с использованием последовательности генома H. medicinalis и была протеомногоанализа ee ССК. Сначала нами проанализирована нуклеотидная последовательность генома H. medicinalis на наличие известных АМП. Базы данных АМП разнообразны по объёму и содержанию, и не всегда информация, находящаяся в каждой из них, достаточна для анализа данных. Поэтому на первом этапе была собрана база данных антимикробных белков и пептидов, объединяющая последовательности из 5 публичных баз данных: APD3, ADAM, UniProtKB, CAMP_{R3} и DADP (Таблица Д.1 (Приложение Д)) [143-145,237,238]. Последовательности белков и пептидов из базы данных UniProtKB были получены с использованием ключевых слов «antimicrobial proteins» и «antimicrobial peptides». Остальные последовательности были загружены непосредственно из баз данных, доступных онлайн. Итоговая база данных антимикробных белков и пептидов содержала 10245 3216 последовательностей, которых последовательностей были среди природного происхождения. С помощью обобщённой базы данных был проведён BLAST-анализ всех белок-кодирующих последовательностей генома *H. medicinalis* [239].

В результате анализа было обнаружено 9 последовательностей-гомологов уже известных антимикробных пептидов медицинских пиявок. Были найдены четыре гомолога Hmнейромацина (Рисунок 15(Б)) [196]. Два пептида Nrm1 и Nrm2 отличаются между собой и

62

референсной последовательностью Hm-Nrm всего на один аминокислотный остаток. Гомология всех трёх пептидов составила 98,3%. Пептиды Nrl1 и Nrl2 не обладают высокой степенью гомологии с известным нейромацином, процент сходства составил 67,2% и 63,8% соответственно. Также были обнаружены четыре гомологичных пептида Hm-люмбрицину, Lbr, PP12, PP13, PP14, для которых процент гомологии составил 64,9%, 64,9%, 62,3% и 100% соответственно (Рисунок 15(А)) [196]. Последний пептид Thm является гомологом Hm-теромацина, процент схожести составляет всего 29% (Рисунок 15(В)) [192].



Рисунок 15 - Выравнивание аминокислотных последовательностей АМП *H. medicinalis* (А) люмбрицинов, (Б) нейромацинов, (В) теромацинов. Lbr, PP12, PP13, PP14 – гомологи люмбрицина, Nrl1, Nrl2, Nrm1, Nrm2 – гомологи нейромацинов, Thm – гомолог теромацина, Hm-Lbr - Hm-люмбрицин, Lr-Lbr - Lr-люмбрицин, Mt-PP_1 – антимикробный пептид PP1 (*Metaphire tschiliensis*), Hm-Nrm — Hm-нейромацин, Ef-Nrm – нейромацин (*Eisenia foetida*), Hmc — гирудомацин, Hm-Th — Hm-теромацин, Tt-Th — Tt-теромацин.

Так как цель исследования заключалась в поиске новых АМП *H. medicinalis*, был оптимизирован алгоритм поиска АМП и применён для анализа полученных геномных и протеомных данных *H. medicinalis*.

3.5. Использование биоинформатического алгоритма для поиска новых антимикробных пептидов *H. medicinalis*

Для поиска новых АМП *H. medicinalis* был построен алгоритм идентификации антимикробных молекул в на основе разработанного ранее подхода для анализа транскриптомных данных таракана *Periplaneta americana* и многоножки *Scolopendra subspinipes mutilans* [240,241]. Идея алгоритма поиска заключалась в том, чтобы найти среди

аминокислотных последовательностей короткие положительно заряженные пептиды с тенденцией к образованию α-спиральных структур (Рисунок 16). Этот вычислительный алгоритм включает в себе онлайн-сервисы по расчёту структурных особенностей пептидов и их склонности к агрегации (TANGO, AGGRESCAN), а также несколько онлайн предикторов, направленных на определение антимикробного потенциала пептида (AMPA, ADAM, CAMP_{R3}) [144,145,219–221]. Алгоритм был дополнен созданным нами методом расчёта физикохимических свойств аминокислотных последовательностей, написанным с использованием пакетов языка программирования R: protr, seqinr, Peptides [216,217,242]. Для разбиения аминокислотных последовательности на пептиды определённой длины со сдвигом на один аминокислотный остаток была написана функция на языке программирования Python.



Рисунок 16 - Схематичное описание биоинформатического анализа нуклеотидной последовательности генома *H. medicinalis* с целью поиска новых АМП

Ha биоинформатического белок-кодирующие первом этапе анализа все последовательности *H. medicinalis* были разбиты на участки длиной 50 а.а. со сдвигом на один аминокислотный остаток, что в результате составило 7891765 молекул. Затем из всего списка соединений были отобраны те, что по своим физико-химическим свойствам соответствовали Анализ физико-химических свойств (суммарный известным AMΠ. заряд, значения изоэлектрической точки, индекс гидрофобности, тенденция к формированию вторичной структуры) позволил сократить количество анализируемых аминокислотных последовательностей до 370807 молекул. Затем с помощью онлайн предикторов для каждой аминокислотной последовательности были определены те участки последовательности, которые могут являться антимикробным пептидом. На финальной стадии анализа была проведена верификация последовательностей относительно протеомных данных ССК Н. medicinalis. В анализе были оставлены только те последовательности, для которых в протеоме были обнаружены пептиды. Далее, среди выбранных пептидов, было отобрано 12

последовательностей с наилучшими показателями антимикробного потенциала согласно онлайн предикторам (Таблица 6).

Название пептида	Аминокислотная последовательность	Длина, а.о.	Заряд	Молекулярный вес, Да
756	IKIVAPKKIKL	11	+4	1250,7
3063	WKKLVPYKIHAGI	13	+3	1552,9
3967	FRIMRILRVLKL	12	+4	1558,1
8557	WKIFKLKLRMLW	12	+4	1662,2
9332	LRWLGTKVGILK	12	+3	1383,7
12530	KFKKVIWKSFL	11	+4	1423,8
19347_1	LIVLTCRKKKKPF	13	+5	1574,1
19347_2	RPILIRVRRIRVI	13	+5	1660,1
536_1	RWRLVCFLCRRKKV	14	+6	1863,4
536_2	GFIVKRFKILV	11	+3	1319,7
8361_1	CLIMKVRRKK	10	+5	1274,7
8361_2	ISRACLIMKVRRKK	14	+6	1702,2

Таблица 6 – Кандидатные антимикробные пептиды, идентифицированные при анализе нуклеотидной последовательности генома *H. medicinalis*

Для отобранных пептидов с помощью онлайн-сервисов анализировали их вторичную структуру. Анализ проекций пространственного расположения боковых цепей аминокислотных остатков АМП *H. medicinalis* с использованием пакета modlAMP показал, что все пептиды содержат как гидрофобные, так и гидрофильные аминокислотные остатки и, в целом, все структуры продемонстрировали способность к формированию амфипатической структуры (Рисунок 17) [125]. В соответствии с предсказательным алгоритмом, применённым в нашем исследовании, все потенциальные АМП должны образовывать α-спиральную структуру. Тем не менее, согласно рассчитанным вторичным структурам, полученным с помощью онлайн-сервисов I-TASSER-MR и PEP-FOLD3, только 4 пептида (3967, 8557, 12530 и 8361_2) из 12 способны формировать α-спираль (Рисунок 18, 19) [222,223].



Рисунок 17 - Анализ амфипатичности кандидатных антимикробных пептидов *H. medicinalis* Для каждого пептида представлена проекция пространственного расположения боковых цепей его аминокислотных остатков. Серым цветом обозначены гидрофобные, чёрным – гидрофильные, белым - остальные аминокислотные остатки. Иллюстрация разработана автором с использованием программного пакета modlAMP [125].

66



Рисунок 18 - Предсказанные вторичные структуры АМП *Н. medicinalis*, полученные с помощью онлайн-сервисов PEP_FOLD3 (А) и I-TASSER (Б)

Красным цветом обозначены гидрофобные аминокислотные остатки, белым - нейтральные аминокислотные остатки, серым - гидрофильные аминокислотные остатки. Иллюстрация разработана автором с использованием графической программы РуМОL (версия 1.2.2.3) [67].

Таким образом, в ходе данной работы был оптимизирован биоинформатический метод поиска в геномных данных кандидатных белковых последовательностей, обладающих антимикробным потенциалом. Применённый к геномным и протеомным данным H. medicinalis сконструированный алгоритм позволил обнаружить 12 кандидатных AMΠ. Биоинформатический метод анализа в дальнейшем может быть использован для исследования других организмов, что позволит идентифицировать новые перспективные активные молекулы, обладающие антибактериальным действием. Для окончательной верификации сформированного подхода по определению противомикробных соединений, была изучена биологическая активность отобранных пептидов.

3.6. Изучение биологической активности кандидатных антимикробных пептидов *H. medicinalis*

Отобранные кандидатные пептиды были синтезированы химически с применением Fmoc стратегии на микроволновом пептидном синтезаторе Liberty Blue (CEM, CША). На первом этапе оценивали антимикробную активность отобранных пептидов в отношении трёх бактерий: грамположительной *B. subtilis* и грамотрицательных *E. coli*, *C. trachomatis*. МИК пептидов в отношении *B. subtilis* и *E. coli* определяли с помощью методики двухкратных серийных разведений [224]. Пептиды в различной концентрации инкубировали с суспензией бактериальных клеток в течение 24 часов. Антимикробную активность пептидов в отношении тестируемых бактерий оценивали по оптической плотности раствора. За значение МИК принимали минимальную концентрацию пептида, которая полностью ингибирует рост микроорганизма. Минимальную подавляющую концентрацию пептидов для штамма *C. thrachomatis* определяли согласно методике, разработанной ранее в нашей лаборатории [203]. Пептиды в различной концентрации инкубировали с хламидийными ЭТ. Затем полученной смесью заражали клетки McCoy. Через 48 часов инкубации оценивали уровень заражения. За значение МПК90 принимали минимальные концентрации пептидов, при которых количество включений *C. thrachomatis* уменьшалось на 90% по сравнению с контрольными образцами.

Измеренные значения МИК и МПК90 пептидов для каждого вида бактерий приведены в Таблице 7. Мелиттин, антимикробный пептид из яда пчелы, использовали в качестве положительного контроля. Согласно проведённым экспериментам два пептида, 3967 и 536_1, проявили антимикробную активность против всех тестируемых штаммов при низких значениях МИК. Для пептида 3967 значение МИК для *B. subtilis, E. coli* составило 10,3 мкМ, для *C. trachomatis* 3,1 мкМ, для пептида 536_1 — 8,6 мкМ, 17,2 мкМ и 6,3мкМ соответственно. Пять других пептидов, 8557, 12530, 19347 1, 8361 1 и 8361 2, ингибировали рост *B. subtilis* и *C.*

trachomatis, а пептид 19347_2 был активен в отношении только *B. subtilis*. Остальные пептиды 756, 3063, 9332 и 536_2 с незначительной или отсутствующей способностью к ингибированию бактериального роста были определены как неактивные. Интересно отметить, что значения МИК для всех пептидов ниже в отношении *B. subtilis* и *C. trachomatis*, чем для *E. coli*. Вероятно, такие отличия вызваны структурными особенностями клеточных стенок разных бактериальных штаммов.

Таким образом, мы подтвердили, что найденные кандидатные АМП обладают антимикробной активностью. Основным недостатком известных АМП является их цитотоксичность по отношению к эукариотическим клеткам. Поэтому необходимо было оценить гемолитическую и цитотоксическую активности отобранных пептидов.

	МИК (м	МПК90 (мкМ)	
название пептида	B. subtilis 168 HT	<i>E. coli</i> MG1655	C. trachomatis D/UW-3/Cx
756	>102	>102	>100
3063	>82	>82	>100
3967	10,3	10,3	3,1
8557	38,5	77	50
9332	>92	>92	>100
12530	45	90	12,5
19347_1	40,7	>81	12,5
19347_2	9,7	77,1	>100
536_1	8,6	17,2	6,3
536_2	97	>97	>100
8361_1	12,6	100,4	12,5
8361_2	9,4	37,6	12,5
мелиттин	1,5	5,7	2

Таблица 7 - Антимикробная активность кандидатных АМП *H. medicinalis*

Гемолитическую активность АМП оценивали по проценту гемолиза человеческих эритроцитов для шести разных концентраций соединений, от 3,125 до 100 мкМ после одночасовой инкубации с пептидами. В качестве положительного контроля использовали 0,1%-ый раствор Triton X-100 в однократном фосфатно-солевом растворе (1xPBS), который лизировал 100% всех эритроцитов в растворе. Чистый 1xPBS использовали в качестве отрицательного контроля. Большая часть пептидов обладают низкой гемолитической активностью (лизируют <10% эритроцитов при обработке пептидом в концентрации 100 мкМ) (Таблица 8). Два пептида, 536_1 и 536_2, лизируют приблизительно 30% эритроцитов в концентрации 100 мкМ. Только два пептида, 8557 и 12530, проявляют высокую гемолитическую активность, лизируя приблизительно 50% эритроцитов в концентрациях 6,25 и 50 мкМ соответственно. Для сравнения мелиттин, который является токсичным для клеток млекопитающих вызывал 100%-ый лизис эритроцитов в концентрации 1 мкМ.

Название	Процент гемолиза, %							
пептида	0	3,125 мкМ	6,25 мкМ	12,5 мкМ	25 мкМ	50 мкМ	100 мкМ	
3967	-	2,3±1,2	1,8±3,4	2,7±2,9	2,8±1,7	$5,8{\pm}4,6^{*}$	11,1±5,1*	
12530	-	4,1±3*	4,3±4,7*	$7,2\pm 2,5^*$	6,5±3,3*	13,1±5,3*	32,9±6,5*	
756	-	1,6±1,1	1,3±1,1	3,5±1,1*	2,8±2,7	4,4±3,1	4,6±4,5*	
8361_1	-	1,1±0,8	0,8±1,9	2±1,7	2,6±1,6	2,8±2,8	$6{\pm}5,9^{*}$	
3063	-	2,4±2,2	$4,8{\pm}2,8^*$	9,3±2,6*	24,2±4,9*	55,8±7,1*	87,3±4,4*	
536_1	-	2,1±3,7	$6,9{\pm}6,6^{*}$	6,9±4,4*	$7,4{\pm}4,5^{*}$	13±2,8*	27,6±6,4*	
536_2	-	2,3±3,8	2,8±2,6	4,4±2,5*	6,9±2,3*	13,6±4*	35,7±8,6*	
8557	-	$35,2{\pm}3,5^*$	$45,6\pm3,2^*$	$60,3\pm3,3^*$	79,5±12*	94±6,2*	$98,4{\pm}7,9^*$	
9332	-	1±3,1	1,8±4,3	3,2±4,1*	4,1±2,6*	$5,8\pm 3,5^{*}$	10,8±6,9*	
1X PBS	0,1±2,7	-	-	-	-	-	-	
0,1%-ый раствор Triton X-100	99,8±2,4*	-	-	-	-	-	-	
мелиттин, 1мкМ	99±3,4*	-	-	-	-	-	-	

Таблица 8 - Гемолитическая активность АМП H. medicinalis

* — Статистически значимые различия между контрольной и экспериментальной группами, p < 0.05. Результаты представлены в виде средних значений \pm S.D. (n = 3). Статистически значимые различия между контрольной и экспериментальной группами были определены с помощью непараметрического теста Крускала-Уоллиса с последующим сравнительным анализом с использованием критерия Данна.

Цитотоксичность пептидов по отношению к клеткам млекопитающих оценивали по проценту выживших мышиных фибробластов McCoy после 24-часовой инкубации клеток с соединениями в концентрациях, соответствующих 1/2×, 1×, 2× и 4× значению МИК. Жизнеспособность клеток оценивали при помощи двойного окрашивания клеток флуоресцентными красителями, кальцеином-АМ и йодидистым пропидием, которые позволяют измерять активность эстеразы и целостность мембраны, и флуоресцентной микроскопии. Практически во всех случаях жизнеспособность клеток составляла 100% и существенно не отличалась от контроля (Таблица 9). Пептид 19347_2 в концентрации 4×МИК незначительно снижал количество жизнеспособных клеток по сравнению с контролем. Пептиды 536 2 и

19347_1 в концентрации 4×МИК приводили к существенному снижению процента выживших клеток по сравнению с контрольными клетками (0% и 10,5% выживших клеток для пептидов 536_2 и 19347_1 соответственно).

Harrawya wawayya	Процент жизнеспособных клеток, %						
пазвание пептида	0	1/2xМИК	1хМИК	2хМИК	4хМИК		
3967	-	99,8±0,6	99,7±0,9	99,9±0,4	99,9±0,3		
12530	-	99,7±0,7	99,8±0,3	99,9±0,2	99,4±0,8		
19347_2	-	99,8±0,4	99,6±0,4	99,4±0,6	99,3±1		
8361_1	-	98,9±1,6	99,5±0,7	98,9±1,3	99,1±1		
8361_2	-	99,7±0,5	99,8±0,6	99,4±1	99,9±0,3		
536_1	-	99,8±0,9	99,7±0,6	99,8±0,7	99,6±0,9		
536_2	-	99,9±0,4	99,7±0,8	99,3±1,4	$0\pm0^*$		
19347_1	-	99,7±0,6	99,8±0,4	98,6±1,9*	10,5±17,3*		
среда DMEM	99,1±1,7	-	-	-	-		

Таблица 9 - Жизнеспособность мышиных фибробластов МсСоу после инкубации с идентифицированными пептидами *H. medicinalis* в течение 24 часов

* — Статистически значимые различия между контрольной и экспериментальной группами, p < 0.05. Результаты представлены в виде средних значений ± S.D. (n = 4). Статистически значимые различия между контрольной и экспериментальной группами были определены с помощью непараметрического теста Крускала-Уоллиса с последующим сравнительным анализом с использованием критерия Данна.

