

Федеральное государственное автономное образовательное учреждение
высшего образования
«Южный федеральный университет»

На правах рукописи



Куликов Максим Павлович

**Молекулярно-генетический подход к поиску новых нерибосомально синте-
зируемых противогрибковых метаболитов бактерий**

1.5.7. Генетика (биологические науки)

Диссертация на соискание учёной степени
кандидата биологических наук

Научный руководитель
доктор биологических наук,
заведующий научной лабораторией
Празднова Евгения Валерьевна

Ростов-на-Дону — 2026

ОГЛАВЛЕНИЕ

ПЕРЕЧЕНЬ СОКРАЩЕНИЙ, УСЛОВНЫХ ОБОЗНАЧЕНИЙ, СИМВОЛОВ, ЕДИНИЦ И ТЕРМИНОВ.....	6
ВВЕДЕНИЕ.....	8
1. ОБЗОР ЛИТЕРАТУРЫ.....	16
1.1. Характеристика нерибосомальных пептид-синтетаз.....	16
1.1.1. Механизм работы и строение НРПС.....	18
1.1.2. Домены НРПС.....	20
1.1.2.1. А-домен.....	20
1.1.2.2. РСР-домен.....	23
1.1.2.3.С-домен и его вариации.....	24
1.1.2.4. Тиоэстеразные ТЕ и ТЕП домены.....	27
1.1.2.5. R-домены.....	28
1.1.2.6. Домены модификации субстрата.....	29
1.1.2.7. Некаталитические домены.....	32
1.1.3. Типы НРПС.....	33
1.1.3.1. Линейные НРПС.....	34
1.1.3.2. Итеративные НРПС.....	35
1.1.3.3. Нелинейные НРПС.....	35
1.1.3.4. Поликетидсинтетазы (ПКС) и гибридные НРПС/ПКС.....	37
1.1.4. Регуляция нерибосомального синтеза.....	39
1.2. Аннотирование НРПС.....	41
1.3. Основные подходы к инженерии НРПС.....	45
1.4. Отличительные особенности нерибосомальных пептидов.....	47
1.5. НРП с противогрибковыми свойствами.....	49

1.5.1 Итурины	49
1.5.2. Фенгицины	50
1.5.3. Сурфактин	50
1.5.4. Полимиксины	51
1.5.5. HSAF	52
1.5.6. Сирингомицин и сирингопептин	52
1.5.7. Тилоперин и фузаристатин	53
1.6. Практическое применение	53
2. МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ	56
2.1. Характеристика штаммов и видов биологических объектов	56
2.2. Выделение метаболитов и масс-спектрометрия	56
2.2.1. Выделение и очистка липопептидов	57
2.2.2. Масс-спектрометрический анализ (MALDI-TOF)	58
2.3. Выделение ДНК	58
2.3.1. Подготовка биомассы	58
2.3.2. Лизис клеток	59
2.3.3. Очистка ДНК	59
2.4. Секвенирование	60
2.4.1. Секвенирование на платформе Illumina	60
2.4.2. Нанопоровое секвенирование (Oxford Nanopore Technologies)	60
2.5. Анализ геномов	61
2.5.1. Сборка геномов из коротких прочтений (Illumina)	61
2.5.2. Сборка геномов из нанопоровых прочтений и гибридная сборка	61
2.5.3. Таксономическая идентификация штаммов	62
2.6. База данных для машинного обучения	62

2.7. Биоинформатический анализ и машинное обучение.....	63
2.7.1. Обработка данных и извлечение признаков.....	63
2.7.2. Интерпретация моделей и статистический анализ.....	66
2.7.3. Автоматизация аннотации и предсказания.....	67
2.8. Тестирование антагонистической и противогрибковой активности <i>in vitro</i>	68
2.8.1. Первичный скрининг антагонистической активности методом двой- ных культур.....	68
2.8.2. Оценка активности культуральной жидкости.....	69
2.9. Дизайн модельного эксперимента на ячмене (<i>Hordeum vulgare</i>).....	69
2.10. Дизайн модельного эксперимента на стерляди (<i>Acipenser ruthenus</i>).....	70
2.11. Анализ экспрессии генов.....	70
3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ.....	72
3.1. Секвенирование и сборка геномов новых штаммов из коллекции лабора- тории.....	72
3.2. Сравнение результатов масс-спектрометрии метаболитов, выделенных из штаммов, с предсказанием AntiSMASH.....	72
3.3. Разработка улучшенного алгоритма AFPredict.....	87
3.3.1. Создание базы данных и первичный анализ БГК.....	87
3.3.2. Машинное обучение.....	94
3.4. Аннотация геномов новых штаммов с помощью AFPredictor.....	101
3.5. Повторный анализ штаммов с использованием алгоритма AFPredictor..	103
3.6. Подтверждение предсказанной противогрибковой активности <i>in vitro</i> ..	105
3.7. Анализ биосинтетического потенциала штаммов-пробиотиков для аква- культуры.....	106
3.8. Изучение активности штаммов-продуцентов НРП <i>in vivo</i>	108

3.9. Изучение стимулирующей рост активность на растениях.....	108
3.10. Изучение пробиотической активности на объектах аквакультуры.....	114
ЗАКЛЮЧЕНИЕ.....	116
ВЫВОДЫ.....	118
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ.....	119
Приложение 1.....	137
Приложение 2.....	148
Приложение 3.....	152
Приложение 4.....	157

ПЕРЕЧЕНЬ СОКРАЩЕНИЙ, УСЛОВНЫХ ОБОЗНАЧЕНИЙ, СИМВОЛОВ, ЕДИНИЦ И ТЕРМИНОВ

А-домен – домен аденилирования

БГК – биосинтетический генный кластер / кластеры биосинтетических генов

БКЖ – бесклеточная культуральная жидкость

КОЕ – колониеобразующая единица

МПА – мясо-пептонный агар

НРП – нерибосомальные пептиды

НРПС – нерибосомальные пептидсинтазы

ПКС – поликетидсинтазы / поликетидсинтазы

ПЦР-РВ – полимеразная цепная реакция в реальном времени

С-домен – домен конденсации

СЖК – синтазы жирных кислот

УЗВ – установка замкнутого водоснабжения

ФПТ – 4'-фосфопантетеинилтрансферазы

ЦТАБ – цетилтриметиламмонийбромид

ЭДТА – этилендиаминтетрауксусная кислота

АСР – белок-переносчик ацила

ANI – среднее нуклеотидное тождество (Average Nucleotide Identity)

AP – средняя точность (Average Precision)

АТ – ацилтрансфераза

СОМ – коммуникационные домены

Су-домен – домен гетероциклизации

DD – докинг-домены

DH – дегидратаза

Е-домен – домен эпимеризации

ER – еноилредуктаза

F-домен – домен формилирования

FDR – доля ложных открытий

FPR – доля ложно положительных ответов (False Positive Rate)

GTDB – Genome Taxonomy Database

HSFAF – термостабильный противогрибковый фактор (Heat-stable antifungal factor)

KR-домен – кетоацилредуктазный домен / кеторедуктаза

KS – кетоацилсинтаза

M-домен – домен метилирования

m/z – соотношение массы к заряду

MALDI-TOF – времяпролетная масс-спектрометрия с матрично-активированной лазерной десорбцией/ионизацией

MCC – Коэффициент корреляции Мэтьюса (Matthews Correlation Coefficient)

MLP – MbtH-подобные белки

MRSA – метициллин-резистентный *Staphylococcus aureus*

OOOF – Out-of-Fold предсказания

Ох-домен – оксидазный домен (зависимый от флавинмонопнуклеотида)

PCP – белок-переносчик пептидилов

PU-обучение – Positive-Unlabeled Learning

R-домен – редуктазный домен

ROC AUC – площадь под ROC-кривой (Receiver Operating Characteristic Area Under Curve)

SHAP – векторы Шепли (SHapley Additive exPlanations)

T-домен – домен тиолирования

TE-домен – тнэстеразный домен (включая TEII)

TPR – доля истинно положительных ответов (True Positive Rate)

ВВЕДЕНИЕ

Актуальность темы

Нерибосомально синтезируемые продукты бактерий – это вещества преимущественно с антимикробным и противогрибковым действием, хотя среди них встречаются и противовирусные, противораковые и иммуномодулирующие вещества. Особый механизм их синтеза (без участия рибосом) и особая модульная структура производящих их ферментов делает эти вещества перспективным объектом биотехнологии и биоинженерии, поскольку понимание этой структуры позволит в будущем перейти к производству веществ с заданными свойствами.

Наиболее известны противогрибковые нерибосомальные пептиды, и они наиболее широко представлены в биоинформатических базах данных, поэтому в данной работе они используются как модель для проверки работоспособности создаваемого алгоритма. Однако в дальнейшем этот алгоритм можно будет применить для поиска и разработки новых антибиотиков и других биоактивных веществ.

Серьезную угрозу для сельскохозяйственных животных, в частности для мировой аквакультуры, представляют оомицеты рода *Saprolegnia*. Данные патогены вызывают сапролегниоз, характеризующийся тяжелым поражением тканей, осмотическим дисбалансом и последующей массовой гибелью пресноводных рыб (особенно ценных лососевых видов) и их икры (Провоторов и др., 2022; Shreves et al., 2024; Тангатарова и др., 2024). Поражая внешние покровы и жабры, инфекция стремительно нарушает осморегуляцию и дыхательные функции, что ведет к быстрой гибели хозяина и огромным экономическим потерям для рыбных хозяйств, сопоставимым с ущербом от фузариоза в агросекторе. Использование некогда наиболее эффективных синтетических препаратов против них, таких как малахитовый зеленый, повсеместно запрещено из-за их доказанной токсичности, мутагенности и канцерогенности. В то же время к разрешенным химическим альтернативам возбудители проявляют высокую способность к адаптации, что делает поиск новых, экологически безопасных антими-

кробных средств одной из наиболее приоритетных задач отрасли (Lindholm-Lehto, Pylkkö, 2024; Tang и др., 2025).

Говоря о грибах, которые часто вредят сельскому хозяйству в целом, нельзя не упомянуть грибы рода *Fusarium*. Все злаковые культуры подвержены фузариозу, однако наиболее часто встречаются фузариозы пшеницы, которые могут приводить более чем 50% прямым потерям урожая, а также снижению качества собранного зерна (Nicholson et al., 2007; Khaytbayeva et al., 2021; Щеклеина, Шешегова, 2024). Кроме того, данные грибы при этом способны нанести серьезный вред здоровью сельскохозяйственным животным и людям-потребителям сельхоз продукции, из-за выработки большого спектра микотоксинов, из которых наиболее токсичными являются трихотецены, фумонизины и зеараленон (Munkvold, 2017; Гагкаева и др., 2016). При этом *Fusarium* также постепенно развивает устойчивость к широко используемым противогрибковым препаратам, а некоторые штаммы проявляют мультирезистентность – так, известны случаи устойчивости к азолам, эхинокандам, амфотерицину и вориконазолу (D’Agostino и др., 2018; Taj-Aldeen, 2017; Соколова, Глинушкин, 2020).

По причине все возрастающей устойчивости грибковых патогенов требуются все новые и новые соединения борьбы с ними, и перспективным источником подобных соединений являются бактерии, в особенности представители *Firmicutes*, так как они имеют широкий спектр кластеров генов, ответственных за синтез вторичных метаболитов, среди которых много противогрибковых веществ. Согласно геномным исследованиям, 31 % представителей типа *Firmicutes* имеют кластеры генов, кодирующих разнообразные вторичные метаболиты, а 70 % этих кластеров кодируют нерибосомальные пептид-синтетазы (НРПС) (Wang и др., 2014).

Однако даже сейчас, в большинстве случаев, поиск новых соединений проводится функциональным скринингом бактерий-продуцентов. Этот подход позволил обнаружить подавляющее большинство из ныне известных классов антибиотиков (Aminov, 2017). Однако, он во многом себя исчерпал. В частности, переоткрытие давно известных веществ стало настоящей проблемой

(Lewis, 2013). Особенно ярко эта проблема заметна на примере нерибосомальных пептидов, многие из которых либо повторно открываются, если скрининг проводится без использования молекулярных методов, на основе одной лишь противогрибковой активности. Также нередки случаи, когда новые метаболиты выделены, но гены, ответственные за их синтез, неизвестны (He et al., 2022), либо наоборот, гены синтеза метаболитов известны лишь гипотетически по данным секвенирования (Agrawal и др., 2017; Blin и др., 2021; Medema и др., 2011), но их продукты не выделены и функции неизвестны, а и их структура не подтверждена ни экспериментально, ни расчетно. Эти случаи объясняются тем, что, хотя биосинтетический потенциал бактерий огромен, но по крайней мере, три четверти специализированных метаболитов не производятся микроорганизмами постоянно и, за редким исключением, их продукцию невозможно обнаружить в условиях стандартного лабораторного культивирования.

Разработка алгоритма для поиска генов нерибосомального синтеза и предсказания противогрибковых свойств метаболитов на основе их структуры позволит значительно ускорить и удешевить разработку новых противогрибковых препаратов, а в будущем также антибиотиков и других полезных соединений

Проработанность исследуемой проблематики в науке

Несмотря на то, что исследование нерибосомального синтеза началось в 1950-х годах, технические ограничения долгое время не позволяли определить строение нерибосомальных пептид-синтетаз и точный механизм их действия. Фактически, структурные исследования начались только в начале 21 века. Поэтому можно сказать, что данная проблема проработана довольно односторонне. Достаточно хорошо изучены основные структурные элементы наиболее распространенных нерибосомальных пептид-синтетаз - домены аденилирования, пептидил-переноса, конденсации, эпимеризации и тимоэстеразный домен, изучены механизмы активации и переноса аминокислот, а также формирования пептидных связей, установлены связи между структурой генов и производимыми пептидами (Challis, Naismith, 2004; Weissman, 2015; Dowling и др., 2016; Marahiel,

2016; Полуэктова и др., 2022). Существуют базы данных известных нерибосомальных пептид-синтетаз (Zdouc и др., 2025). Также разработано программное обеспечение, позволяющее находить генные кластеры ответственные за нерибосомальный синтез в бактериальных геномах и частично предсказывать их структуру (Blin и др., 2021; Skinnider и др., 2017). Активно изучаются экологические функции НРПС (Ongena, Jacques, 2008; Lee et al., 2013; Tang et al., 2014; Sumi и др., 2015; Aleti, 2016; Tajbakhsh и др., 2017). В частности, наш коллектив ранее продемонстрировал антагонистическую активность штаммов *Bacillus* и *Paenibacillus* в отношении грибов рода *Fusarium*, обусловленную нерибосомальными пептидами, а также изменение экспрессии генов НРПС у штаммов при совместном культивировании с фузариумом (Vasilchenko et al., 2022).

Тем не менее, в рамках существующих исследований остается ряд значимых методических и концептуальных пробелов. Существующие биоинформатические подходы традиционно ориентированы на скрининг на уровне целых генных кластеров. Однако такой подход показал свою ограниченность при работе со сложносоставными биосинтетическими генными кластерами (БГК), не позволяя эффективно идентифицировать и разделять множественные нерибосомальные пептидсинтазы в их составе. Проблема выявления отдельных НРПС внутри крупных локусов требует перехода к более детализированному, субкластерному уровню скрининга. Существующие алгоритмы фокусируются на структуре, в то время как влияние строения нерибосомальных метаболитов на их специфические биологические функции (включая противогрибковую активность) изучено недостаточно.

В ранее выполненной дипломной работе изучалось влияние изменения экспрессии плеiotропных и специфических регуляторных генов на экспрессию генов нерибосомального синтеза и бактерий родов *Bacillus* и *Paenibacillus* при совместном культивировании с фузариумом (Васильченко Н.Г., 2023). В рамках выполнения данной работы уже выделено и отсекуено множество штаммов продуцентов противогрибковых веществ и определена нерибосомальная природа этих соединений с помощью MALDI-TOF масс-спектрометрии.

Цель исследования:

Разработать методику для выявления и прогнозирования структуры новых противогрибковых метаболитов, синтезируемых нерибосомальным путем, на основе созданной нами модели машинного обучения с применением подходов молекулярной генетики и биоинформационного анализа.

Задачи исследования:

1. Сформировать базу данных известных нерибосомальных пептидов с подтвержденной противогрибковой активностью и кластеров генов их синтеза.
2. Разработать модель с применением машинного обучения для предсказания противогрибковых свойств у нерибосомально синтезируемых метаболитов бактерий на основе геномных данных.
3. Провести отбор штаммов, проявляющих антагонистическую активность против родов *Fusarium* и *Saprolegnia*, патогенных для растений и животных, и провести их биохимический и генетический анализ.
4. Выполнить валидацию разработанного алгоритма на целевой группе штаммов-продуцентов, обладающих подтвержденной противогрибковой активностью *in vivo*.

Научная новизна и значимость ожидаемых результатов исследования

Впервые был разработан и обоснован новый комплексный подход к прогнозированию функциональных (противогрибковых) свойств нерибосомальных пептидов (НРП) на основе геномных данных. Подход базируется на интеграции данных о доменной архитектуре синтетаз, аминокислотной специфичности доменов аденилирования по сайту aa34 и контексте вспомогательных генов БГК. Созданная модель машинного обучения позволяет предсказывать противогрибковые свойства с точностью до 80%, и впервые дает возможность оценить функциональность кластера генов по данным секвенирования.

Обоснована целесообразность проведения биоинформатического скрининга биосинтетических генных кластеров на субкластерном уровне. Показано,

что такой подход позволяет успешно идентифицировать отдельные нерибосомальные пептид-синтетазы в сложносоставных кластерах биосинтетических генов, что не было достижимо ранее существовавшими методами. При аннотации геномов 69 штаммов было идентифицировано 1096 соединений (на 104 больше, чем antiSMASH), при этом у 54 штаммов обнаружены БГК с несколькими скрытыми подкластерами. Впервые выявлены устойчивые соседствующие в одном БГК пары соединений (фенгицин и микосубтилин у *Bacillus atrophaeus*; бациллибактин и амилоциклин, а также бацилломицин D и фенгицин у *B. velezensis*; китациннамицины и скилламицины у *Streptomyces anulatus*). В результате применения метода обнаружены как скрытые кластеры известных метаболитов в изученных геномах, так и ряд новых нерибосомальных пептид-синтетаз. Найден новый лихенизин-подобный противогрибковый метаболит, синтезируемый штаммом *Bacillus safensis* AG46.

Положения, выносимые на защиту

1. Субстратная специфичность доменов аденилирования нерибосомальных пептид-синтетаз критически важна для предсказания их противогрибковых свойств их продуктов.
2. Скрининг бактериальных геномов на уровне субкластеров, в отличие от применяемого до сих пор скрининга на уровне кластеров, является эффективным подходом для обнаружения и идентификации множественных нерибосомальных пептидсинтаз, кодируемых в пределах одного генного кластера.
3. Предсказание противогрибковых свойств нерибосомальных пептидов на основе геномных данных возможно с использованием методов машинного обучения, при условии комплексного анализа трех элементов: доменной организации синтаз, последовательности сайта домена аденилирования aa34, и контекста вспомогательных генов биосинтеза.
4. Разработанный алгоритм для анализа геномов микроорганизмов позволяет обнаруживать новые нерибосомальные метаболиты с противогрибковыми свойствами, что подтверждено масс-спектрометрически и *in vitro*.

Апробация результатов исследований

Материалы, положенные в основу работы, были представлены на Международной научной конференции XIV Международной научно-практической конференции «ИНТЕРАГРОМАШ 2021» (2021 г.), «BIO Web of Conferences» (2024 г.), XIX Международной научно-практической конференции («ИнтерАгро 2026») (2026 г.) и II конференции Ростовского общества генетиков и селекционеров «Фундаментальные и прикладные аспекты генетики и селекции» (2026 г.).

Публикации

По материалам диссертации опубликовано 6 работ, в том числе 4 статьи в журналах, рецензируемых ВАК Минобрнауки РФ, из них 3 в журналах, индексируемых в БД Scopus/WOS.

Объем и структура работы

Диссертация изложена на 164 страницах машинописного текста и состоит из введения, обзора литературы, экспериментальной части и обсуждения результатов, выводов, списка литературы (156 источник). Иллюстрационный материал включает 35 рисунков, 7 таблиц, 4 приложения.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках госзадания (Южный федеральный университет, проекты №№ 0852-2020-0029, FENW-2023-0008, FENW-2026-0019), а также в рамках программы «Приоритет-2030» в Молодежной лаборатории «Молекулярная генетика микробных консорциумов» и лаборатории «Биотехнологический кластер» Академии биологии и медицины им. Д. И. Ивановского Южного федерального университета (АБИМ ЮФУ). Экспериментальные работы в области аквакультуры, выделение штаммов-пробиотиков для аквакультуры и их секвенирование проводились в Научно-исследовательской лаборатории «Центр агробиотехнологий» в Донском государственном техническом университете (НИЛ ЦАБТ ДГТУ), при поддержке проекта РНФ № 23-76-30006.

1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Характеристика нерибосомальных пептид-синтетаз

Изучение синтеза нерибосомальных пептидов (НРП) началось в начале 1960-х годов с исследований биосинтеза циклических пептидов, содержащих D-аминокислоты, таких как тироцидин, производимый видами *Bacillus*. Татум и соавторы показали, что на биосинтез тироцидина в клетках не влияют ингибиторы рибосом, и выдвинули гипотезу, что его синтез происходит по механизму, отличному от трансляции белка. В дальнейшем Фриц Липманн и Сёрен Лаланд продемонстрировали, что активация аминокислот осуществляется через образование тиоэфирных связей с 4'-фосфопантетеиновыми кофакторами в АТФ-зависимой реакции, сопровождающейся выделением пирофосфата и АМР. Ковалентное связывание аминокислот с ферментами открыло представление о роли Ррант-переносчика в каждом модуле НРПС, а пошаговый характер сборки — о модульной организации комплекса.

С конца XX века интенсивно развивались молекулярно-биологические методы: клонирование и экспрессия генов нерибосомальных пептидсинтаз/синтетаз (НРПС) *in vitro*, направленная инактивация отдельных компонентов, а затем полногеномное секвенирование микроорганизмов-продуцентов. В 1988 году группа Марахиеля расшифровала нуклеотидную последовательность гена *tusA*, кодирующего тироцидин-синтетазу, что стало первым генетическим подтверждением структуры НРПС. Уже через год ими был описан оперон биосинтеза грамицидина С, включающий гены *grsA* и *grsB*. Тем самым была окончательно подтверждена модульная организация этих многофункциональных энзимов, что заложило основу для предсказания структуры нерибосомных пептидов по данным секвенирования.

Нерибосомальные пептид-синтетазы представляют собой мультидоменные ферментативные комплексы, которые катализируют синтез структурно разнообразных пептидов без участия рибосом и могут включать широкий спектр субстратов, включая стандартные и непротеиногенные аминокислоты, жирные

кислоты, α -гидроксикислоты, α -кетокислоты и гетероциклы, что обеспечивает химическое и структурное разнообразие нерибосомальных пептидов (Krättschmar, Krause, Marahiel, 1989; Mach, Reich, Tatum, 1963; Weckermann, Fürbass, Marahiel, 1988; Süssmuth, Mainz, 2017; Stanišić, Kries, 2019). Сходный принцип также применяется бактериями для синтеза поликетидов с помощью поликетидсинтетаз (ПКС), только в их случае субстратом служат малонил-КоА и его производные. Некоторые биосинтетические кластеры содержат гибридные НРПС-ПКС системы, сочетающие оба механизма синтеза в единой ферментативной архитектуре.

Молекулярная масса НРПС может варьироваться в широких пределах. Так, ферменты, производящие небольшие пептиды имеют массу около 100 кДа, а крупнейшая из известных непрерывных бактериальных НРПС - синтетаза колоссина А - состоит из 15 модулей и 45 доменов имеет молекулярную массу 1,8 мДа. (Bode и др., 2015; Iacovelli, Bovenberg, Driessen, 2021; Cai, Zhao, Bode, 2023).

Хотя нерибосомальные пептиды продуцируются различными группами бактерий, основными филумами, ответственными за производство большинства бактериальных НРП, являются *Proteobacteria*, *Actinobacteria*, *Firmicutes* и *Cyanobacteria*. . Накопление кластеров генов НРПС в этих четырёх филумах может частично отражать их сложный экологический образ жизни и потребность в механизмах химической защиты и коммуникации .

В рамках *Proteobacteria* ключевыми продуцентами выступают *Gammaproteobacteria* (особенно псевдомонады, синтезирующие сидерофоры и циклические липопептиды) и *Deltaproteobacteria* (преимущественно миксобактерии, обладающие разнообразным вторичным метаболизмом), тогда как *Betaproteobacteria* рассматриваются как потенциальный источник новых НРП (Baldim и др., 2017; Léger et al., 2021; Kooner, Soby, 2022; Wang и др., 2014; Borsetto et al., 2019). *Actinobacteria*, в частности стрептомицеты, демонстрируют высокую продуктивность благодаря крупным геномам, содержащим множе-

ственные кластеры НРПС и поликетидсинтетаз, что обуславливает синтез значительной доли известных антибиотиков и биоактивных соединений (Komaki и др., 2015; Sharma, Kaur, Salwan, 2021; Jian и др., 2023). Представители *Firmicutes*, такие как *Bacillus*, продуцируют структурно разнообразные пептиды, включая антибиотики (грамицидин, бацитрацин) и липопептиды. *Cyanobacteria* синтезируют биотоксичные НРП (например, микроцистины), играющие критическую роль в экологии водных систем. Таким образом, данные филумы представляют собой важнейшие источники НРП с широким спектром биологической активности и значительным фармакологическим потенциалом.

Благодаря своему структурному разнообразию и высокой биологической активности нерибосомальные пептиды нашли широкое применение в медицине и сельском хозяйстве. В медицинской сфере НРП применяются в качестве антибиотиков, иммунодепрессантов и противоопухолевых агентов. В сельском хозяйстве они используются как агенты биологического контроля. Изучаются возможности применения нерибосомально синтезируемых сидерофоров для селективного извлечения металлов из сложных промышленных растворов, что стало бы более экологичным и эффективным по сравнению с традиционными методами при низких концентрациях целевых катионов.

В данной главе произведена систематизация имеющейся информации о молекулярно-генетических механизмах нерибосомального синтеза у прокариот, определение малоизученных областей, перспектив развития данной области исследований и потенциально новых областей применения продуктов нерибосомального синтеза.

1.1.1. Механизм работы и строение НРПС

Синтез пептидов нерибосомальными пептид-синтетазами осуществляется посредством образования амидных связей между аминокислотными или другими аминокислотными мономерами. В отличие от рибосомального синтеза, где все аминокислоты добавляются на одном и том же активном центре, НРПС функционируют по модульному принципу: для взаимодействия с каждой аминокислот-

ным остатком задействован отдельный набор ферментативных доменов. Совокупность таких доменов образует модуль. Выделяют 3 типа модулей: элонгации, инициации и терминации.

Модуль инициации, с которого начинается синтез пептида, как правило, включает домен аденилирования (А) и домен тиолирования (Т), также известный как домен белка-переносчика пептидилов (РСР). За модулем инициации следуют один или несколько модулей элонгации, отличающихся наличием домена конденсации (С) расположенным перед доменами А и РСР.