По результатам оценки функциональных характеристик АМП *H. medicinalis* два пептида, 3967 и 536_1, из 12 соединений, идентифицированных с помощью биоинформатического алгоритма поиска, обладают антимикробным действием в отношении бактерий, низкой гемолитической и цитотоксической активностями. Эти пептиды являются наиболее перспективными антимикробными агентами для дальнейшего исследования. С целью определения вероятных механизмов действия идентифицированных АМП *H. medicinalis* была изучена их вторичная структура.

3.7. Определение вторичной структуры антимикробных пептидов

Для анализа вторичной структуры методом ЯМР-спектроскопии были выбраны пять пептидов: 3967, 536_2, 12530, 756 и 9332. Пространственные структуры выбранных пептидов были исследованы с использованием стандартных методов ЯМР. Сначала были определены резонансы С α , С β , HN в спектрах ROESY, 1H-13C HSQC и 1H-15N HSQC и вторичная структура исследуемых пептидов определялась с использованием программы TALOS-N [226].

Анализ ЯМР-спектров пептидов в водных растворах показал, что структура пептидов неупорядочена в воде, что соответствует ранее опубликованным данным об известных АМП [33,243]. При добавлении суспензии додецилфосфатидилхолиновых (ДФХ) мицелл в каждый образец для имитации мембранного окружения, схожего с бактериальной клеточной мембраной, наблюдались значительные изменения в химических сдвигах для всех пептидов, что указывало на конформационные изменения.

Согласно расчётам по полученным химическим сдвигам, в растворе ДФХ мицелл пептид 536_2 образует α-спираль (остатки Phe2-Leu10) (Рисунок 19), пептид 3967 принимает αспиральную конформацию (остатки Ile3-Val9) (Рисунок 20). Пептид 12530 имеет С-концевой αвиток (Trp7-Phe10), в то время как остальная часть остова не упорядочена (Рисунок 21). Несмотря на заметную разницу между спектрами ЯМР в воде и растворе ДФХ мицелл, пептиды 756 и 9332 не имеют упорядоченной вторичной структуры как в водном, так и в мицеллярном окружении.



Рисунок 19 - Пространственная структура пептида 536_2 в растворе ДФХ мицелл (A) Набор из 20 ЯМР-структур, совмещенных по атомам основной α-спирали Phe2-Leu10 (ID PDB: 6RRO). Основные и боковые цепи показаны чёрным и светло-серым соответственно. (Б) Пространственная структура с наименьшей расчётной энергией. (В) Карта гидрофобности по шкале Кайта – Дулитла: от серого наиболее гидрофильного до жёлтого наиболее гидрофобного. * N-концевой аминокислотный остаток пептида; ** С-концевой аминокислотный остаток пептида. Иллюстрация разработана автором с использованием программы Chimera [244].


Рисунок 20 - Пространственная структура пептида 3967 в растворе ДФХ мицелл (A) Набор из 20 ЯМР-структур, совмещенных по атомам основной α-спирали Ile3-Val9 (ID PDB: 6RRL). Основные и боковые цепи показаны чёрным и светло-серым соответственно. (Б) Пространственная структура с наименьшей расчётной энергией. (В) Карта гидрофобности по шкале Кайта – Дулитла: от серого наиболее гидрофильного до жёлтого наиболее гидрофобного. * N-концевой остаток пептида; ** С-концевой остаток пептида. Иллюстрация разработана автором с использованием программы Chimera [244].



Рисунок 21 - Пространственная структура пептида 12530 в растворе ДФХ мицелл (A) Набор из 20 ЯМР-структур, совмещенных по атомам основной α-спирали Ile3-Val9 (ID PDB: 6RRL). Основные и боковые цепи показаны чёрным и светло-серым соответственно. (Б) Пространственная структура с наименьшей расчётной энергией. (В) Карта гидрофобности по шкале Кайта – Дулитла: от серого наиболее гидрофильного до жёлтого наиболее гидрофобного. * N-концевой остаток пептида; ** С-концевой остаток пептида. Иллюстрация разработана автором с использованием программы Chimera [244].

Методом ЯМР-спектроскопии была подтверждена α-спиральная структура для двух пептидов, 3967, 536_2, и было показано, что пептид 12530 на С-конце образует α-спираль в мембранно-миметических мицеллах ДФХ. Результаты анализа вторичной структуры с помощью применённых в работе предсказательных алгоритмов не согласуются с экспериментальными данными. Согласно онлайн-сервисам 4 пептида (3967, 8557, 12530 и 8361 2) из 12 способны формировать α-спираль.

Таким образом, полученные результаты свидетельствуют в пользу того, что алгоритм идентификации АМП должен быть доработан. Предполагается, что в процессе анализа данных необходимо рассчитывать вторичную структуру пептидов точнее. Тем не менее, созданный алгоритм идентификации АМП позволил обнаружить два перспективных антимикробных пептида, 3967 и 536_1. Для определения вероятного механизма действия этих молекул было изучено их взаимодействие с бактериями и липосомами, мимикрирующими мембраны эукариотических и бактериальных клеток.

3.8. Сканирующая электронная микроскопия клеток *E. coli* и *B. subtilis* после инкубации с антимикробными пептидами

Для амфипатических α-спиральных АМП известно, что, в основном, при взаимодействии патогеном, связываются бактериальной мембраной с пептилы с посредством электростатических взаимодействий с отрицательно заряженными липидными И полисахаридными компонентами бактериальных мембран [42,245]. Затем гидрофобные аминокислотные остатки АМП внедряются в липидный бислой, вызывая необратимые повреждения мембран за счёт образования пор. Описанное нарушение целостности клеточной мембраны приводит к деполяризации, утечке основных метаболитов из клетки. Этот механизм действия был продемонстрирован для мелиттина, магайнина 2 и кателицидинов [20,246]. Мы предположили, что схожим механизмом действия могут обладать АМП *H. medicinalis*. В начале было изучено влияние пептидов на бактериальные клеточные стенки.

Структурные изменения клеток, при инкубации с идентифицированными пептидами, оценивали с помощью сканирующей электронной микроскопии. Бактерии *E.coli* и *B. subtilis* инкубировали с пептидами 3967 и 536_1 в конечной концентрации 1/2×МИК и 1×МИК в течение 8 ч и 24 ч. Затем клетки отбирали, выполняли химическое высушивание образцов с использованием гексаметилдисилазана и анализировали с помощью микроскопа Zeiss Merlin, оснащенного GEMINI II Electron Optics (Zeiss, Германия).





Клетки инкубировали с пептидами в течение 8 ч при концентрации, соответствующей 1/2×МИК для каждого пептида (В, Г, Д, Е). В качестве интактного контроля использовались *B. subtilis* и *E. coli*, инкубируемые со средой МНВ в течение 8 часов (А, Б). Характерный масштаб - 1 мкм.

После 8 часов инкубации клеток с пептидами 3967 и 536_1 при концентрации 1/2×МИК значительно снижается количество бактерий и наблюдаются существенные структурные изменения клеток (Рисунок 22). Нарушение целостности клеточной мембраны приводило к образованию бактериальных агрегатов, в которых одиночные бактерии окружены клеточным дебрисом. Интересно, что при воздействии пептида 3967 на *B. subtilis* наблюдается резкое снижение количества клеток без видимых морфологических изменений клеточных мембран или образования клеточного дебриса. При концентрации пептидов, равной 1×МИК, не наблюдалось отличий между 8 и 24 часами инкубации.

3.9. Взаимодействие антимикробных пептидов с липосомами

Липосомы (замкнутый бислой, образующий сферическую частицу), имитирующие биологические мембраны, используются в качестве модельных объектов во многих исследованиях. Для изучения взаимодействия АМП с мембранами клеток, были получены большие моноламеллярные везикулы (БМВ), состоящие из липидных молекул и нагруженные флуоресцентной меткой — карбоксифлуоресцеином (СF). Полученные БМВ по составу имитировали клеточные мембраны прокариот (ПОФС, ПОФХ: ДОФГ (80:20), ПОФХ: ПОФС (80:20), ПОФХ: ПОФС (50:50), ПОФХ: ПОФС (75:25)) и эукариот (ПОФХ). БМВ инкубировали с пептидами в различных концентрациях и измеряли интенсивность флуоресценции в реальном времени. Мы наблюдали быстрое высвобождение СF из БМВ, состоящих из ПОФС, но медленное при взаимодействии с цвиттерионными везикулами (ПОФХ, ПОФС). Этот эффект свидетельствует о высокой скорости воздействия АМП на бактериальные клетки и медленном взаимодействии с клетками эукариот (Рисунок 23).



Рисунок 23 - Высвобождение карбоксифлуоресцеина в реальном времени из больших БМВ различного липидного состава, инкубируемых с пептидами 3967, 536_1 и мелиттином (А) везикулы из чистого ПОФС (Б) везикулы из ПОФХ: ПОФС (50:50) (В) везикулы из ПОФХ: ПОФС (80:20). На графике представлено соотношение F/F_0 , где F — интенсивность флуоресценции БМВ, обработанных пептидом, F_0 — интенсивность флуоресценции необработанных БМВ. Значения указаны как среднее значение \pm SD (n = 3).

Пермеабилизирующую активность определяли путём количественной оценки высвобождения CF через 1 мин после добавления пептида (Рисунок 24). Пептиды 3967 и 536_1 проявляют зависимую от концентрации повреждающую активность в отношении везикул с анионными липидами (ДОФГ, ПОФС), но не взаимодействуют с цвиттерионными везикулами (чистый ПОФХ), которые моделируют мембрану эукариотических клеток. Способность пептидов воздействовать на везикулы разного липидного состава сравнима со способностью известных катионных АМП, для которых показано, что они обладают мембранолитическим механизмом действия.



Рисунок 24 - Высвобождение карбоксифлуоресцеина из больших БМВ различного липидного состава, после добавления пептида 3967 (А, Б) и пептида 536_1 (В, Г) (синий кружок: чистый ПОФС; оранжевый ромб: чистый ПОФХ; зеленый квадрат: ПОФХ: ДОФГ (80:20); оранжевый треугольник: ПОФХ: ПОФС (80:20); голубой квадрат: ПОФХ: ПОФС (50:50); фиолетовый квадрат: ПОФХ: ПОФС (75:25)). КФ - карбоксифлуоресцеин.

Таким образом, с помощью СЭМ было показано, что пептиды 3967 и 536_1 в концентрации равной 1/2хМИК при инкубации с грамположительными и грамотрицательными бактериями приводят к дезинтеграции их клеточных мембран. При инкубации пептидов с липосомами, содержащими флуоресцентную метку, было продемонстрировано, что

исследуемые пептиды имеют большее сродство к липосомам, состав которых сходен с составом бактериальных клеточных мембран, а не эукариотических. Полученные результаты косвенно объясняют отсутствие цитотоксического эффекта пептидов в отношении клеток млекопитающих. Кроме того, результаты свидетельствуют в пользу того, что пептиды проявляют мембранолитический механизм действия при взаимодействии с бактериями.

ГЛАВА 4. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Системный анализ живых организмов с применением вычислительных методов и алгоритмов в настоящее время представляется весьма многообещающей стратегией для поиска и изучения новых биологически активных веществ. В частности, такой подход уже был применён для поиска АМП. Например, анализ генома и транскриптома головастиков жабы Rana [Lithobates] catesbeiana позволил определить семь АМП, обладающих терапевтическим потенциалом [247]. Идентифицированные пептиды проявляли бактериостатическую и бактерицидную активность в отношении микобактерий, но не грамотрицательных или грамположительных бактерий. Также, метод поиска АМП in silico был применён к полученному de novo транскриптому таракана P. americana [240]. Такой подход позволил обнаружить 86 кандидатных АМП. Одиннадцать из идентифицированных пептидов показали сильную антимикробную активность против дрожжей, грамположительных и грамотрицательных бактерий. Эти пептиды также проявили незначительную цитотоксическую и гемолитическую активности.

В последнее время внимание исследователей привлекают гематофаги, животные, питающиеся кровью других животных и человека. Известно, что многие гематофаги являются переносчиками ряда заболеваний, вызываемых вирусами, бактериями, простейшими и гельминтами, сами при этом не подвергаясь воздействию этих патогенов [248]. В практическом плане системные исследования гематофагов имеют большое значение для медицины и ветеринарии [249–251].

H. medicinalis - единственные животные среди гематофагов, которые, используются в лечебной практике на протяжении тысячелетий [156,252]. Гирудотерапия применяется при лечении воспалительных и сердечно-сосудистых заболеваний, тромбозов глубоких вен, остеоартрита, осложнений сахарного диабета [253,254]. Также гирудотерапия широко используется в пластической и реконструктивной хирургии. Известно более 20 белков ССК Н. medicinalis, а именно, гирудин, гиалуронидаза, калин, дестабилаза, антистазин, апираза, эглины, декорсин, бделлины, гуамерин, саратин, гирустатин, ингибиторы триптазы И гистаминоподобные вещества, ингибиторы комплемента, ингибиторы карбоксипептидазы А [157,160,165,185-190]. ССК *Н. medicinalis* обладает тромболитическим, антикоагулянтным, сосудорасширяющим, противовоспалительным действием, проявляет анальгетические свойства, ингибирует адгезию и агрегацию тромбоцитов, регулирует функции тромбина, а также вызывает разрушение внеклеточного матрикса [149,156]. Показано, что ССК *H. medicinalis* обладает противомикробным действием [183].

АМП H. medicinalis защищают пищеварительную систему пиявки от негативного

воздействия патогенов, попавших внутрь при кормлении. АМП выполняют роль антисептиков и консервантов крови, дезинфицируя и сохраняя кровь жертвы. *Н. medicinalis* после кормления способна в течение длительного времени (вплоть до шести месяцев) хранить кровь интактной. Таким образом, можно предположить, что существующие АМП пиявки, вероятно, обладают пониженной гемолитической активностью и токсичностью. Кроме того, АМП *Н. medicinalis*, возможно, контролируют внутреннюю микробиоту червя.

Для поиска новых АМП *H. medicinalis* нами была выбрана стратегия системного анализа организма с использованием данных высокопроизводительного секвенирования и протеомного анализа. На данный момент изучены транскриптомы слюнных клеток нескольких пиявок и транскриптомы нервной системы медицинских пиявок [1–3,7]. Ранее, исследователи выполнили попытку получить и расшифровать нуклеотидную последовательность генома *H. robusta*, но полноценного исследования, аннотации и анализа генной онтологии не было проведено [10]. В нашей лаборатории ранее впервые была получена последовательность митохондриального генома *M. verbana* и *H. medicinalis* [9]. Впервые проведено секвенирование и аннотация генома *H. medicinalis* с последующим анализом протеома секрета слюнных клеток медицинской пиявки [215].

На первом этапе была определена нуклеотидная последовательность генома *H. medicinalis*. Геном медицинской пиявки был получен методом гибридной сборки на основе секвенирования трёх ДНК-библиотек: ДНК-библиотеки коротких прочтений (Ion Proton, #SRR6926478¹), парной ДНК-библиотеки с длиной прочтений 3-6 т.п.н. (Ion TrueMate, #SRR6926477¹) и парной ДНК-библиотеки с длиной прочтений 8-12 т.п.н. (Illumina TruSeq, #SRR6926477¹) с использованием геномного сборщика SPAdes [206] (Таблица 5, Таблица А.1 (Приложение А)).

ДНК-библиотека коротких прочтений с низким уровнем ошибок использовалась для посторения графа сборки, на который потом выравнивались длинные прочтения парных ДНКбиблиотек. Таким образом, была решена проблема пробелов в геноме и неопределённостей, вызванных повторяющимися элементами. Затем была проведена процедура биннинга, которая представляет собой анализ и кластеризацию для дифференцировки эукариотических и прокариотических контигов. Результаты биннинга представлены на двумерной диаграмме в пространстве покрытия прочтений и GC-состава ДНК-фрагментов (Рисунок 10). В ходе сравнения экспериментальных данных было показано, что контиги с наибольшим покрытием, принадлежат геному медицинской пиявки, в то время как контиги с более высоким уровнем

¹ – идентификационный номер нуклеотидной последовательности в Архиве последовательностей прочтений (Sequence Read Archive, SRA)

ошибок образуют кластер ДНК-фрагментов прокариот, которые были затем исключены из дальнейшего исследования. Таким образом, были выделены только эукариотические контиги, которые подверглись процедуре скаффолдинга - упорядочивании контигов, разделённых пробелами в покрытии. В результате сборки было получено 14042 скаффолдов, которые содержали 14596 белок-кодирующих генов (Таблица А.2 (Приложение А)). Было отсеквенировано 187,5 Мп.н., что составляет 85% теоретического размера генома (http://www.genomesize.com/result_species.php?id=1061). Сравнение генома *H. medicinalis* с геномами гематофагов и спиральнодробящихся животных в целом показало, что геном *H. medicinalis* является наименьшим как по размеру, так и по количеству генов (Рисунок 12). Тем не менее средняя длина генов у *H. medicinalis* почти в два раза больше, чем у *H. robusta* и *C. teleta*, но меньше, чем у других представителей спиральнодробящихся.

Впервые была проведена аннотация генома медицинской пиявки. С помощью программы OrthoFinder был проведён анализ ортологичных классов генов группы кольчатых червей (*H. medicinalis, H. robusta, C. teleta*), группы плеченогих (*L. anatina*), группы моллюсков (*O. bimaculoides, A. californica, C. gigas*) и контрольной группы, группы позвоночных (*D. rerio* и *H. sapiens*) [235]. Была проведена кластеризация генов всех групп с помощью программного обеспечения OrthoMCL [255]. Как видно на Рисунке 13, все четыре группы имеют 7520 общих генов. Диаграмма демонстрирует близость родственных таксонов пиявок и спиральнодробящихся и их отдалённость от группы позвоночных.

Полученные результаты имеют важное значение для понимания роли структуры генома в регуляции экспрессии генов, связанных с образом жизни *H. medicinalis* и эволюционной адаптации пиявки к питанию кровью теплокровных животных. В результате аннотации генома *H. medicinalis* получены нуклеотидные последовательности генов, кодирующих уникальные, ранее не изученные белки, которые могут быть использованы для разработки перспективных лекарственных препаратов.

На следующем этапе был выполнен анализ протеома ССК медицинской пиявки. Метод сбора ССК медицинской пиявки, разработанный нами ранее, позволил эффективно и с минимальным загрязнением собрать секрет. Известно, что слюнные клетки медицинской пиявки выделяют сложный по своему составу секрет, компоненты которого обладают различными функциями [149,256]. В ССК *Н. medicinalis* содержатся ферменты, ингибиторы протеаз, белки, участвующие в адгезии, вспомогательные белки (Рисунок 14, Таблица В.1 (Приложение В)).

Самой многочисленной группой белков являются цистеин-богатые секреторные белки (англ. <u>cysteine rich secretory protein</u>, *CRISP*). Представители суперсемейства CRISP широко

81

распространены среди всех основных таксонов ядовитых животных, от змей до улитокхищников конусов и книдарий, включая многощетинковых червей [257-260]. Немногие из этих белков функционально охарактеризованы, но те, что исследованы, проявляют различную активность. CRISP чаще всего являются нейротоксинами, способны блокировать ионные каналы, подавлять ангиогенез, увеличивать проницаемость сосудов и способствовать воспалительным реакциям [257,261]. Среди белков ССК были обнаружены цистеин-богатые антистазины, известные ингибиторы сериновых протеаз. Антистазины являются сильными антикоагулянтами и играют ключевую роль в ингибирования системы свёртывания крови жертвы [160,165,186]. Помимо антистазинов, были обнаружены и другие ингибиторы протеаз. Цистатины ингибируют цистеиновые протеазы и играют важную роль в иммунитете пиявок [262,263]. Эглин-подобные белки являются ингибиторами сериновых протеаз, а именно ахимотрипсина, химазы, субтилизина, эластазы и катепсина G [156,162,165]. В результате эглины нарушают хемотаксис нейтрофилов и запускают защитные реакции, которые способствуют воспалению. Ещё одним важным белком, является альфа-2-макроглобулин, ингибитор различных протеаз [264]. Альфа-2-макроглобулин подавляет компоненты системы комплемента и системы гемостаза [265].

Отдельно хотелось бы упомянуть белки ССК, вовлечённые в процесс регулирования системы гемостаза жертвы и переваривания крови. Карбоангидраза участвует в регуляции рН крови, пищеварительного трака и других тканей, а также у гематофагов поддерживает оптимальные условия для переваривания крови [266–268]. Карбоангидраза вызывает местное усиление ацидоза в месте укуса, снижая активность факторов свёртывания крови. Фиколин является компонентом врождённой иммунной системы и запускает лектин-зависимый путь активации комплемента [269]. Фиколин-подобные белки у рептилий способны вызывать агрегацию тромбоцитов и активировать процесс свёртывания крови [270]. Кальмодулин, консервативный многофункциональный кальций-связывающий белок, регулирует коллагензависимую агрегацию тромбоцитов [271].

Существенная часть идентифицированных белков играет важную роль в антикоагуляции и взаимодействии с жертвой [272,273]. Долгое время считалось, что основными молекулами мишенями для компонентов ССК являются тромбин и фактор Ха [175,186], но, как и у других гематофагов, в ССК обнаруживается сложная смесь антигемостатических молекул, ориентированных на ингибирование вторичного и первичного гемостаза. Также были обнаружены консервативные белки, обладающие широким спектром действия. Эти белки участвуют в межклеточных взаимодействиях, опосредуют сигнализацию и адгезию клеток. Важно отметить, что *H. medicinalis* в процессе питания приходится взаимодействовать с

жертвой продолжительное время [149,274]. Долгий процесс кормления должен быть незаметным для жертвы и его иммунной системы. Вовлечение в процесс компонентов иммунной системы организма жертвы может привести к воспалительному процессу и изменению состава крови, что сделает её непригодной для питания. Эффективное ингибирование факторов свёртывания и иммунной системы жертвы имеет прямое отношение к процессу хранения крови и её переваривания. Белки ССК *Н. medicinalis* смешиваются с кровью, тем самым подготавливая её для дальнейшего хранения и консервации в кропе червя.

Таким образом, анализ протеома ССК *Н. medicinalis* подтвердил наличие высоко консервативных белков, с помощью которых червь воздействует на процессы гемостаза и иммунитета жертвы. Как видно из Рисунка 14 в ССК содержатся ранее не изученные молекулы, роль и функции которых не определены.

Цель данной работы заключалась в поиске новых антимикробных пептидов Н. medicinalis. Ранее было охарактеризовано семь пиявочных антимикробных пептидов (Таблица 3), среди которых только три пептида были обнаружены у *H. medicinalis*: Hm-теромацин, Hm-Нт-люмбрицин [192,193,196]. Поэтому был нейромацин И сначала выполнен целенаправленный поиск известных генов АМП в геноме *H. medicinalis*. В результате анализа было обнаружено 9 последовательностей-гомологов, а именно, четыре гомолога Нтнейромацина, четыре гомолога Hm-люмбрицина и гомолог Hm-теромацина (Рисунок 15). Среди найденных последовательностей только два пептида-гомолога Нт-нейромацина показали процент гомологии равный 98,3%, для остальных пептидов он составил 67,2% и ниже.

Использование вычислительных методов для предсказания новых АМП сокращает время и стоимость экспериментального анализа. На данный момент существует большое разнообразие алгоритмов и методов для идентификации новых активных молекул [117,275,276]. Доступные онлайн базы данных регулярно обновляются, что позволяет не только систематизировать полученную информацию, но и также служит необходимым базисом для построения прогностических алгоритмов [143–145,277]. В последнее время все больше исследований посвящены анализу компонентного состава секретов и ядов организмов, которые являются перспективными источниками прототипов лекарственных средств [278,279]. Для идентификации новых АМП был выполнен биоинформатический анализ последовательности генома *H. medicinalis*.

Анализ проводили по алгоритму, который является модификацией метода, описанного в исследованиях по поиску новых антимикробных молекул в транскриптомных данных таракана *P. americana* и многоножки *S. subspinipes mutilans* [240,241]. Поиск антимикробных пептидов заключался в обнаружении катионных амфифильных молекул, физико-химические свойства

83

которых соответствуют таковым известных антимикробных пептидов. Вычислительный алгоритм включает в себя расчёт физико-химических свойств, структурных особенностей, склонности к агрегации, а также определение антимикробного потенциала пептида. Последовательность алгоритма продемонстрирована на Рисунке 16. Было предположено, что αспиральные АМП, идентифицированные в медицинской пиявке, в отличие от других организмов, могут быть менее токсичными из-за эволюционной приспособленности *H. medicinalis* к питанию кровью.

Для идентификации последовательностей потенциальных АМП в геноме *H. medicinalis*, выполненной в работе, были получены все кодирующие последовательности. Аминокислотные последовательности были разделены на пептиды и были определены их физико-химические свойства с помощью алгоритма, написанного нами на языке программирования R. В анализе оставили только те пептидные последовательности, по физико-химическим характеристикам соответствующие известным АМП. Затем, применяя онлайн алгоритмы и предикторы для идентификации последовательностей обладающих антимикробным потенциалом, были отобраны предположительные АМП. Далее пептиды были верифицированы относительно протеома ССК медицинской пиявки. Были отобраны только те пептиды, последовательности которых были найдены в протеоме ССК *H. medicinalis*. На заключительном этапе биоинформатического анализа для последующего химического синтеза были отобраны двенадцать кандидатных последовательностей АМП с самым высоким противомикробным потенциалом согласно примененным в алгоритме фильтрам. Последовательности пептидов и некоторые их физико-химические характеристики представлены в Таблице 6.

На следующем этапе работы необходимо было оценить функциональную активность идентифицированных АМП *H. medicinalis* для проверки сконструированного алгоритма поиска АМП. Антимикробную активность АМП *H. medicinalis* определяли в отношении грамположительных *B. subtilis* и грамотрицательных бактерий *E. coli* и внутриклеточного патогена *C. trachomatis* (Таблица 7). Было обнаружено, что тестируемые пептиды различаются по своей способности ингибировать рост этих бактерий, и, в целом, все пептиды являются более активными в отношении бактерий *B. subtilis* и *C. trachomatis*. Два пептида - 3967 и 536_1, продемонстрировали антимикробную активность против всех трёх видов бактерий. Значение МИК для пептида 3967 в отношении *B. subtilis, E. coli* составило 10,3 мкМ, в отношении *C. trachomatis* 3,1 мкМ. Для пептида 536_1 — 8,6 мкМ, 17,2 мкМ и 6,3мкМ соответственно. Стоит отметить, что антимикробная активность АМП *H. medicinalis*, сравнима с таковой известных АМП. Например, ранатуерин-1, антимикробный пептид из лягушки *R. catesbeiana*, ингибирует рост *E.coli* и *S. aureus* при значениях МИК 48 и 97 мкМ соответственно [247]. Для липопептида

даптомицина, выделенного из почвенного микроорганизма *Streptomyces roseosporus*, одобренного FDA для лечения взрослых с бактериемией *S. aureus*, значение МИК для *S. aureus* составляет 0,7 мкМ [61]. Пексиганан, аналог антимикробного пептида магаинина-2, найденного на коже африканской лягушки *X. laevis*, убивает *Klebsiella pneumonia* при значении МИК 6,5 мкМ, а концентрация пептида, необходимая для ингибирования роста метициллинрезистентных *S. aureus*, составляет 26 мкМ [280]. Идентифицированные пептиды *H. medicinalis* 3967 и 536_1 продемонстрировали также низкую гемолитическую и цитотоксическую активности (Таблица 8, 9).

Для проверки способности биоинформационного алгоритма идентифицировать аминокислотные последовательности, формирующие α-спиральную структуру, была исследована вторичная структура АМП *Н. medicinalis*. Известно, что небольшие линейные АМП обычно не структурированы в физиологическом растворе, и первоначально они взаимодействуют с клеточными мембранами электростатически из-за их положительного заряда [72,73]. Вторичная структура АМП *Н. medicinalis* была проанализирована с помощью ЯМР-спектроскопии растворе В водном растворе И с добавлением додецилфосфатидилхолиновых мицелл для имитации бактериальной мембраны. Все пептиды не структурированы в водном растворе, но в мицеллярном растворе пептиды 536 2 и 3967 принимают α-спиральную конформацию (Рисунок 19, 20. Однако, ЯМР-анализ вторничной структуры пептида 3967 противоречит некоторым предыдущим исследованиям, в которых утверждается, что АМП, обладающие высокой гемолитической активностью, в основном имеют α-спиральную структуру [281,282]. Исследования АМП такого класса с использованием эукариотических мембран или имитирующих их цвиттерионных липосом, позволяют предположить, что α-спиральная структура необходима для реализации антимикробного действия катионных АМП, механизм действия которых заключается в формировании пор в бактериальных мембранах. Но существуют и обратные примеры, такие как, плеуроцидин, антимикробный пептид из камбалы Pseudopleuronectes americanus, который проявляет сильную антибактериальную активность в отношении грамположительных и грамотрицательных бактерий [283]. принимает α-спиральную Сам пептид структуру в мицеллах додецилфосфохолина [284]. Было показано, что пептиды, полученные из плеуроцидина, ингибируют рост E. coli за счёт нарушения синтеза белков клеточной стенки [285,286]. Идентифицированный нами пептид 3967, обладающий заметной антимикробной активностью, принимает α-спиральную конформацию, в то же время проявляет низкую гемолитическую и цитотоксическую активность.

В соответствии с вычислительным алгоритмом поиска все потенциальные АМП должны

принимать α-спиральную конформацию (Рисунок 18). Согласно онлайн-сервисам, I-TASSER и PEP-FOLD, использованным нами для предсказания вторичной структуры AMΠ *H. medicinalis*, только 4 пептида (3967, 8557, 12530 и 8361_2) из 12 кандидатных способны формировать αспираль [219,220]. Результаты анализа пространственной структуры методом ЯМРспектроскопии подтвердили α-спиральную структуру только для двух пептидов, 3967 и 536_2. Пептид 12530 имеет С-концевой α-виток, в то время как остальная часть остова не упорядочена (Рисунок 21). Таким образом, был сделан вывод, что расчёт физико-химических свойств пептидов для определения вторичной структуры, реализованный в вычислительном алгоритме, не обладает достаточной эффективностью. Было высказано предположение, что для предсказания вторичной структуры необходимы более сложные подходы, такие как молекулярно-динамическое моделирование [287,288].

В соответствии с механизмом действия амфипатических α-спиральных АМП, пептиды связываются с бактериальной мембраной посредством электростатических взаимодействий с отрицательно заряженными липидными и полисахаридными компонентами на поверхности бактериальных клеток [20,289,290]. Затем гидрофобные аминокислотные остатки АМП внедряются в липидный бислой, вызывая разрушение бактериальной мембраны за счёт образования пор или каналов. Это нарушение приводит к деполяризации мембраны, утечке основных метаболитов из клетки. Так действуют многие АМП, например, мелиттин, магайнин 2 и кателицидин [20,246].

Анализ с использованием сканирующей электронной микроскопии подтвердил, что пептиды 3967 и 536 1 действуют на мембраны бактерий в основном путем нарушения её целостности, что приводит к гибели бактериальных клеток. Более того, это предположение было подтверждено анализом мембранно-повреждающей способности пептидов против больших моноламеллярных везикул (БМВ), моделирующих мембраны эукариотических и бактериальных клеток (Рисунок 23, 24). Пептиды 3967 и 536 1 не взаимодействуют с цвиттерионными везикулами, составленными из ПОФХ, что довольно типично для катионных АМП. Тем не менее, пептиды повреждают отрицательно заряженные липосомы. Было обнаружено, что способность пептидов 3967 и 536 1 разрушать мембраны сравнима со способностью других катионных АМП, для которых доказана мембрано-повреждающая активность. Например, пептид темпорин вызывает 100%-ую утечку карбоксифлуоресцеина из липосом состава ДОФХ/ДОФГ (1:1) при концентрации 3 мкМ [291]. Наблюдаемое явление подтверждает общий действия амфипатических пептидов, основанный способ на внутримолекулярных электростатических взаимодействиях катионных остатков АМП с анионными липидами мембран бактериальных клеток. Предполагается, что высокая амфипатичность приводит к увеличению антимикробной и цитотоксической активности, но именно токсичность является основным препятствием для клинического применения АМП. Тем не менее, в нескольких исследованиях было показано, что несовершенные амфифильные пептиды более активны, чем соответствующие совершенные амфипатические пептиды [42,292]. Однако, токсичность обнаруженных АМП не зависела от величины амфипатичности. Несмотря на обширные исследования антимикробных пептидов, в клинической практике сейчас используется не более 10 противомикробных пептидов [34,293]. Это в первую очередь связано с высокой токсичностью антимикробных пептидов и их нестабильностью. Несмотря на то, что практическая применимость идентифицированных и изученных в работе антимикробных пептидов должна быть подтверждена путём дальнейшего тестирования *in vivo*, данная работа представляет собой первый шаг в процессе создания АМП с пониженной гемолитической и цитотоксической активностью, но которые проявляют высокую антимикробную активность.

ЗАКЛЮЧЕНИЕ

АМП природного происхождения являются перспективными соединениями для разработки лекарственных препаратов для лечения заболеваний, вызванных антибиотикоустойчивыми бактериями. Основным недостатком АМП является их токсичность по отношению к клеткам эукариот. Мы предположили, что медицинская пиявка *H. medicinalis*, ССК которой обладает полифункциональным, в том числе и антимикробным действием, может послужить источником новых АМП с пониженной цитотоксичностью. Для поиска и описания новых биологически активных веществ мы применили системный «омиксный» анализ Н. medicinalis. Впервые была получена нуклеотидная последовательность генома H. medicinalis, была проведена аннотация генома *H. medicinalis*, выполнен анализ белкового состава ССК *H.* medicinalis. Полученные данные представляют собой базу данных нуклеотидных последовательностей, кодирующих уникальные белки, для многих из которых функции еще не изучены. Дальнейшее изучение белков ССК Н. medicinalis позволит более полно охарактеризовать механизмы с помощью которых гематофаг воздействует на процессы гемостаза и иммунитета жертвы.

В настоящей работе был оптимизирован биоинформатический алгоритм для идентификации АМП в геноме *H. medicinalis*. Алгоритм нацелен на поиск нуклеотидных последовательностей, кодирующих короткие положительно заряженные пептиды со способностью к образованию α-спиральных структур. Вычислительный алгоритм объединяет расчёт физико-химических свойств, вторичной структуры пептида, а также определение его антимикробного потенциала. С помощью алгоритма в геноме *H. medicinalis* были идентифицированы двенадцать последовательностей, кодирующих потенциальные АМП. Они были химически синтезированы, был проведён анализ их антимикробной, гемолитической и цитотоксической активности.