Синтез пептида начинается с активации субстрата в домене аденилирования, который может распознавать и активировать как протеиногенные, так и непротеиногенные аминокислоты, такие как D-формы протеиногенных аминокислот, орнитин, иминокислоты и т. д (Walsh, O'Brien, Khosla, 2013; Miller, Gulick, 2016; Iacovelli, Bovenberg, Driessen, 2021; Saiyam и др., 2024; Kalyvas et al., 2025; Hartnell и др., 2020; Wang и др., 2014; Felnagle и др., 2008; Agrawal и др., 2017; Kaspar, Neubauer, Gimpel, 2019; Ranjan et al., 2023; Sieber и др., 2024; Waongo и др., 2025). На данном этапе в присутствии Mg-АТФ А-домен аденилирует аминокислоту, образуя аминокацил-АМФ, который затем взаимодействует с тиоловой группой 4'-фосфопантетеина (РР) домена РСР, с образованием комплекса пептидил-РСР. Пептидил-РСР взаимодействует с донорным туннелем домена конденсации следующего модуля, образуя комплекс РСР-С, в результате чего происходит активация А-домена этого модуля и описанный ранее процесс повторяется в этом модуле, вплоть до момента образования пептидил-РСР, который взаимодействует сначала с акцепторным туннелем собственного С-домена, образуя комплекс РСР-С-РСР. С-домен катализирует образование пептида с образованием дипептидного промежуточного продукта, связанного с Т-доменом второго модуля и переносится в донорный туннель домена конденсации следующего модуля. Затем модуль инициации может загрузить другую аминокислоту субстрата и начать сборку следующего пептида (Shi и др., 2020; Corpuz, Sanlley, Burkart, 2022; Mansour, Gauld, 2024).

Промежуточный пептидилтиоэфир передается от одного модуля к следу-

ющему, при этом в каждом модуле добавляется одна аминокислота. Цикл повторяется, пока пептид не достигнет модуля терминации, содержащего тиюэстеразный домен (ТЕ), который либо катализирует реакцию гидролиза и высвобождает линейный пептид, либо катализирует реакцию циклизации для высвобождения циклического пептида.

Помимо основных доменов, модули НРПС могут содержать дополнительные домены, например оксидазные, редуктазные, эпимеризационные и метилтрансферазные, ответственные за модификацию пептида. Также, в ряде случаев терминальный этап синтеза может осуществляться не тиюэстеразным, а редуктазным доменом, или вообще неферментативным путем.

Один из примеров неферментативного высвобождения пептида обнаружен во время биосинтеза капурамицина. Поскольку ТЕ домен в его синтетазе отсутствовал, первоначально была выдвинута гипотеза, что домен С катализирует внутримолекулярную лактамизацию, как это происходит в СТ-доменах грибковых НРПС, однако, используя синтетические тиюэферы L-лизина в качестве контролей, было продемонстрировано, что лактамизация происходит спонтанно в присутствии воды. Результаты также поднимают интригующий вопрос о том, как другие НРПС, участвующие в биосинтезе НРП, содержащих нециклизованный L-лизин, избегают лактамизации во время процесса удлинения пептидной цепи.

На геномном уровне гены, ответственные за нерибосомальный синтез, организованы в биосинтетические генные кластеры (БГК), которые содержат не только гены кодирующие модули НРПС, но и дополнительные ферменты, транспортеры, а также регуляторные гены.

1.1.2. Домены НРПС

1.1.2.1. А-домен

Домен аденилирования представляет собой каталитический домен, осуществляющий селективное распознавание, активацию и загрузку чрезвычайно разнообразного спектра субстратов, включающего более 500 различных мономеров.

ров, таких, как непротеиногенные аминокислоты в L- и D-конфигурациях, α -, β - и δ -аминокислоты, 2-аминоизомасляную кислоту (Aib), гидроксифенилглицин (Hpg), дигидроксibenзоат (Dhb) и множество других структурно разнообразных соединений нерибосомного синтеза пептидов. При этом некоторые А-домены обладают сниженной специфичностью по отношению к субстратам, например, один и тот-же А-домен может аденилировать как аргинин, так и лизин. Такая вариативность приводит к синтезу сразу нескольких вариантов пептидов, что, вероятно, представляет собой механизм адаптации к изменяющимся условиям окружающей среды, позволяя синтезировать несколько аналогов в зависимости от доступности субстратов .

А-домен характеризуется значительными размерами, составляющими приблизительно 500-550 аминокислотных остатков, и демонстрирует сложную двухсубдоменную архитектуру, включающую более крупный N-концевой субдомен Acore и меньший C-концевой субдомен Asub, которые разделены широкой щелью. Субстрат-связывающий карман локализован в пределах субдомена Acore и образован десятью ключевыми аминокислотными остатками, локализованными между мотивами $\alpha 3$ и $\alpha 6$, которые составляют так называемую последовательность, определяющую субстратную специфичность. Два особенно значимых консервативных остатка формируют критические водородные связи с α -аминогруппой и α -карбоксильной группой субстрата: Asp235 и каталитический остаток Lys517, тогда как остальные восемь остатков взаимодействуют с боковой цепью субстрата, обеспечивая его специфическое распознавание и правильное позиционирование в активном центре .

Современные структурные и биоинформатические модели демонстрируют, что для высокоточного распознавания лиганда требуется анализ более широкого функционального паттерна, состоящего из 34 аминокислотных остатков. Этот ансамбль выстилает полную топологию связывающей полости, охватывая все остатки, расположенные в радиусе примерно 8 Å от аминоксил-аденилата.

Принципиальное ограничение предсказания по aa34-сигнатуре заключается в следующем: в отличие от aa10-кода, по которому для хорошо представленных аминокислот удаётся однозначно идентифицировать субстрат А-домена, 34-позиционный дескриптор не всегда позволяет с высокой уверенностью предсказать конкретную аминокислоту. Для таких А-доменов разработаны методы предсказания физико-химических свойств субстрата: гидрофобности, полярности, заряда, наличия ароматического кольца, карбоксильной и гидроксильной групп - что позволяет сузить пространство возможных мономеров даже при невозможности однозначной идентификации (Crawford и др., 2011; Jian и др., 2023; Miller, Gulick, 2016; Mongia и др., 2023; Nemoto и др., 2025; Röttig и др., 2011; Zhang et al., 2024; McErlean, Overbay, Van Lanen, 2019; Stanišić, Kries, 2019).

Каталитический механизм А-доменов реализуется через строго регулируемую последовательность конформационных изменений, включающую четыре основные стадии, определяемые каталитическим состоянием и позиционированием субдоменов A_{sub} и PCP относительно каталитической платформы, образованной С-доменом и субдоменом A_{core}. На первой стадии А-домен принимает открытую конформацию, позволяющую диффузию АТФ и субстрата в активный центр, где взаимодействие с остатками связывающего кармана обеспечивает правильное позиционирование функциональных групп.

На второй стадии субдомен A_{sub} поворачивается приблизительно на 30°, закрываясь по направлению к субдомену A_{core}, что приводит каталитический лизин, расположенный на гибкой петле, внутрь активного центра и подготавливает систему к осуществлению реакции аденилирования. Заряженный каталитический лизин взаимодействует одновременно с АТФ и кислотным субстратом, сближая их в пространстве и в конечном итоге стимулируя атаку карбоксилата на α-фосфат АТФ, что приводит к образованию ацил-АМФ интермедиата (активированного субстрата) и высвобождению неорганического пирофосфата.

Третья стадия включает дальнейшее движение субдомена A_{sub}, который поворачивается приблизительно на 140° в горизонтальной плоскости, представ-

ляя теперь противоположную поверхность субдомену Acore и перетаскивая РСР-домен над активным центром. Кофактор фосфопантетеин может теперь проникнуть в связывающий карман и присоединить субстрат посредством тиоэтерификации, высвобождая АМФ в процессе и подготавливая субстрат для транспорта к соседним доменам.

На заключительной четвертой стадии субдомен Asub снова поворачивается приблизительно на 180° и отдаляется от субдомена Acore, согласованно с поворотом РСР-домена, что позволяет последнему преодолеть необходимое расстояние для достижения С-домена и обеспечения субстрата для реакции конденсации, в то время как А-домен возвращается в исходную открытую конформацию, готовую к началу нового цикла .

1.1.2.2. РСР-домен

Для обретения каталитической активности апо-РСР-домены должны подвергаться посттрансляционной модификации специфическими ферментами, называемыми 4'-фосфопантетеинилтрансферазами (PPTases), которые катализируют перенос 4'-фосфопантетеинового кофактора, производного кофермента А, на этот выступающий остаток серина, с образованием фосфопантеиновой «руки», способной к значительным конформационным изменениям позволяет РСР-домену взаимодействовать с множественными окружающими доменами (Iacovelli, Bovenberg, Driessen, 2021; Duban, Cociancich, Leclère, 2022; Zhang et al., 2024).

Механизм функционирования РСР-домена включает несколько критически важных этапов: после аденилирования субстрата А-доменом происходит трансэтерификация активированного аминоксил-АМФ на сульфгидрильную группу фосфопантетеина с образованием тиоэфирной связи и высвобождением АМФ. Затем РСР-домен транспортирует ковалентно связанный субстрат к активному центру домена конденсации, где происходит образование пептидной связи между донорным и акцепторным субстратами.

Несмотря на то, что А-домены рассматриваются как основные "привратники" в системах нерибосомного пептидного синтеза, накапливающиеся экспе-

риментальные данные свидетельствуют о том, что РСР-домены также могут проявлять определенную степень специфичности по отношению к своим субстратам, что добавляет дополнительный уровень контроля качества в процесс биосинтеза.

Некоторые исследования демонстрируют, что РСР-домены из модулей, содержащих эпимеризационные домены, значительно отличаются по своей первичной структуре от стандартных РСР-доменов, обычно связанных с конденсационными доменами, что, вероятно, отражает специализированную коммуникацию с модифицирующими доменами.

1.1.2.3. С-домен и его вариации

Домены С чаще всего располагаются в начале каждого модуля НРПС и катализируют образования пептидной связи между двумя субстратами, как правило, соседних модулей. Домен конденсации состоит из двух больших субдоменов, обозначаемых как N-концевой субдомен (CNTD) и С-концевой субдомен (CCTD). Эти структурные элементы образуют характерную V-образную конфигурацию. Каждый из субдоменов содержит большую центральную β -складчатую область, окруженную несколькими α -спиральными элементами. Между двумя лобами располагается центральный туннель, в котором находится каталитический активный сайт, содержащий консервативный мотив NHxxxDG , ответственный за образование пептидной связи между субстратами с помощью нуклеофильной атаки α -аминогруппы акцепторного субстрата на карбонильный углерод донорного тиоэфира. Согласно современным представлениям, консервативная гистидиновая группа в мотиве NHxxxDG осуществляет правильное позиционирование субстрата через образование водородных связей с α -аминогруппой акцепторного субстрата, что способствует нуклеофильной атаке на донорный карбонил и замена этого остатка на аланин полностью элиминирует каталитическую способность домена.

Интересной структурной адаптацией является модификация канонического мотива NHxxxDG в некоторых С-доменах, например, вариант NHxxxDE , кото-

рый специфически адаптирован для работы с малыми акцепторными субстратами, такими, как глицин. Глутаминовая кислота в этом варианте мотива стабилизирует и ориентирует аминокгруппу акцепторного субстрата для обеспечения эффективной нуклеофильной атаки на донорный тиоэфир.

Домен конденсации обладает двумя различными сайтами связывания субстратов, расположенными на противоположных сторонах центрального туннеля. Акцепторный сайт предназначен для связывания аминокцил-тиоэфира, присоединенного к соответствующему РСР-домену данного модуля, в то время как донорный сайт взаимодействует с аминокцил-тиоэфирным субстратом, переданным от предыдущего модуля. Несмотря на отсутствие четко определенных карманов субстратной специфичности, С-домены тоже участвуют в процессе корректировки синтеза, предотвращающем выполнение реакции конденсации при несоответствии субстрата, связанного с РСР-доменом, при этом акцепторный сайт демонстрирует значительно более высокую субстратную специфичность (распознавание D/L-аминокислот, N-метильных остатков, оксазолин/тиазолин-акцепторов), тогда как донорный сайт более толерантен к вариациям субстрата (Iacovelli, Bovenberg, Driessen, 2021; Izoré et al., 2021; Chu Yuan Kee et al., 2022; Folger et al., 2024; Corpuz, Sanlley, Burkart, 2022; Hur, Vickery, Burkart, 2012; Patteson и др., 2022).

Следует отметить большое разнообразие вариантов доменов конденсации, выполняющих разнообразные функции в нерибосомальных пептид-синтетазах. LCL- и DCL-домены представляют собой две самые распространенные подгруппы С-доменов. LCL катализирует образование пептидной связи между двумя L-конфигурированными аминокислотами. DCL-домены используют D-донор и L-акцептор. Исследования мутаций в С-доменах GrsB1 и TysB1 продемонстрировали, что замена глутамата на лейцин в позиции 386 (E386L) приводит к смене субстратной специфичности фермента с D-аминокислот на L-аминокислоты. Фактически это трансформирует DCL-домен в LCL-домен, в свою

очередь, обратная мутация в С-домене SrfAC обеспечивает обратный переход LCL-домена в DCL-домен .

Некоторые С-домены (С/Е-домены) обладают бифункциональной активностью, выполняя как конденсацию, так и эпимеризацию. Их архитектура представляет собой гибрид из Е-домена встроенного между N- и С-субдоменами классического С-домена, с двумя независимыми активными центрами. Эти домены особенно характерны для биосинтеза липопептидов у бактерий родов *Pseudomonas*, *Burkholderia* и *Xanthomonas*.

Стартовые С-домены (Cstart) в иницирующих модулях НРПС обеспечивает присоединение нестандартных или ацильных стартерных групп (ацил-СоА, β -гидроксикислот и др.) к первым аминокислотам пептидной цепи. Эти домены отличаются от стандартных С-доменов расширенным донорным сайтом, способным принимать ацильные доноры .

Во многих модулях НРПС С-домен заменяется родственным ему, доменом гетероциклизации (Су-домен), который сначала катализирует образование амидной связи между донорным и акцепторным субстратами, а затем внутримолекулярную циклодегидратацию с образованием оксазолинового или тиазолинового гетероцикла из боковых цепей Ser/Thr или Cys соответственно. Кристаллографические и мутационные исследования показали, что в циклодегидратации участвует ранее недооценённая каталитическая диада Asp–Thr.

Х-домены являются эволюционно родственными С-доменам, но каталитически неактивными из-за мутаций в консервативном мотиве NHxxxDG. Они служат платформами для рекрутирования монооксигеназ цитохрома P450, ответственных за циклизацию связанного с Т-доменом растущего пептида в синтезе гликопептидов, таких как ванкомицин и теикопланин .

В исследовании биосинтеза нокардицина А был найден новый специализированный С-домен, который осуществляет интеграцию β -лактамного кольца непосредственно в процессе пептидной элонгации. Этот уникальный С-домен способен преобразовать серилтетрапептид в пентапептидный продукт, содержа-

щий интегрированное β -лактамное кольцо, через сложный механизм, включающий дегидратацию серинового остатка с последующей внутримолекулярной циклизацией (Gaudelli, Long, Townsend, 2015; Long, Townsend, 2021, 2018; Peng и др., 2024; Zhong и др., 2021; Haslinger et al., 2015; Wang et al., 2022; Duerfahrt и др., 2004).

1.1.2.4. Тиоэстеразные ТЕ и ТЕII домены

Оба типа доменов относятся к классу сериновых гидролаз, что определяется наличием каталитической триады (Ser-His-Asp/Glu), делающей возможным как гидролиз, так и, при необходимости, лактамизацию или макроциклизацию продукта. ТЕ-домены, как правило, располагаются внутри модуля терминации и отвечают за гидролитическое высвобождение или макроциклизацию пептидной цепи, а ТЕII представляют из себя отдельные белки, выполняющие функцию удаления ошибочно загруженных или аномальных остатков с РСР-доменов, например, ацил-групп, полипептидов неправильной длины для поддержания работоспособности модулей. В отдельных случаях обнаружены тандемы ТЕ: первый катализирует макроциклизацию, второй — высвобождение продукта (Kohli, Takagi, Walsh, 2002; Kotowska, Pawlik, 2014; Patel и др., 2023).

В процессе синтеза растущий пептид, ковалентно связанный с фосфопантеиновой "рукой" РСР-домена, переносится на каталитический серин ТЕ-домена с образованием временного ковалентного ацил-ТЕ промежуточного комплекса (О-ацил-сериновая связь). Этот комплекс затем подвергается нуклеофильной атаке. Атака молекулой воды приводит к гидролизу и высвобождению линейного пептида. Если же атака происходит внутримолекулярным нуклеофилом, таким как α -аминогруппа N-конца, аминогруппа боковой цепи лизина или гидроксигруппа серина/треонина, то это приводит к циклизации с образованием лактамного или лактонного кольца соответственно.

1.1.2.5. R-домены

Редуктазные домены представляют собой альтернативный механизм высвобождения пептидного продукта в нерибосомных пептид-синтетазах, заменяя тио-

эстеразные (TE) домены. В отличие от TE-доменов, которые катализируют гидролиз тиоэфирной связи, R-домены осуществляют NAD(P)H-зависимое восстановительное высвобождение, что приводит к образованию продуктов со свободным С-концом, лишённым отрицательных зарядов. Это свойство открывает больше возможностей для последующих модификаций, таких как гликозилирование. Кроме того, образующиеся альдегидные фрагменты могут непосредственно обуславливать биологическую активность конечного продукта, как в случае мощного ингибитора протеасом — феллутамида В. Помимо образования линейных спиртов или альдегидов, некоторые R-домены способны катализировать макроциклизацию с образованием структур по типу пиразинов, как это происходит при биосинтезе ауреузимины и ностоциклопептидов (Patel и др., 2023; Tietze и др., 2020; Süßmuth, Mainz, 2017).

Трёхмерная структура R-домена организована в два функциональных субдомена: N-концевой кофактор-связывающий субдомен с консервативной складкой Россмана для связывания NAD(P)H и С-концевой субстрат-связывающий субдомен, формирующий гидрофобный активный центр для взаимодействия с пептидил-тиоэфирным субстратом.

Каталитический механизм R-доменов включает $2e^-$ или $4e^-$ восстановление. Первоначальный акт $2e^-$ восстановления представляет собой шаг высвобождения, в результате которого образуется альдегидный продукт, он может быть конечным, как в случае синтеза предшественника сафрамицина А, однако в большинстве случаев он подвергается второму циклу $4e^-$ восстановления с потреблением ещё одной молекулы NAD(P)H, что приводит к образованию первичного спирта.

Второй восстановительный этап служит механизмом детоксикации, предотвращая накопление в клетке реактивных альдегидов. В некоторых системах НРПС эту вторую стадию катализирует отдельный фермент альдо/кеторедуктаза .

1.1.2.6. Домены модификации субстрата

Дополнительные домены модификации в бактериальных НРПС включа-

ют функциональные модули, которые изменяют конфигурацию, окислительно-восстановительное состояние и химические группы мономеров на РСР, тем самым расширяя структурное и функциональное разнообразие продуктов. К ним относятся домены эпимеризации (E), метилирования (M), формилирования (F), окисления (Ox) и кеторедуктазы (KR).

Домен эпимеризации в нерибосомальных пептид-синтетазах катализирует превращение L-аминокислотного остатка, в соответствующий D-энантиомер, что принципиально изменяет конформационные предпочтения, устойчивость к протеолизу и взаимодействия с мишенями у конечного пептида. Эволюционно E-домены являются потомками доменов конденсации и потому обладают схожим строением, включая консервативный мотив NhxxxDG, однако, акцепторный сайт у E-доменов закрыт дополнительными остатками, которые блокируют вход в реакционный центр и препятствуют связыванию вышестоящего T-домена, поэтому, в отличие от домена C, которому нужны два субстрата для образования пептидной связи, домен E использует только один субстрат для эпимеризации (Iacovelli, Bovenberg, Driessen, 2021; Manavalan и др., 2010).

Домены метилирования представляют собой компактные каталитические модули (около 45 кДа), ответственные за введение метильных групп в структуру нерибосомальных пептидов на этапе их сборки. Метилирование пептидного остова значительно повышает гидрофобность пептида, повышая его устойчивость к протеолизу и проницаемость через клеточные мембраны. В НРПС они могут встречаться как отдельными доменами, так и быть интегрированы в домен A, между двумя основными мотивами A8 и A9. Наиболее распространенным типом модификации опосредованное M доменом является N-метилирование пептидного остова. Помимо модификации остова, M-домены могут катализировать и другие, более редкие типы метилирования: N-метилирование боковой цепи, например, диаминопропионовой кислоты в биосинтезе пениламицина; O-метилирование, как в случае генерации остатка O-Me-Ser в биосинтезе куцнеридов; S-метилирование в биосинтезе тиокоралина. C-метилирование в биосинтезе иерсиниабактина, где домен M, встроенный между доменами Cy и T, катализи-

рует метилирование по С α -атому тиазолинового кольца (Süssmuth, Mainz, 2017).

Домены формилирования преимущественно функционируют на стартовом этапе синтеза нерибосомального пептида и располагаются в модуле инициации в порядке F–A–PCP. Каталитический цикл начинается с активации аминокислоты А-доменом и переноса её на тиол PCP-домена. Затем PCP доставляет субстрат к каталитическому центру F-домена, который использует формилтетрагидрофолат (fTHF) в качестве кофактора для переноса формильной группы на α -аминогруппу субстрата, эта модификация устраняет положительный заряд на N-конце конечного продукта, что может быть чрезвычайно важно для конечных свойств продукта. Прекрасным примером служит линейный грамицидин, аминокислотная последовательность которого содержит исключительно гидрофобные остатки, N-конец блокирован формилированием валина а С-конец блокирован этаноламином. Эти три особенности объясняют высокую нерастворимость в воде, но очень хорошую растворимость в различных органических растворителях, поскольку грамицидин не способен принимать суммарный заряд или образовывать цвиттер-ион при любом pH. Обычно две молекулы грамицидина димеризуются, образуя либо двойную спираль (форма «поры»), либо спиральный димер (форма «канала»), внедряются в мембрану и специфично транспортируют одновалентные катионы через билипидный слой, тем самым снижая трансмембранный ионный потенциал, при этом димеризация осуществляется «голова к голове», то есть N-концами, что было бы невозможно без формилирования (Kessler и др., 2004; Schoenafinger и др., 2006).

Домены Ох, зависимые от флавинмононуклеотида (FMN), отвечают за окисление гетероциклических промежуточных продуктов. Подобно доменам М, они обычно встроены в С-концевой субдомен А домена между консервативными мотивами А8 и А9. Ох-домены катализируют стереоселективное окисление пятичленных тиазолиновых или оксазолиновых колец, образованных доменами циклизации (Су), до ароматических тиазолов или оксазолов соответственно.

Эти домены очень консервативны и имеют приблизительно 40% идентичности последовательностей друг с другом (Schneider, Shen, Walsh, 2003; Süssmuth, Mainz, 2017).

Кетоацилредуктазные домены (KR), широко распространённые в поликетид-синтазах, в редких случаях обнаруживаются и в НРПС. KR-домены в НРПС катализируют NADPH-зависимое стереоспецифическое восстановление α -кетогруппы активированного субстрата до гидроксильной, превращая α -кетокислоту в α -гидроксикислоту. Эта реакция является ключевой для биосинтеза депсипептидов — соединений, содержащих в своей цепи как амидные, так и сложноэфирные связи. Классическими примерами таких продуктов являются цереулид и валиномицин. KR-домены в НРПС могут располагаться между A и PCP доменами, а могут находиться внутри A-домена, как домены M и Oх. В этом случае A-домен сначала активирует α -кетокислоту, а затем встроенный KR-домен восстанавливает её до α -гидроксикислоты до того, как субстрат поступит на следующий этап сборки (Ehling-Schulz и др., 2006; Alonzo, Schmeing, 2020; Süssmuth, Mainz, 2017).

1.1.2.7. Некаталитические домены

К группе некаталитических доменов относятся домены, которые не осуществляют прямого химического превращения субстратов, но обеспечивают важные функции, такие как обеспечение межмодульной коммуникации и регуляции каталитической активности. К основным некаталитическим доменам относятся MbtH-подобные белки (MLP) и докинг-домены (DD), также называемые коммуникационными доменами (COM).

MbtH-подобные белки (MLP) представляют собой небольшие (~70-80 аминокислот, молекулярная масса около 10 кДа) высококонсервативные белки, которые обнаруживаются в сотнях бактериальных НРПС, но полностью отсутствуют в геномах грибов и других эукариотических организмов (Felnagle и др., 2010). Кристаллографические исследования выявили компактную глобулярную структуру, состоящую из четырех β -нитей, образующих β -шпильку, и двух α -

спиралей, при этом два строго консервативных триптофановых остатка (Trp-25 и Trp-35) формируют важный гидрофобный кластер на поверхности белка .

Биохимические исследования выявили, что MLP помогают в сборке и поддержании правильной структуры А доменов и существенно стимулируют их каталитическую активность, присутствие MLP может повышать эффективность аденилирования в 5-10 раз, при этом эффект варьирует в зависимости от комбинации MLP и А домена. Особенно интересным является феномен функционального кросс-взаимодействия MLP, при котором MLP из одного биосинтетического пути может стимулировать активность НРПС из совершенно другой системы. Более того, геномные исследования выявили существование 50 "орфанных" MLP из различных бактериальных таксонов, которые кодируются генами, не локализованными вблизи каких-либо НРПС-кодирующих биосинтетических генных кластеров, что предполагает их участие в более широких регуляторных сетях. Наиболее примечательным открытием стала идентификация MLP MXAN_3118 из *Muxococcus xanthus* DK1622, который естественно взаимодействует с НРПС по крайней мере семи различных биосинтетических путей, включая системы биосинтеза ДК-ксантена, миксохроминов, миксовиресцинов и других вторичных метаболитов (Zhang и др., 2010; Esquilín-Lebrón и др., 2018; Herbst и др., 2013).

Докинг-домены обеспечивают специфическую и направленную сборку многобелковых НРПС через формирование линейных массивов из белковых субъединиц. Структурная классификация докинг-доменов включает N-терминальные докинг-домены (N-DD), обладающий β -слоем с характерными петлевыми регионами, которые располагаются перед конденсационными доменами и функционируют как акцепторные элементы, и C-терминальные докинг-домены (C-DD), обладающие α -спиральной архитектурой, локализованные после РСР доменов и выполняющие донорные функции. Взаимодействие DD опосредовано связыванием « β -руки» акцепторной субъединицы с C-концевой α -спиралью донорной субъединицы . Специфичность коммуникационных доменов опреде-

ляется комбинацией электростатических и гидрофобных взаимодействий между комплементарными поверхностями донорных и акцепторных доменов. Замена специфических аминокислотных остатков приводит к изменению партнёрской специфичности и, как следствие, к изменению спектра продуктов (Fage и др., 2021; Hahn, Stachelhaus, 2004).

1.1.3. Типы НРПС

Нерибосомальные пептид-синтетазы можно разделить по организации каталитических доменов, на 2 группы: модульные многодоменные ферменты (тип I) и автономные (тип II), то есть синтетазы, каталитические домены которых организованы не по модульному принципу, а представлены отдельными белками. Другая классификация НРПС основана на алгоритме сборки пептидной цепи и включает три основных типа: линейные (тип A), итеративные (тип B) и нелинейные (тип C). Отдельно выделяются гибридные ПКС-НРПС системы, включающие в себя не только модули нерибосомального синтеза, но и модули синтеза поликетидов.

Считается, что модульные НРПС составляют подавляющее большинство всех нерибосомальных пептидных синтетаз, и поэтому именно их структура была описана выше как наиболее типичная. Согласно имеющимся данным, модульные системы встречаются значительно чаще автономных систем типа II. Так, исследование Ванга с соавторами показало, что только 18% кластеров генов НРПС прокариот принадлежали к типу II (Iacovelli, Bovenberg, Driessen, 2021; Jaremko и др., 2020; Wang и др., 2014; Duban, Cociancich, Leclère, 2022; Süssmuth, Mainz, 2017).

1.1.3.1. Линейные НРПС

Линейные НРПС представляют собой классическую модель нерибосомального синтеза, где каждый модуль используется однократно в строго последовательном порядке. В линейных системах пептидная цепь удлиняется согласно принципу колинеарности: порядок модулей в составе НРПС непосредственно соответствует последовательности аминокислот в синтезируемом продукте. Линей-

ные НРПС могут состоять из одного белка, содержащего все модули и домены, необходимые для выполнения полного процесса синтеза, либо, из нескольких белков, каждый из которых предоставляет активированные субстраты или промежуточные продукты для поэтапной сборки конечного пептидного продукта (Sundlov и др., 2012; Watzel et al., 2021). Примером НРПС, состоящей из одного белка, является упомянутая ранее синтетаза колоссина А. Этот линейный пентадекапептид целиком синтезируется одним гигантским ферментом с молекулярной массой 1,8 МДа, кодируемым одним геном (Bode и др., 2015; Süssmuth, Mainz, 2017). Однако, значительно чаще, даже меньшие пептиды синтезируются несколькими белками. Например, сурфактин синтезируется опероном *srfA*, состоящим из трех генов: *srfAA*, *srfAB* и *srfAC*. Каждый ген кодирует отдельную субъединицу НРПС с разным числом модулей, так субъединицы *SrfAA* и *SrfAB* содержат по три модуля, а субъединица *SrfAC* один. Эти субъединицы совместно работают над синтезом циклического липогептапептида сурфактина .

Важно различать линейный режим биосинтеза и конечную структуру пептида, которая может быть либо линейной, как у колоссина А, либо циклической, как у сурфактина, хотя оба они синтезированы НРПС с линейным типом синтеза.

1.1.3.2. Итеративные НРПС

Итеративные НРПС типа В характеризуются многократным использованием одних и тех же модулей для синтеза повторяющихся структурных элементов в пептидной цепи. Накопление и хранение промежуточных продуктов в итеративных НРПС осуществляется либо Т-доменом (синтетазы конгоцидина и энниатина), способном удерживать растущий пептид посредством фосфопантетеиновой руки, либо непосредственно терминальным доменом (синтетаза грамицидина С), промежуточный продукт находится на хранении до тех пор, пока пептид не достигнет длины, подходящей для высвобождения, инициируемого Те-доменом. Классическим примером итеративного биосинтеза служит грамицидин С, циклический декапептид, синтезируемый синтетазой содержащей пять

модулей. Повторное использование домена-ТЕ позволяет дважды пройти полный цикл элонгации, создавая симметричный декапептид. Итеративный способ синтеза позволяет избегать создания чрезмерно крупных и сложных мегасинтетаз, подверженных неправильному сворачиванию и протеолизу. Так, синтетаза грамицидина С имеет массу 280 кДа, тогда как линейно функционирующая система была бы в два раза больше. Таким образом, итеративность обеспечивает экономию ресурсов и сохранение точности сборки сложных макроциклических структур (Hoyer, Mahlert, Marahiel, 2007; Kleinkauf, Gevers, Lipmann, 1969; Xu et al., 2023).

1.1.3.3. Нелинейные НРПС

Нелинейные НРПС представляют собой наиболее сложную и разнообразную группу нерибосомальных пептидных синтетаз. Отличительной чертой этих систем является нестандартное расположение хотя бы одного основного домена, например отсутствие А-домена в модуле, или наличие отдельных доменов, не организованных в модули. В отличие от линейных и итеративных систем, нелинейные НРПС характеризуются многократным использованием отдельных доменов, а не целых модулей, обеспечивая тем самым гораздо более сложные архитектуры и разнообразие конечных пептидов. Наиболее часто в таком режиме функционируют домены аденилирования, которые активируют и поставляют аминоксил-АМФ не только для своего родственного Т-домена, но и для Т-доменов других модулей в составе того же или других белков. Такое взаимодействие обеспечивается различными как междоменными взаимодействиями, так и вспомогательными факторами, такими как тиюэстеразы II типа (ТЕII) и специализированные транспортные белки (Fiedler и др., 2024).

Классическим примером нелинейного биосинтеза служит система синтеза маннопептимицина. В процессе сборки этого липогликопептидного антибиотика один А-домен аденилирует β -гидроксиэндурацидидин не только для собственного Т-домена, но и для Т-домена следующего модуля, в котором нет собственного А-домена. Такая организация доменов с топологией С-А-Т-С-Т приводит к

дупликации остатка эндурацидидина в растущей пептидной цепи (Magarvey и др., 2006).

В недавно описанной синтетазе фимсбактина А у *Acinetobacter baumannii* нелинейность распространяется на многофункциональный Су-домен, который катализирует разветвление пептидных цепей и вторичное конденсационное событие между двумя дипептидными промежуточными продуктами, образуя тетрапептидный тиоэфир с ветвлением, тем самым обеспечивая специфическую разветвленную структуру фимсбактина А (Yang и др., 2025).

Нелинейные НРПС демонстрируют замечательную гибкость и универсальность в структуре и механизме, что делает их перспективными объектами для биоинженерии и создания новых биологически активных соединений.

1.1.3.4. Поликетидсинтетазы (ПКС) и гибридные НРПС/ПКС

Значительная доля природных биологически активных соединений синтезируется не мономодульными НРПС или ПКС, а гибридными мегасинтетазами, в которых модули нерибосомальных пептидных синтетаз и поликетидсинтаз (ПКС) функционируют в едином сборочном конвейере. По данным масштабного геномного исследования 2699 геномов, из 3339 генных кластеров, кодирующих НРПС и ПКС, треть (1147) принадлежала гибридному типу (Wang и др., 2014). Структурное и каталитическое сходство между НРПС и ПКС создаёт принципиальную возможность комбинирования их модулей, что в ходе эволюции реализовалось в многочисленных природных гибридных системах (Walsh, O'Brien, Khosla, 2013; Blin и др., 2021).

Гибридные НРПС-ПКС-системы подразделяют на два класса. К первому относятся соединения, в биосинтезе которых НРПС- и ПКС-модули не вступают в прямое функциональное взаимодействие: в таких случаях оба типа ферментов действуют независимо, а окончательная структура продукта формируется посредством посттрансляционных модификаций или конъюгации (Cane, Walsh, Khosla, 1998; Walsh, O'Brien, Khosla, 2013). Ко второму классу принадлежат системы, в которых НРПС- и ПКС-модули непосредственно взаимодействуют,

обеспечивая последовательное наращивание пептидно-поликетидной цепи: именно эти системы наиболее интересны с биотехнологической и комбинаторной точки зрения (Konz, Marahiel, 1999). Примерами продуктов таких гибридных систем служат эпотилон, блеомицин, эпоксомидин и ряд других соединений с выраженной биологической активностью (Felnagle и др., 2008; Williams, 2013).

ПКС относятся к мегасинтезам, родственным синтазам жирных кислот (СЖК). Подобно НРПС, модульные ПКС I типа организованы как «сборочные конвейеры», в которых каждый модуль отвечает за одну реакцию удлинения поликетидной цепи (Staunton, Weissman, 2001; Williams, 2013). Минимальный каталитически функциональный модуль ПКС состоит из трёх обязательных доменов: кетоацилсинтазы (KS), ацилтрансферазы (AT) и белка-переносчика ацила (ACP).

Домен KS (кетоацилсинтаза) — каталитическое ядро модуля ПКС, осуществляющее реакцию Клайзена: конденсацию тиоэфирного интермедиата с новым удлиняющим субстратом. В гибридных НРПС-ПКС системах домены KS имеют уникальные особенности, обеспечивающие эффективный приём субстратов от соседних НРПС-модулей (Walsh, O'Brien, Khosla, 2013).

Домен AT (ацилтрансфераза) — определяет субстратную специфичность модуля по отношению к удлиняющим единицам (малонил-КоА, метилмалонил-КоА, этилмалонил-КоА и др.). Мутации в сайте связывания субстрата AT позволяют изменять состав поликетидов (Williams, 2013).

Домен ACP (белок-переносчик ацила) — аналог домена PCP (T) НРПС. Несёт ковалентно связанный простетический кофактор 4'-фосфопантетеин, к которому через тиоэфирную связь присоединены как субстраты, так и нарастающая поликетидная цепь. Посттрансляционная активация ACP осуществляется 4'-фосфопантетеинилтрансферазами (ФПТ) с широкой специфичностью к белкам-носителям, что принципиально важно для функционирования гибридных систем (Konz, Marahiel, 1999).

Помимо трёх обязательных доменов, модули ПКС могут содержать факкультативные редуکتивные домены, определяющие степень восстановления β -кетогруппы в каждом цикле удлинения цепи:

Домен KR (кеторедуктаза) — восстанавливает β -кетогруппу до гидроксильной, используя НАДФН в качестве кофактора. Стереоспецифичность восстановления (А- или В-тип KR) определяет конфигурацию формируемого гидроксильного центра (Williams, 2013).

Домен DH (дегидратаза) — катализирует дегидратацию β -гидроксилтиоэфира с образованием двойной связи.

Домен ER (еноилредуктаза) — восстанавливает двойную связь до насыщенной метиленовой группы. Совместное присутствие KR, DH и ER обеспечивает полное восстановление углеродного фрагмента — аналогично тому, как это происходит при синтезе жирных кислот.

Таким образом, каждый модуль ПКС реализует различные степени модификации поликетидной цепи в зависимости от набора присутствующих доменов. Разнообразие субстратов АТ и комбинации редуکتивных доменов в разных модулях определяют богатое структурное разнообразие поликетидов и их гибридных производных (Staunton, Weissman, 2001; Walsh, O'Brien, Khosla, 2013).

Ключевую роль в поддержании межмодульной коммуникации в гибридных НРПС-ПКС системах играют специфические линкерные последовательности, обеспечивающие передачу нарастающего интермедиата между соседними модулями разных типов, а также фосфопантетеинилтрансферазы с широкой субстратной специфичностью, необходимые для активации как АСР- (ПКС), так и РСР-доменов (НРПС) (Konz, Marahiel, 1999; Blin и др., 2021).

1.1.4. Регуляция нерибосомального синтеза

Регуляция НРПС происходит преимущественно на уровне транскрипции генов, кодирующих эти синтетазы, и включает сложные регуляторные сети, охватывающие до 12% всего генома у некоторых микроорганизмов, таких как *Streptomyces*.

У спорообразующих бактерий, таких как бактерии родов *Bacillus* и *Paenibacillus*, синтез НРПС тесно связан с глобальными регуляторами спорообразования, биопленкообразования, состояния компетенции и других жизненно важных процессов в бактериальной клетке.

Так, например, белок Spo0A является регулятором спорообразования и одновременно ключевым активатором многих генов вторичного метаболизма, включая НРПС. Этот белок напрямую связывается с промоторными областями генов НРПС и активирует их транскрипцию в стационарной фазе роста. У *Paenibacillus polymyxa* было показано, что Spo0A как напрямую, так и через влияние на другие гены-регуляторы активирует экспрессию генов полимиксинового кластера (Cui и др., 2023; Liu и др., 2023; Yang и др., 2025).

Ген AbrB функционирует как регулятор перехода к стационарной фазе роста, контролирующей множество генов, включая многие НРПС. AbrB также связывается с промоторными областями, но при этом блокирует транскрипцию. Его активность максимальна в экспоненциальной фазе роста, а при переходе в стационарную фазу уровень AbrB снижается, что приводит к активации синтеза нерибосомальных пептидов. Сам abrB при этом контролируется Spo0A, который подавляет abrB, связываясь с консенсусными Spo0A-боксами в его регуляторном регионе. Такая регуляторная система обеспечивает производство нерибосомальных пептидов в наиболее подходящее время жизненного цикла бактерии — стационарной фазе, когда уровни AbrB снижаются, а активность Spo0A возрастает. Нарушения в гене abrB приводит к сверхэкспрессии генов нерибосомального синтеза в экспоненциальной фазе (Banse и др., 2008; Li и др., 2021; Wang et al., 2024).

ComP-ComA система регулирует синтез сурфактина у бацилл в зависимости от плотности популяции. ComA фосфорилируется ComP в ответ на достижение необходимой концентрации сигнального пептида ComX и связывается с ComA-боксами в промоторе srfA оперона, иницируя транскрипцию (Roggiani, Dubnau, 1993; Piazza, Tortosa, Dubnau, 1999; Zhi, Wu, Xu, 2017).

DegS-DegU система играет центральную роль в регуляции многих НРПС, особенно у грамположительных бактерий. Белок DegU является регулятором ответа, который активируется фосфорилированием сенсорной киназой DegS. Различные уровни фосфорилирования DegU определяют его специфичность: низкие уровни фосфорилирования активируют гены подвижности и компетентности, а высокие уровни инициируют транскрипцию генов внеклеточных протеаз и НРПС. При этом фосфорилирование самого DegU регулируется несколькими различными факторами транскрипции: репрессор SinR ингибирует *degU*, связываясь с его промоторной областью и предотвращая инициацию транскрипции, антирепрессорный белок SlrR действует в качестве деререпрессора, способного связываться с SinR и нейтрализовать его репрессорную активность (Ogura, Yoshikawa, Chibazakura, 2014). При этом белок DegQ функционирует как активатор фосфорилирования DegU, стимулируя активность киназы DegS и способствуя накоплению фосфорилированной формы DegU в клетке (Miras, Dubnau, 2016; Lilge et al., 2021).

При избытке азота глутаминсинтетаза GlnA находится в ингибированной форме и связываясь с TnrA подавляя его активность, при дефиците азота GlnA переходит в активную форму, активируя TnrA который, в свою очередь, связывает с промоторной областью *degU* и индуцирует его транскрипцию (Yasumura, Abe, Tanaka, 2008).

CsrA, глобальный регулятор транскрипции, который действует как репрессор или реже как активатор большого количества генов в зависимости от наличия предпочтительных источников углерода, в присутствии глюкозы он связывается с CRE-сайтом в промоторной области гена *degU* и активирует его транскрипцию (Ishii, Tanaka, Ogura, 2013).

Также, важную роль играют ион-зависимые механизмы, которые регулируют транскрипцию нерибосомально синтезируемых сидерофоров. Ключевыми для регуляции являются железозависимые репрессоры Fur/DtxR/IdeR, медь-чувствительные репрессоры CsoR/RicR а также цинк-регуляторы Zur/SmtB, опо-

средованно влияющие на металлотропные метаболиты (Schmitt и др., 1995; Chillappagari и др., 2010; Kallifidas и др., 2010; Troxell, Hassan, 2013; Chen et al., 2019; Lyu и др., 2022).

1.2. Аннотирование НРПС

Стремительное развитие технологий полногеномного секвенирования открыло новую эру в изучении вторичного метаболизма микроорганизмов. Сегодня стала возможной идентификация кластеров биосинтетических генов и предсказание структур синтезируемых ими соединений непосредственно по геномным данным, минуя трудоемкий химический скрининг. Благодаря биоинформатическим инструментам геномного майнинга исследователи могут систематически картировать биосинтетический потенциал микробных геномов и расставлять приоритеты для экспериментальной верификации (Blin и др., 2021; Medema и др., 2011). Центральное место среди таких инструментов занимает antiSMASH (Antibiotics and Secondary Metabolite Analysis Shell) - автономный программный пакет и популярный веб-сервер. С момента своего первого релиза в 2011 году программа прошла через серию крупных обновлений, достигнув к 2025 году версии 8.0, что позволило значительно расширить спектр детектируемых типов БГК и глубину их аннотации (Blin и др., 2021, 2025). Алгоритм анализирует нуклеотидные последовательности геномов (в форматах GenBank, FASTA, EMBL) или хромосомные контиги, генерируя подробный отчет. В нем представлена карта всех идентифицированных кластеров с цветовой кодировкой генов по их функциям (биосинтетические, транспортные, регуляторные), а также информация о гомологичных охарактеризованных кластерах из базы данных MIBiG (Blin и др., 2021).

Одной из ключевых задач при анализе таких кластеров, в частности нерибосомных пептидных синтетаз, является определение их субстратной специфичности. Она во многом зависит от ограниченного набора аминокислотных остатков, формирующих связывающий карман аденирилирующего (A) домена. Опираясь на данные рентгеноструктурного анализа домена активации фенилалани-

на GrsA из *Bacillus brevis*, Стахельхаус и соавторы выделили десять позиций активного центра, определяющих эту специфичность: Asp235, Ala236, Trp239, Thr278, Ile299, Ala301, Thr322, Ile330, Cys331 и Lys517 (Stachelhaus, Mootz, Marahiel, 1999). Этот компактный дескриптор, получивший название «код Стахельхауса» (или «aa10-код»), позволяет однозначно предсказывать включаемую аминокислоту для многих изученных А-доменов (Stachelhaus, Mootz, Marahiel, 1999; Challis, Ravel, Townsend, 2000). Тем не менее, чувствительность aa10-кода падает при анализе доменов с нестандартными или редкими субстратами. Для решения этой проблемы был разработан расширенный набор из 34 позиций (aa34-код), охватывающий дополнительные остатки субдоменного кармана. Хотя увеличение числа переменных повышает качество машинного обучения, оно лишает предсказание однозначности. Поэтому современные алгоритмы (например, SANDPUMA), применяющие aa34-признаки, прогнозируют не конкретную аминокислоту, а распределение вероятностей по набору возможных субстратов или их физико-химические свойства — гидрофобность, заряд, размер боковой цепи и способность к циклизации (Kunyavskaya et al., 2021; Medema et al., 2015). Таким образом, предсказание по aa34 позволяет отнести неизвестный А-домен к определенной функциональной категории, что критически важно для дерепликации и приоритизации новых кластеров даже без точного установления мономера.

Для практической реализации таких сложных предсказаний был создан специализированный инструмент NERPA (Nonribosomal Peptide Recognizer). В то время как antiSMASH фокусируется на первичной аннотации БГК, NERPA решает задачу верификации гипотетического продукта. Программа сопоставляет аминокислотный состав, предсказанный по А-доменам кластера, со структурами известных нерибосомных пептидов (НПП) из баз данных, таких как NP Atlas или MIBiG. В актуальной версии NERPA 2 используется вероятностная модель на основе скрытых марковских моделей (НММ), что позволяет эффективно работать с неколлинейными сборочными конвейерами, вставками и делециями модулей, а также различными циклическими конформациями пептидов

(Olkhovskii et al., 2026). Оценка каждого А-домена базируется на матрице признаков, включающей aa10- и aa34-сигнатуры, которую алгоритм случайного леса (Random Forest) использует для построения апостериорного распределения вероятностей по 42 возможным мономерам. Затем алгоритм Витерби находит оптимальный путь в НММ, вычисляя итоговую оценку как отношение правдоподобия. На практике NERPA органично интегрируется в пайплайны с antiSMASH: первый алгоритм идентифицирует кластеры и извлекает сигнатуры А-доменов, а второй применяет их для выявления новых структур (Blin и др., 2021; Kunyavskaya et al., 2021) В эту же биоинформатическую экосистему интегрируются и другие инструменты (NRPSPredictor2, PRISM), а также специализированные базы данных (antiSMASH-DB), обеспечивающие всесторонний сравнительный анализ (Röttig и др., 2011; Medema et al., 2015).

Параллельно с совершенствованием методов структурной аннотации развивались подходы к машинному предсказанию биологической активности синтезируемых соединений. Так, масштабный анализ более 1000 НПП из базы Norine, проведенный Caboche и соавторами (2010), выявил корреляцию между наличием определенных непротеиногенных мономеров и конкретными классами активности — антибиотической, антифунгальной, иммуномодулирующей или сидерофорной. Предложенный ими статистический метод вычислял Р-значение для каждого класса активности на основе мономерного состава. Однако этот подход имеет существенное ограничение: он требует точного знания химической структуры продукта, что делает его малоприменимым для анализа неохарактеризованных кластеров. Принципиально иную, более перспективную стратегию предложили Walker и Clardy (2021). В качестве входных данных они использовали непосредственно нуклеотидные последовательности самих БГК, обучив классификаторы (логистическую регрессию, метод опорных векторов и случайный лес) на выборке охарактеризованных кластеров с известной активностью. Ключевым методологическим решением стало использование суб-РFAM-доменов (выделенных из сетей сиквенс-сходства) в качестве предикторов. Подобный подход открывает широкие возможности для прямого прогнози-

рования биологических свойств метаболитов на основе сырых геномных данных, однако показал относительно низкую точность предсказаний для противогрибковых свойств.

1.3. Основные подходы к инженерии НРПС

Сайт-специфический мутагенез аденилирующих доменов - наиболее стабильная, точная и хорошо изученная методика генетической инженерии НРПС. Она основана на модификации десяти ключевых аминокислотных остатков, которые определяют субстратную специфичность и расположены в активном центре А-домена, что позволяет с высокой точностью перепрограммировать распознавание аминокислот. Так, исследователи успешно модифицировали А домен модуля 10 в CdaPS3, введя единичную мутацию K278Q, которая изменила субстратное предпочтение с глутаминовой кислоты на глутамин. В другом случае единичная мутация W239S в А-домене GrsA грамицидин S синтетазы обеспечила смену специфичности с фенилаланина на О-пропаргил-тирозин. Третий А-домен FusA, в норме неспецифичный, принимает множественные аминокислоты (Tyr, Val, Ile, Phe), однако сайт-направленный мутагенез (S239W, L299I, G322A, V330I) перенаправил специфичность к фенилаланину, увеличив продукцию наиболее биоактивного варианта фузарицидина в три раза (Williams, 2013; Winn et al., 2016; Zhang et al., 2024).

Замещение доменов и модулей представляет собой более радикальный подход, включающий обмен целых доменов или модулей между различными НРПС для изменения продукта. Так в масштабных исследованиях даптомицина компанией Cubist Pharmaceuticals, с помощью обмена целых субъединиц НРПС между биосинтетическими путями даптомицина, A54145 и кальций-зависимых антибиотиков (CAD) были получены более 120 аналогов даптомицина с сохраненной или улучшенной антибактериальной активностью.

Замещение субдоменов представляет более тонкий подход, включающий обмен только субдоменов А-домена, ответственных за связывание субстрата, без нарушения междоменных взаимодействий. Этот подход был опробован на синтета-

зах грамицидина и гормаомицина, позволив создать химерные А-домены с новой специфичностью .

Вырезание и вставка целых модулей из НРПС позволяет получить пептиды с новой структурой и принципиально иными биологическими активностями, что было наглядно продемонстрировано при инженерии сурфактинсинтетазы. Целенаправленная делеция модуля 2 из трехбелковой НРПС системы сурфактина с последующим прямым соединением модулей 1 и 3 через оптимизированные линкерные последовательности приводила к успешному образованию укороченного гексапептидного варианта сурфактина без лейцинового остатка, который сохранил при этом свои поверхностно-активные свойства . При инженерии балимицинсинтетазы была продемонстрирована возможность удлинения пептидной цепи через вставку специально сконструированного модуля между существующими модулями 4 и 5 белка BspB, что привело к генерации как октапептидных, так и гептапептидных продуктов, содержащих три остатка гидроксифенилглицина вместо исходных двух (Hu, Liu, Li, 2019; Winn et al., 2016; Zhang et al., 2024; Kries, 2016). Следует заметить, что радикальные модификации на уровне доменов и модулей неизбежно сопровождаются значительным снижением общей эффективности биосинтеза (Williams, 2013; Winn et al., 2016; Kries, 2016; Zhang et al., 2024).

Комбинация мутагенеза и селекции является одним из наиболее эффективных методов оптимизации работы НРПС. Яркими примерами успешного применения этой методики служат работы с энтеробактин синтетазой, где замена родного А-домена EntF на домен сирингицин-синтетазы SyrE, привела к 30 кратному снижению эффективности, но после селекции эффективность синтеза у штамма была восстановлена практически до первоначального уровня (Kries, 2016).

Инженерия вспомогательных и модифицирующих ферментов представляет собой альтернативную стратегию расширения химического разнообразия продуктов НРПС, через интеграцию дополнительных ферментативных активных элементов, способных модифицировать как предшественники, так и конечные про-

дукты биосинтеза, не затрагивая основные каталитические домены. Интеграция галогеназ из разных путей в биосинтез эндурацидина изменила паттерны хлорирования, в результате чего образовались моно-, ди- и трихлорированные варианты, а интеграция маннозилтрансферазы из синтетазы рамопланина в штамм, продуцирующих эндурацидин, привела к появлению новых мономаннозилированных вариантов эндурацидина (Winn et al., 2016).

1.4. Отличительные особенности нерибосомальных пептидов

Нерибосомальные пептиды обладают рядом структурных и химических свойств, принципиально отличающих их от рибосомных белков и обуславливающих уникальные биологические активности. Эти особенности являются прямым следствием необычной ферментативной природы их синтеза - вне рибосомального аппарата, без информационной РНК в качестве матрицы.

Одной из наиболее характерных черт НРП является включение в их состав не протеиногенных аминокислот — субстратов, которые не входят в стандартный набор из 20 аминокислот, кодируемых генетическим кодом. А-домены НРПС способны активировать более 500 различных аминокислот, включая D-энантиомеры стандартных аминокислот, β -аминокислоты, α -метиламинокислоты, гидроксикислоты, орнитин, гомоаргинин, 2-аминоизомасляную кислоту (Aib), оксиаминокислоты, а также разнообразные циклические и гетероциклические строительные блоки (Stachelhaus, Mootz, Marahiel, 1999; Walsh, O'Brien, Khosla, 2013). Кроме того, посттрансляционные модифицирующие ферменты вводят в структуру НРП дополнительные функциональные группы путём гликозилирования, галогенирования, N- и O-метилирования, гидроксирования и формирования гетероциклических колец — оксазолинов, тиазолинов, пирролинов и их окисленных аналогов (Iacovelli, Bovenberg, Driessen, 2021; Süssmuth, Mainz, 2017).

Включение не протеиногенных аминокислот кардинально меняет физико-химические свойства пептида по сравнению с его рибосомным аналогом. D-аминокислоты и N-метилированные остатки создают стерическое препятствие для

пептидаз, обеспечивая выраженную устойчивость к протеолизу (Süssmuth, Mainz, 2017). Такая устойчивость принципиально важна для биологической активности НРП в агрессивных внеклеточных средах, в которых небольшой рибосомный пептид был бы немедленно деградирован протеазами.

По совокупности структурных особенностей НРП демонстрируют более высокую химическую стабильность, чем линейные рибосомные пептиды. Циклическая или разветвлённо-циклическая топология молекулы — характерная для большинства НРП - резко снижает гибкость пептидного скелета и существенно ограничивает доступность амидных связей для протеолитических ферментов, поскольку таковые требуют для связывания субстрата протяжённых линейных участков (Ding et al., 2020; Walsh, O'Brien, Khosla, 2013). Кроме того, D-конфигурация аминокислотных остатков, встречающаяся практически во всех семействах НРП, также препятствует эффективному взаимодействию с сайтами связывания L-специфичных протеаз (Süssmuth, Mainz, 2017).