В результате, два пептида, 3967 (FRIMRILRVLKL) и 536_1 (RWRLVCFLCRRKKV), ингибируют рост грамположительных (*B. subtilis*) и грамотрицательных бактерий (*E. coli, C. thrachomatis*) при значении МИК порядка 10 мкМ, обладают низкой гемолитической активностью и не токсичны для клеток линии McCoy. Методом ЯМР-спектроскопии было показано, что пептиды 536_2 и 3967 принимают α -спиральную конформацию в мембраномиметических мицеллах ДФХ. Для наиболее активных АМП *Н. medicinalis* был изучен их вероятный механизм действия. Показано, что пептиды 3967 и 536_1 проявляют низкое сродство к бислойным мембранам и способны полностью разрушать анионные липосомы, что даёт возможность предположить, что эти пептиды проявляют мембранолитический механизм действия и снижают выживаемость клеток *E. coli* и *B. subtilis*, в основном, за счёт нарушения целостности бактериальных мембран. Данные исследования активности найденных АМП *H. medicinalis* подтверждают хорошую прогностическую способность биоинформатического алгоритма поиска новых АМП. В дальнейшем, предсказательный алгоритм может быть использован для анализа геномов других организмов. Идентифицированные в ходе работы пептиды 3967 и 536_1, благодаря широкому спектру действия и отсутствию токсических эффектов на клетки эукариот, являются перспективной основой для разработки антимикробного препарата широкого спектра действия.

выводы

1. Определена нуклеотидная последовательность генома *H. medicinalis*, размер которого составил 187,5 Мп.н. и включает в себя 14596 белок-кодирующих генов. Анализ состава ССК *H. medicinalis* позволил идентифицировать в секрете 189 белковых молекул.

2. Оптимизирован алгоритм поиска АМП в геномных данных, заключающийся в идентификации коротких положительно заряженных пептидов с тенденцией к образованию αспиральных структур. С использованием алгоритма было отобрано 12 потенциальных антимикробных пептидов.

3. Анализ антимикробной активности пептидов показал, что восемь из двенадцати обладают различной антимикробной активностью в отношении *E. coli*, *B. subtilis* и *C. thrachomatis*. Пептиды 3967 и 536_1 обладали наибольшей антимикробной активностью. Все пептиды не цитотоксичны и проявили низкую гемолитичную активность, за исключением пептидов 8557 и 3063.

4. Методом ЯМР-спектроскопии α-спиральная структура подтверждена для пептидов 3967, 536 2 и 12530. Пептиды 756 и 9332 не структурированы.

5. Пептиды 536_1 и 3967 вызывают дезинтеграцию бактериальных клеточных стенок и проявляют низкое сродство к цвиттерионным везикулам, но полностью разрушают анионные везикулы.

БЛАГОДАРНОСТИ

Автор диссертации выражает благодарность научному руководителю, Лазареву Василию Николаевичу за его помощь и неоценимые советы. Автор также благодарен коллегам, которые участвовали в совместных исследованиях и обсуждениях, без которых данная работа не была бы возможной. Отдельную благодарность автор выражает своей семье и близким друзьям, которые поддерживали его в каждый момент этого длинного пути.

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

АДФ - аденозин-5'-дифосфата (АДФ)

АМП - антимикробный пептид

БМВ - большие моноламеллярные везикулы

ВИЧ - вирус иммунодефицита человека

ВЭЖХ - высокоэффективная жидкостная хроматография

ГТАБ - гексадецилтриметиламмонийбромид

ДНК - дезоксирибонуклеиновая кислота

ДОФГ - 1,2-диолеоил-глицеро-3-фосфоглицерол

ДФХ - додецилфосфатидилхолин

КОЕ/мл - колониеобразующих единиц в мл

ЛПС - липополисахарид

МИК - минимальная ингибирующая концентрация

МПК90 - минимальная концентрация пептидов, при которой уровень инфекции был меньше 90% от уровня в контрольных образцах

ПОФС - 1-пальмитоил-2-олеоил-глицеро-3-фосфо-L-серин

ПОФХ - 1-пальмитоил-2-олеоил-глицеро—3-фосфохолин

РНК - рибонуклеиновая кислота

ССК - секрет слюнных клеток

УФ - ультрафиолет

ЭДТА - этилендиаминтетрауксусная кислота

ЭТ - элементарные тельца

ЯМР - ядерный магнитный резонанс

ACN - (англ. acetonitrile) ацетонитрил

ADAM - (англ. A Database of Anti-Microbial peptides)

APD - (англ. Antimicrobial Peptide Database) база данных антимикробных пептидов

BLAST - (англ. *Basic Local Alignment Search Tool*) семейство компьютерных программ для поиска локальных соответствий между последовательностями

САМР_{R3} - (англ. Collection of Anti-Microbial Peptides) коллекция антимикробных пептидов

CF - (англ. carboxyfluorescein) карбоксифлуоресцеин

CRISP - (англ. cysteine rich secretory protein) цистеин-богатые секреторные белки

DADP - (англ. *Database of Anuran Defense Peptides*) база данных антимикробных пептидов земноводных

DCNa - (англ. sodium deoxycholate) дезоксихолат натрия

DIC - (англ. *N*,*N'-diisopropylcarbodiimide*) N, N'-диизопропилкарбодиимид

DMEM - (англ. Dulbecco's Modified Eagle Medium) модифицированная по способу Дульбекко среда Игла

DMF - (англ. *N*,*N*-dimethyl formamide) N,N-диметилформамид

DTT - (англ. dithiothreitol) дитиотреитол

FASP - (англ. *filter-aided sample preparation*) пробоподготовка с использованием центрифужных фильтров

FDA - (англ. *Food and Drug Administration*) управление по контролю качества пищевых продуктов и лекарственных препаратов США

FDR - (англ. false discovery rate) уровень ложноположительных результатов

Fmoc - (англ. 9-fluorenylmethyloxycarbonyl) 9-флуоренилметилоксикарбонил

FITC - (англ. fluorescein isothiocyanate) флуоресцеинизотиоционат

HBSS - (англ. Hank's balanced salt solution) сбалансированный солевой раствор Хэнкса

HMDS - (англ. hexamethyldisilazane) гексаметилдисилазан

HmN - Нт-нейромацин, выделенный из пиявки H. medicinalis

HmTh - Hm-теромацин, выделенный из пиявки H. medicinalis

Hmc - гирудомацин, выделенный из пиявки H. medicinalis

IAA - (англ. iodacetic acid) йодуксусная кислота

IFU - (англ. inclusion forming units) включениеобразующие единицы

LDTI - (англ. leech-derived tryptase inhibitor) специфический ингибитор триптазы

LSTM - (англ. *generative long short-term memory*) рекуррентные нейронные сети с генеративной долгой краткосрочной памятью

LTR - (англ. long terminal repeat, LTR) длинные концевые повторы

4M-pip - (англ. 4-methyl piperidine) 4-метилпиперидин

МНВ - (англ. Mueller Hinton Broth) питательный бульон Мюллера-Хинтона

MEM - (англ. Minimum Essential Medium) минимальная среда, среда Игла

NCBI - (англ. National Center for Biotechnology Information, NCBI) Национальный центр биотехнологической информации

NOESY - (англ. nuclear Overhauser effect spectroscopy) - спектроскопия ядерного эффекта Оверхаузера

non-LTR - (англ. *long interspersed nuclear element, LINE* и *short interspersed nuclear element, SINE*) длинные и короткие диспергированные ядерные элементы

OXYMA - (англ. ethyl cyanohydroxyiminoacetate) этил-(гидроксиимино)-цианоацетат

PBS - (англ. phosphate-buffered saline) фосфатно-солевой раствор

PDB - (англ. Protein Database) База данных структуры белков

ROESY - (англ. rotating-frame nuclear Overhauser effect correlation spectroscopy) спектроскопия ядерного эффекта Оверхаузера во вращающейся системе координат

SDS - (англ. sodium dodecylsulfate) додецилсульфат натрия

SPG - (англ. sodium phosphate glutamate) сахарозо-фосфат-глутаматный буфер

SRA - (англ. Sequence Read Archive) Архив последовательностей прочтений

TFA - (англ. trifluoroacetic acid) трифторуксусная кислота

TOCSY - (англ. total correlation spectroscopy) полная корреляционная спектроскопия

TrisHCl - (англ. Tris, THAM) — кислая соль трис(гидроксиметил)аминометана

TRITC - (англ. tetramethylrodamine B isothiocyanate) - тетраметилродамин

TtTh - Tt-теромацин, выделенный из пиявки Theromyzon tessulatum

UniProtKB - открытая база данных последовательностей белков

СПИСОК ЛИТЕРАТУРЫ

- 1. Amorim A.M.X.P. et al. Transcripts involved in hemostasis: Exploring salivary complexes from Haementeria vizottoi leeches through transcriptomics, phylogenetic studies and structural features // Toxicon. 2015. Vol. 106. P. 20–29.
- 2. Lu Z. et al. Transcriptomic analysis of the salivary gland of medicinal leech Hirudo nipponia // PLoS One. 2018. Vol. 13, № 10. P. 1–14.
- Khan M.S. et al. Transcriptomics and differential gene expression in Whitmania pigra (Annelida: Clitellata: Hirudinida: Hirudinidae): Contrasting feeding and fasting modes // Ecol. Evol. 2019. Vol. 9, № 8. P. 4706–4719.
- 4. Kvist S. et al. Pyrosequencing the salivary transcriptome of Haemadipsa interrupta (Annelida: Clitellata: Haemadipsidae): Anticoagulant diversity and insight into the evolution of anticoagulation capabilities in leeches // Invertebr. Biol. 2014. Vol. 133, № x. P. 74–98.
- 5. Min G.-S., Sarkar I.N., Siddall M.E. Salivary Transcriptome of the North American Medicinal Leech, Macrobdella decora // J. Parasitol. 2010. Vol. 96, № 6. P. 1211–1221.
- 6. Hibsh D. et al. De novo transcriptome assembly databases for the central nervous system of the medicinal leech // Sci. Data. 2015. Vol. 2. P. 150015.
- Northcutt A.J. et al. An annotated CNS transcriptome of the medicinal leech, Hirudo verbana: De novo sequencing to characterize genes associated with nervous system activity // PLoS One. 2018. Vol. 13, № 7. P. 1–33.
- Oceguera-Figueroa A. et al. Comparative mitogenomics of leeches (Annelida: Clitellata): Genome conservation and placobdella-specific trnD gene duplication // PLoS One. 2016. Vol. 11, № 5. P. 1–16.
- 9. Nikitina A. et al. Draft mitochondrial genomes of Hirudo medicinalis and Hirudo verbana (Annelida, Hirudinea) // Mitochondrial DNA Part B Resour. 2016. Vol. 1, № 1. P. 254–256.
- 10. Simakov O. et al. Insights into bilaterian evolution from three spiralian genomes // Nature. 2013. Vol. 493, № 7433. P. 526–531.
- 11. Maltz M.A. et al. Metagenomic analysis of the medicinal leech gut microbiota. // Front. Microbiol. 2014. Vol. 5, № April. P. 151.
- 12. Ott B.M. et al. Characterization of shed medicinal leech mucus reveals a diverse microbiota. // Front. Microbiol. 2014. Vol. 5, № January. P. 757.
- 13. Worthen P.L., Gode C.J., Graf J. Culture-independent characterization of the digestive-tract microbiota of the medicinal leech reveals a tripartite symbiosis // Appl. Environ. Microbiol. 2006. Vol. 72, № 7. P. 4775–4781.
- M. Zasloff. Antimicrobial peptides of multicellular organisms. // Nature,. 2002. Vol. 415. P. 389–395.
- 15. Uvell H., Engström Y. A multilayered defense against infection: combinatorial control of insect immune genes // Trends Genet. 2007. Vol. 23, № 7. P. 342–349.
- 16. Nair D.G. et al. Antimicrobial activity of omwaprin, a new member of the waprin family of snake venom proteins // Biochem. J. 2007. Vol. 402, № 1. P. 93–104.
- 17. Budagavi D.P., Chugh A. Antibacterial properties of Latarcin 1 derived cell-penetrating peptides // Eur. J. Pharm. Sci. Elsevier B.V, 2018. Vol. 115, № 2017. P. 43–49.
- 18. Ponnappan N., Chugh A. Cell-penetrating and cargo-delivery ability of a spider toxin-derived peptide in mammalian cells // Eur. J. Pharm. Biopharm. Elsevier B.V., 2017. Vol. 114. P. 145–153.
- 19. Brogden K.A. et al. Antimicrobial peptides in animals and their role in host defences // Int. J. Antimicrob. Agents. 2003. Vol. 22, № 5. P. 465–478.
- 20. Brogden K.A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? // Nat. Rev. Microbiol. 2005. Vol. 3, № 3. P. 238–250.

- 21. Müller C.A. et al. Human α-defensins HNPs-1, -2, and -3 in renal cell carcinoma: Influences on tumor cell proliferation // Am. J. Pathol. 2002. Vol. 160, № 4. P. 1311–1324.
- 22. Lima S.M.F. et al. Antimicrobial and immunomodulatory activity of host defense peptides, clavanins and LL-37, in vitro : An endodontic perspective // Peptides. Elsevier, 2017. Vol. 95, № July. P. 16–24.
- 23. Steinstraesser L. et al. Host defense peptides in wound healing // Mol. Med. 2008. Vol. 14, № 7-8. P. 528-537.
- 24. Yang D. et al. Multiple Roles of Antimicrobial Defensins, Cathelicidins, and Eosinophil-Derived Neurotoxin in Host Defense // Annu. Rev. Immunol. 2004. Vol. 22, № 1. P. 181–215.
- 25. Davidson D.J. et al. The Cationic Antimicrobial Peptide LL-37 Modulates Dendritic Cell Differentiation and Dendritic Cell-Induced T Cell Polarization // J. Immunol. 2004. Vol. 172, № 4. P. 2704.2-2704.
- 26. Bowdish D.M.E. et al. Impact of LL-37 on anti-infective immunity // J. Leukoc. Biol. 2004. Vol. 77, № 4. P. 451–459.
- 27. Fleischmann J., Selsted M.E., Lehrer R.I. Opsonic activity of MCP-1 and MCP-2, cationic peptides from rabbit alveolar macrophages // Diagn. Microbiol. Infect. Dis. 1985. Vol. 3, № 3. P. 233–242.
- 28. Biragyn A. et al. Mediators of Innate Immunity That Target Immature, But Not Mature, Dendritic Cells Induce Antitumor Immunity When Genetically Fused with Nonimmunogenic Tumor Antigens // J. Immunol. 2001. Vol. 167, № 11. P. 6644–6653.
- 29. Eliasson M., Egesten A. Antibacterial chemokines-actors in both innate and adaptive immunity // Contrib Microbiol. Basel, Karger, 2008. Vol. 15. P. 101-117.
- 30. Ding X. et al. Enfuvirtide (T20)-Based Lipopeptide Is a Potent HIV-1 Cell Fusion Inhibitor: Implications for Viral Entry and Inhibition // J. Virol. 2017. Vol. 91, № 18. P. 1–20.
- 31. Qian S., Sharma V.K. Aurein 1.2, a Short and Potent Antimicrobial Peptide, Changes Charged Lipid Distribution and Lipid Dynamics in Bilayer // Biophys. J. Elsevier Ltd, 2019. Vol. 116, № 3. P. 86a.
- 32. Palffy R. et al. On the Physiology and Pathophysiology of Antimicrobial Peptides // Mol. Med. 2008. Vol. 15, № 1–2. P. 1.
- 33. Yeaman M.R., Yount N.Y. Mechanisms of antimicrobial peptide action and resistance. // Pharmacol. Rev. 2003. Vol. 55, № 1. P. 27–55.
- 34. Mookherjee N. et al. Antimicrobial host defence peptides: functions and clinical potential // Nat. Rev. Drug Discov. Springer US, 2020. Vol. 19, № 5. P. 311–332.
- 35. Yuan Z., Tam V.H. Polymyxin B: A new strategy for multidrug-resistant Gram-negative organisms // Expert Opin. Investig. Drugs. 2008. Vol. 17, № 5. P. 661–668.
- 36. Zavascki A.P. et al. Polymyxin B for the treatment of multidrug-resistant pathogens: A critical review // J. Antimicrob. Chemother. 2007. Vol. 60, № 6. P. 1206–1215.
- 37. Gallardo-Godoy A. et al. Activity and Predicted Nephrotoxicity of Synthetic Antibiotics Based on Polymyxin B // J. Med. Chem. 2016. Vol. 59, № 3. P. 1068–1077.
- 38. Severino P. et al. Antimicrobial activity of polymyxin-loaded solid lipid nanoparticles (PLX-SLN): Characterization of physicochemical properties and in vitro efficacy // Eur. J. Pharm. Sci. Elsevier, 2017. Vol. 106, № May. P. 177–184.
- 39. Cotter P.D., Hill C., Ross R.P. Bacteriocins: developing innate immunity for food // Nat. Rev. Microbiol. 2005. Vol. 3. P. 777–788.
- 40. Maher S., McClean S. Investigation of the cytotoxicity of eukaryotic and prokaryotic antimicrobial peptides in intestinal epithelial cells in vitro // Biochem. Pharmacol. 2006. Vol. 71, № 9. P. 1289–1298.