Циклические НРП, как правило, проявляют повышенную термостабильность по сравнению с линейными пептидами. Жёсткость кольцевой конформации снижает энтропийные потери при свёртывании и повышает температуру дестабилизации молекулы. Помимо этого, дополнительные ковалентные поперечные сшивки - тиоэфирные мостики, оксазолиновые кольца, лактонные сложноэфирные связи - формируют дополнительные ограничения подвижности, повышая устойчивость структуры к денатурации под воздействием температуры или химических агентов (Iacovelli, Bovenberg, Driessen, 2021; Ding et al., 2020). Перечисленные свойства делают НРП привлекательными прототипами для разработки химически стабильных терапевтических агентов.

НРП принадлежат к трём топологическим классам: линейным, циклическим (в том числе лактамным и лактонным) и разветвлённо-циклическим. Циклизация осуществляется тиюэстеразным (ТЕ) доменом по механизму макролактонизации или макролактамизации — образования сложноэфирной или амидной связи между концом растущей цепи и гидроксильной или аминогруппой одного из

внутренних остатков (Miller, Gulick, 2016; Süssmuth, Mainz, 2017). Разветвлённые структуры возникают при нелинейных реакциях конденсации, включая редкие случаи образования β -лактамного ядра — как в биосинтезе пенициллинов (Iacovelli, Bovenberg, Driessen, 2021). Именно многообразие топологий в сочетании с включением непротеиногенных мономеров определяет исключительное химическое разнообразие НРП и широкий спектр их биологических функций — от антибиотической и противогрибковой до иммуносупрессивной и сидерофорной.

1.5. НРП с противогрибковыми свойствами

Среди известных нерибосомальных пептидов значительная часть обнаруживает противогрибковую активность, что делает их важными объектами исследований в области создания антимикотических препаратов. Механизмы антифунгального действия НРП разнообразны: нарушение целостности клеточной мембраны, ингибирование синтеза клеточной стенки, подавление репликации ДНК и индукция апоптоза (Karpiński, Adamczak, 2018; Ongena, Jacques, 2008). Ниже охарактеризованы основные семейства НРП с противогрибковой активностью.

1.5.1 Итурины

Итурины - семейство липопептидов, продуцируемых бактериями рода *Bacillus* (прежде всего *B. subtilis*, *B. amyloliquefaciens*, *B. velezensis*). Структурно итурины представляют собой циклические гептапептиды, связанные с β -гидроксижирнокислотным хвостом с 13–17 атомами углерода (Ongena, Jacques, 2008). К семейству итуринов относятся iturin A, iturin C, mycosubtilin, bacillomycin D, F и L. Iturin A — наиболее изученный представитель семейства - включает семь аминокислотных остатков в последовательности Asn-Tyr-Asn-Gln-Pro-Ser-Asn, из которых часть находится в D-конфигурации (Romero и др., 2007; Yang и др., 2024). Противогрибковая активность итуринов направлена исключительно против грибов; бактерии они практически не ингибируют. Механизм действия связан с вставкой гидрофобного жирнокислотного хвоста в мембрану гриба с последующим нарушением электрохимического градиента через образование

ионных каналов (каналов для K^+ и H^+), что приводит к осмотическому дисбалансу и гибели клетки (Yang и др., 2024; Ongena, Jacques, 2008). Генный кластер, кодирующий биосинтез итурина у *B. subtilis*, включает гены *ituA–ituD*, кодирующие четыре субъединицы мегасинтетазы НРПС; активность кластера регулируется транскрипционными факторами *Spo0A* и *DegU* в ответ на лимитирование питательных веществ (Markelova, Chumak, 2025).

1.5.2. Фенгицины

Фенгицины (плипастатины) - ещё одно семейство циклических липопептидов *Bacillus spp.* Их молекула состоит из 10 аминокислотных остатков (β -лактонный цикл образован тирозином в положении 3 и С-концевым изолейцином) и β -гидрокси-жирнокислотного хвоста с 14–18 атомами углерода, связанного по N-концу пептида. Различают фенгицин А (содержащий Ala в положении 6) и фенгицин В (с Val в той же позиции) (Del Carratore и др., 2022). Противогрибковая активность фенгицинов ограничена нитчатыми грибами; по отношению к дрожжам они практически инактивны, тогда как итурины подавляют и те, и другие (Romero и др., 2007). Механизм действия фенгицинов включает: (1) деградацию клеточной стенки; (2) нарушение целостности клеточной мембраны; (3) индукцию реакций оксидативного стресса и апоптоза (Del Carratore и др., 2022; Desmyttere и др., 2019). Биосинтез фенгицина кодируется генами *fenA–fenE*, а его экспрессия координирована с биосинтезом сурфактина. Существенным преимуществом фенгицинов является низкая гемолитическая активность при высокой биodeградируемости, что определяет их перспективность в качестве биофунгицидов (Del Carratore и др., 2022).

1.5.3. Сурфактин

Сурфактин - циклический липогептапептид, биосинтезируемый мультимодульной НРПС-мегасинтетазой *Bacillus subtilis*, кодируемой оперонами *srfAA*, *srfAB*, *srfAC* (Xu et al., 2023; Ongena, Jacques, 2008). Молекула сурфактина включает семь аминокислотных остатков (L-Glu, L-Leu, D-Leu, L-Val, L-Asp, D-Leu, L-Leu) и β -гидрокси-жирнокислотный хвост (C12–C15), связанных в коль-

цевую лактонную структуру. Сурфактин является одним из наиболее мощных биоповерхностно-активных веществ природного происхождения, что обусловлено его амфифильной структурой (Martínez-Núñez, López, 2016). По отношению к грибам собственная противогрибковая активность сурфактина относительно невысока, однако он играет ключевую роль в синергетическом взаимодействии с итуринами и фенгицинами: встраиваясь в мембрану, он разрушает нанодомены и повышает её проницаемость для других антифунгальных липопептидов (Romero и др., 2007; Desmyttere и др., 2019). Кроме того, сурфактин способствует образованию биоплёнок и является обязательным кофактором генетической компетентности *B. subtilis* через систему ComP–ComA (Martínez-Núñez, López, 2016).

1.5.4. Полимиксины

Полимиксины - семейство циклических липодекапептидов, продуцируемых почвенной бактерией *Paenibacillus polymyxa* (ранее *Bacillus polymyxa*) с помощью крупных НРПС-мегасинтетаз (Khabthani, Rolain, Merhej, 2021; Martínez-Núñez, López, 2016). Их молекула включает характерное поликатионное гептапептидное кольцо и линейный N-концевой трипептидный хвост с жирнокислотным остатком. Наиболее известны полимиксин В и полимиксин Е (колистин). Хотя полимиксины преимущественно известны как антибактериальные препараты резерва против грамотрицательных бактерий с множественной лекарственной устойчивостью, они проявляют и противогрибковую активность в отношении некоторых патогенных грибов - вероятно, за счёт неизбирательного нарушения любых мембран с отрицательно заряженными фосфолипидами (Khabthani, Rolain, Merhej, 2021). Механизм действия полимиксинов связан с вытеснением ионов Ca^{2+} и Mg^{2+} из мембраны за счёт электростатического взаимодействия поликатионных диаминамасляных кислот (DAB) с отрицательно заряженными головными группами фосфолипидов, что влечёт нарушение структуры мембраны и лизис клетки (Martínez-Núñez, López, 2016). Применение полимиксинов в клинической практике ограничено нефро- и нейротоксич-

ностью, однако нарастание устойчивости к другим классам антибиотиков обусловило их ренессанс в качестве «последней линии» терапии.

1.5.5. HSAF

Heat-stable antifungal factor (HSAF) - уникальный поли-ен макролактам, продуцируемый бактерией *Lysobacter enzymogenes*. Его биосинтез осуществляется гибридной ПКС-НППС-мегасинтетазой: ПКС часть формирует полиеновый макроциклический скелет, тогда как НППС часть встраивает остаток орнитина, участвующего в макролактамизации (Lou и др., 2011). Механизм действия HSAF связан с нарушением биосинтеза сфинголипидов через ингибирование церамидсинтазы — принципиально иной мишени по сравнению с полиеновыми макролидами (Li и др., 2009). Это делает HSAF особенно ценным соединением для борьбы с устойчивыми грибами.

1.5.6. Сирингомицин и сирингопептин

Сирингомицин (SytE) и сирингопептин (SyrA) — семейства циклических липодепептидов, синтезируемых фитопатогенной бактерией *Pseudomonas syringae*. Сирингомицины - небольшие нонапептиды (9 аминокислотных остатков) с жирнокислотным хвостом, содержащие редкую аминокислоту 4-хлор-threonine. Сирингопептины представляют собой более крупные пептиды (22 или 25 аминокислотных остатков). Оба класса проявляют фитотоксическую активность и ингибируют широкий спектр грибов и грамположительных бактерий через механизм формирования ионных каналов в мембране, что вызывает пассивный трансмембранный ток ионов и деполяризацию мембраны (Hutchison, Gross, 1997). Доказана прямая роль сирингомицина и сирингопептина в вирулентности *P. syringae*: мутанты, дефектные по биосинтезу обоих токсинов, демонстрируют снижение вирулентности на 76% (Scholz-Schroeder, Soule, Gross, 2003).

1.5.7. Тилоперин и фузаристатин

Помимо перечисленных основных семейств, описан ряд НПП с избирательной противогрибковой активностью меньшей молекулярной массы. Так, фузариста-

тин - ингибитор митохондриального дыхания грибов, продуцируемый *Streptomyces spp.* (Süssmuth, Mainz, 2017). Висконсинамиды — тетрапептиды из *Bacillus spp.* — обнаруживают активность против *Phytophthora* и *Pythium* (Ongena, Jacques, 2008). Нарастание резистентности патогенных грибов к стандартным антимикотикам диктует необходимость поиска НРП с ещё не охарактеризованными мишенями и механизмами действия, что подчёркивает актуальность геномного майнинга и биохимической верификации новых НРПС-кластеров (Blin и др., 2021; D'Agostino и др., 2018).

1.6. Практическое применение

Фармацевтическая промышленность активно использует бактериальные НРП в качестве антибиотиков, причем несколько основных классов этих соединений стали ключевыми средствами лечения тяжелых бактериальных инфекций. В последнее время даптомицин, продуцируемый *Streptomyces roseosporus*, стал критически важным антибиотиком «последней линии» для борьбы с метициллин-резистентным *Staphylococcus aureus* (MRSA) и другими мультирезистентными патогенами. Ванкомицин, еще один важный антибиотик на основе бактериального НРП, остается незаменимым для лечения тяжелых грамположительных инфекций, особенно устойчивых к другим антибиотикам.

Разработка новых антибиотиков на основе бактериальных НРП становится все более актуальной в условиях глобального распространения антимикробной резистентности, при этом исследования сосредоточены как на открытии новых соединений, так и на модификации существующих для преодоления механизмов устойчивости.

Однако применение бактериальных нерибосомальных пептидов выходит далеко за рамки антибиотиков, играя ключевую роль в различных областях медицины. Блеомицин, продуцируемый *Streptomyces verticillus*, является мощным противоопухолевым агентом, используемым для лечения рака яичек, лимфомы Ходжкина и других злокачественных новообразований. Этот препарат действует за счет генерации активных форм кислорода, повреждающих ДНК, что дела-

ет его особенно эффективным против быстро делящихся раковых клеток (Jiang и др., 2023; Khabthani, Rolain, Merhej, 2021; Li и др., 2018; Karpiński, Adamczak, 2018; Kilianova и др., 2024).

Бактериальные сидерофоры нашли широкое применение в сельском хозяйстве в качестве стимуляторов роста растений и агентов биоконтроля. Эти соединения улучшают усвоение железа растениями, преобразуя менее доступное трехвалентное железо (Fe^{3+}) в биодоступные формы, что значительно повышает урожайность и качество культур. Штаммы *Pseudomonas*, продуцирующие сидерофор пиовердин, продемонстрировали высокую эффективность в стимуляции роста сахарного тростника, райграса и огурцов. Аналогичные эффекты были отмечены у *Streptomyces*, выделенных из рисовых растений, которые увеличивали биомассу корней и побегов (Musialowski и др., 2023; Sharma и др., 2019; Styczynski и др., 2022; Timofeeva, Galyamova, Sedykh, 2022). Сидерофор-продуцирующие бактерии, такие как *Pseudomonas putida* и *Bacillus subtilis*, эффективно подавляют вредные микроорганизмы, включая *Fusarium oxysporum* и *Sclerotium rolfsii*, снижая потребность в синтетических пестицидах (Abo-Zaid et al., 2023; Nerek et al., 2022; Zou и др., 2022; Moon, Sang, 2024; Сидорова, Асатурова, Хомяк, 2018).

Комбинированное применение различных штаммов *B. subtilis*, продуцирующих разные типы липопептидов, создает аддитивный эффект в защите растений. Например, совместное использование штаммов, продуцирующих сурфактин и итурин, значительно повышает устойчивость пшеницы к *Schizaphis graminum* через активацию салицилат- и этилен-зависимых защитных генов (Rumyantsev et al., 2023), а штамм *B. subtilis* QST 713 продемонстрировал способность ингибировать рост микотоксин-продуцирующих грибов, таких как *Aspergillus flavus*, *Fusarium graminearum* и *Alternaria alternata*, снижая радиальный рост до 86% и полностью подавляя продукцию афлатоксина В1, дезоксиниваленола и зеараленона (Bertuzzi и др., 2022).

2. МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

2.1. Характеристика штаммов и видов биологических объектов

В ходе работы были отобраны и отсекувенированы 69 штаммов бактерий, представляющих 12 различных таксономических родов. Основную часть выборки составили 50 представителей рода *Bacillus*, среди которых 23 штамма *Bacillus velezensis*, 12 штаммов *Bacillus subtilis*, 7 штаммов *Bacillus halotolerans*, 3 штамма *Bacillus atrophaeus*, по 2 штамма видов *Bacillus licheniformis* и *Bacillus safensis*, а также 1 штамм *Bacillus pumilis*. Также были секвенированы 8 штаммов *Paenibacillus polymyxa*, 2 штамма *Achromobacter* sp. и по одному изоляту, относящемуся к таксонам *Acinetobacter baumannii*, *Citrobacter sedlakii*, *Cytobacillus horneckiae*, *Peribacillus frigiditolerans*, *Pseudomonas alloputida*, *Rhodococcus opacus*, *Staphylococcus saprophyticus*, *Streptomyces anulatus* и *Sutcliffiella* sp.

Для экспериментальной валидации были выбраны две функциональные группы штаммов: антагонисты грибковых фитопатогенов (штаммы AG46 и AG8) и потенциальные пробиотики для аквакультуры (MT14, MT42, MT48 и MT74).

В качестве антагонистических видов грибов использовались штаммы *Fusarium* Fus1 (из коллекции Академии биологии и медицины, выделен с пшеничного поля, пораженного фузариозом), и *Saprolegnia* spp. (из коллекции НИЛ ЦАБТ ДГТУ, выделен из тканей рыбы, пораженной сапролегниозом). Экспериментальные растения и животные относились к видам *Hordeum vulgare* и *Acipenser ruthenus* соответственно.

2.2. Выделение метаболитов и масс-спектрометрия

Для активации экспрессии генов нерибосомных пептид-синтетаз (НРПС) бактериальные штаммы предварительно пересекали за шесть часов до начала эксперимента, чтобы получить активно растущие культуры, свободные от спор. Затем их совместно культивировали на чашках Петри с фитопатогенным грибом *Fusarium graminearum* штамма Fus1 по ускоренному протоколу в течение 3–5 суток при температуре 25 °C до появления четких зон подавления роста

гриба. Собранную бактериальную биомассу переносили в жидкую среду для получения инокулята, в состав которой входили: пептон (10 г/л), хлорид натрия (10 г/л), дрожжевой экстракт (5 г/л), глюкоза (5 г/л) и бульон с солодовым экстрактом (8,5 г/л). Культивирование инокулята проводили от 18 до 24 часов в шейкере-инкубаторе при 180 об/мин и температуре 30 °С. Впоследствии 5% (по объему) полученного инокулята вносили в модифицированную среду Лэнди, имеющую следующий состав: глюкоза (20 г/л), глутаминовая кислота (5 г/л), дрожжевой экстракт (1 г/л), $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ (0,5 г/л), KCl (0,5 г/л), MnSO_4 (0,005 г/л), $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (0,00016 г/л), $\text{Fe}_2(\text{SO}_4)_3 \cdot 7\text{H}_2\text{O}$ (0,00015 г/л), KH_2PO_4 (1 г/л), фенилаланин (0,002 г/л) и триптофан (0,016 г/л). Бактерии культивировали в этой среде для продукции пептидов в течение трех суток (72 часа) при постоянном перемешивании (180 об/мин) и 30 °С.

2.2.1. Выделение и очистка липопептидов

Этап выделения и очистки липопептидов начинался с отделения бесклеточного культурального супернатанта путем центрифугирования при $6000 \times g$ в течение 30 минут. Далее супернатант подвергали ультрафильтрации с использованием систем Stericup (Millipore), при этом небольшую аликвоту стерильного супернатанта резервировали для последующего анализа противогрибковой активности методом диффузии в лунки. Для осаждения целевых пептидов уровень pH оставшегося супернатанта доводили до значения 2,0 и оставляли для преципитации на ночь при 4 °С. Образовавшийся пептидный осадок собирали центрифугированием при $12000 \times g$ в течение 10 минут и дважды промывали стерильной деионизированной водой, также подкисленной до pH 2,0. После удаления надосадочной жидкости осадок подвергали ночной экстракции метанолом. На следующее утро метанольный супернатант переносили в чистую пробирку, а оставшийся осадок повторно экстрагировали метанолом на протяжении еще одной ночи для максимального извлечения соединений. Все полученные таким образом метанольные экстракты объединяли и упаривали досуха на ротаторном вакуумном испарителе при температуре 45–50 °С, после чего взве-

шивали сухой остаток с точностью до микрограммов и растворяли его в 96%-ном этаноле.

2.2.2. Масс-спектрометрический анализ (MALDI-TOF)

Полученную смесь очищенных пептидов анализировали методом время-пролетной масс-спектрометрии с матрично-активированной лазерной десорбцией/ионизацией (MALDI-TOF) на масс-спектрометре (Bruker Daltonics, Германия). В качестве матрицы использовали α -циано-4-гидроксикоричную кислоту (HCAA). Для идентификации соединений в качестве коммерческих стандартов применяли: фенгицин (fengycin), итурин (iturin), сурфактин (surfactin), полимиксин В (Sigma-Aldrich, США), а также сульфат полимиксина М (Киевмедпрепарат, Украина). Потенциально схожие пептиды в дальнейшем дополнительно исследовались путем сравнения с базами данных nAtlas, Norine и PubChem, а также посредством предиктивного анализа AFPredictor. При этом спектр поиска был строго сфокусирован на соединениях, соотношение массы к заряду (m/z) которых превышало 124.

2.3. Выделение ДНК

Выделение геномной ДНК из клеток бактериальных штаммов осуществляли методом ЦТАБ (с применением цетилтриметиламмонийбромид) с предварительной обработкой лизоцимом и протеиназой К (Doyle, Doyle, 1987; Murray, Thompson, 1980).

2.3.1. Подготовка биомассы

Штаммы предварительно пересеивали из коллекционных чашек в 100 мл жидкой питательной среды LB и инкубировали на шейкере при 200 об/мин и температуре 37 °C в течение 24–72 часов, в зависимости от скорости роста штамма. По 1 мл клеточной суспензии из каждой колбы переносили в 6 микроцентрифужных пробирок объемом 1,5 мл. Клетки осаждали центрифугированием при 8000 g в течение 15 мин, супернатант удаляли. Осадок ресуспендировали в 500 мкл TE-bio буфера (10 mM Трис-HCl, 5 mM ЭДТА, 0,9% NaCl, pH 7,6–7,8). При необходимости процедуру промывки повторяли 2–3 раза до обесцвечивания супернатанта. Для образцов почвенного происхождения, содержа-

щих гуминовые кислоты и полифенолы, в состав TE-bio буфера дополнительно вносили поливинилпирролидон (ПВП) до концентрации 2%.

2.3.2. Лизис клеток

К 200 мкл TE-bio с клеточной биомассой добавляли 200 мкл раствора лизоцима в TE-буфере (10 мг/мл) с добавлением РНКазы А (1 мкл РНКазы, 20 мг/мл, на 1000 мкл TE-лизосима) и инкубировали 30–60 мин при 37 °С. Затем вносили 200 мкл 5 М NaCl и тщательно перемешивали. После этого добавляли 10 мкл раствора протеиназы К (2 мг/мл в TE-буфере). Далее добавляли 70 мкл 10% ЦТАБ, смесь осторожно перемешивали переворачиванием и инкубировали не менее 60 мин при 56 °С. Раствор ЦТАБ предварительно прогревали на водяной бане при 56–60 °С для растворения кристаллов.

2.3.3. Очистка ДНК

К лизату добавляли 500 мкл хлороформа, интенсивно перемешивали и центрифугировали при максимальных оборотах в течение 5–10 мин. Водную фазу аккуратно переносили в новую пробирку, стараясь не захватить интерфазу. Экстракцию повторяли с 0,5 объёма хлороформа. После двукратной очистки к водной фазе добавляли 1,5 объёма изопропанола, перемешивали и инкубировали 10 мин при комнатной температуре. ДНК осаждали центрифугированием при 4 °С при максимальных оборотах в течение 30–60 мин, либо образцы оставляли при –20 °С на ночь с последующим центрифугированием 15 мин. Супернатант удаляли, осадок дважды промывали 500 мкл 80% этанола. После удаления этанола пробирки с открытой крышкой оставляли для высыхания на 10 мин при комнатной температуре или на 2–5 мин при 60 °С. ДНК растворяли в 50 мкл TE-НК буфера (10 мМ Трис-HCl, 1 мМ ЭДТА, pH 7,6–7,8). Концентрацию и качество выделенной ДНК оценивали спектрофотометрически на приборе NanoPhotometer P330 (Implen, США) по соотношению A260/A280, а также методом электрофореза в 1,5% агарозном геле. Рекомендуемый объём клеточного осадка для эффективного лизиса составляет 10–20 мкл; проведение цикла замораживания–оттаивания образцов перед выделением повышает выход конечного продукта.

2.4. Секвенирование

2.4.1. Секвенирование на платформе Illumina

Подготовку библиотек для секвенирования на платформе Illumina проводили с использованием набора QIAseq FX DNA Library kit (Qiagen, Германия). Качество и распределение фрагментов по размеру контролировали на биоанализаторе Agilent Bioanalyzer 2100 с применением чипов High Sensitivity DNA (Agilent, США); хроматограммы наносимых библиотек демонстрировали целевые пики, соответствующие фрагментам необходимого размера. Секвенирование осуществляли на инструменте Illumina HiSeq 2500 в режиме парных прочтений длиной 100 п.о. Полученные данные конвертировали из формата BCL в стандартный формат FASTQ с использованием программы bcl-convert. Контроль качества исходных прочтений проводили с помощью программ FastQC и MultiQC. Для большинства штаммов доля нуклеотидов с качеством прочтения $Q > 30$ составляла от 85,4 до 88,1%.

2.4.2. Нанопоровое секвенирование (Oxford Nanopore Technologies)

Высокомолекулярную геномную ДНК для нанопорового секвенирования выделяли с использованием набора MagAttract HMW DNA Kit (Qiagen, Германия), обеспечивающего получение фрагментов длиной свыше 3 тыс. п.о. Концентрацию и целостность ДНК определяли с применением флуориметра Qubit 3.0; размер фрагментов контролировали электрофорезом в агарозном геле. Очистку ДНК от фрагментов малого размера проводили с помощью магнитных шариков Agencourt Ampure XP (Beckman Coulter, США).

Подготовку библиотек для нанопорового секвенирования выполняли с применением набора Rapid Barcoding Sequencing Kit V14 (SQK-RBK114; Oxford Nanopore Technologies, Великобритания) либо набора Native Barcoding Genomic DNA (EXP-NBD104 + SQK-LSK109; Oxford Nanopore Technologies) для мультиплексирования образцов. Секвенирование проводили на приборе MinION (Oxford Nanopore Technologies); исходные данные получали в формате POD5. Перевод необработанного электрического сигнала в нуклеотидные последова-

тельности формата FASTQ осуществляли с помощью бейсколлера *dorado* в режиме высокой точности. Первичный контроль качества нанопоровых прочтений проводили с применением программ *NanoPlot v1.42.0* и *pycoQC*. Низкокачественные и короткие прочтения удаляли с использованием *Filtlong v0.2.1*.

2.5. Анализ геномов

2.5.1. Сборка геномов из коротких прочтений (Illumina)

Адаптерные последовательности и низкокачественные нуклеотиды (*Phred* < 20) с концов прочтений *Illumina* удаляли с применением *Cutadapt v2.10*. *De novo* сборку геномов выполняли с использованием сборщика *SPAdes* (Nurk и др., 2013) в режиме сборки изолятных геномов. Качество полученных сборок оценивали программой *QUAST* (Gurevich и др., 2013) по числу контигов, значению *N50* и суммарному размеру сборки.

2.5.2. Сборка геномов из нанопоровых прочтений и гибридная сборка

Адаптерные последовательности нанопоровых прочтений удаляли программой *Porechop v0.2.4*. Низкокачественные нуклеотиды (*Phred* < 20) обрезали *Cutadapt v2.10*. Коррекцию ошибок выполняли с помощью *Canu v2.1.1*. *De novo* сборку геномов проводили из подготовленных нанопоровых прочтений с использованием *Flye v2.9.5* (Kolmogorov et al., 2019). Полноту и уровень контаминации сборок оценивали с применением *CheckM v1.2.3* и *CheckM2 v1.0.2*. Топологию хромосом (образование кольцевых сборок) визуализировали в *Bandage v0.9.0*.

Для гибридного улучшения сборок (*polishing*) прочтения *Illumina* картировали на нанопоровую геномную сборку с использованием *bwa mem v0.7.17*. Исправление инсерций и делеций, типичных для нанопоровых прочтений, проводили с применением *Pilon v1.23* (Walker et al., 2014). Такой гибридный подход сочетает преимущества длинных нанопоровых прочтений для разрешения повторов и высокую точность прочтений *Illumina* для точечных коррекций.

2.5.3. Таксономическая идентификация штаммов

Таксономическое определение штаммов по данным полногеномного секвенирования выполняли несколькими независимыми методами. Предварительную идентификацию по наиболее длинным и высококачественным прочтениям проводили с помощью Kraken v2.1.04 с применением стандартной базы данных RefSeq (выпуск 2020-10-13) и референсного генома человека GRCh38.p13. Уточнённую видовую идентификацию проводили на основании среднего нуклеотидного тождества (ANI, Average Nucleotide Identity) к геномам из базы GTDB (Genome Taxonomy Database) с использованием протокола classify_wf программного пакета GTDB-Tk v4.2.0 (Chaumeil и др., 2022); параллельно строили филогенетическое дерево методом de_novo_wf на основе конкатенированного выравнивания по 120 филогенетически информативным маркерам с визуализацией в iTOL. Результаты дополнительно верифицировали через модуль ANI-TAX в составе PGAP NCBI (Tatusova и др., 2016).

2.6. База данных для машинного обучения

Для построения предиктивных моделей машинного обучения была сформирована специализированная база данных, включающая 623 биосинтетических генных кластера (БГК), кодирующих ферменты, осуществляющие нерибосомальный синтез. Разметка данных проводилась на основе анализа существующих литературных источников. Соединения, для которых экспериментально подтверждена противогрибковая активность, формировали целевую группу (маркер класса «1»). Остальные соединения были отнесены к контрольной группе неразмеченных данных (маркер класса «0»). Важно отметить, что присвоение метки «0» свидетельствует исключительно об отсутствии задокументированной информации о противогрибковой активности вещества на момент сбора данных, а не о фактическом отсутствии у него таких свойств. Полный перечень соединений представлен в **Приложении 1**.

Структурно-функциональная аннотация отобранных генных кластеров проводилась с использованием программного комплекса antiSMASH версии 8.0.0. Для формирования признакового пространства базы данных из получен-

ных файлов аннотации автоматически извлекались следующие характеристики: идентификаторы белковых доменов (*pfam_id*); идентификаторы кластеров ортологичных генов вторичного метаболизма (*smCOG_id*); предсказанные мономеры соединений, определенные на основе анализа специфичности активных центров кетосинтазных (KS) и аденилирующих (A) доменов; Информация об организации синтетаз на доменном уровне; предсказанные физико-химические классы аминокислот, выступающих в качестве мономеров для нерибосомных пептид-синтетаз (НРПС).

2.7. Биоинформатический анализ и машинное обучение

2.7.1. Обработка данных и извлечение признаков

Для построения признакового пространства (feature space) из результатов аннотации извлекались следующие характеристики биосинтетических кластеров: идентификаторы белковых семейств (*pfam_id*), кластеры ортологичных генов вторичного метаболизма (*smcogs_id*), функциональные домены (*domains*), предсказанные мономеры (*predicted_monomers*) и предсказанные классы аминокислот (*aa_classes*).

Текстовые данные объединялись в единый вектор и преобразовывались в бинарную матрицу присутствия/отсутствия признаков с использованием алгоритма CountVectorizer (библиотека scikit-learn). Для отсева неинформативного шума и универсальных маркёров применялась фильтрация: в итоговую выборку попадали только те признаки, которые встречались не менее чем в 4% (*min_df*=0.04) и не более чем в 80% (*max_df*=0.8) кластеров.

Поскольку в данных присутствуют только подтвержденные позитивные (противогрибковые) и неразмеченные примеры, задача классификации решалась в парадигме PU-обучения (Positive-Unlabeled Learning) с учетом дисбаланса классов. Были использованы и протестированы следующие алгоритмы:

1. Случайный лес (Random Forest) — классический алгоритм случайного леса со взвешиванием классов (*class weight*).
2. Сбалансированный случайный лес (Balanced Random Forest) — модификация случайного леса из библиотеки *imblearn*, выполняющая сбалансиро-

ванное бутстрэп-сэмплирование для устранения смещения в сторону мажоритарного класса.

3. Сбалансированный классификатор на основе бэггинга (Balanced Bagging Classifier) — композитный метод на основе деревьев решений с автоматическим балансированием подвыборок.

4. Финальная ансамблевая модель, усредняющая предсказанные вероятности (soft voting) сбалансированного классификатора на основе бэггинга и сбалансированного случайного леса для повышения обобщающей способности и робастности предсказаний.

Для оценки качества в процессе поиска гиперпараметров применялась кастомная PU-метрика (PU F-score), модифицированная для оценки recall на известном позитивном классе и контроля уровня ложноположительных срабатываний на неразмеченном пуле.

Для объективной оценки моделей применялась стратегия стратифицированной вложенной кросс-валидации (5 внешних и 5 внутренних фолдов). Стратификация выборки с разбиением 9 к 1 для обучающей и тестовой выборки проводилась с учетом как активности, так и типа продукта (НРПС, ПКС, Гибрид НРПС/ПКС).

Для предотвращения смещения оценок (эффекта «утечки данных» или data leakage) и обеспечения высокой обобщающей способности алгоритмов, процесс настройки и валидации моделей был организован по схеме вложенной кросс-валидации (Nested Cross-Validation). Во внутреннем цикле кросс-валидации осуществлялся непосредственный поиск оптимального вектора гиперпараметров. Для этого использовался метод случайного поиска (RandomizedSearchCV).

Стандартный порог бинаризации непрерывных вероятностных предсказаний моделей $T = 0.5$ часто является неоптимальным, особенно в условиях несбалансированного распределения классов или несимметричной функции потерь. Для нахождения оптимальной классификации применялась стратегия

Out-of-Fold (OOF) предсказаний. Оптимальный порог определялся путем максимизации метрики F1 на векторе OOF-вероятностей:

$$T = \arg \max_T F_1(T) \quad (1)$$

Метрика F1 представляет собой гармоническое среднее между точностью (Precision) и полнотой (Recall), что позволяет сбалансировать ошибки I и II рода:

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2)$$

Где базовые компоненты рассчитываются через элементы матрицы ошибок (Confusion Matrix):

$$\text{Precision} = \frac{TP}{TP+FP} \text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

(где TP — истинно положительные, FP — ложно положительные, FN — ложно отрицательные ответы).

Оценка финального качества моделей на внешнем отложенном контуре валидации носила комплексный характер и включала метрики, инвариантные к порогу (оценивающие качество ранжирования вероятностей), метрики жесткой классификации (после применения порога T^*), а также метрику калибровки.

Площадь под ROC-кривой оценивает качество ранжирования объектов моделью. ROC AUC численно равна вероятности того, что случайно выбранный положительный объект выборки получит предсказанную вероятность выше, чем случайно выбранный отрицательный объект. Математически выражается как интеграл зависимости доли истинно положительных ответов (True Positive Rate, TPR) от доли ложно положительных ответов (False Positive Rate, FPR):

$$\text{ROC AUC} = \int \text{TPR}(\text{FPR}) d\text{FPR}; \text{TPR} = \frac{TP}{TP+FN}, \text{FPR} = \frac{FP}{TN+FP} \quad (4)$$

Средняя точность (Average Precision, AP) - альтернатива ROC AUC, более чувствительная к качеству алгоритма в условиях сильного дисбаланса классов (когда целевой класс является редким). AP аппроксимирует площадь под кривой точности-полноты (Precision-Recall Curve) и рассчитывается как

взвешенная сумма точностей на каждом векторе порогов, где весом выступает приращение полноты:

$$AP = \sum P_n (R_n - R_{n-1}) \quad (5)$$

где P_n и R_n — точность и полнота на n -м пороге сортировки предсказаний

Коэффициент корреляции Мэтьюса (Matthews Correlation Coefficient, MCC) - метрика качества дискретных предсказаний. В отличие от Accuracy, MCC учитывает пропорции всех четырех квадрантов матрицы ошибок (TP, TN, FP, FN) и возвращает значение в диапазоне $[-1; +1]$, где $+1$ означает идеальное предсказание, 0 — случайное угадывание, а -1 — полное обратное несоответствие. Формула имеет вид:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (6)$$

Сбалансированная точность (Balanced Accuracy) применяется для получения несмещенной оценки доли правильных ответов на несбалансированных выборках. Рассчитывается как среднее арифметическое полноты для положительного и отрицательного классов (чувствительности и специфичности):

$$\text{Balanced Accuracy} = \frac{\frac{TP}{TP+FN} + \frac{TN}{TN+FP}}{2} \quad (7)$$

(где TN — истинно отрицательные ответы, а TNR — доля истинно отрицательных ответов / специфичность)

2.7.2. Интерпретация моделей и статистический анализ

Для обеспечения прозрачности (explainability) алгоритмов машинного обучения применялся метод SHAP (SHapley Additive exPlanations). С помощью TreeExplainer и PermutationExplainer рассчитывались глобальные (вклад признака во всей выборке) и локальные (триггеры для конкретного кластера) значения SHAP, что позволило выявить биомаркеры, ассоциированные с противогрибковой активностью. Оценка важности признаков дополнительно подтверждалась анализом снижения критерия Джини (Mean Decrease in Impurity).

Поиск статистически значимых различий в частоте встречаемости признаков между противогрибковыми и неразмеченными соединениями проводил-

ся с помощью точного критерия Фишера. Для контроля доли ложных открытий (FDR) при множественном тестировании применялась поправка Бенджамини-Хохберга. Различия считались статистически значимыми при скорректированном значении $p < 0.05$.

2.7.3. Автоматизация аннотации и предсказания

На первом этапе проводилось первичное, ускоренное сканирование геномных сборок с использованием платформы antiSMASH, где для поиска открытых рамок считывания применялся алгоритм prodigal, а ресурсоемкие модули в целях оптимизации вычислительных ресурсов были отключены. Главной задачей этого шага являлась базовая локализация потенциальных биосинтетических генных кластеров и их компонентов в масштабах всего генома.

Далее, на основе полученных JSON-отчетов и GenBank-файлов, выполнялась процедура таргетной подкластеризации, учитывающая вложенность структур. В ходе этого процесса из рассмотрения исключались соседствующие регионы, не входящие в состав ядра кластера, при этом для одиночных протокластеров координаты извлекались строго по их установленным границам, тогда как для остальных типов дополнительно определялись границы основных биосинтетических генов с фиксированным отступом в 20 000 пар оснований. Также применялась процедура фильтрации избыточных гибридных структур: алгоритм сканировал специфические типы кластеров, и если ядро такого подкластера полностью поглощалось границами более крупного кластера-кандидата, он распознавался как вложенный элемент и исключался из итоговой выборки во избежание дублирования.

Отфильтрованные уникальные подкластеры изолировались в виде самостоятельных GenBank-файлов, после чего на втором этапе подвергались повторной, изолированной обработке в antiSMASH с включением полного набора аналитических модулей, включая clusterhmmer для поиска белковых доменов, ss-mibig и cb-knownclusters для сравнения с референсными базами. Подобная изолированная обработка позволила максимально точно картировать внутренние структурные элементы подкластеров без внесения информационного шума

от смежных геномных регионов. По завершении финальной машинной аннотации специализированный парсер автоматически извлекал необходимые структурные и функциональные признаки из выходных отчетов, формируя признаковую матрицу, которая затем подавалась на вход предварительно обученному ансамблю алгоритмов машинного обучения. Итоговым этапом работы пайплайна являлось ранжирование потенциально перспективных новых кластеров, не имеющих прямых подтвержденных хитов в базе MIBiG, на основе предсказанной моделью вероятности проявления ими противогрибковой активности.

2.8. Тестирование антагонистической и противогрибковой активности *in vitro*

2.8.1. Первичный скрининг антагонистической активности методом двойных культур

Первичный скрининг антагонистической активности бактериальных штаммов в отношении фитопатогенных грибов осуществляли на агаризованных питательных средах (картофельно-декстрозный агар или солодовый агар). Суспензию бактериальных клеток (либо агаровый блок с культурой) наносили на чашку Петри на расстоянии 2–3 см от инокулята тест-гриба. Инкубацию проводили при 27 °C в течение 7–14 суток. Антагонистическую активность оценивали путем измерения ширины зоны ингибирования роста мицелия (в миллиметрах) — от края колонии гриба до границы зоны подавления роста. Эксперименты проводили в трехкратной биологической повторности. В качестве отрицательного контроля служила стерильная питательная среда.

2.8.2. Оценка активности культуральной жидкости

Для предварительной оценки противогрибковой активности БКЖ использовали метод диффузии в агар. На поверхность агаризованной среды Сабу-ро в чашках Петри наносили суспензию тест-гриба и равномерно распределяли шпателем до полного высыхания влаги. В толще агара стерильным пробочным сверлом (или наконечником дозатора) формировали лунки, в которые вносили по 100 мкл тестируемой культуральной жидкости. Антимикотическую активность метанольных экстрактов оценивали луночным методом на среде МПА с

добавлением сусл. Суспензию культуры *Candida albicans* или *Fusarium oxysporum*, приготовленную в стерильной дистиллированной воде, в объеме 100 мкл наносили на поверхность агара и распределяли сплошным газоном с помощью стерильного шпателя. В центре чашки вырезали лунку, в которую вносили 150 мкл исследуемого образца. Культуры инкубировали в термостате при 37 °С в течение 24 часов, после чего оценивали наличие и диаметр зон задержки роста тест-микроорганизма.

2.9. Дизайн модельного эксперимента на ячмене (*Hordeum vulgare*)

В исследовании использовались штаммы-фитостимуляторы *Bacillus atrophaeus* AG8 и *Bacillus safensis* AG46. Ночную культуру бактерий (жидкая среда LB, 37°C, 140 об/мин) стандартизировали денситометром до 1 по Макфарланду, трижды отмывали стерильной водой путем центрифугирования (12 000 g, 5 мин) и готовили серийные разведения. Оценку стимулирующей рост активности проводили на семенах ячменя обыкновенного (*Hordeum vulgare*, сорт Медикум ОС 157) в условиях высокого инфекционного фона по *Fusarium* ($1,7 \cdot 10^5$ КОЕ/семя при смыве физиологическим раствором на среду Чапека; более 5% пораженных семян).

В модельном опыте по оценке всхожести *in vitro* семена (по 5 шт. на чашку, в трех повторностях) проращивали на стерильной фильтровальной бумаге, пропитанной бактериальной суспензией; после появления корней в темноте растения инкубировали под фитолампой в течение 10 суток, затем измеряли длину проростков, а ткани замораживали при -80°C.

В вегетационном опыте обработанные семена выращивали в перлите под фитолампой в течение 14 суток (в трех повторностях) с последующим измерением длины стеблей и корней. Для анализа экспрессии генов базового иммунного ответа ячменя (PR-1, PR-2, PR-3) растительные ткани фиксировали в реагенте IntactRNA, гомогенизировали и выделяли тотальную РНК.

2.10. Дизайн модельного эксперимента на стерляди (*Acipenser ruthenus*)

В исследовании использовали штаммы *Bacillus velezensis* (MT14, MT42) и *B. subtilis* (MT48, MT74). Их противогрибковую (против *Saprolegnia spp.*) активность определяли *in vitro* методом совместного культивирования на среде Сабуро (25°C, 7 суток) в трех повторностях с расчетом средних значений, измеряя зоны задержки роста мицелия. Для эксперимента *in vivo* 519 особей молоди стерляди (*Acipenser ruthenus*, 5,21–5,65 г) без клинических признаков заболеваний случайным образом распределили в три бассейна УЗВ (по 173 особи). Фоновое содержание *Saprolegnia spp.* в воде, определенное трехкратным высевом на агар Сабуро, составило $3,2 \pm 1,1$ КОЕ/100 мл. В течение 64 суток контрольная группа получала базовый рацион (54% белка, 12% жира), а опытные — корм с добавлением 0,1% пробиотиков: группа 1 получала штаммы MT14 и MT42 ($5,2 \cdot 10^6$ КОЕ/г), группа 2 — MT48 и MT74 ($6,1 \cdot 10^6$ КОЕ/г). Для анализа экспрессии генов отбирали образцы мышечной ткани.

2.11. Анализ экспрессии генов

Тотальную РНК из зафиксированных в IntactRNA мышечной ткани стерляди и замороженных при -80 °C образцов ячменя выделяли набором «ExtractRNA» (Евроген). Одноцепочечную кДНК синтезировали с использованием набора «MMLV RT kit» (Евроген). Количественную ПЦР проводили на амплификаторе RotorGene Q (Qiagen) с реагентами «5X qPCR-Mix-HS» (Евроген). Для тканей рыб оценивали экспрессию генов-биомаркеров состояния и иммунитета (HSP 70, IL-1 β , IL-10, IGF-1, GST) с АКТВ в качестве референсного гена, для растений — генов патогенез-ассоциированных белков (PR-1, PR-2, PR-3), в качестве референсного использовался ген β -тубулина. Праймеры представлены на таблице 1.

Таблица 1. Последовательности праймеров для ПЦР-РВ.

Ген	Прямой праймер (5'-3')	Обратный праймер (5'-3')	Длина продукта (п.н.)	Температура отжига (°C)
IL-1 β	TGACGGTTCCT	ATCCCCTCCGCT	132	F: 56.80,

Ген	Прямой праймер (5'-3')	Обратный праймер (5'-3')	Длина продукта (п.н.)	Температура отжига (°C)
	CCTCTTGA	CTTCTATT		R: 56.02
IL-10	TGTCCTACTTAT TTGACAGCCA	TTGCAGCCTACA CACGTCAT	581	F: 57.11, R: 59.57
IGF-1	TGCTCCTGTGTG TTCTATGCCT	GGAAGCAGCAC TCATTACAG	189	F: 59.26, R: 57.91
GST	GAAGGTTCTTGA GGGTGAGC	AGGAGGCACTTC ACAGTACAAT	580	F: 60.51, R: 59.61
ACTB	TTGGCGCTTGAC TCAGGATT	CCTCAACTCTCT CTACGGCTCA	429	F: 59.96, R: 60.94
PR1	CCTACAGGAAG AAGACCGGC	GTGACCCTCCTC GCCTAGTA	484	F: 59.82, R: 60.11
PR2	GGCCAACGCTAG GTTCTACA	TAGGGACTTAAC GAGCACGA	213	F: 59.75, R: 58.17
PR3	TGGCTGCAGCAA TGGAAGTA	AAAGCCAGAAT GCCCTCCTC	145	F: 59.96, R: 60.03
β- Тубулин	GAGTTCACTGAG GCCGAGAG	TCCCTACGAAAA CAAGCTCCA	193	F: 59.83, R: 59.30

Относительные уровни экспрессии рассчитывали методом $2^{(-\Delta\Delta Ct)}$, а статистическую достоверность результатов оценивали с применением t-критерием Стьюдента (для нормального распределения) или критерием Уилкоксона (при отклонении от нормального распределения).

3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3.1. Секвенирование и сборка геномов новых штаммов из коллекции лаборатории

Полногеномное секвенирование, сборка геномов и таксономическая идентификация 69 бактериальных штаммов были выполнены с использованием платформ Illumina MiSeq и Oxford Nanopore MinION. Детальные метрики качества полученных сборок (включая полноту, уровень контаминации, значения N50, размер геномов, число кодирующих последовательностей и общее количество контигов) приведены в Приложении 2. Результаты таксономической идентификации с использованием базы данных GTDB представлены в Приложении 3.

3.2. Сравнение результатов масс-спектрометрии метаболитов, выделенных из штаммов, с предсказанием AntiSMASH

Был проведен скрининг 8 штаммов *Paenibacillus polymyxa*, *Bacillus velezensis* V3.14 и *Bacillus velezensis* R4.6, обладающих доказанной противогрибковой активностью, на наличие известных биосинтетических кластеров генов с помощью antiSMASH.

Профили двух штаммов *B. velezensis* (R4.6 и V3.14) оказались полностью идентичны и включают бациллен, бациллибактин, бацилизин, диффицидин, фенгицин, макролактин Н, плантозолицин и сурфактин.

Для всех штаммов *P. polymyxa* общим является синтез фузарицидин В и тридекаптин, а почти для всех, кроме K1.14 - группы аурантининов. Помимо этого базового набора, ряд штаммов пенибацилл выделяется уникальными соединениями: у K1.14 предсказаны пельгипептины и скабихелин; у R3.13 — фрейразин, пенибациллин и пеницидин А; у R4.24 — октапептин C4; у R4.5 — богорол А и пеницидин В. Остальные изоляты не имеют строго уникальных соединений: O1.27, O2.11 и R6.14 образуют группу с одинаковым профилем (дополнительно включают бовинимид А, пенилан и полимиксин), а штамм R5.31 характеризуется лишь комбинацией соединений, разделяемых с другими изолятами (пенилан, пенилогептин, пенинодин и полимиксин).

Известно, что эффективному выпадению в осадок при кислотном осаждении за счет изоэлектрической нейтрализации и гидрофобных взаимодействий подвержены липопептиды, содержащие избыток дикарбоновых аминокислот (таких как глутаминовая кислота), у которых в кислой среде исчезает отрицательный заряд. Напротив, поликатионные молекулы с высоким содержанием остатков диаминомасляной кислоты или лизина в этих условиях приобретают сильный положительный заряд за счет протонирования аминогрупп. Из-за взаимного электростатического отталкивания они остаются растворимыми. При анализе наших данных короткие неконденсированные метаболиты, состоящие всего из 1–3 мономеров (например, свободные малонильные производные или комплексы неизвестного мономера с малонилом), были классифицированы как заведомо неосаждаемые из-за слишком малого размера, препятствующего образованию крупных белковых агрегатов при центрифугировании.

Для найденных соединений была спрогнозирована склонность к кислотному осаждению, а также определена их масса на основе баз данных *npAtlas*, *Norine* и *PubChem* (Таблица 2).

Таблица 2. Масса и подверженность кислотному осаждению предсказанных соединений

Соединение	Осаждение при pH 2.0	Масса / Диапазон масс (Да)
тридекаптин	Нет	1546.9, 1547.9, 1532.9
сурфактин	Да	1007.6, 1021.7, 1035.7
скабихелин	Возможно	646.3
полимиксин	Нет	1154.8, 1168.8, 1202.8
плантазолицин	Да	1334.6
пелгипептин A/B/C/D	Нет	1057.6, 1071.6, 1085.6
пенилодин	Нет	2326.1

Соединение	Осаждение при pH 2.0	Масса / Диапазон масс (Да)
пенилогептин	Возможно	1099.6
пенилан	Нет	2509.2
пеницидин А, В	Нет	3374.5
пенибациллин	Нет	194.5
октапептин С4	Нет	1023.7
макролактин Н	Нет	402.2
фузарицидин В	Да	881.5, 895.5, 909.5
фенгицин	Да	1462.8, 1476.8, 1490.8
диффицидин	Возможно	544.3
бовинимид А	Возможно	710.4
богорол А	Нет	1583.9
бац依лизин	Нет	270.1
бац依ллибактин	Нет	882.2
бац依ллен	Нет	580.3
аурантинин В/С/Д	Возможно	780.4, 766.4

Был проведен масс-спектрометрический анализ экстрактов, полученных после культивации бактерий на среде Ланди.

В ходе МС-анализа контрольных коммерческих пептидов фенгицина, итурина, сурфактина и полимиксинов В и М были получены спектры ионизации, представленные на рисунках 1-5. Так в ходе проведенного анализа было выявлено, что фенгцину и его изоформам соответствует m/z в пределах от 1450 до 1540 с пиковым значением 1478 и 1496 m/z (Рисунок 1).

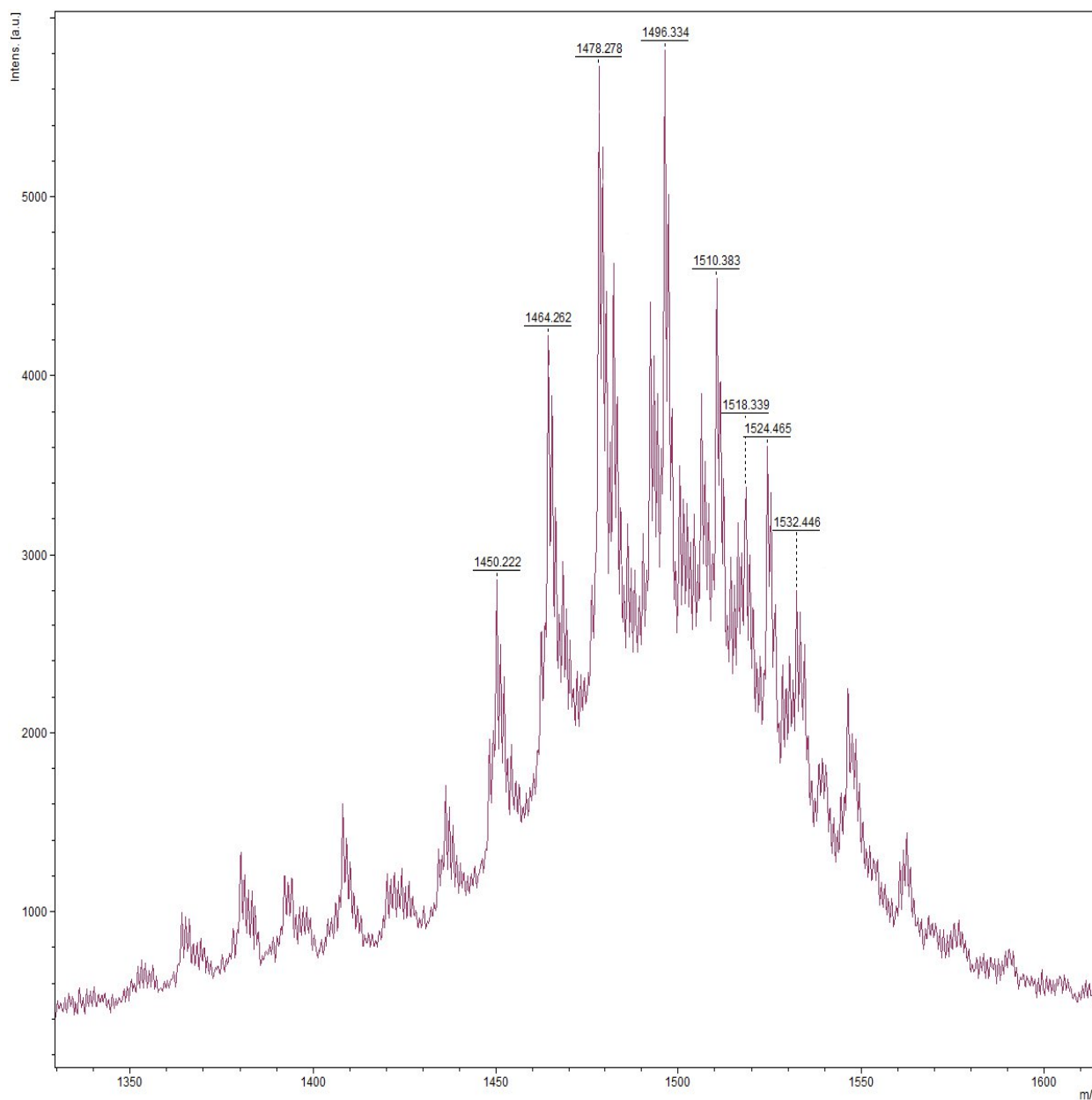


Рисунок 1. MALDI-TOF MS анализ фенгцидина (SMB00292 Sigma Aldrich).

В свою очередь итурину А и его изоформам соответствует m/z в пределах от 1043 до 1095 с пиковым значением 1065 и 1082 m/z (Рисунок 2).