- 41. Shin J.M. et al. Biomedical applications of nisin // J. Appl. Microbiol. 2016. Vol. 120, № 6. P. 1449–1465.
- 42. Gomes B. et al. Designing improved active peptides for therapeutic approaches against infectious diseases // Biotechnol. Adv. Elsevier, 2018. Vol. 36, № 2. P. 415–429.
- 43. Felício M.R. et al. Peptides with Dual Antimicrobial and Anticancer Activities // Front. Chem. 2017. Vol. 5, № February. P. 1–9.
- 44. Sierra J.M. et al. An overview of antimicrobial peptides and the latest advances in their development // Expert Opin. Biol. Ther. Taylor & Francis, 2017. Vol. 0, № 0. P. 14712598.2017.1315402.
- 45. Mahlapuu M. et al. Antimicrobial Peptides: An Emerging Category of Therapeutic Agents // Front. Cell. Infect. Microbiol. 2016. Vol. 6, № December. P. 1–12.
- 46. van der Weide H. et al. Antimicrobial activity of two novel antimicrobial peptides AA139 and SET-M33 against clinically and genotypically diverse Klebsiella pneumoniae isolates with differing antibiotic resistance profiles // Int. J. Antimicrob. Agents. Elsevier B.V., 2019. Vol. 54, № 2. P. 159–166.
- 47. Wang X. et al. Candidacidal mechanism of the arenicin-3-derived peptide NZ17074 from Arenicola marina // Appl. Microbiol. Biotechnol. 2014. Vol. 98, № 17. P. 7387–7398.
- 48. Ciandrini E. et al. Antimicrobial Activity of Different Antimicrobial Peptides (AMPs) Against Clinical Methicillin-resistant Staphylococcus aureus (MRSA) // Curr. Top. Med. Chem. 2018. Vol. 18, № 24. P. 2116–2126.
- 49. Lamb H.M., Wiseman L.R. Pexiganan acetate // Drugs. 1998. Vol. 56, № 6. P. 1047–1052.
- 50. Lipsky B.A., Holroyd K.J., Zasloff M. Topical versus Systemic Antimicrobial Therapy for Treating Mildly Infected Diabetic Foot Ulcers: A Randomized, Controlled, Double-Blinded, Multicenter Trial of Pexiganan Cream // Clin. Infect. Dis. 2008. Vol. 47, № 12. P. 1537–1545.
- 51. Vakharia P.P., Silverberg J.I. New therapies for atopic dermatitis: Additional treatment classes // J. Am. Acad. Dermatol. Elsevier Inc, 2018. Vol. 78, № 3. P. S76–S83.
- 52. Rijsbergen M. et al. Results of phase 2 trials exploring the safety and efficacy of omiganan in patients with human papillomavirus-induced genital lesions // Br. J. Clin. Pharmacol. 2019. P. 0–2.
- 53. Saravolatz L.D. et al. In vitro activities of LTX-109, a synthetic antimicrobial peptide, against methicillin-resistant, vancomycin-intermediate, vancomycin-resistant, daptomycin-nonsusceptible, and linezolid-nonsusceptible Staphylococcus aureus // Antimicrob. Agents Chemother. 2012. Vol. 56, № 8. P. 4478–4482.
- 54. Nilsson A.C. et al. Ltx-109 is a novel agent for nasal decolonization of methicillin-resistant and -sensitive staphylococcus aureus // Antimicrob. Agents Chemother. 2015. Vol. 59, № 1. P. 145–151.
- 55. Flamm R.K. et al. In vitro spectrum of pexiganan activity; bactericidal action and resistance selection tested against pathogens with elevated MIC values to topical agents // Diagn. Microbiol. Infect. Dis. Elsevier Inc., 2016. Vol. 86, № 1. P. 66–69.
- 56. Das B. et al. Telavancin: a novel semisynthetic lipoglycopeptide agent to counter the challenge of resistant Gram-positive pathogens // Ther. Adv. Neurol. Disord. 2017. Vol. 4, № 2. P. 49–73.
- 57. Britt N.S. et al. Telavancin for refractory MRSA bacteraemia in intermittent haemodialysis recipients // J. Antimicrob. Chemother. 2018. Vol. 73, № 3. P. 764–767.
- 58. Barroso S. et al. Metabolic, mitochondrial, renal and hepatic safety of enfuvirtide and raltegravir antiretroviral administration: Randomized crossover clinical trial in healthy volunteers // PLoS One. 2019. Vol. 14, № 5. P. 1–14.
- 59. Khalilieh S. et al. Clinical Pharmacology Profile of Boceprevir, a Hepatitis C Virus NS3 Protease Inhibitor: Focus on Drug–Drug Interactions // Clin. Pharmacokinet. 2015. Vol. 54, № 6. P. 599–614.

- 60. Bouza E. et al. Dalbavancin in the treatment of different gram-positive infections: a real-life experience // Int. J. Antimicrob. Agents. 2018. Vol. 51, № 4. P. 571–577.
- 61. Heidary M. et al. Daptomycin // Journal of Antimicrobial Chemotherapy. 2018. Vol. 73, № 1. P. 1–11.
- 62. Poulakou G. et al. New treatments of multidrug-resistant Gram-negative ventilator-associated pneumonia // Ann. Transl. Med. 2018. Vol. 6, № 20. P. 423–423.
- 63. Cong T.X. et al. From pathogenesis of acne vulgaris to anti-acne agents // Arch. Dermatol. Res. Springer Berlin Heidelberg, 2019. Vol. 311, № 5. P. 337–349.
- 64. Schulz L.T. et al. Multiple-Dose Oritavancin Evaluation in a Retrospective Cohort of Patients with Complicated Infections // Pharmacotherapy. 2018. Vol. 38, № 1. P. 152–159.
- 65. Jenssen H., Hamill P., Hancock R.E.W. Peptide Antimicrobial Agents // Clin. Microbiol. Rev. 2006. Vol. 19, № 3. P. 491–511.
- 66. Lee T., Hall K.N., Aguilar M. Antimicrobial Peptide Structure and Mechanism of Action : A Focus on the Role of Membrane Structure // Curr. Top. Med. Chem. 2016. Vol. 16. P. 25–39.
- 67. DeLano W. L. Pymol: An open-source molecular graphics tool // {CCP4} Newsletter On Protein Crystallography. 2002. T. 40. C. 1–8.
- 68. Yi H.Y. et al. Insect antimicrobial peptides and their applications // Appl. Microbiol. Biotechnol. 2014. Vol. 98, № 13. P. 5807–5822.
- 69. Ageitos J.M. et al. Antimicrobial peptides (AMPs): Ancient compounds that represent novel weapons in the fight against bacteria // Biochem. Pharmacol. Elsevier Inc., 2017. Vol. 133. P. 117–138.
- 70. Mangoni M.L. et al. Naturally occurring peptides from Rana temporaria: Antimicrobial properties and more. 2016. Vol. 16, № 1. P. 54–64.
- 71. Iwanaga S., Lee B.L. Recent advances in the innate immunity of invertebrate animals. // J. Biochem. Mol. Biol. 2005. Vol. 38, № 2. P. 128–150.
- Tossi A., Sandri L., Giangaspero A. Amphipathic, α -Helical Antimicrobial Peptides // Biopolym. (Peptide Sci. 2000. Vol. 55. P. 4–30.
- 73. Gagnon M.C. et al. Influence of the Length and Charge on the Activity of α-Helical Amphipathic Antimicrobial Peptides // Biochemistry. 2017. Vol. 56, № 11. P. 1680–1695.
- 74. Sher Khan R. et al. Plant defensins: types, mechanism of action and prospects of genetic engineering for enhanced disease resistance in plants // 3 Biotech. Springer International Publishing, 2019. Vol. 9, № 5. P. 1–12.
- 75. Sinha S. et al. Structure and Interactions of A Host Defense Antimicrobial Peptide Thanatin in Lipopolysaccharide Micelles Reveal Mechanism of Bacterial Cell Agglutination // Sci. Rep. Springer US, 2017. Vol. 7, № 1. P. 17795.
- 76. Harris F., Dennison S., Phoenix D. Anionic Antimicrobial Peptides from Eukaryotic Organisms // Curr. Protein Pept. Sci. 2009. Vol. 10, № 6. P. 585–606.
- Konno K. et al. Anoplin, a novel antimicrobial peptide from the venom of the solitary wasp Anoplius samariensis // Biochim. Biophys. Acta Protein Struct. Mol. Enzymol. 2001. Vol. 1550, № 1. P. 70–80.
- 78. Wiesner J., Vilcinskas A. Antimicrobial peptides: the ancient arm of the human immune system. // Virulence. 2010. Vol. 1, № 5. P. 440–464.
- 79. Hancock R.E.W., Sahl H.G. Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies // Nat. Biotechnol. 2006. Vol. 24, № 12. P. 1551–1557.
- 80. Li C. et al. Centrocins: Isolation and characterization of novel dimeric antimicrobial peptides from the green sea urchin, Strongylocentrotus droebachiensis // Dev. Comp. Immunol. Elsevier Ltd, 2010. Vol. 34, № 9. P. 959–968.

- 81. Bulet P. et al. A novel inducible antibacterial peptide of Drosophila carries an O- glycosylated substitution // J. Biol. Chem. 1993. Vol. 268, № 20. P. 14893–14897.
- 82. Simmaco M., Mignogna G., Barra D. Antimicrobial peptides from amphibian skin: What do they tell us? // Biopolymers. 1998. Vol. 47, № 6. P. 435–450.
- Zhao L., Lu W. Defensins in innate immunity // Curr. Opin. Hematol. 2014. Vol. 21, № 1. P. 37–42.
- 84. Chen J. et al. Melittin, the Major Pain-Producing Substance of Bee Venom // Neurosci. Bull. Springer Singapore, 2016. Vol. 32, № 3. P. 265–272.
- 85. Huang L. et al. Dermaseptin-PH: A Novel Peptide with Antimicrobial and Anticancer Activities from the Skin Secretion of the South American Orange-Legged Leaf Frog, Pithecopus (Phyllomedusa) hypochondrialis // Molecules. 2017. Vol. 22, № 11. P. 1805.
- 86. Wang G. Human Antimicrobial Peptides and Proteins // Pharmaceuticals. 2014. Vol. 7, № 5. P. 545–594.
- 87. Vassilevski A. a, Kozlov S. a, Grishin E. V. Antimicrobial peptide precursor structures suggest effective production strategies. // Recent Pat. Inflamm. Allergy Drug Discov. 2008. Vol. 2, № 1. P. 58–63.
- 88. Park C.B. et al. Structure-activity analysis of buforin II, a histone H2A-derived antimicrobial peptide: the proline hinge is responsible for the cell-penetrating ability of buforin II. // Pnas. 2000. Vol. 97, № 15. P. 8245–8250.
- 89. García-Montoya I.A. et al. Lactoferrin a multiple bioactive protein: An overview // Biochim. Biophys. Acta Gen. Subj. Elsevier B.V., 2012. Vol. 1820, № 3. P. 226–236.
- 90. Zasloff M. Magainins, a class of antimicrobial peptides from Xenopus skin: isolation, characterization of two active forms, and partial cDNA sequence of a precursor. // Proc. Natl. Acad. Sci. U. S. A. 1987. Vol. 84, № 15. P. 5449–5453.
- 91. Casteels-Josson K. et al. Apidaecin multipeptide precursor structure: a putative mechanism for amplification of the insect antibacterial response. // EMBO J. 1993. Vol. 12, № 4. P. 1569–1578.
- Baumann G., Mueller P. A molecular model of membrane excitability // J. Supramol. Cell. Biochem. 1974. Vol. 2, № 5–6. P. 538–557.
- 93. Bechinger B. Structure and functions of channel-forming peptides: Magainins, cecropins, melittin and alamethicin // J. Membr. Biol. 1997. Vol. 156, № 3. P. 197–211.
- 94. Chen F.Y., Lee M.T., Huang H.W. Evidence for membrane thinning effect as the mechanism for peptide-induced pore formation // Biophys. J. 2003. Vol. 84, № 6. P. 3751–3758.
- 95. Huang H.W. Molecular mechanism of antimicrobial peptides: The origin of cooperativity // Biochim. Biophys. Acta Biomembr. 2006. Vol. 1758, № 9. P. 1292–1302.
- 96. Campagna S. et al. Structure and mechanism of action of the antimicrobial peptide piscidin // Biochemistry. 2007. Vol. 46, № 7. P. 1771–1778.
- 97. Ludtke S.J. et al. Membrane pores induced by magainin // Biochemistry. 1996. Vol. 35, № 43. P. 13723–13728.
- 98. Matsuzaki K. et al. Transbilayer transport of ions and lipids coupled with mastoparan X translocation // Biochemistry. 1996. Vol. 35, № 25. P. 8450–8456.
- 99. Sengupta D. et al. Toroidal pores formed by antimicrobial peptides show significant disorder // Biochim. Biophys. Acta Biomembr. 2008. Vol. 1778, № 10. P. 2308–2317.
- 100. Shai Y. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by α-helical antimicrobial and cell non-selective membrane-lytic peptides // Biochim. Biophys. Acta Biomembr. 1999. Vol. 1462, № 1–2. P. 55–70.
- 101. Gazit E. et al. Structure and Orientation of the Mammalian Antibacterial Peptide Cecropin P1 within Phospholipid Membranes peptide-membrane interaction; toxic mechanism // J. Mol. Biol.

1996. Vol. 258. P. 860-870.

- 102. Wu Q., Patočka J., Kuča K. Insect Antimicrobial Peptides, a Mini Review // Toxins (Basel). 2018. Vol. 10, № 11. P. 1–17.
- Huang H.W. Action of antimicrobial peptides: Two-state model // Biochemistry. 2000. Vol. 39, № 29. P. 8347–8352.
- 104. Silva O.N. et al. An anti-infective synthetic peptide with dual antimicrobial and immunomodulatory activities // Sci. Rep. 2016. Vol. 6. P. 35465.
- 105. Lohner K., Hilpert K. Antimicrobial peptides: Cell Membrane and Microbial Surface Interactions // Biochim. Biophys. Acta - Biomembr. Elsevier B.V., 2016. Vol. 1858, № 5. P. 915– 917.
- 106. Carlsson A. et al. Attacin an insect immune protein binds LPS and triggers the specific inhibition of bacterial outer-membrane protein synthesis // Microbiology. 1998. Vol. 144, № 8. P. 2179–2188.
- 107. Grein F., Schneider T., Sahl H.G. Docking on Lipid II—A Widespread Mechanism for Potent Bactericidal Activities of Antibiotic Peptides // J. Mol. Biol. Elsevier Ltd, 2019. № June.
- 108. Laridi R. et al. Liposome encapsulated nisin Z: Optimization, stability and release during milk fermentation // Int. Dairy J. 2003. Vol. 13, № 4. P. 325–336.
- 109. Graf M. et al. Proline-rich antimicrobial peptides targeting protein synthesis // Nat. Prod. Rep. Royal Society of Chemistry, 2017. Vol. 14. P. 1188–1190.
- 110. Mardirossian M. et al. The host antimicrobial peptide Bac71-35 binds to bacterial ribosomal proteins and inhibits protein synthesis // Chem. Biol. Elsevier Ltd, 2014. Vol. 21, № 12. P. 1639–1647.
- Park C.B., Kim H.S., Kim S.C. Mechanism of action of the antimicrobial peptide buforin II: buforin II kills microorganisms by penetrating the cell membrane and inhibiting cellular functions. // Biochem. Biophys. Res. Commun. 1998. Vol. 244, № 1. P. 253–257.
- 112. Friedrich C.L. et al. Structure and Mechanism of Action of an Indolicidin Peptide Derivative with Improved Activity against Gram-positive Bacteria // J. Biol. Chem. 2001. Vol. 276, № 26. P. 24015–24022.
- 113. Edwards I.A. et al. Structure-Activity and -Toxicity Relationships of the Antimicrobial Peptide Tachyplesin-1 // ACS Infect. Dis. 2017. Vol. 3, № 12. P. 917–926.
- 114. Taniguchi M. et al. Pyrrhocoricin, a proline-rich antimicrobial peptide derived from insect, inhibits the translation process in the cell-free Escherichia coli protein synthesis system // J. Biosci. Bioeng. Elsevier Ltd, 2016. Vol. 121, № 5. P. 591–598.
- 115. Nadezhdin K.D. et al. Molecular insight into mechanism of antimicrobial action of the βhairpin peptide arenicin: Specific oligomerization in detergent micelles // Biopolymers. 2007. Vol. 89, № 5. P. 455–464.
- 116. Choi H., Hwang J.-S., Lee D.G. Identification of a novel antimicrobial peptide, scolopendin 1, derived from centipede Scolopendra subspinipes mutilans and its antifungal mechanism // Insect Mol. Biol. 2014. Vol. 23, № 6. P. 788–799.
- 117. Cardoso M.H. et al. Computer-aided design of antimicrobial peptides: are we generating effective drug candidates? // Front. Microbiol. 2020. Vol. 10, № January. P. 1–15.
- 118. Gendrault Y. et al. Computer-aided design in synthetic biology. 2011. P. 1–7.
- 119. Lionta E. et al. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances // Curr. Top. Med. Chem. 2014. Vol. 14. P. 1923–1938.
- Wang C. et al. Current strategies and applications for precision drug design // Front. Pharmacol. 2018. Vol. 9, № JUL. P. 1–19.
- 121. Kore P.P. et al. Computer-Aided Drug Design: An Innovative Tool for Modeling // Open J.

Med. Chem. 2012. Vol. 02, № 04. P. 139–148.