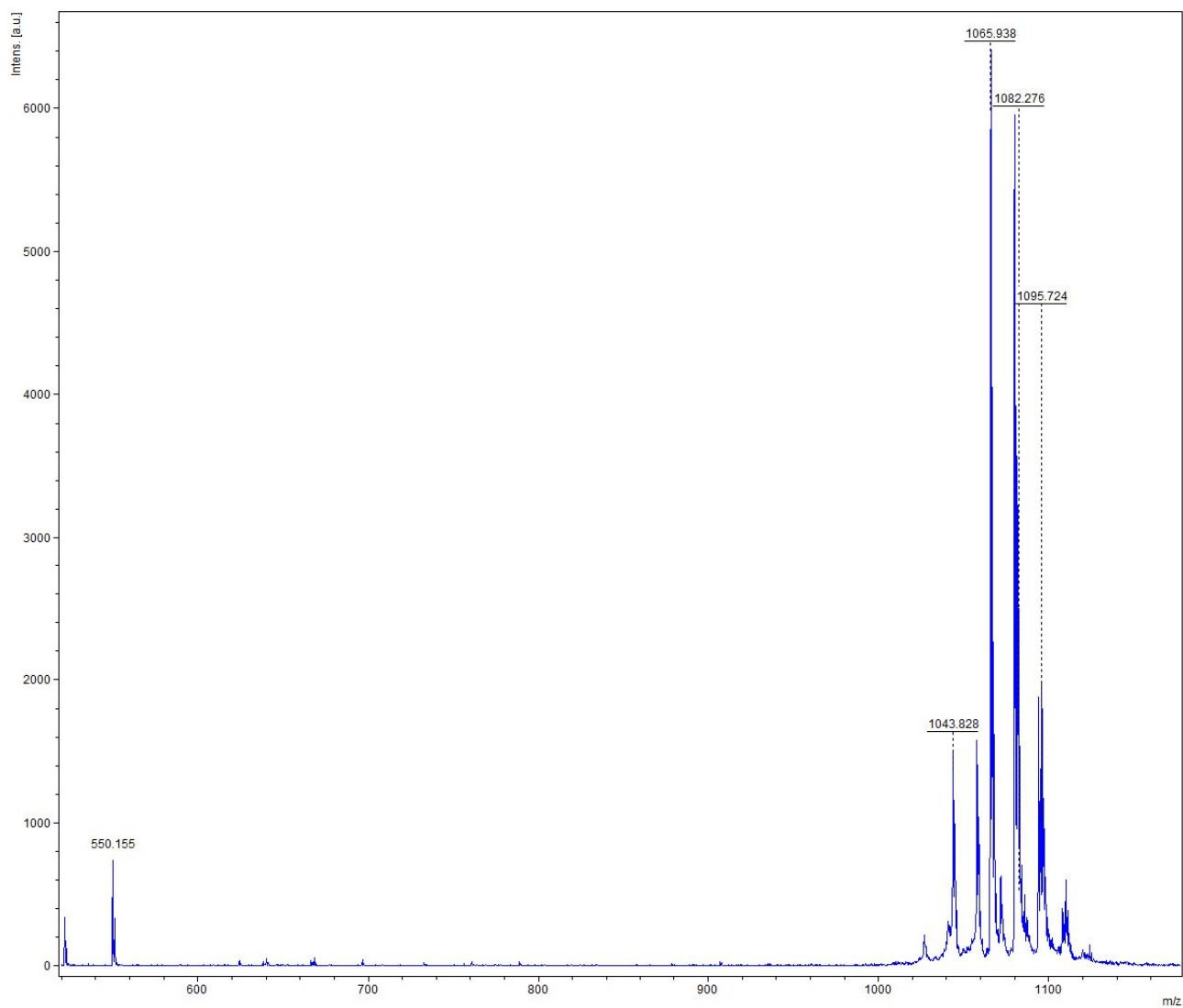


Рисунок 2. MALDI-TOF MS анализ итурина А (I1774 Sigma Aldrich).

Сурфактину и его изоформам соответствует m/z в пределах от 1030 до 1074 с пиковым значением 1044 и 1058 m/z (Рисунок 3).

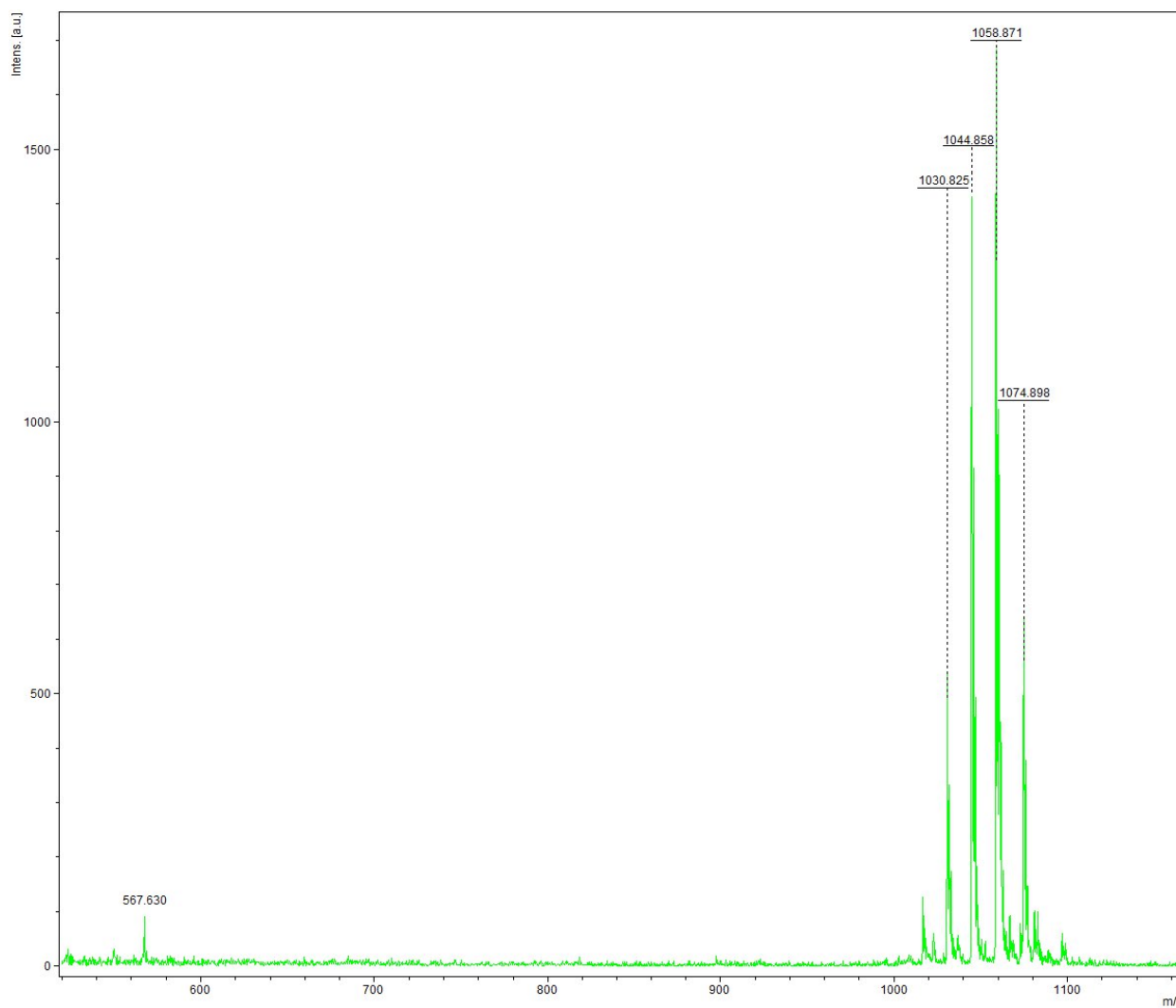


Рисунок 3. MALDI-TOF MS анализ сурфактина (S3523 Sigma Aldrich).

Полимиксину В соответствует m/z в пределах от 1171 до 1230 с пиковым значением 1196 и 1203 m/z (Рисунок 4).

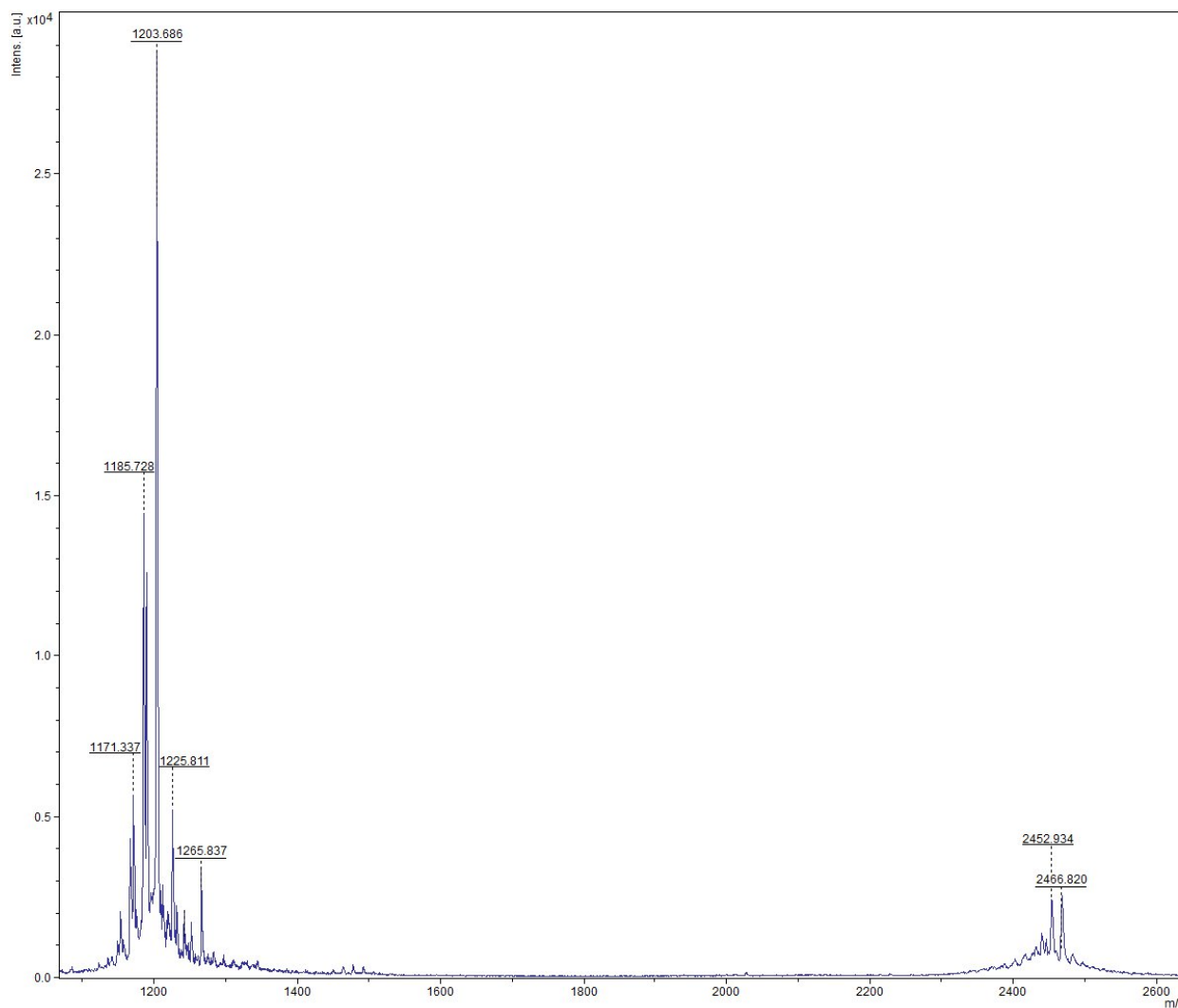


Рисунок 4. MALDI-TOF MS анализ полимиксина В (92283 Sigma Aldrich).

Полимиксину М соответствует m/z в пределах от 1157 до 1220 с пиковым значением 1205 и 1219 m/z (Рисунок 5). Кроме того, для обоих полимиксинов были характерны пики ионизации в пределах от 2300 до 2500 m/z .

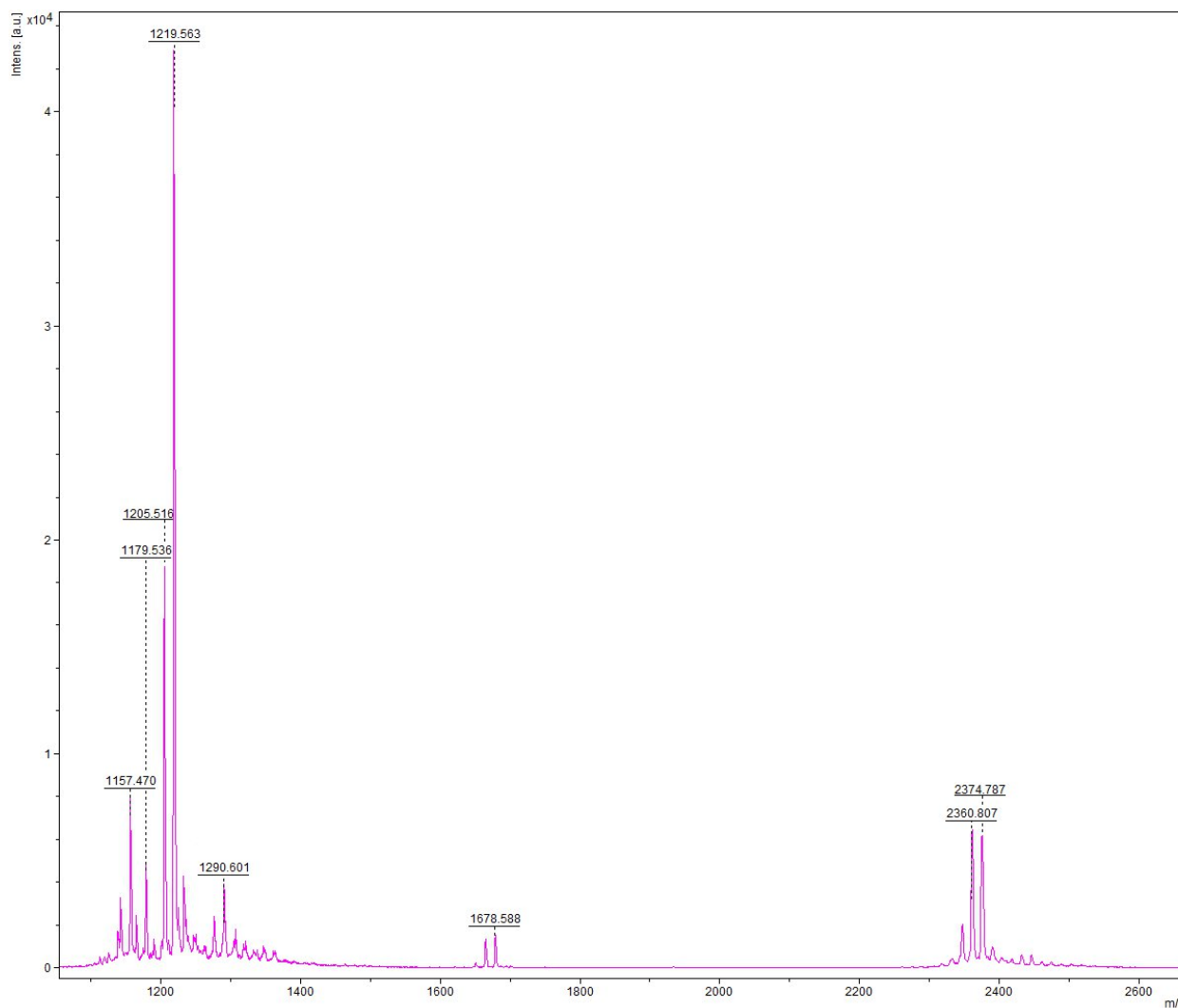


Рисунок 5. MALDI-TOF MS анализ полимиксина М

В ходе предварительного анализа чистых пептидов были получены стандарты относительно которых проводили анализ пептидов, продуцируемых отобранными штаммами бактерий-антагонистов.

Так у штамма *Paenibacillus polymyxa* К1.14 (Рисунок 6) были обнаружены лишь соединения с m/z 124.639, 132.477, 160.067, 177.775, 199.706, 221.391, 227.510, 237.283, 255.351, 291.440, 313.368, 319.360, 343.325.

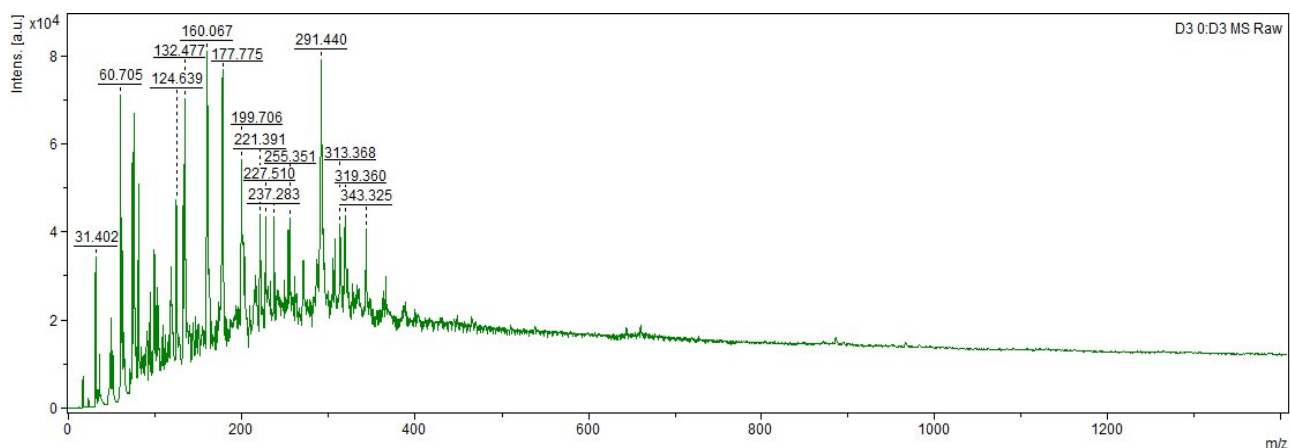


Рисунок 6. MALDI-TOF MS анализ метаболитов, продуцируемых штаммом *Paenibacillus polymyxa* K1.14

У штамма *Paenibacillus polymyxa* O1.27 (Рисунок 7) были обнаружены соединения с m/z 160.192, 199.891, 243.579, 291.590, 343.538, 872.520, 886.656, 904.802, 936.897, 951.039, 959.190, 969.201, 2780.356, 2803.484, 2819.503, 2840.905.

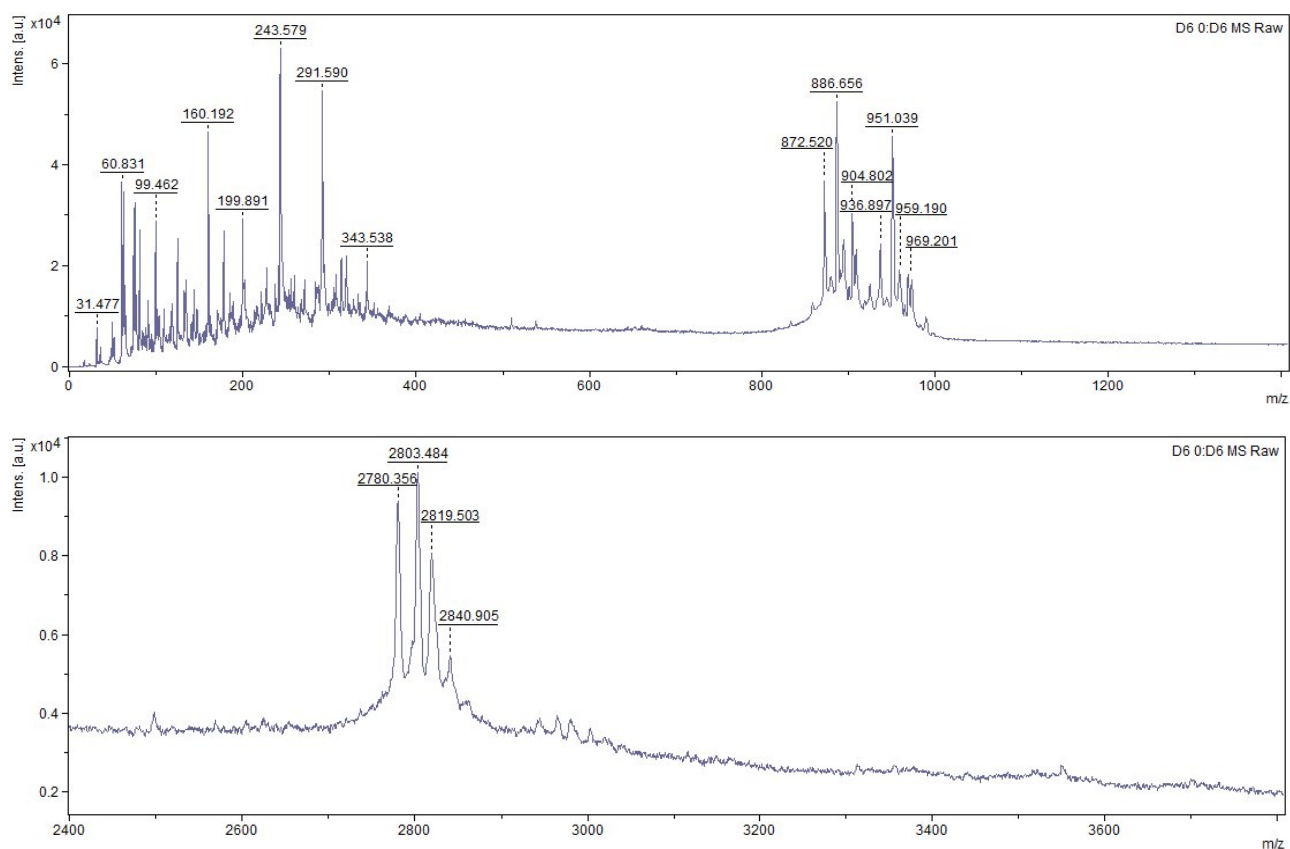


Рисунок 7. MALDI-TOF MS анализ метаболитов, продуцируемых штаммом *Paenibacillus polymyxa* O1.27

Проведенный МС-анализ метаболитов штамма *Paenibacillus polymyxa* O2.11 (Рисунок 8) позволил обнаружить соединения с m/z 160.214, 177.900, 199.873, 255.501, 291.558, 307.467, 313.523, 319.515, 343.482, 886.741 и 1257.205.

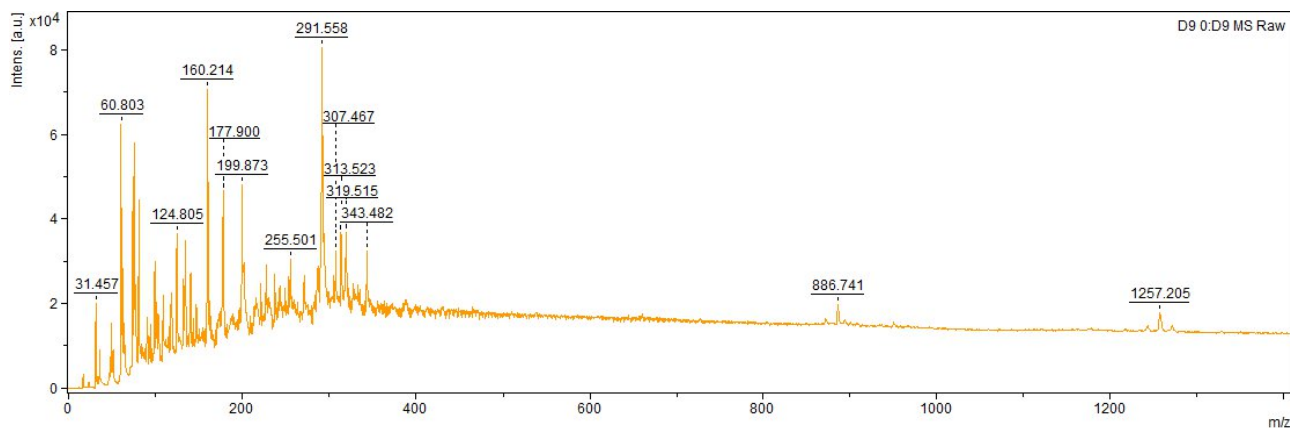


Рисунок 8. MALDI-TOF MS анализ метаболитов, продуцируемых штаммом *Paenibacillus polymyxa* O2.11

В свою очередь у штамма *Paenibacillus polymyxa* R3.13 (Рисунок 9) были обнаружены соединения с m/z 160.159, 177.906, 199.800, 227.628, 255.453, 291.497, 307.412, 313.461, 319.410, 343.442, 872.363, 936.829.

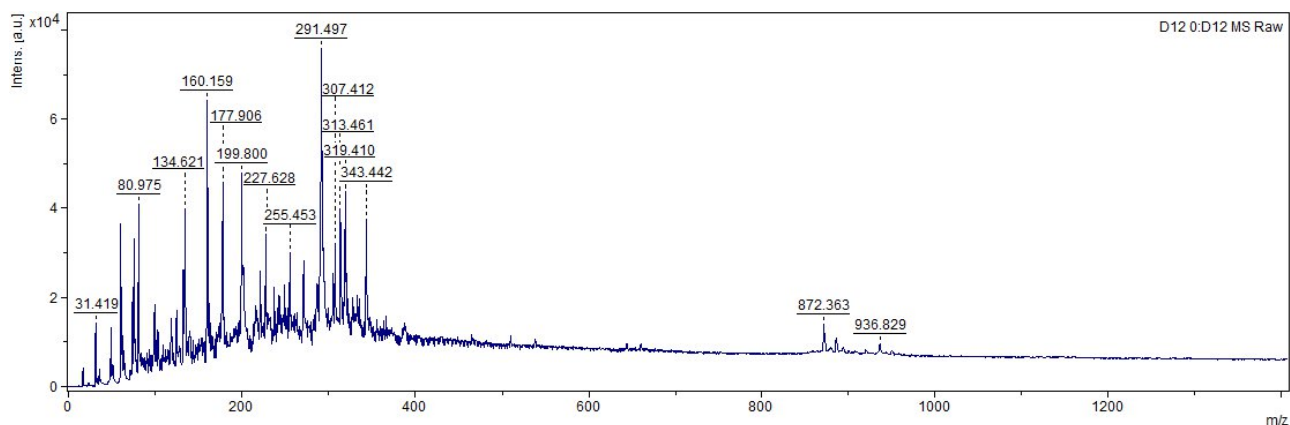


Рисунок 9. MALDI-TOF MS анализ метаболитов, продуцируемых штаммом *Paenibacillus polymyxa* R3.13

У штамма *Paenibacillus polymyxa* R4.5 (Рисунок 10) были обнаружены соединения с m/z 160.045, 199.735, 227.570, 243.464, 291.475, 307.422, 343.418, 858.566, 872.698, 886.849, 905.170, 936.962, 951.101, 958.946, 969.302.

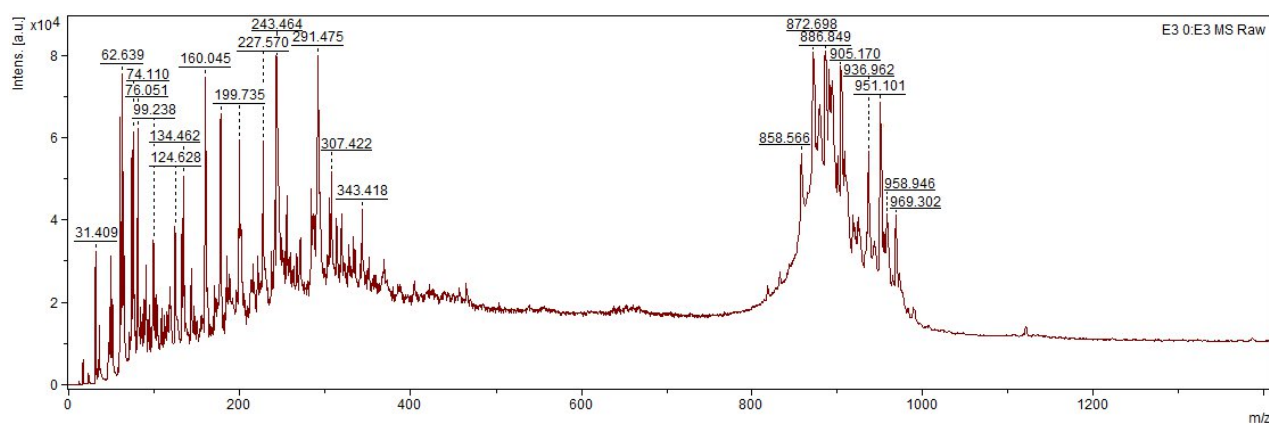


Рисунок 10. MALDI-TOF MS анализ метаболитов, продуцируемых штаммом *Paenibacillus polymyxa* R4.5

Анализ данных масс-спектрометрии метаболитов штамма *Paenibacillus polymyxa* R4.24 (Рисунок 11) позволил обнаружить соединения с m/z 124.713, 134.530, 160.122, 177.867, 199.838, 227.641, 243.513, 255.550, 291.578, 307.516, 313.526, 319.522, 343.551, 872.709, 886.859, 905.061, 937.180, 951.300.