- 122. Beltran J.A., Aguilera-Mendoza L., Brizuela C.A. Optimal selection of molecular descriptors for antimicrobial peptides classification: An evolutionary feature weighting approach // BMC Genomics. BMC Genomics, 2018. Vol. 19, № Suppl 7.
- 123. Leszczynski J. et al. Handbook of Computational Chemistry. 2nd ed. Springer Nature, 2017.
- 124. Todeschini R., Consonni V. Molecular Descriptors for Chemoinformatics // Recent Advances in QSAR Studies. 2010. 29–103 p.
- Müller A.T. et al. modlAMP: Python for antimicrobial peptides // Nanoscale Suppl. 2015. Vol. 33, № 17. P. 2753–2755.
- 126. Chen Z. et al. IFeature: A Python package and web server for features extraction and selection from protein and peptide sequences // Bioinformatics. 2018. Vol. 34, № 14. P. 2499–2502.
- 127. Dai J. et al. Computer-aided drug discovery: Novel 3,9-disubstituted eudistomin U derivatives as potent antibacterial agents // Eur. J. Med. Chem. 2018. Vol. 157. P. 333–338.
- 128. Lee E.Y., Wong G.C.L., Ferguson A.L. Machine learning-enabled discovery and design of membrane-active peptides // Bioorg. Med. Chem. 2017.
- 129. Porto W.F. et al. In silico optimization of a guava antimicrobial peptide enables combinatorial exploration for peptide design // Nat. Commun. Springer US, 2018. Vol. 9, № 1.
- 130. Supady A., Blum V., Baldauf C. First-Principles Molecular Structure Search with a Genetic Algorithm // J. Chem. Inf. Model. 2015. Vol. 55, № 11. P. 2338–2348.
- 131. Lima A.N. et al. Use of machine learning approaches for novel drug discovery // Expert Opin. Drug Discov. Taylor & Francis, 2016. Vol. 11, № 3. P. 225–239.
- 132. Yoshida M. et al. Using Evolutionary Algorithms and Machine Learning to Explore Sequence Space for the Discovery of Antimicrobial Peptides // Chem. 2018. Vol. 4, № 3. P. 533–543.
- 133. Lee E.Y. et al. Mapping membrane activity in undiscovered peptide sequence space using machine learning // Proc. Natl. Acad. Sci. U. S. A. 2016. Vol. 113, № 48. P. 13588–13593.
- 134. Lecun Y., Bengio Y., Hinton G. Deep learning // Nature. 2015. Vol. 521, № 7553. P. 436–444.
- 135. Veltri D., Kamath U., Shehu A. Deep learning improves antimicrobial peptide recognition // Bioinformatics. 2018. Vol. 34, № 16. P. 2740–2747.
- 136. Müller A.T., Hiss J.A., Schneider G. Recurrent Neural Network Model for Constructive Peptide Design // J. Chem. Inf. Model. 2018. Vol. 58, № 2. P. 472–479.
- 137. Schneider P. et al. Hybrid Network Model for "Deep Learning" of Chemical Data: Application to Antimicrobial Peptides // Mol. Inform. 2017. Vol. 36, № 1. P. 1–8.
- 138. Melo M.C.R. et al. NAMD goes quantum: An integrative suite for hybrid simulations // Nat. Methods. 2018. Vol. 15, № 5. P. 351–354.
- 139. Humphrey W., Dalke A., Schulten K. VMD: Visual Molecular Dynamics // J. Mol. Graph. 1996. Vol. 14, № October 1995. P. 33–38.
- 140. Hočevar T., Demčar J. Computation of Graphlet Orbits for Nodes and Edges in Sparse Graphs // J. Stat. Softw. 2016. Vol. 71, № 10.
- 141. Vishnepolsky B. et al. Predictive Model of Linear Antimicrobial Peptides Active against Gram-Negative Bacteria: research-article // J. Chem. Inf. Model. American Chemical Society, 2018. Vol. 58, № 5. P. 1141–1151.
- 142. Veltri D., Kamath U., Shehu A. Improving Recognition of Antimicrobial Peptides and Target Selectivity through Machine Learning and Genetic Programming // IEEE/ACM Trans. Comput. Biol. Bioinforma. 2017. Vol. 14, № 2. P. 300–313.
- 143. Wang G., Li X., Wang Z. APD3: The antimicrobial peptide database as a tool for research and education // Nucleic Acids Res. 2016. Vol. 44, № D1. P. D1087–D1093.
- 144. Waghu F.H. et al. CAMP R3: a database on sequences, structures and signatures of

antimicrobial peptides // Nucleic Acids Res. 2015. Vol. 44, № D1. P. D1094–D1097.

- 145. Lee H.-T. et al. A Large-Scale Structural Classification of Antimicrobial Peptides. // Biomed Res. Int. 2015. Vol. 2015. P. 475062.
- Sawyer R. Leech biology and behaviour // Oxford Univ. Press. New York, New York. 1986.
 Vol. II. P. 1065.
- 147. Utevsky S. et al. Chromosome numbers for three species of medicinal leeches (Hirudo spp.) // Syst. Parasitol. 2009. Vol. 74, № 2. P. 95–102.
- 148. Utevsky S. et al. Distribution and status of medicinal leeches (genus Hirudo) in the Western Palaearctic: anthropogenic, ecological, or historical effects? // Aquat. Conserv. Mar. Freshw. Ecosyst. 2010. Vol. 20. P. 198–210.
- 149. Hildebrandt J.P., Lemke S. Small bite, large impact-saliva and salivary molecules in the medicinal leech, Hirudo medicinalis // Naturwissenschaften. 2011. Vol. 98, № 12. P. 995–1008.
- 150. Dong H. et al. Chinese Medicinal Leech: Ethnopharmacology, Phytochemistry, and Pharmacological Activities // Evidence-based Complement. Altern. Med. Hindawi Publishing Corporation, 2016. Vol. 2016.
- 151. Houschyar K.S. et al. Medical leech therapy in plastic reconstructive surgery // Wiener Medizinische Wochenschrift. 2015. Vol. 165, № 19–20. P. 419–425.
- 152. Nawa Y. et al. A case of conjunctival leech infestation // Jpn. J. Ophthalmol. 2006. Vol. 50, № 1. P. 64–65.
- 153. Rados C. Beyond bloodletting: FDA gives leeches a medical makeover. // FDA Consum. 2004.
 Vol. 38, № 5. P. 9.
- 154. Baskova I.P. et al. Proteins and peptides of the salivary gland secretion of medicinal leeches Hirudo verbana, H. medicinalis, and H. orientalis. // Biochemistry. (Mosc). 2008. Vol. 73, № 3. P. 315–320.
- 155. Baskova I.P. et al. Protein profiling of the medicinal leech salivary gland secretion by proteomic analytical methods. // Biochemistry. (Mosc). 2004. Vol. 69, № 7. P. 770–775.
- 156. Sig A.K. et al. Medicinal leech therapy—an overall perspective // Integr. Med. Res. Korea Institute of Oriental Medicine, 2017. Vol. 6, № 4. P. 337–343.
- 157. Rigbi M. et al. The saliva of the medicinal leech Hirudo medicinalis-1. Biochemical characterization of the high molecular weight fraction // Comp. Biochem. Physiol. -- Part B Biochem. 1987. Vol. 87, № 3. P. 567–573.
- 158. Diamandis E.P., Yousef G.M. Human tissue kallikrein gene family: A rich source of novel disease biomarkers // Expert Rev. Mol. Diagn. 2001. Vol. 1, № 2. P. 182–190.
- 159. Steranka L.R. et al. Bradykinin as a pain mediator: Receptors are localized to sensory neurons, and antagonists have analgesic actions // Proc. Natl. Acad. Sci. U. S. A. 1988. Vol. 85, № 9. P. 3245–3249.
- 160. Söllner C. et al. Isolation and characterization of hirustasin, an antistasin-type serine-proteinase inhibitor from the medical leech Hirudo medicinalis // Eur. J. Biochem. 1994. Vol. 219, № 3. P. 937–943.
- 161. Braun N.J., Schnebli H.P. Interaction of Eglin c with Polymorphonuclear Cells: Evidence for Binding to the Cell Surface // Biol. Chem. Hoppe. Seyler. 1987. Vol. 368, № 1. P. 155–162.
- 162. Braun N.J. et al. Kinetic Studies on the Interaction of Eglin c with Human Leukocyte Elastase and Cathepsin G // Biol. Chem. Hoppe. Seyler. 1987. Vol. 368, № 1. P. 299–308.
- 163. Sommerhoff C.P. et al. A Kazal-Type Inhibitor of Human Mast Cell Tryptase: Isolation from the Medical Leech Hirudo medicinalis, Characterization, and Sequence Analysis // Biol. Chem. Hoppe. Seyler. 1994. Vol. 375, № 10. P. 685–694.
- 164. Stubbs M.T. et al. The three-dimensional structure of recombinant leech-derived tryptase

inhibitor in complex with trypsin. Implications for the structure of human mast cell tryptase and its inhibition // J. Biol. Chem. 1997. Vol. 272, № 32. P. 19931–19937.

- 165. Baskova I.P., Zavalova L.L. Proteinase inhibitors from the medicinal leech Hirudo medicinalis // Biokhimiya. 2001. Vol. 66, № 7. P. 869–883.
- 166. Cicardi M. et al. C1 inhibitor: Molecular and clinical aspects // Springer Semin. Immunopathol. 2005. Vol. 27, № 3. P. 286–298.
- 167. Gronwald W. et al. Structure of the Leech Protein Saratin and Characterization of Its Binding to Collagen // J. Mol. Biol. 2008. Vol. 381, № 4. P. 913–927.
- 168. White T.C. et al. The leech product saratin is a potent inhibitor of platelet integrin α2β1 and von Willebrand factor binding to collagen // FEBS J. 2007. Vol. 274, № 6. P. 1481–1491.
- 169. Harsfalvi J. et al. Calin from Hirudo medicinalis, an inhibitor of von Willebrand factor binding to collagen under static and flow conditions // Blood. 1995. Vol. 85, № 3. P. 705–711.
- 170. Deckmyn H. et al. Calin from Hirudo medicinalis, an inhibitor of platelet adhesion to collagen, prevents platelet-rich thrombosis in hamsters // Blood. 1995. Vol. 85, № 3. P. 712–719.
- 171. Fenton J.W. 2nd. Thrombin // Ann N Y Acad Sci. 1986. Vol. 485, № 5. P. 5–15.
- 172. Furie B., Furie B.C. Mechanisms of thrombus formation // N. Engl. J. Med. 2008. Vol. 359, № 9. P. 938–949.
- 173. Nowak G., Schrör K. Hirudin-the long and stony way from an anticoagulant peptide in the saliva of medicinal leech to a recombinant drug and beyond // ThrombHaemost. 2007. Vol. 98. P. 116-119.
- 174. Stone S.R., Betz A., Hofsteenge J. Mechanistic Studies on Thrombin Catalysis // Biochemistry. 1991. Vol. 30, № 41. P. 9841–9848.
- 175. Johnson, Ph.D P.H. Hirudin: Clinical Potential of a Thrombin Inhibitor // Annu. Rev. Med. 1994. Vol. 45, № 1. P. 165–177.
- Weitz J. et al. Clot-bound thrombin isprotected from inhibition by heparin-antithrombin III but is susceptible to inactivation by antithrombin III-independent inhibitors. 1990. Vol. 86, № August. P. 385–391.
- 177. Markwardt F. Hirudin as alternative anticoagulant A historical review // Semin. Thromb. Hemost. 2002. Vol. 28, № 5. P. 405–413.
- 178. Marin E. et al. Crystallographic structure of destabilase from Hirudo medicinalis a small bifunctional enzyme with isopeptidase and lysozyme activities // Febs J. 2017. Vol. 284, № 16. P. 203–204.
- Zavalova L.L., Baskova I.P. Destabilase // Handbook of Proteolytic Enzymes. Elsevier Ltd, 2013. Vol. 3. 3758–3759 p.
- 180. Baskova I.P., Nikonov G.I. Destabilase, the novel isopeptidase with thrombolytic activity // Blood Coagulation and Fibrinolysis. 1991. Vol. 2. P. 167–172.
- 181. Baskova I. et al. Inhibition of induced and spontaneous platelet aggregation by destabilase from medicinal leech // Platelets. 2000. Vol. 11, № 2. P. 83–86.
- 182. Nikonov G.I., Titova E.A. Destabilase complexes Natural liposome produced by medicinal leeches Hirudo medicinalis // Fundam. Clin. Pharmacol. 1999. Vol. 13, № 1. P. 102–106.
- 183. Павлова И.Б. et al. Изучение перспектив использования секрета слюнных клеток медицинской пиявки Hirudo medicinalis и препарата «Пиявит» как антимикробных комплексов, не вызывающих резистентности у микроорганизмов // Современные проблемы науки и образования. 2015. Vol. 2–3. P. 252.
- 184. Reverter D. et al. A carboxypeptidase inhibitor from the medical leech Hirudo medicinalis: Isolation, sequence analysis, cDNA cloning, recombinant expression, and characterization // J. Biol. Chem. 1998. Vol. 273, № 49. P. 32927–32933.

- 185. Snider G.L. et al. Eglin-c, a polypeptide derived from the medicinal leech, prevents human neutrophil elastase-induced emphysema and bronchial secretory cell metaplasia in the hamster // Am. Rev. Respir. Dis. 1985. Vol. 132, № 6. P. 1155–1161.
- 186. Dunwiddie C. et al. Antistasin, a leech-derived inhibitor of factor Xa. Kinetic analysis of enzyme inhibition and identification of the reactive site // J. Biol. Chem. 1989. Vol. 264, № 28. P. 16694–16699.
- 187. Fink E. et al. The Primary Structure of Bdellin B-3 from the Leech Hirudo medicinalis // Biol. Chem. 1986. Vol. 367, № December. P. 1235–1242.
- 188. Kelen E.M.A., Rosenfeld G. Fibrinogenolytic substance (hementerin) of Brazilian bloodsucking leeches (Haementeria lutzi Pinto 1920) // Haemostasis. 1975. Vol. 4. P. 51–64.
- 189. Sawyer R.T., Jones C.P., Munro R. The biological function of hementin in the proboscis of the leech Haementeria ghilianii. // Blood Coagul. Fibrinolysis. 1991. Vol. 2, № 1. P. 153–159.
- 190. Alfred L., Meyer K., Hoffman P. The production of hyaluronate oligosaccharides by leech hyaluronidase and alkali // J. Biol. Chem. 1960. Vol. 235, № 4. P. 924–927.
- 191. Tasiemski A. et al. Molecular characterization of two novel antibacterial peptides inducible upon bacterial challenge in an annelid, the leech Theromyzon tessulatum // J. Biol. Chem. 2004. Vol. 279, № 30. P. 30973–30982.
- 192. Schikorski D. et al. Microbial Challenge Promotes the Regenerative Process of the Injured Central Nervous System of the Medicinal Leech by Inducing the Synthesis of Antimicrobial Peptides in Neurons and Microglia // J. Immunol. 2008. Vol. 181, № 2. P. 1083–1095.
- Tasiemski A. Antimicrobial peptides in annelids // Lab. Neuroimmunol. des Annelides . 2008. P. 75–82.
- 194. Jung S. et al. Hydramacin-1, structure and antibacterial activity of a protein from the basal metazoan hydra // J. Biol. Chem. 2009. Vol. 284, № 3. P. 1896–1905.
- 195. Hung C.-W.W. et al. Determination of disulfide linkages in antimicrobial peptides of the macin family by combination of top-down and bottom-up proteomics // J. Proteomics. Elsevier B.V., 2014. Vol. 103. P. 216–226.
- 196. Jung S. et al. Macin family of antimicrobial proteins combines antimicrobial and nerve repair activities // J. Biol. Chem. 2012. Vol. 287, № 17. P. 14246–14258.
- 197. Ding A. et al. Gene cloning and expression of a partial sequence of Hirudomacin, an antimicrobial protein that is increased in leech (Hirudo nipponica Whitman) after a blood meal // Comp. Biochem. Physiol. Part B Biochem. Mol. Biol. Elsevier, 2019. Vol. 231, № 1. P. 75–86.
- 198. Tasiemski A. et al. Reciprocal immune benefit based on complementary production of antibiotics by the leech Hirudo verbana and its gut symbiont Aeromonas veronii // Sci. Rep. Nature Publishing Group, 2015. Vol. 5, № December. P. 1–13.
- 199. Duan Y. et al. Repair and regeneration of functional synaptic connections: Cellular and molecular interactions in the leech // Cell. Mol. Neurobiol. 2005. Vol. 25, № 2. P. 441–450.
- 200. Salzet M. Innate Immunity in Lophotrochozoans : The Annelids. 2015. № FEBRUARY 2006.
- 201. Kurdyumov A.S. et al. A comparison of the enzymatic properties of three recombinant isoforms of thrombolytic and antibacterial protein-Destabilase-Lysozyme from medicinal leech. // BMC Biochem. BMC Biochemistry, 2015. Vol. 16, № 1. P. 27.
- 202. Zavalova L.L. et al. Antibacterial non-glycosidase activity of invertebrate destabilase-lysozyme and of its helical amphipathic peptides // Chemotherapy. 2006. Vol. 52, № 3. P. 158–160.
- 203. Bobrovsky P. et al. Recombinant human peptidoglycan recognition proteins reveal antichlamydial activity // Infect. Immun. 2016. Vol. 84, № 7. P. 2124–2130.
- 204. Thermo Fisher Scientific. Ion TrueMate TM Library Preparation. 2015. 67 p.
- 205. O'Connell J. et al. NxTrim: Optimized trimming of Illumina mate pair reads // Bioinformatics.

2015. Vol. 31, № 12. P. 2035–2037.

- 206. Bankevich A. et al. SPAdes: A new genome assembly algorithm and its applications to singlecell sequencing // J. Comput. Biol. 2012. Vol. 19, № 5. P. 455–477.
- 207. Boetzer M. et al. Scaffolding pre-assembled contigs using SSPACE // Bioinformatics. 2011. Vol. 27, № 4. P. 578–579.
- 208. Rice P., Longden L., Bleasby A. EMBOSS: The European Molecular Biology Open Software Suite // Trends Genet. 2000. Vol. 16, № 6. P. 276–277.
- 209. Langmead B., Salzberg S.L. Fast gapped-read alignment with Bowtie 2 // Nat. Methods. 2013. Vol. 9, № 4. P. 357–359.
- 210. Quinlan A.R., Hall I.M. BEDTools: A flexible suite of utilities for comparing genomic features // Bioinformatics. 2010. Vol. 26, № 6. P. 841–842.
- Huson D.H. et al. MEGAN analysis of metagenomic data // Genome Res. 2007. Vol. 17, № 3.
 P. 377–386.
- 212. Grabherr M.G. et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data // Nat. Biotechnol. 2013. Vol. 29, № 7. P. 644–652.
- 213. Kent W.J. BLAT-The BLAST-Like Alignment Tool Resource 656 Genome Research // Genome Res. 2002. Vol. 12, № 4. P. 656–664.
- 214. Stanke M., Morgenstern B. AUGUSTUS: A web server for gene prediction in eukaryotes that allows user-defined constraints // Nucleic Acids Res. 2005. Vol. 33, № SUPPL. 2. P. 465–467.
- 215. Babenko V. V. et al. Draft genome sequences of Hirudo medicinalis and salivary transcriptome of three closely related medicinal leeches // BMC Genomics. BMC Genomics, 2020. Vol. 21, № 331. P. 357681.
- 216. Xiao N. et al. protr : R package for generating various numerical representation schemes of protein sequence // Bioinformatics. 2015. Vol. 31, № 11. P. 1857–1859.
- 217. Charif D., Lobry J.R. SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. // Structural approaches to sequence evolution: Molecules, networks, populations / ed. Bastolla U. et al. Springer Verlag, 2007. P. 207–232.
- 218. Osorio D., Rondon-villarreal P., Torres R. Peptides: A package for data mining of antimicrobial peptides // R J. 2015. Vol. 7, № 1. P. 4–14.
- 219. Fernandez-Escamilla A.-M. et al. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. // Nat. Biotechnol. 2004. Vol. 22, № 10. P. 1302–1306.
- 220. Conchillo-Solé O. et al. AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. // BMC Bioinformatics. 2007. Vol. 8, № 1. P. 65.
- 221. Torrent M. et al. AMPA: An automated web server for prediction of protein antimicrobial regions // Bioinformatics. 2012. Vol. 28, № 1. P. 130–131.
- 222. Shen Y. et al. Improved PEP-FOLD approach for peptide and miniprotein structure prediction // J. Chem. Theory Comput. 2014. Vol. 10, № 10. P. 4745–4758.
- 223. Wang Y. et al. I-TASSER-MR: Automated molecular replacement for distant-homology proteins using iterative fragment assembly and progressive sequence truncation // Nucleic Acids Res. 2017. Vol. 45, № W1. P. W429–W434.
- 224. Wiegand I., Hilpert K., Hancock R.E.W. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances // Nat. Protoc. 2008. Vol. 3, № 2. P. 163–175.
- 225. Rochus L.J. Keller. The Computer Aided Resonance Assignment Tutorial. CANTINA Verlag, 1966.
- 226. Shen Y., Bax A. Protein backbone and sidechain torsion angles predicted from NMR chemical

shifts using artificial neural networks // J Biomol NMR. 2013. Vol. 56, № 3. P. 227–241.

- 227. Peter Güntert. Automated NMR Structure Calculation With CYANA // Methods Mol. Biol. 2004. Vol. 278. P. 353–378.
- 228. Sokolova A.I. et al. Imaging human keratinocytes grown on electrospun mats by scanning electron microscopy. 2019. № November 2018. P. 2010–2015.
- 229. Sedlar K., Kupkova K., Provaznik I. Bioinformatics strategies for taxonomy independent binning and visualization of sequences in shotgun metagenomics // Comput. Struct. Biotechnol. J. The Authors, 2017. Vol. 15. P. 48–55.
- 230. Smit AFA, Hubley R G.P. RepeatMasker Open-4.0.
- 231. Bao W., Kojima K.K., Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes // Mob. DNA. Mobile DNA, 2015. Vol. 6, № 1. P. 4–9.
- 232. Thomas Wicker et al. A unified classification system for eukaryotic transposable elements // Nat Rev Genet. 2007. Vol. 8, № 12. P. 973–982.
- 233. McLysaght A. et al. Estimation of Synteny Conservation and Genome Compaction Between Pufferfish (Fugu) and Human // Yeast. 2000. Vol. 1, № 1. P. 22–36.
- 234. Vinogradov A.E. Intron-genome size relationship on a large evolutionary scale // J. Mol. Evol.
 1999. Vol. 49, № 3. P. 376–384.
- 235. Emms D.M., Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy // Genome Biol. Genome Biology, 2015. Vol. 16, № 1. P. 1–14.
- 236. Seymour S.L., Hunter C.L. ProteinPilot TM Software Overview. 2017. P. 1–5.
- 237. Bateman A. et al. UniProt: A hub for protein information // Nucleic Acids Res. 2015. Vol. 43, № D1. P. D204–D212.
- 238. Novković M. et al. DADP: The database of anuran defense peptides // Bioinformatics. 2012. Vol. 28, № 10. P. 1406–1407.
- 239. Camacho C. et al. BLAST plus: architecture and applications // BMC Bioinformatics. 2009. Vol. 10, № 421. P. 1.
- 240. Kim I.-W. et al. De Novo Transcriptome Analysis and Detection of Antimicrobial Peptides of the American Cockroach Periplaneta americana (Linnaeus) // PLoS One. 2016. Vol. 11, № 5. P. e0155304.
- 241. Yoo W.G. et al. Antimicrobial peptides in the centipede Scolopendra subspinipes mutilans // Funct. Integr. Genomics. 2014. Vol. 14, № 2. P. 275–283.
- 242. Torres R., Pedro L. Package ' Peptides .' 2020.
- 243. Wang Y., et al. Spontaneous formation of structurally diverse membrane channel architectures from a single antimicrobial peptide // Nat. Commun. Nature Publishing Group, 2016. Vol. 7. P. 13535.
- 244. Vincent Zoete. 3D structure visualization and high quality imaging . Chimera // Chimera. 2008.
- 245. Casteels P. et al. Apidaecins: antibacterial peptides from honeybees. // EMBO J. 1989. Vol. 8, № 8. P. 2387–2391.
- 246. Rowe-Magnus D.A. et al. Cathelicidin Peptides Restrict Bacterial Growth via Membrane Perturbation and Induction of Reactive Oxygen Species // MBio. 2019. Vol. 10, № 5. P. 1–19.
- 247. Helbing C.C. et al. Antimicrobial peptides from Rana [Lithobates] catesbeiana: Gene structure and bioinformatic identification of novel forms from tadpoles // Sci. Rep. 2019. Vol. 9, № 1. P. 1–12.
- 248. Toh S. et al. Heme and blood-feeding parasites: friends or foes? // Parasit. Vectors. BioMed Central Ltd, 2010. Vol. 3, № 1. P. 108.
- 249. Mesquita R.D. et al. Genome of Rhodnius prolixus, an insect vector of Chagas disease, reveals

unique adaptations to hematophagy and parasite infection // Proc. Natl. Acad. Sci. U. S. A. 2015. Vol. 112, № 48. P. 14936–14941.

- 250. Zepeda Mendoza M.L. et al. Hologenomic adaptations underlying the evolution of sanguivory in the common vampire bat // Nat. Ecol. Evol. 2018. Vol. 2018, № 4. P. 659–668.
- 251. Gulia-Nuss M. et al. Genomic insights into the Ixodes scapularis tick vector of Lyme disease // Nat. Commun. 2016. Vol. 7, № May 2015. P. 10507.
- 252. Aloto D., Eticha E. Leeches: A Review on their Pathogenic and Beneficial Effects // J. Vet. Sci. Technol. 2018. Vol. 09, № 01. P. 1–6.
- 253. Cherniack E.P. Bugs as drugs, part two: worms, leeches, scorpions, snails, ticks, centipedes, and spiders // Altern. Med. Rev. 2011. Vol. 16. P. 50–8.
- 254. Abdualkader A.M., Ghawi A.M., Alaama M., Awang M., Merzouk A. Leech therapeutic applications // Indian. J. Pharm. Sci. 2013. Vol. 75. P. 127–37.
- 255. Li L., Stoeckert C.J.J., Roos D.S. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes -- Li et al. 13 (9): 2178 -- Genome Research // Genome Res. 2003. Vol. 13, № 9. P. 2178– 2189.
- 256. Lemke S. et al. May salivary gland secretory proteins from hematophagous leeches (Hirudo verbana) reach pharmacologically relevant concentrations in the vertebrate host? // PLoS One. 2013. Vol. 8, № 9. P. e73809.
- 257. Tadokoro T. et al. Venomous snakes: An overview of the functional underappreciated superfamily // Toxins (Basel). 2020. Vol. 12, № 175. P. 1–20.
- 258. Madio B., Undheim E.A.B., King G.F. Revisiting venom of the sea anemone Stichodactyla haddoni: Omics techniques reveal the complete toxin arsenal of a well-studied sea anemone genus // J. Proteomics. Elsevier B.V., 2017. Vol. 166. P. 83–92.
- 259. Reumont M. Von et al. A polychaete's powerful punch: venom gland transcriptomics of Glycera reveals a complex cocktail of toxin homologs. // Genome Biol. Evol. 2014. Vol. 6, № 9. P. 2406–2423.
- 260. Rajesh R.P., Franklin J.B. Identification of Conotoxins with Novel Odd Number of Cysteine Residues from the Venom of a Marine Predatory Gastropod Conus leopardus Found in Andaman Sea // Protein Pept. Lett. 2018. Vol. 25, № 11. P. 1035–1040.
- 261. Gonçalves C., Costa P.M. Histochemical detection of free thiols in glandular cells and tissues of different marine Polychaeta // Histochem. Cell Biol. Springer Berlin Heidelberg, 2020. Vol. 154, № 3. P. 315–325.
- 262. Manuscript A., Superfamily C. Cystatin Superfamily. 2011. Vol. 21, № 615. P. 51–70.
- 263. Lefebvre C. et al. Transcriptomic analysis in the leech Theromyzon tessulatum: Involvement of cystatin B in innate immunity // Biochem. J. 2004. Vol. 380, № 3. P. 617–625.
- 264. Borth W. α 2 "Macroglobulin, a multifunctional binding protein with targeting characteristics // FASEB J. 1992. Vol. 6, № 15. P. 3345–3353.
- 265. De Boer J.P. et al. Alpha-2-macroglobulin functions as an inhibitor of fibrinolytic, clotting, and neutrophilic proteinases in sepsis: Studies using a baboon model // Infect. Immun. 1993. Vol. 61, № 12. P. 5035–5043.
- 266. Tripp B.C., Smith K., Ferry J.G. Carbonic Anhydrase: New Insights for an Ancient Enzyme // J. Biol. Chem. 2001. Vol. 276, № 52. P. 48615–48618.
- 267. Carter M.J., Parsons D.S. The isoenzymes of carbonic anhydrase: tissue, subcellular distribution and functional significance, with particular reference to the intestinal tract // J. Physiol. 1971. Vol. 215, № 1. P. 71–94.
- 268. 248.Linser P.J. et al. Carbonic anhydrases and anion transport in mosquito midgut pH regulation // J. Exp. Biol. 2009. Vol. 212, № 11. P. 1662–1671.

- 269. Fujita T., Matsushita M., Endo Y. The lectin-complement pathway Its role in innate immunity and evolution // Immunol. Rev. 2004. Vol. 198. P. 185–202.
- 270. Ompraba G. et al. Identification of a novel family of snake venom proteins veficolins from cerberus rynchops using a venom gland transcriptomics and proteomics approach // J. Proteome Res. 2010. Vol. 9, № 4. P. 1882–1893.
- 271. Gardiner E.E. et al. Role of calmodulin in platelet receptor function // Curr. Med. Chem. Cardiovasc. Hematol. Agents. 2005. Vol. 3, № 4. P. 283–287.
- 272. Kvist S., Min G.S., Siddall M.E. Diversity and selective pressures of anticoagulants in three medicinal leeches (Hirudinida: Hirudinidae, Macrobdellidae) // Ecol. Evol. 2013. Vol. 3, № 4. P. 918–933.
- 273. Tessler M. et al. Marine leech anticoagulant diversity and evolution // J. Parasitol. 2018. № March.
- 274. Singh A.P. Medicinal leech therapy (Hirudotherapy): A brief overview // Complement. Ther. Clin. Pract. Elsevier Ltd, 2010. Vol. 16, № 4. P. 213–215.
- 275. Liu S. et al. Computational resources and tools for antimicrobial peptides // J. Pept. Sci. 2016.
 № July.
- 276. Lee E.Y. et al. What can machine learning do for antimicrobial peptides, and what can antimicrobial peptides do for machine learning? 2017.
- 277. Lata S., Mishra N.K., Raghava G.P.S. AntiBP2: Improved version of antibacterial peptide prediction // BMC Bioinformatics. 2010. Vol. 11, № SUPPLL.1. P. 1–7.
- 278. Gao Y. et al. Targeted Modification of a Novel Amphibian Antimicrobial Peptide from Phyllomedusa tarsius to Enhance Its Activity against MRSA and Microbial Biofilm // Front. Microbiol. 2017. Vol. 8, № April. P. 1–9.
- 279. Marani M.M. et al. Thaulin-1: The first antimicrobial peptide isolated from the skin of a Patagonian frog Pleurodema thaul (Anura: Leptodactylidae: Leiuperinae) with activity against Escherichia coli // Gene. Elsevier B.V., 2016. Vol. 605. P. 70–80.
- 280. Cabello C.M. et al. Structure, Membrane Orientation, Mechanism, and Function of Pexiganan A Highly Potent Antimicrobial Peptide Designed From Magainin // Biochim Biophys Acta. 2009. Vol. 1788, № 8. P. 1680–1686.
- 281. Ruffin M. et al. Quorum-sensing inhibition abrogates the deleterious impact of Pseudomonas aeruginosa on airway epithelial repair // FASEB J. 2016. Vol. 30, № 9. P. 3011–3025.
- 282. Vermote A. et al. Hamamelitannin analogues that modulate quorum sensing as potentiators of antibiotics against staphylococcus aureus // Angew. Chemie Int. Ed. 2016. Vol. 55, № 22. P. 6551–6555.
- 283. Cole A.M., Weis P., Diamond G. Isolation and characterization of pleurocidin, an antimicrobial peptide in the skin secretions of winter flounder // J. Biol. Chem. 1997. Vol. 272, № 18. P. 12008–12013.
- 284. Syvitski R.T. et al. Structural characterization of the antimicrobial peptide pleurocidin from winter flounder // Biochemistry. 2005. Vol. 44, № 19. P. 7282–7293.
- 285. Jia X. et al. Antimicrobial Peptides Protect Coho Salmon from Vibrio anguillarum Infections // Appl. Environ. Microbiol. 2000. Vol. 66, № 5. P. 1928–1932.
- 286. Patrzykat A. et al. Sublethal concentrations of pleurocidin-derived antimicrobial peptides inhibit macromolecular synthesis in Escherichia coli. // Antimicrob. Agents Chemother. 2002. Vol. 46, № 3. P. 605–614.
- 287. Usachev K.S. et al. Oligomerization of the antimicrobial peptide Protegrin-5 in a membraneminicking environment. Structural studies by high-resolution NMR spectroscopy // Eur. Biophys. J. 2017. Vol. 46, № 3. P. 293–300.