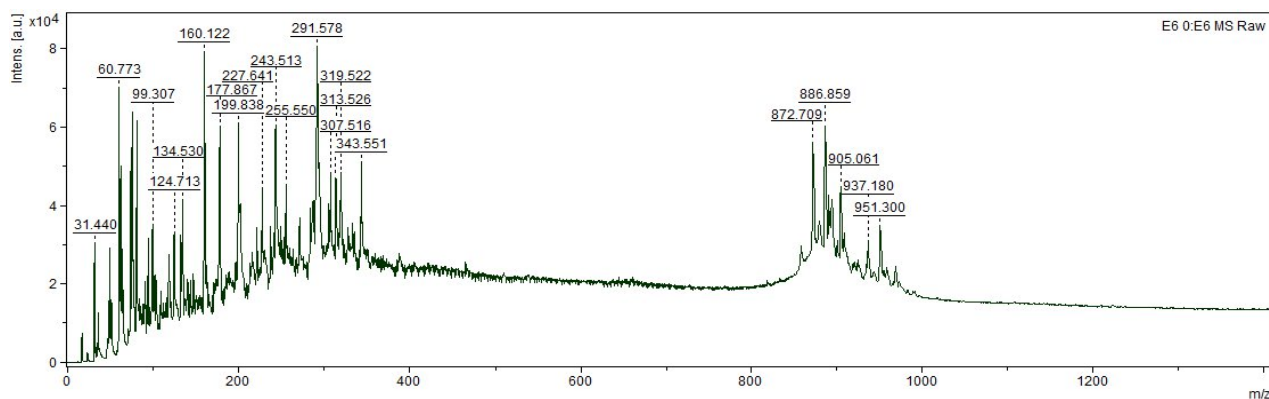


Рисунок 11. MALDI-TOF MS анализ метаболитов, продуцируемых штаммом *Paenibacillus polymyxa* R4.24

Так у штамма *P. polymyxa* R5.31 (Рисунок 12) были обнаружены соединения с m/z 199.904, 243.604, 291.622, 319.530, 343.542, 872.589, 886.782, 904.886, 937.069 и 951.199. Для остальных пептидов найти гомологи не удалось.

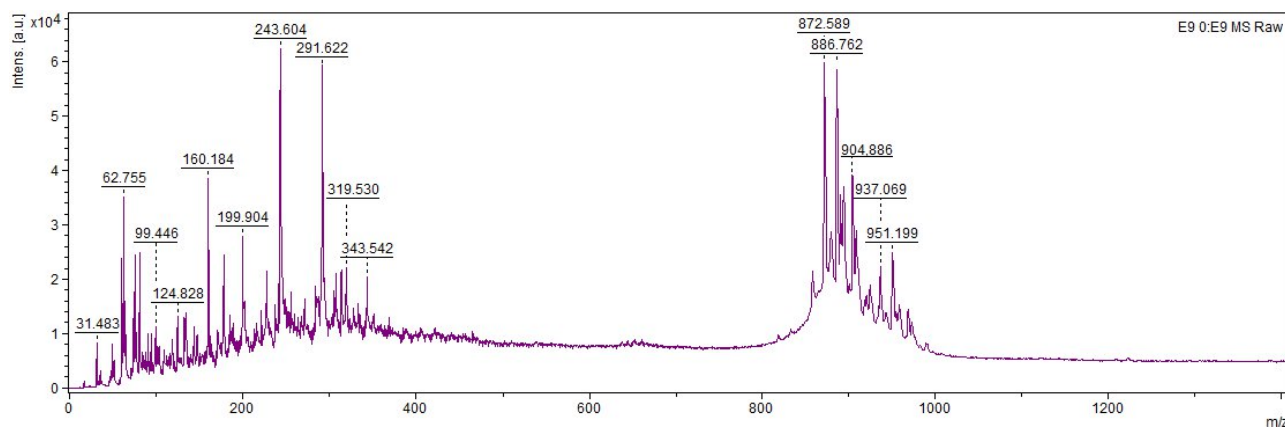


Рисунок 12. MALDI-TOF MS анализ метаболитов, продуцируемых штаммом *Paenibacillus polymyxa* R5.31.

Проведенный МС-анализ метаболитов штамма *Paenibacillus polymyxa* R6.14 (Рисунок 13) позволил обнаружить соединения с m/z 160.157, 199.843, 243.485, 291.473, 313.462, 343.444, 872.325, 886.478, 894.549, 908.691, 936.789.

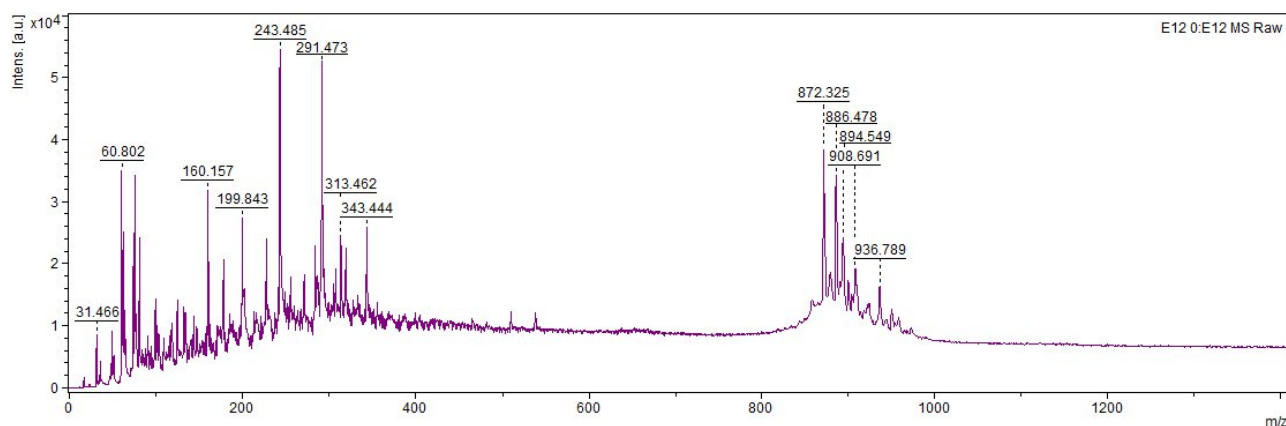


Рисунок 13. MALDI-TOF MS анализ метаболитов, продуцируемых штаммом *Paenibacillus polymyxa* R6.14

В свою очередь у штамма *B. velezensis* V3.14 (Рисунок 14) были обнаружены соединения с m/z 124.818, 160.143, 177.838, 221.422, 291.394, 319.351, 399.308, 537.810, 642.938, 1055.318 и 1069.460, 1083.727, 1097.845, 1345.794, 1448.674, 1463.454, 1477.902, 1492.110, 1506.508, 1521.068, 1535.663.

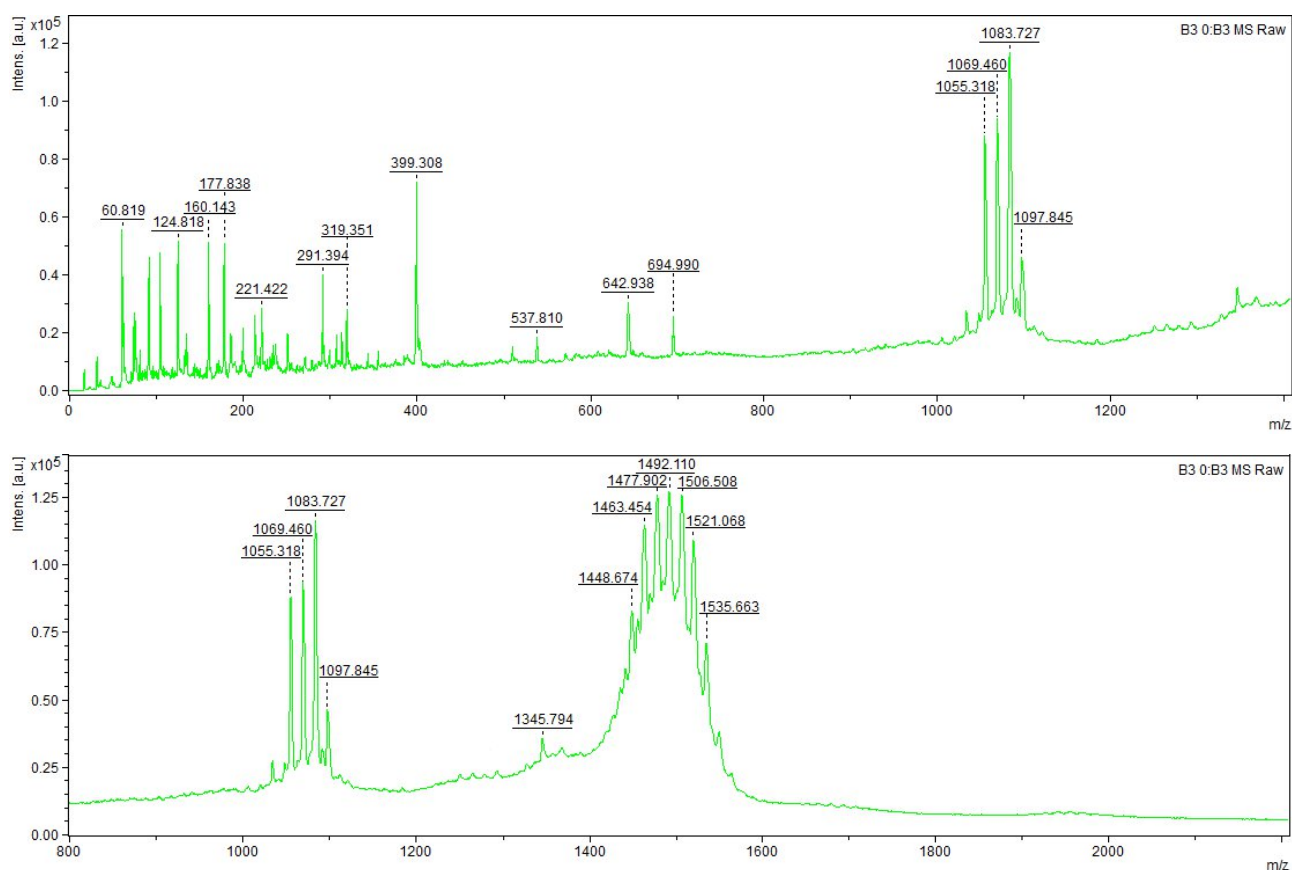
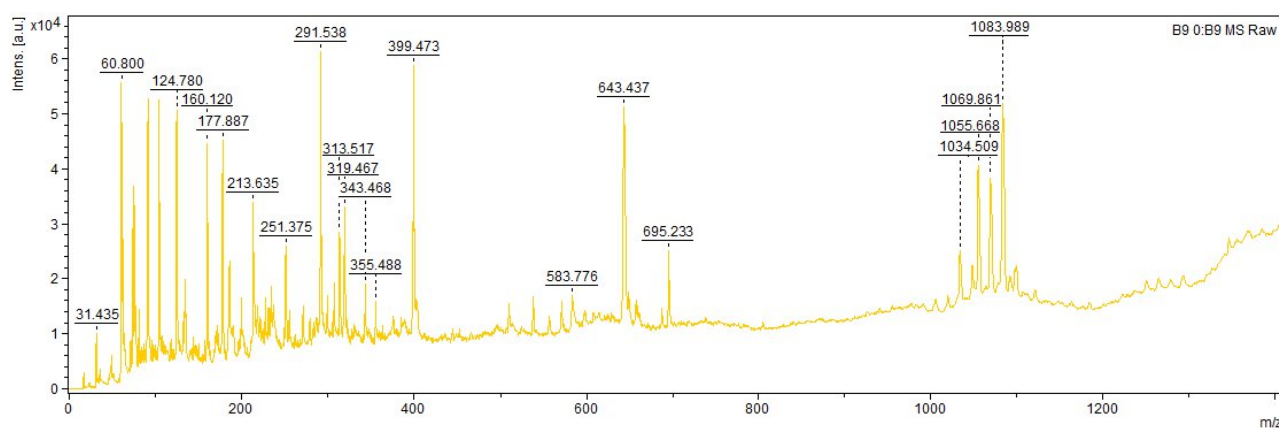


Рисунок 14. MALDI-TOF MS анализ метаболитов, продуцируемых штаммом *Bacillus velezensis* V3.14.

В свою очередь у штамма *B. velezensis* R4.6 (Рисунок 15) были обнаружены соединения с m/z 124.780, 160.120, 177.887, 213.635, 251.375, 291.538, 313.517, 319.467, 343.468, 355.488, 399.473, 583.776, 643.437, 695.233, 1034.509, 1055.668 и 1069.861, 1083.989, 1441.393, 1455.808, 1463.892, 1478.262, 1492.462, 1506.841, 1521.670, 1536.254.



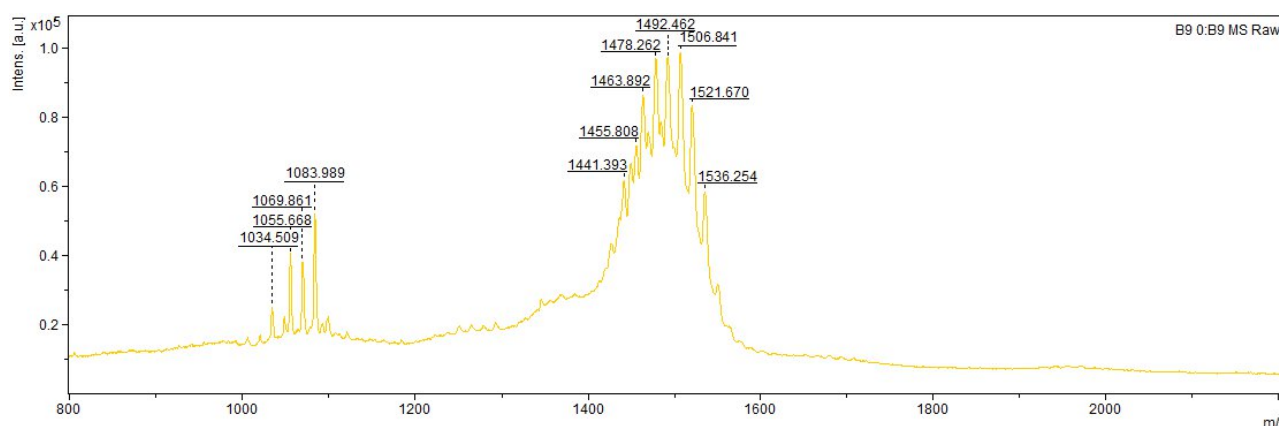


Рисунок 15. MALDI-TOF MS анализ метаболитов, продуцируемых штаммом *Bacillus velezensis* R4.6

Для точной интерпретации данных MALDI-TOF масс-спектрометрии важно учитывать два ключевых правила. Во-первых, прибор регистрирует массу иона, поэтому к теоретической массе молекулы (M) часто добавляется протон $[M+H]^+$ (+1 Да), ион натрия $[M+Na]^+$ (+23 Да) или ион калия $[M+K]^+$ (+39 Да). Во-вторых, липопептиды (к которым относится большинство антимикробных соединений бацилл) образуют гомологичные ряды — их массы отличаются на 14 Да, что соответствует одной метиленовой группе ($-CH_2-$) в липидном хвосте.

Во всех представленных масс-спектрах, в особенности у штамма *P. polymyxa* K1.14, наблюдается большое количество пиков в низкомолекулярном диапазоне (от 100 до 400 m/z). Однако при интерпретации этой части спектра следует учитывать несколько ограничений, связанных с особенностями подготовки проб и проведения масс-спектрометрии. Матрица НССА (α -циано-4-гидроксикоричная кислота) подходит для ионизации пептидов, однако образует собственные кластерные ионы в диапазоне от 100 до 500 m/z. Из-за этого интенсивного «матричного шума» детекция низкомолекулярных метаболитов, таких как бацилизин (270.1 Да), макролактин Н (402.2 Да), диффицидин (544.3 Да) технически затруднена и не может быть достоверно интерпретирована в рамках данного анализа. Помимо этого, кислотная преципитация работает для выделения кислых пептидов; однако, если бактерия синтезирует высокооснов-

ные пептиды, они могут остаться протонированными и растворимыми даже при pH 2.0, и, следовательно, не попасть в анализ.

Тем не менее, в более высокомолекулярной части масс-спектров анализ штаммов *B. velezensis* демонстрирует частичное совпадение с предсказанными данными и контрольными спектрами стандартов. Идентифицированы пики фенгицина в диапазоне 1440–1540 m/z соответствующие фенгичину и его изоформам. Группа пиков в диапазоне 1034–1069 m/z отражает наложение спектров сурфактина, пик 1083 m/z также относится к ним, но пик 1097 m/z не входит в группу сурфактинов. Для штаммов *P. polymyxa* наблюдается специфическая картина, напрямую связанная с физико-химическими свойствами предсказанных пептидов и методикой кислотного осаждения. Согласно расчетам, поликатионные пептиды, такие как полимиксин, тридекаптин и пельгипептины, содержат большое число остатков диаминомасляной кислоты и при кислотной преципитации (pH 2.0) их аминогруппы протонируются, молекулы приобретают сильный положительный заряд и не выпадают в осадок. Именно поэтому ожидаемые интенсивные пики в области 1150–1250 m/z и ~1540 m/z отсутствуют в экстрактах почти всех штаммов *P. polymyxa*.

В спектрах подавляющего большинства штаммов (O1.27, O2.11, R3.13, R4.5, R4.24, R5.31, R6.14) стабильно детектируется характерный кластер масс в диапазоне 870–970 m/z, соответствующие фузарицидину В. Фузарицидины успешно осаждаются в кислой среде и потому доминируют в спектрах. У штамма O1.27 обнаружена группа пиков в диапазоне 2780–2840 m/z. Учитывая, что теоретическая масса пенилана составляет 2509.2 Да, данные сигналы могут являться его тяжелыми аддуктами, либо представлять собой димерные комплексы других пептидов, образующиеся при ионизации. Для пика 1257 m/z у штамма O2.11 совпадений среди предсказанных antiSMASH соединений не найдено.

Таким образом, из данного сопоставления очевидна необходимость разработки улучшенного алгоритма для более точного *in silico* прогнозирования спектра соединений.

3.3. Разработка улучшенного алгоритма AFPredict

3.3.1. Создание базы данных и первичный анализ БГК

Для проведения исследования был сформирован и предварительно обработан специализированный набор данных, содержащий информацию о биосинтетических генных кластерах 574 соединений. На этапе первичной обработки данных была проведена стандартизация признакового пространства: выполнена очистка от пропущенных значений в профилях предсказанных мономеров, а также унификация номенклатуры классов аминокислот. Был проведен анализ распределения данных в обучающем датасете. Результаты представлены в таблице 3.

Таблица 3. Распределение данных в обучающем датасете

Тип	Всего	Противогрибковые (1)	Прочие (0)	Доля (1) (%)
НРПС	177	38	139	21.5%
ПКС	195	54	141	27.7%
НРПС/ПКС	202	53	149	26.2%
Итого	574	145	429	25,3%

Можно заметить, что наблюдается незначительный дисбаланс классов: так, класс (1) в общей выборке составляет 25,3%, а противогрибковая активность чаще встречается в группах ПКС или гибридных соединений на 28,8% и 21,8% соответственно. Для выявления уникальных и общих структурных компонентов (мономеров, доменов) между пулом противогрибковых соединений и контрольной группой был реализован анализ пересечения множеств. Во избежание включения в анализ шумовых и редких признаков применялся глобальный фильтр частоты встречаемости признаков, в 1% от размера выборки. Оценка степени перекрытия структурных характеристик проводилась с использованием диаграмм Венна — для попарного сравнения распределения конкретных

признаков между активными и неактивными метаболитами, результат представлен на рисунке 16.

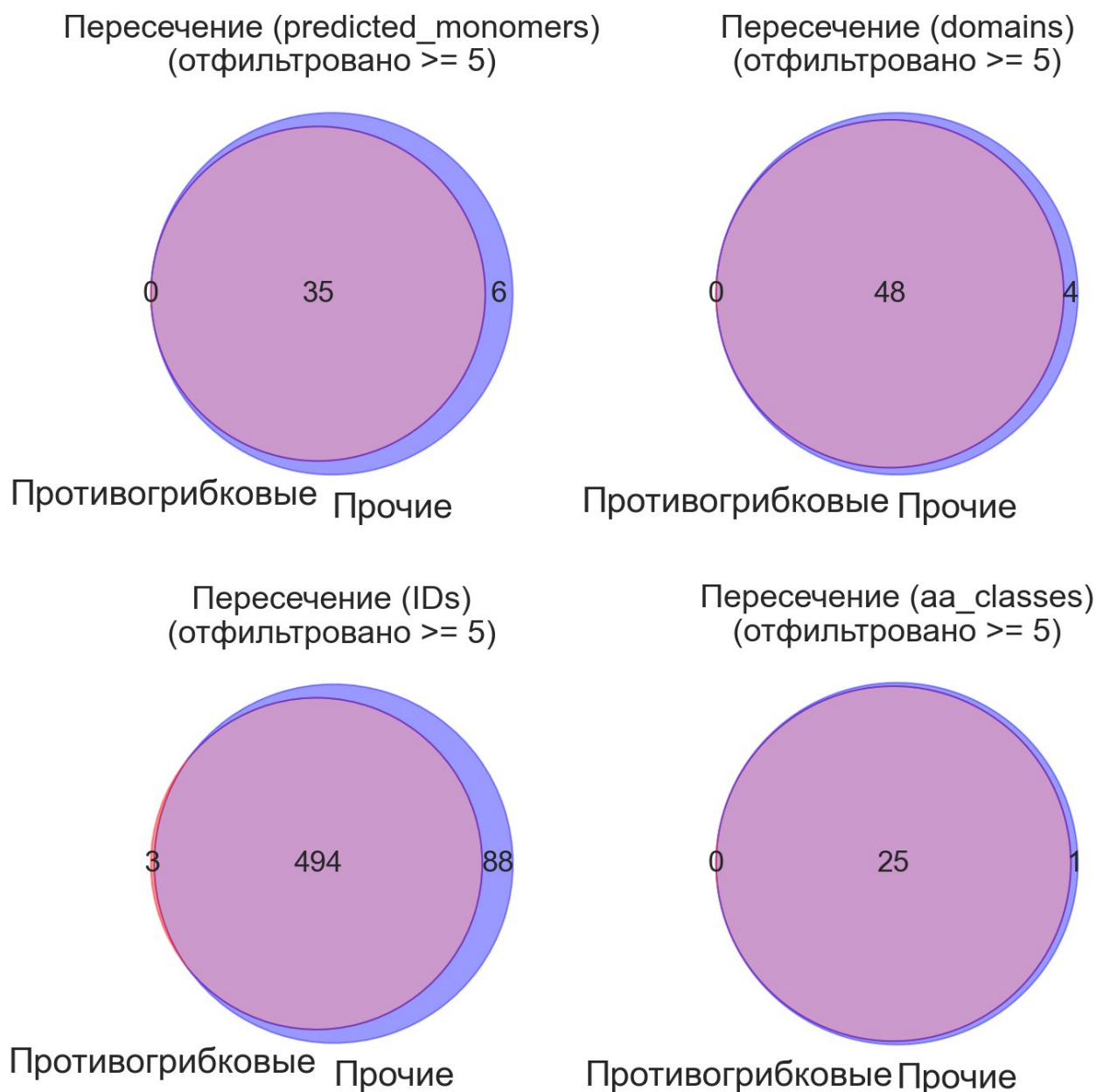


Рисунок 16. Диаграммы Венна для мономеров нерибосомальных пептидов (predicted_monomers), доменов синтетаз, идентификаторов семейств белков и физико-химических классов аминокислот.

На диаграммах видно, что уникальные для противогрибковых соединений признаки присутствуют только среди идентификаторов семейств белков, остальные признаки полностью пересекаются между противогрибковыми и прочими соединениями.

По результатам исследования мономерного состава было установлено, что из 51 исходного уникального мономера в структурах с подтвержденной противогрибковой активностью встречаются 38. Анализ показал, что основу данных соединений составляет активное использование пула стандартных протеиногенных аминокислот. Наряду со стандартными структурными компонентами, выявлено присутствие широкого спектра непротеиногенных мономеров, относящихся к различным структурным классам: D-изомеры: представлены D-аланином (D-Ala) и D-2-гидроксиизовалериановой кислотой (D-Hiv); Алло-формы (диастереомеры): включают алло-треонин (aThr); Укороченные гомологи лизина: представлены орнитином (Orn) и 2,4-диаминомасляной кислотой (Dab); Гидроксильированные производные: включают бета-гидрокситирозин (bOH-Tyr) и гидроксипролин (Hpr); Специфические ароматические блоки: представлены 3-амино-5-гидроксibenзойной кислотой (АНВА). При этом мономерный состав поликетидов и поликетидных фрагментов не отличается между противогрибковыми и непротивогрибковыми соединениями. Основные поликетидные строительные блоки представлены преимущественно метилмалонилом (mmal) и малонилом (mal), а также метоксималонил (mxmal) и этилмалонил (emal).

Для каждого уникального признака оценивалась разница в частоте встречаемости между целевой и контрольной группами. Результаты анализа для каждого типа соединений (НРПС, ПКС, Hybrids) представлены на рисунках 17, 18 и 19.

В целом можно обобщить, что алифатические и разветвленные гидрофобные аминокислоты на 38% чаще встречаются среди противогрибковых соединений. А аминокислоты с гидоксибензольным радикалом. Это подчеркивает методологическое преимущество оценки физико-химических свойств субстрата (на основе 34-аминокислотного паттерна, aa34) по сравнению с традиционным предсказанием конкретных мономеров (по 10-аминокислотному коду, aa10). В отличие от жесткого aa10-кода, физико-химическая классификация на уровне aa34-сигнатур предоставляет более гибкий инструмент для анализа, доказывая,

что общий профиль аминокислотного состава играет важную роль в предсказании противогрибковых свойств НРП.