- 288. Unubol N. et al. Peptide Antibiotics Developed by Mimicking Natural Antimicrobial Peptides // Clin. Microbiol. Open Access. 2017. Vol. 06, № 04.
- 289. Melo M.N., Ferre R., Castanho M.A.R.B. Antimicrobial peptides: Linking partition, activity and high membrane-bound concentrations // Nat. Rev. Microbiol. 2009. Vol. 7, № 3. P. 245–250.
- 290. Melo M.N., Castanho M.A.R.B. The Mechanism of Action of Antimicrobial Peptides: Lipid Vesicles vs. Bacteria // Front. Immunol. 2012. Vol. 3, № August. P. 1–4.
- 291. Mizukami S. et al. Enzyme-triggered compound release using functionalized antimicrobial peptide derivatives // Chem. Sci. Royal Society of Chemistry, 2017. Vol. 8, № 4. P. 3047–3053.
- 292. Lee J. et al. Effect of side chain hydrophobicity and cationic charge on antimicrobial activity and cytotoxicity of helical peptoids // Bioorg. Med. Chem. Lett. 2017.
- 293. Mishra B. et al. Host defense antimicrobial peptides as antibiotics: design and application strategies // Curr. Opin. Chem. Biol. Elsevier Ltd, 2017. Vol. 38. P. 87–96.
приложения

Приложение А. Статистика сборки генома H. medicinalis

Таблица А.1 - Статистика сборки генома *H. medicinalis*, полученной с использованием программы SPAdes на основе нуклеотидной последовательности генома *H. medicinalis*

	Общая сборка без биннинга	Геном H. medicinalis	Бактериальный метагеном
Суммарная длина, п.н.	232003367	154802040	48830001
Количество контигов	168624	29764	4094
N50	12947	13786	31787
Максимальная длина контига	1675099	188484	1675099
Минимальная длина контига	128	500	500

Таблица А.2 - Статистика скаффолдов эукариотических контигов *H. medicinalis*, полученных с использованием программы Sspase на основе нуклеотидной последовательности генома *H. medicinalis*

Суммарная длина, п.н.	187580802
Общее количество скаффолдов	14042
N50	97775
Минимальная длина скаффолда	500
Средняя длина скаффолда	13358
Количество прочтений	32778762
Суммарная длина, п.н.	154802040

Таблица А.3 - Статистика аннотации генома *H. medicinalis*

Характеристика	Количество	Характеристика	Количество
Общая длина последовательности	187580802	Длина самой короткой кодирующей последовательности	201
Количество генов	14596	Длина самого длинного гена	111002
Количество мРНК	15728	Длина самой длинной мРНК	111002
Количество экзонов	129359	Длина самого длинного экзона	10212
Количество интронов	113631	Длина самого длинного интрона	42547
Количество кодирующих последовательностей	15728	Длина самой длинной кодирующей последовательности	26013
Количество перекрывающихся генов	389	Средняя длина гена	7002
Общая длина генов	102198381	Средняя длина мРНК	6997
Общая длина мРНК	110044980	Средняя длина экзона	224
Общая длина экзонов	28951546	Средняя длина интрона	716
Общая длина интронов	81320696	Средняя длина интрона	1393
Общая длина кодирующих последовательностей	21901875	% генома, покрытого генами	54,5
Длина самого короткого гена	255	% генома, покрытого кодирующими последовательностями	11,7
Длина самой короткой мРНК	255	Среднее количество мРНК на ген	1
Длина самого короткого экзона	3	Среднее количество экзонов на мРНК	8
Длина самого короткого интрона	42	Среднее количество интронов на мРНК	7

Приложение Б. Геномы организмов, использованных в работе

i dettindu bil e i di tito i i								
Организм	Рег. номер BioProject ⁴	Размер генома, п.н.	Кол-во генов	Средняя длина гена, п.н.	Кол-во экзонов	Средняя длина экзона, п.н.	Кол-во интронов	Средняя длина интрона, п.н.
Hirudo medicinalis	PRJNA257563	187580802	14596	7002	129359	224	113631	716
Helobdella robusta	PRJNA521030	235376169	23426	3919	143175	203	119749	526
Capitella teleta	PRJNA175705	333283208	31978	2802	172792	215	140814	374
Anopheles gambiae	PRJNA254152	265027044	14723	5695	62328	411	47605	1476
Cules quinquefasciatus	PRJNA18751	579042118	22629	4746	74841	343	52212	1568
Cimex lectularius	PRJNA167477	650477627	25930	26672	250604	242	224674	4158
Ixodes scapularis	PRJNA16232	1765382190	24770	8966	93772	213	69002	2931
Myotis lucifugus	PRJNA16951	2034575300	46428	30016	487796	256	441368	4966
Linqula anatina	PRJNA286275	425494968	47016	8525	445078	267	398062	1048
Crassostrea gigas	PRJNA70283	557735934	45806	7965	476046	239	430240	1003
Aplysia californica	PRJNA13635	927310431	29094	20957	257038	316	227944	2643
Octopus bimaculoides	PRJNA270931	2338188782	26592	39379	237798	251	211206	6001

Табщица Б 1 -	Статистика	TELLOMOR H	modicinalis	прецстарителей	гематофагов и	I onhotrochozoa
таолица Б.т -	Claincinka		. meaicinaiis	, представителен	10 mature mat	$_{20}pnon ocno20a$

⁴ - регистрационный номер в базе данных NCBI

Соединение	Pfam/InterPro идентификатор	ИД ⁵ гена	ИД⁵ скаффолда	Начало гена	Конец гена	Длина гена (п.н.)	Кол-во экзонов	Длина (a.o.)
		g11304.t1	scaffold1202 size46801	20181	23845	3664	9	237
Бделлин А		g11305.t1	scaffold1202 size46802	25435	29835	4400	5	164
	PF02822 (Antistasin)	g11306.t1	scaffold1202 size46803	41426	46622	5196	6	389
		g4529.t1	scaffold210 size345909	124016	125423	1407	7	133
		g6012.t1	scaffold335 size63276	54026	62233	8207	5	106
Антистазин-	PF02822 (Antistasin)	g4239.t1	scaffold188 size67202	9556	12796	3240	10	391
		g6343.t1	scaffold364 size217826	196186	197474	1288	3	187
	PF00050	g6344.t1	scaffold364 size217827	197994	199424	1430	3	249
Блеплин В3	(Kazal-type serine	g6345.t1	scaffold364 size217827	199887	201284	1397	4	304
	protease inhibitor	g6346.t1	scaffold364 size217828	202703	204214	1511	3	287
	domain)	g6347.t1	scaffold364 size217829	206025	207066	1041	3	171
		g6348.t1	scaffold364 size217828	208945	210916	1971	3	258
		g11560.t1	scaffold1303 size13824	1414	5265	3851	3	258
	PF00280	g11561.t1	scaffold1303 size13825	5893	9585	3692	4	163
Эглин-С		g12915.t1	scaffold2531 size16879	6063	16605	10542	6	297
	(Polato inhibitor I family)	g11258.t1	scaffold1188 size60920	4236	5605	1369	5	201
		g11720.t1	scaffold1379 size30790	2356	5157	2801	5	305
	PF00059	g11865.t1	scaffold1454 size22262	8296	10271	1975	5	462
Лектоксин	(Lectin C type domain)	g13772.t1	scaffold4397 size2052	576	1775	1199	3	253
	(Lectin C-type domain)	g12022.t1	scaffold1543 size37248	31836	32983	1147	3	269
Саратин	NA	g9274.t1	scaffold725 size82469	58496	60431	1935	4	211
Cupurini		g6796.t1	scaffold404 size70608	56461	58105	1644	4	235
Ингибитор	IPR02406	g7223.t1	scaffold451 size223322	71666	72309	643	3	168
Гирулины	PF00713 (Hirudin)	g9136.t1	scaffold705 size226376	6617	7515	898	4	128
1 mp / dumpi		hirudin2	scaffold2168 size5612	121	5347	5226	-	-

Таблица Б.2 - Гены *Н. medicinalis*, кодирующие белки, вовлеченные во взаимодействие с жертвой.

Гирудин-	IPR011061	g9138.t1	scaffold705 size226377	12662	13565	903	4	106
подобный фактор	(Hirudin/antistatin)	g9139.t1	scaffold705 size226378	14431	16375	1944	5	287
		hirustasin	scaffold1202 size46801	22441	23034	593	-	-
		g4529.t1	scaffold210 size345909	124016	125423	1407	7	134
Гирустазин	PF02822 (Antistasin)	g11304.t1	scaffold1202 size46801	20181	23845	3664	9	237
		g11305.t1	scaffold1202 size46801	25435	29835	4400	5	164
		g11306.t1	scaffold1202 size46801	41426	46622	5196	6	389
		g8052.t1	scaffold551 size129727	90668	95785	5117	1	402
Ингилбитор		g8052.t2	scaffold551 size129728	90668	95785	5117	3	531
эластазы	PF00079 (Serpin)	g8052.t3	scaffold551 size129729	93413	95785	2372	3	623
		g8053.t1	scaffold551 size129730	102789	103805	1016	1	339
		g8054.t1	scaffold551 size129731	104600	109715	5115	4	665
		g6528.t1	scaffold380 size114101	66756	67848	1092	4	209
		g8541.t2	scaffold619 size54468	29596	35082	5486	4	190
		g8542.t1	scaffold619 size54469	44785	53915	9130	5	224
Лестабилаза	PF05497 (Destabilase)	253	scaffold253 size172121	-	-	-	-	-
		481	scaffold481 size52494	-	-	-	-	-
		701	scaffold701 size261053	-	-	-	-	-
		619_1	scaffold619 size54468	-	-	-	-	-
		619_4	scaffold619 size54468	-	-	-	-	-
	PF00031 (Cystatin	g11713.t1	scaffold1375 size41249	24999	25915	916	3	225
Цистатин	domain)	g10530.t1	scaffold991 size34421	4186	6605	2419	4	209
	,	g9499.t3	scaffold770 size110308	27016	30656	3640	6	370
Leech-derived tryptase inhibitor C	PF00050 (Kazal-type serine protease inhibitor domain)	Leech DTI	scaffold11 size165743	8937	9555	618	-	-

⁵ - идентификационный номер

Приложение В. Геномы организмов, использованных в работе

Номер кластера	Кол-во ортологов	ИД ⁶ протеомной базы	ИД ⁶ PFAM/ UniProt	Описание	ИД6 гена	Предполагаемая функция белка
16	5	hm comp13725_c0_seq1_m.28437; hm comp13725_c0_seq2_m.28442	PF07703	Alpha-2-macroglobulin family N-terminal region	g12051.t1	Альфа-2-макроглобулин
38	4	hm comp10037_c0_seq1_m.14144	PF07678	A-macroglobulin complement component	g7979.t1	Альфа-2-макроглобулин
134	2	hm comp10247_c0_seq1_m.14767	PF02822	Antistasin family	g4768.t1	Антистазин
3	8	hm comp7679_c1_seq1_m.8660; hm comp7679_c1_seq2_m.8661	PF02822	Antistasin family	g9922.t1	Антистазин
20	4	-	N/A	Therostasin	-	Антистазин
116	2	-	N/A	Therostasin	-	Антистазин
48	3	hm comp6253_c0_seq1_m.6064	N/A	WAP Cys-rich	g3476.t1	Антистазин
135	2	hm comp13086_c0_seq1_m.25222	PF13499	EF-hand domain pair	g10003.t1	Кальмодулин
11	6	hm comp7246_c0_seq1_m.7755 ; hm comp7246_c0_seq2_m.7756	PF00194	Eukaryotic-type carbonic anhydrase	g7193.t1	Карбоангидраза
115	2	hm comp11142_c0_seq4_m.17572	PF00188	Cysteine-rich secretory protein family	-	CRISP
51	3	hm comp17910_c0_seq1_m.33641	PF00188	Cysteine-rich secretory protein family	g10878.t1	CRISP

Таблица В.1 - Кластеры ортологичных	белков, илентифицирован	ных с помошью анализа про	теома ССК <i>H. medicinalis</i>
Turner op reporter in the	••••••••••••••••••••••••••••••••••••••		

42	3	hm comp9835_c0_seq2_m.13648; hm comp9835_c0_seq1_m.13647	PF00188	Cysteine-rich secretory protein family	g7000.t1	CRISP
8	6	hm comp9437_c0_seq1_m.12642; hm comp9437_c0_seq2_m.12643	PF00188	Cysteine-rich secretory protein family	g4751.t1	CRISP
75	3	hm comp8954_c0_seq1_m.11418	PF00188	Cysteine-rich secretory protein family	g7001.t1	CRISP
72	3	hm comp14944_c0_seq1_m.32021	PF00754	F5/8 type C domain	g7916.t1	F5/8 type C domain like
58	3	hm comp9278_c0_seq2_m.12219	P14328	Spore coat protein SP96/lectin	g11933.t1	F5/8 type C domain like
26	4	hm comp11442_c1_seq2_m.18518; hm comp11442_c1_seq4_m.18519	PF00147	Fibrinogen beta and gamma chains, C-terminal globular domain	g14072.t1	Фиколин
122	2	hm comp5522_c0_seq1_m.4924	PF00042	Globin	g3446.t1	Глобин
126	2	hm comp4318_c0_seq1_m.3402	PF03662	Glycosyl hydrolase family 79, N-terminal domain	g5922.t1	Гиалуронидаза
29	4	hm comp6781_c0_seq1_m.6916	N/A	Glycosyl hydrolase family 79, N-terminal domain	g5922.t1	Гиалуронидаза
45	3	hm comp11138_c0_seq3_m.17552; hm comp11138_c0_seq1_m.17550	PF14295	PAN domain	g6052.t1	PAN-domane
139	2	hm comp2891_c0_seq1_m.2042	PF08277	PAN-like domain	g976.t1	PAN-domane
61	3	hm comp8764_c0_seq1_m.10994	N/A	Peptidase family M28	g10407.t1	Протеаза М28
143	2	hm comp8764_c0_seq1_m.10995	PF04389	Peptidase family M28	g10407.t1	Протеаза М28

11	16
----	----

125	2	hm comp6953_c0_seq1_m.7241	PF00652	Ricin-type beta-trefoil lectin domain	g7423.t1	Лектин типа R
53	3	hm comp9223_c0_seq1_m.12083	PF00080	Copper/zinc superoxide dismutase (SODC)	g7707.t1	Супероксиддизмутаза
59	3	hm comp7740_c0_seq1_m.8805	PF00335	Tetraspanin family	-	Тетраспанин
37	4	hm comp11399_c0_seq1_m.18370	PF00092	von Willebrand factor type A domain	g6733.t1	vWFA-like
74	3	hm comp14077_c0_seq1_m.30596	PF08742, PF01826	C8 domain, Trypsin Inhibitor like cysteine rich domain	g10240.t1	vWFD contain
2	9	hm comp14077_c0_seq1_m.30595; hm comp14077_c0_seq3_m.30600	PF00094	von Willebrand factor type D domain/C8/	g10240.t1	vWFD contain
22	4	hm comp14077_c0_seq1_m.30597	N/A	Protease ingibitor 15	g10240.t1	vWFD contain
43	3	hm comp14077_c0_seq4_m.30601; hm comp14077_c0_seq5_m.30603	N/A	N/A 21	g10240.t1	vWFD contain
28	4	hm comp11962_c0_seq1_m.20426	N/A	N/A 10	g11991.t1	Не определена
36	4	hm comp11521_c0_seq1_m.18840	N/A	N/A 12	g5191.t1	Не определена
39	4	hm comp11521_c0_seq2_m.18842	N/A	N/A 13	g5191.t1	Не определена
62	3	hm comp11444_c0_seq1_m.18524	N/A	N/A 15	g7928.t1	Не определена
64	3	hm comp11444_c0_seq1_m.18525	N/A	N/A 16	g7928.t1	Не определена
12	6	hm comp11962_c0_seq1_m.20427	N/A	N/A 7	g11991.t1	Не определена
60	3	hm comp10250_c0_seq1_m.14781	N/A	N/A 14	g1441.t1	Не определена

70	3	hm comp12147_c2_seq1_m.21111	N/A	N/A 17	g14193.t1	Не определена
124	2	hm comp13252_c0_seq1_m.26031	N/A	N/A 19	g7938.t1	Не определена
123	2	hm comp13132_c0_seq3_m.25451	N/A	N/A 24	g2705.t2	Не определена
131	2	hm comp10776_c1_seq1_m.16393	N/A	N/A 25	-	Не определена
144	2	hm comp4621_c0_seq1_m.3703	N/A	N/A 27	g5894.t1	Не определена
113	2	hm comp9492_c0_seq1_m.12768	N/A	N/A 28	-	Не определена
40	3	-	N/A	N/A 29	-	Не определена
57	3	-	N/A	N/A 30	-	Не определена
111	2	-	N/A	N/A 31	-	Не определена
24	4	hm comp7698_c0_seq2_m.8710 ; hm comp7698_c0_seq1_m.8709	N/A	N/A 9	g6936.t1	Не определена
67	3	hm comp11821_c0_seq1_m.19872	PF00171	Aldehyde dehydrogenase family	g3624.t1	Другие
68	3	hm comp11821_c0_seq1_m.19870	PF00171	Aldehyde dehydrogenase family	g3624.t1	Другие
133	2	hm comp13015_c0_seq1_m.24875	PF00012	Hsp70 protein	-	Другие
140	2	hm comp9614_c0_seq2_m.13067	PF00244	14-3-3 protein	g3791.t1	Другие
46	3	hm comp11805_c0_seq2_m.19813; hm comp9614_c0_seq3_m.13068	PF00244	14-3-3 protein	g3990.t1, g11995.t1	Другие
6	6	hm comp9678_c0_seq1_m.13225	PF00022	Actin	g4376.t1	Другие

110

114	2	hm comp5751_c0_seq1_m.5242	PF00306, PF00006	ATP synthase alpha/beta chain, C terminal domain, ATP synthase alpha/beta family, nucleotide- binding domain	g2850.t1	Другие
27	4	hm comp12822_c1_seq2_m.23878	PF03114	BAR domain	g4525.t1	Другие
132	2	hm comp13803_c0_seq3_m.28868	PF07716	Basic region leucine zipper	g2990.t1	Другие
117	2	hm comp6941_c0_seq1_m.7226	PF15511	Centromere kinetochore component CENP-T histone fold	g9617.t1	Другие
41	3	hm comp14028_c0_seq2_m.30226	PF00241	Cofilin/tropomyosin-type actin- binding protein	g7441.t1	Другие
47	3	hm comp12490_c0_seq3_m.22497; hm comp12490_c0_seq1_m.22494	PF00012	Hsp70 protein	g7896.t1	Другие
121	2	hm comp13774_c0_seq3_m.28694	PF12777, PF12775, PF12780	Microtubule-binding stalk of dynein motor, P-loop containing dynein motor region D3, P-loop containing dynein motor region D4	g433.t1	Другие
65	3	hm comp12041_c0_seq1_m.20713	PF00071	Ras family	g7577.t1	Другие
141	2	hm comp5511_c0_seq1_m.4909	PF01020, PF00240	Ribosomal L40e family, Ubiquitin family	g4151.t1	Другие
118	2	hm comp6601_c0_seq1_m.6595	PF01599, PF00240	Ribosomal protein S27a, Ubiquitin family	-	Другие

⁶ - идентификационный номер

Приложение Д. Описание базы данных антимикробных пептидов, собранной в работе

База данных	Количество исходных последовательностей	Количество кластеров АМП
ADAM	7007	2493
APD3	2619	1910
UniProtKB	1781	451
DADP	1923	592
CAMP _{R3}	8164	269

Таблица Д.1 - Базы данных антимикробных пептидов