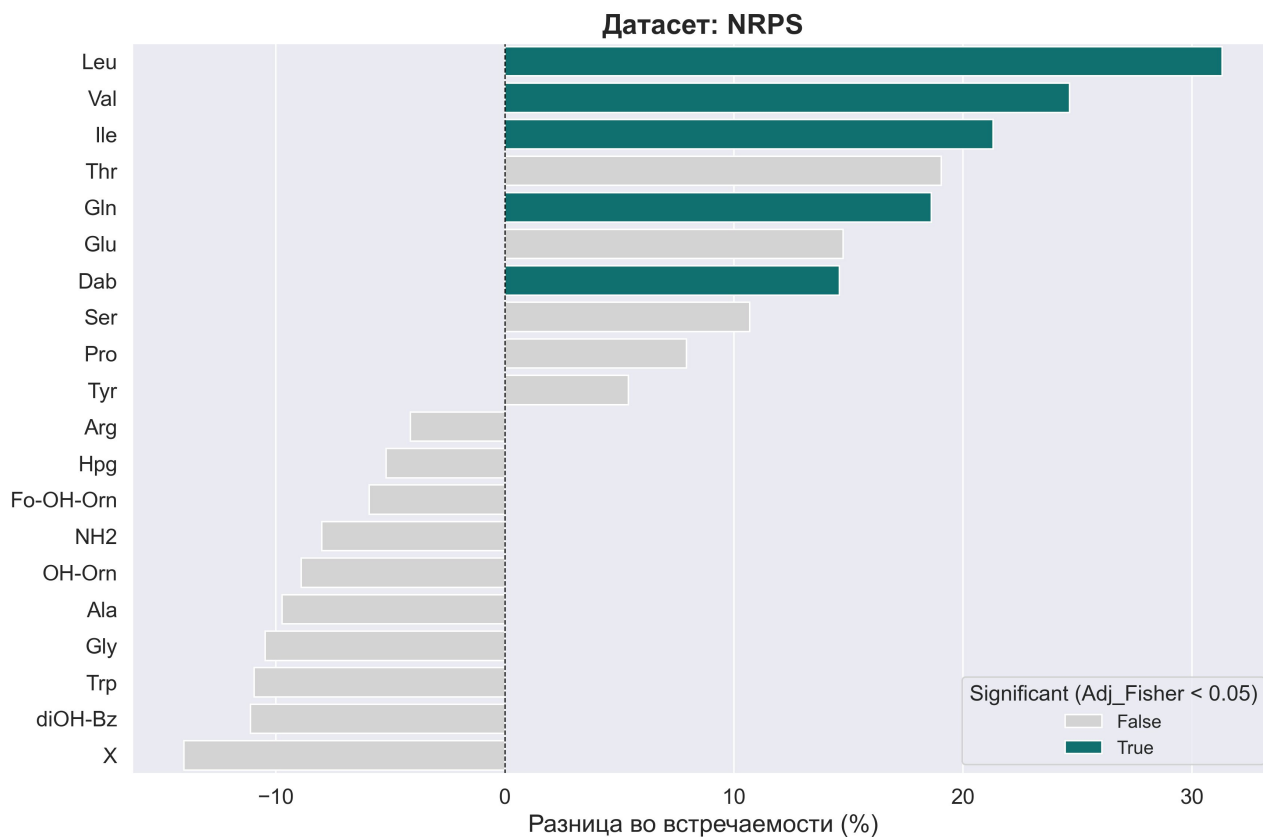


Рисунок 17. Частота встречаемости признака в зависимости от того, является соединение противогрибковым (правая часть рисунка) или нет (левая).

Аминокислоты Leu, Val, Ile, Gln и Dab статистически значимо чаще встречаются среди противогрибковых соединений, а именно на 31, 24, 21, 19 и 14%.

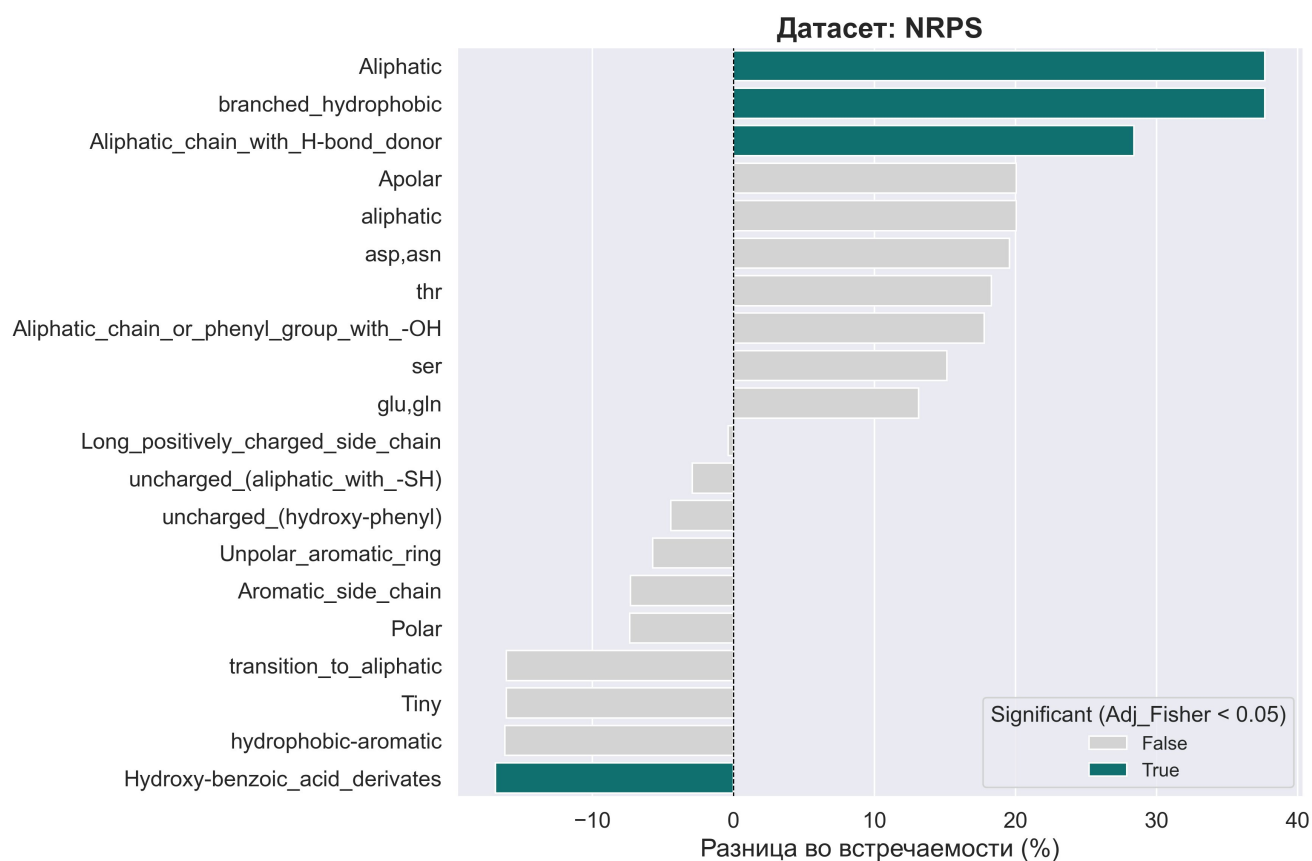


Рисунок 18. Частота встречаемости признака в зависимости от того является соединение противогрибковым (правая часть рисунка) или нет (левая).

Далее проводили сравнение встречаемости частот семейств доменов (рис. 3).

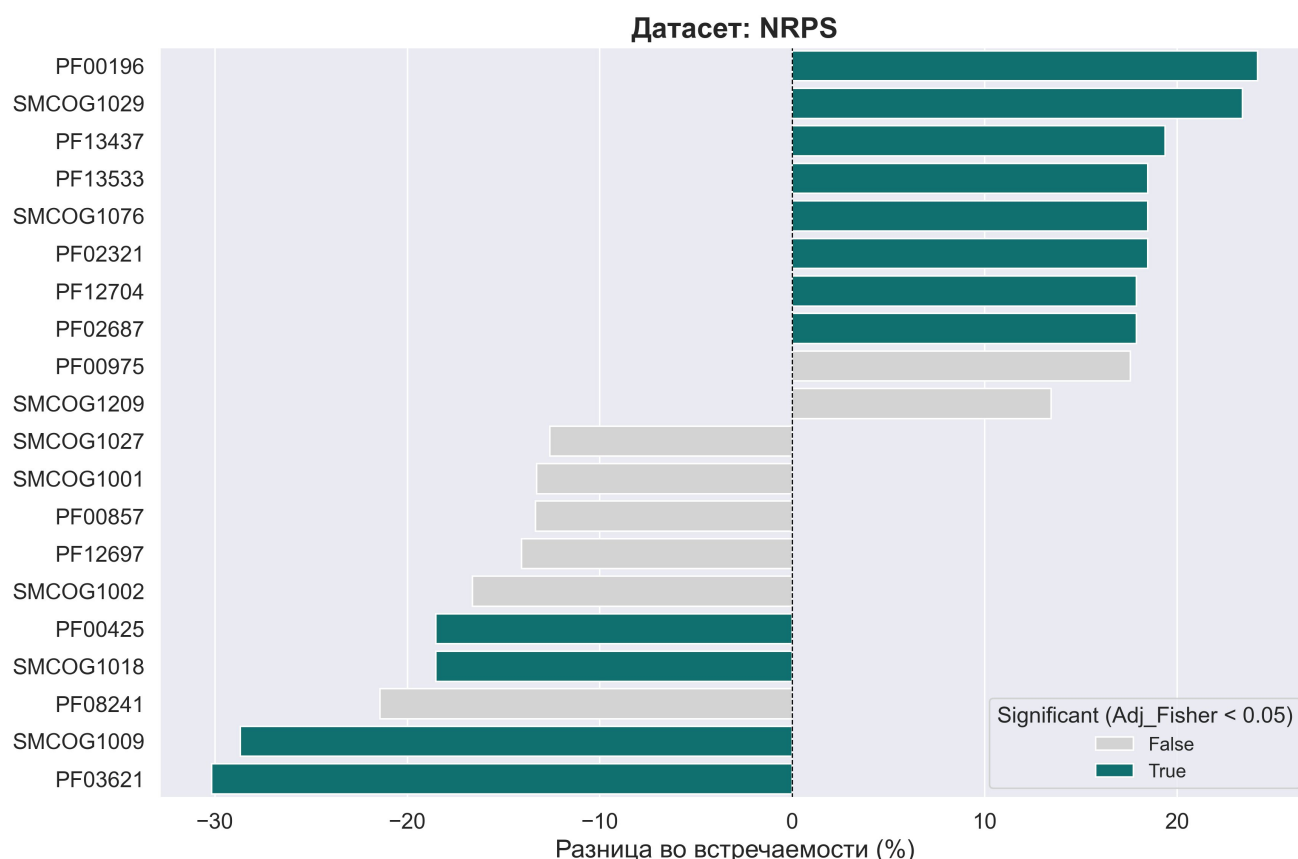


Рисунок 19. Частота встречаемости признака в зависимости от того, является соединение противогрибковым (правая часть рисунка) или нет (левая).

При анализе встречаемости отдельных белковых доменов выделяется группа маркеров, наличие которых имеет выраженную отрицательную корреляцию с противогрибковой активностью метаболита. В частности, можно заметить, что домены PF03621 и smCOG1009 статистически значимо реже встречаются среди противогрибковых соединений — на 30% и 29% соответственно. Оба этих идентификатора относятся к MbtH-подобным белкам, ассоциированным с металл-связывающей активностью соединения, а разница в частоте встречаемости связана с разными референсными генами и особенностями аннотации. Столь сильная отрицательная связь свидетельствует о том, что архитектура кластеров, синтезирующих противогрибковые липопептиды, эволюционно реже требует MbtH-активации по сравнению с другими вторичными метаболитами. Более того, эта закономерность функционально дополняется присутствием в этой же зоне графика доменов PF00425 и SMCOG1018, связанных с превращением хоризмата. В совокупности эти признаки безошибочно маркируют биосинтетические генные кластеры, специализирующиеся на производстве

металлосвязывающих сидерофоров, а не фунгицидных агентов. Таким образом, модель выявляет, что отсутствие MbtH-подобных белков и ферментов синтеза изохоризмата является таким же важным дискриминирующим фактором при предиктивном скрининге антимикотиков, как и наличие развитого транспортного аппарата.

Напротив, домены с наиболее выраженной положительной связью (+18–24%) формируют единую функциональную кассету, отвечающую за секрецию метаболита и регуляцию его синтеза. Большинство из них (SMCOG1029, PF13437, PF12704 и др.) кодируют структурные элементы сложных трехкомпонентных эффлюксных помп (например, RND или MacAB-TolC). Наличие такого развитого транспортного аппарата непосредственно в составе БГК необходимо для эффективного вывода крупных гидрофобных липопептидов во внешнюю среду. Кроме того, присутствие регуляторного домена семейства LuxR (PF00196, +24%) указывает на кворум-зависимый контроль.

Также, оценивалась совместная встречаемость стартовых конденсационных доменов (Condensation_Starter), алифатических (aliphatic) и разветвленных гидрофобных (branched_hydrophobic) структур. Оценивалась как полная коэкспрессия данной триады признаков, так и специфические исключения (например, наличие гидрофобного радикала при отсутствии стартового домена), что критически важно для понимания нетипичных механизмов биосинтеза фунгицидных агентов. Изолированный анализ предсказанных аминокислотных субстратов показал, что присутствие алифатических и разветвленных гидрофобных аминокислот является базовой характеристикой вторичного метаболизма в целом и не может служить строгим дискриминирующим фактором. Так, частота встречаемости алифатических мономеров в противогрибковой группе сопоставима с контролем (61.4% против 67.6%). Разветвленные гидрофобные аминокислоты также демонстрируют близкие значения (26.9% в активной группе и 29.6% в контрольной).

Принципиально иная картина наблюдается при оценке совместной встречаемости признаков. Анализ показал, что одновременное наличие всей триады

признаков в структуре соединения встречается у противогрибковых метаболитов в 2.2 раза чаще (17.2%), чем в контрольной группе (7.7%). Это свидетельствует о том, что именно совместная встречаемость специфических аминокислотных классов и стартового домена формирует уникальный структурный паттерн (вероятно, липопептидной природы), критически важный для проявления фунгицидного действия. Ситуация, при которой в биосинтетическом кластере закодирован разветвленный гидрофобный радикал, но при этом отсутствует стартовый домен, нехарактерна для противогрибковых метаболитов (9.7%). Напротив, в контрольной выборке подобные «неполные» комбинации встречаются более чем в два раза чаще (21.9%).

Таким образом, результаты комбинированного скрининга доказывают, что предиктивная ценность биосинтетических признаков раскрывается только при анализе их совместной встречаемости, отражающей целостный механизм действия метаболита.

3.3.2. Машинное обучение

Было разработано несколько моделей машинного обучения, способных предсказывать противогрибковые свойства нерибосомально синтезируемых соединений (рис. 20).

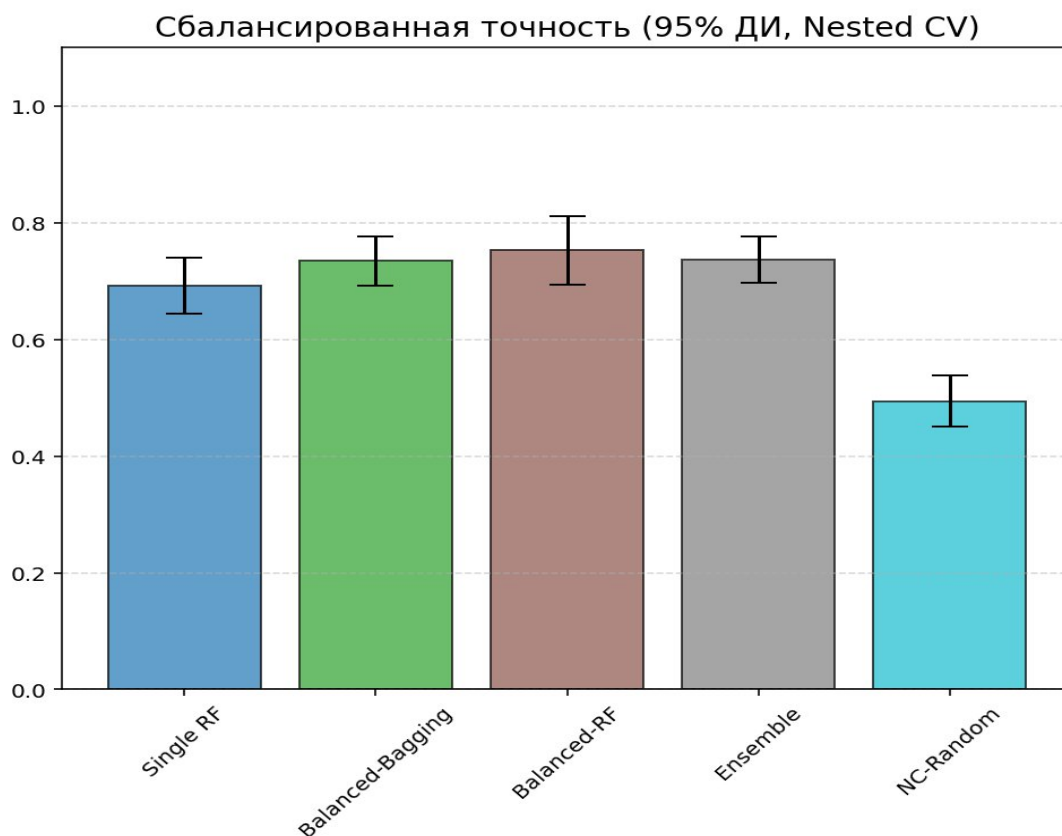


Рисунок 20. Оценка качества моделей

Статистически значимых различий между эффективностью моделей обнаружено не было, для дальнейшей работы была выбрана модель сбалансированного случайного леса (Balanced-RF) способная классифицировать противогрибковые соединения с точностью $76 \pm 5\%$ (модель AFPredictor).

Для оценки надежности предсказанных вероятностей был проведен анализ калибровочных кривых (Reliability Diagram). Данный метод позволяет установить, насколько статистическая уверенность алгоритма соответствует реальной частоте обнаружения целевого признака. Калибровочная кривая строится путем разбиения всего диапазона предсказанных моделью вероятностей на равные интервалы. Внутри каждого интервала вычисляется средняя предсказанная вероятность (ось абсцисс) и сопоставляется с фактической долей истинно положительных объектов в этой же группе (ось ординат). Идеально откалиброванная модель образует на графике строго диагональную линию. Любые отклонения от нее позволяют диагностировать излишнюю «самоуверенность» или, напротив, избыточную консервативность классификатора. Понимание этих от-

клонений критически важно для выбора адекватных порогов отсечения при биоинформатическом скрининге новых молекул.

Как видно на рисунке 21, базовая модель случайного леса (Single RF) проявляет существенные недостатки: она слишком сильно занижает оценки для явно положительных объектов, достигая доли истинно положительных соединений в 100% уже при предсказанной вероятности 0.63. Главный недостаток Single RF заключается в том, что ее кривая обрывается на отметке примерно 0.75. Это свидетельствует о том, что стандартный алгоритм, чувствительный к дисбалансу классов, в принципе не способен генерировать прогнозы с высокой уверенностью (выше 0.8) для целевого класса. В задачах скрининга это делает модель непригодной для жесткой фильтрации топ-кандидатов.

В противовес этому, использование сбалансированных подходов решает данную проблему. Ключевым практическим результатом является то, что для ансамбля моделей при высокой предсказанной вероятности (более 0.82) доля истинно положительных соединений достигает 100%. Это позволяет использовать данный порог для высоконадежного отбора перспективных кандидатов с противогрибковыми свойствами, практически полностью исключая риск ложноположительных срабатываний.

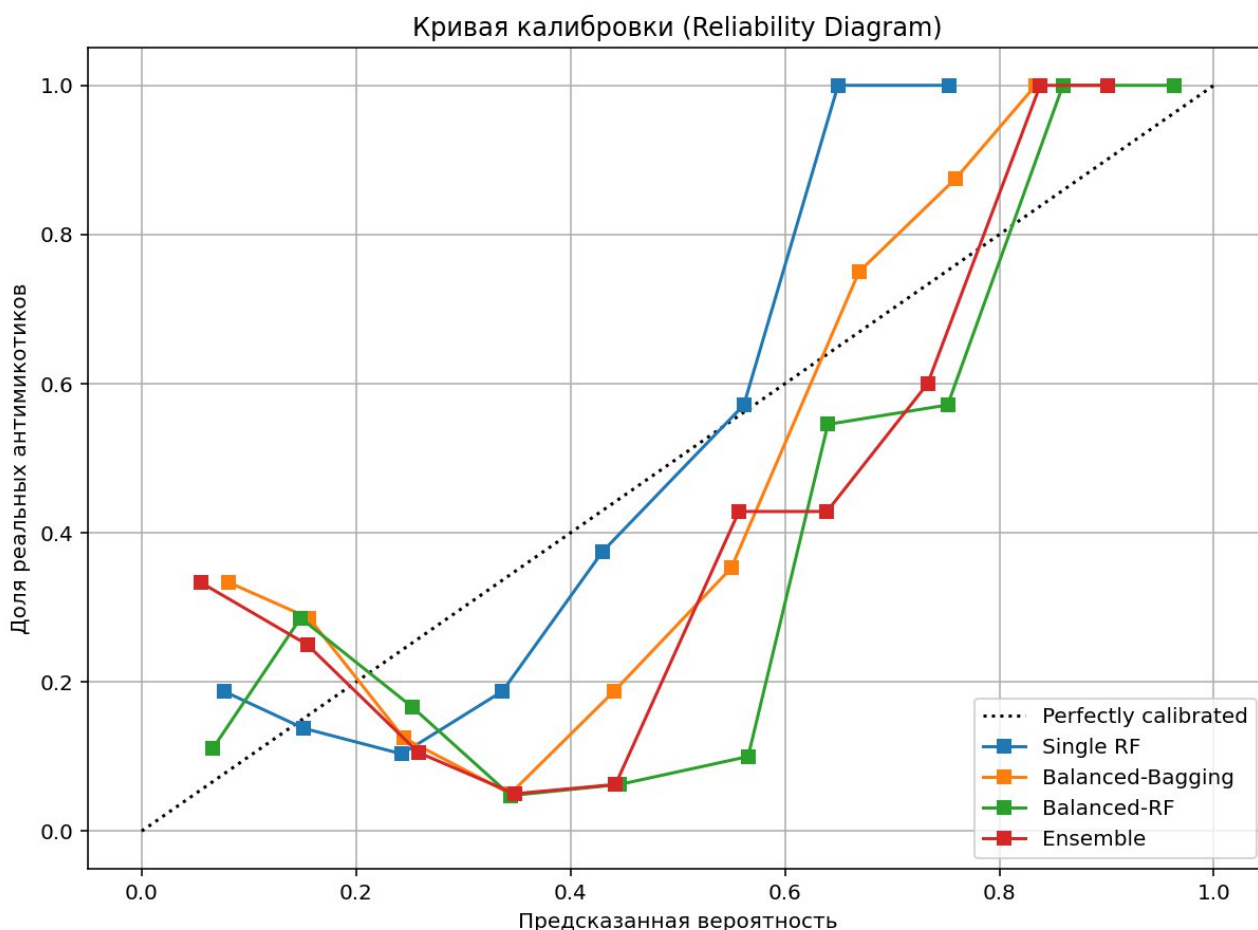


Рисунок 21. Калибровочная кривая вероятностей правильного предсказания класса вещества

С целью интерпретации результатов классификации, оценки качества модели и выявления биологически значимых закономерностей были использованы два взаимодополняющих подхода: оценка важности признаков на основе среднего уменьшения загрязнения (Mean Decrease in Impurity, Gini) и анализ векторов Шепли (SHAP Values) с использованием алгоритма TreeExplainer.

Критерий Gini оценивает, насколько эффективно конкретный биосинтетический домен разделяет целевые классы в узлах решающих деревьев. Этот метод дает надежную глобальную оценку того, насколько признак полезен для модели в целом (Рисунок 22), однако не позволяет определить направление его влияния.

Метод SHAP решает эту проблему, вычисляя индивидуальный вклад каждого признака в итоговое предсказание. Визуализация результатов в виде гра-

фигов SHAP Summary Plots позволила не только идентифицировать топ-20 наиболее значимых биосинтетических доменов, но и определить вектор их влияния, а именно — положительный или отрицательный вклад в противогрибковый потенциал исследуемого кластера генов (Рисунок 23).

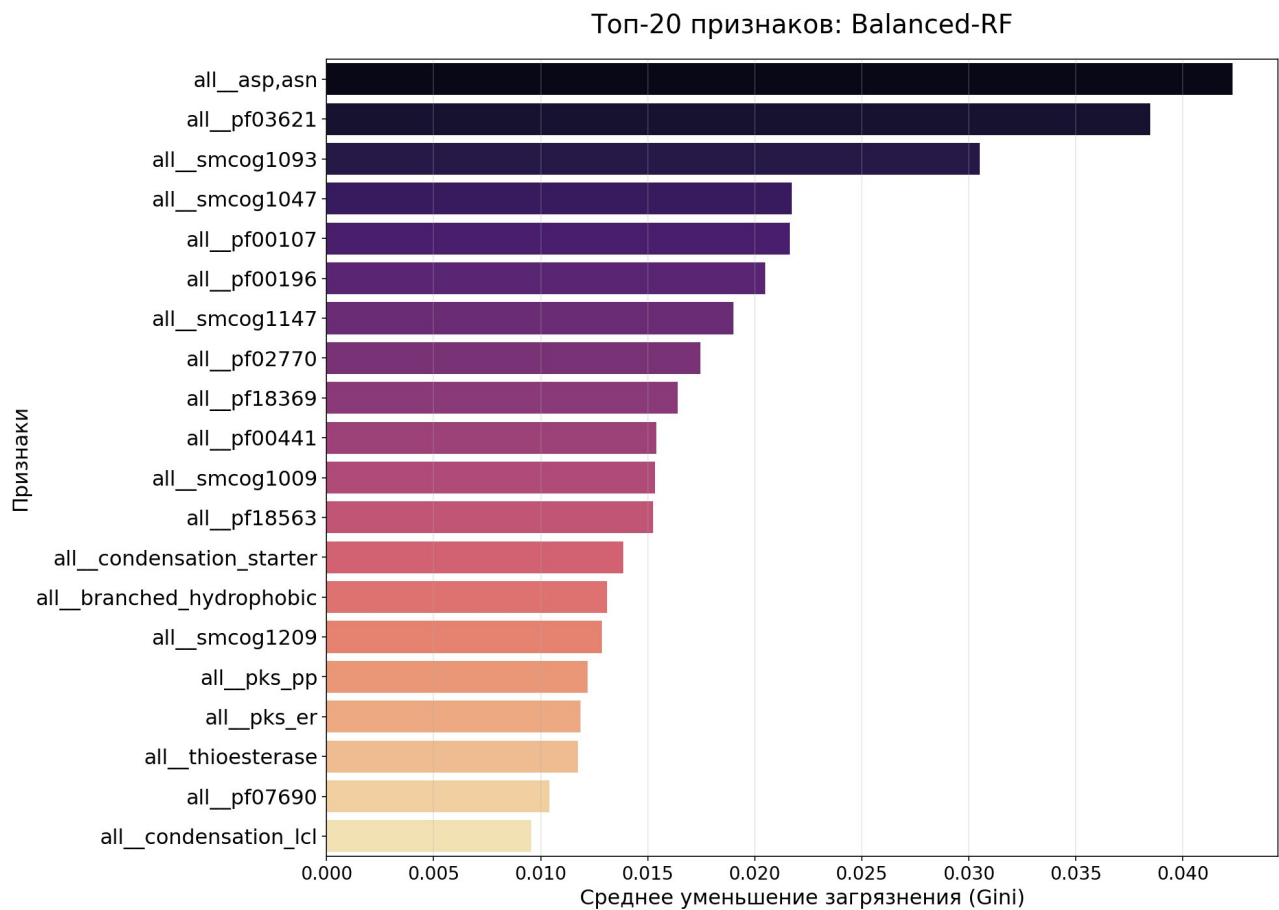


Рисунок 22. Топ-20 признаков по критерию Gini

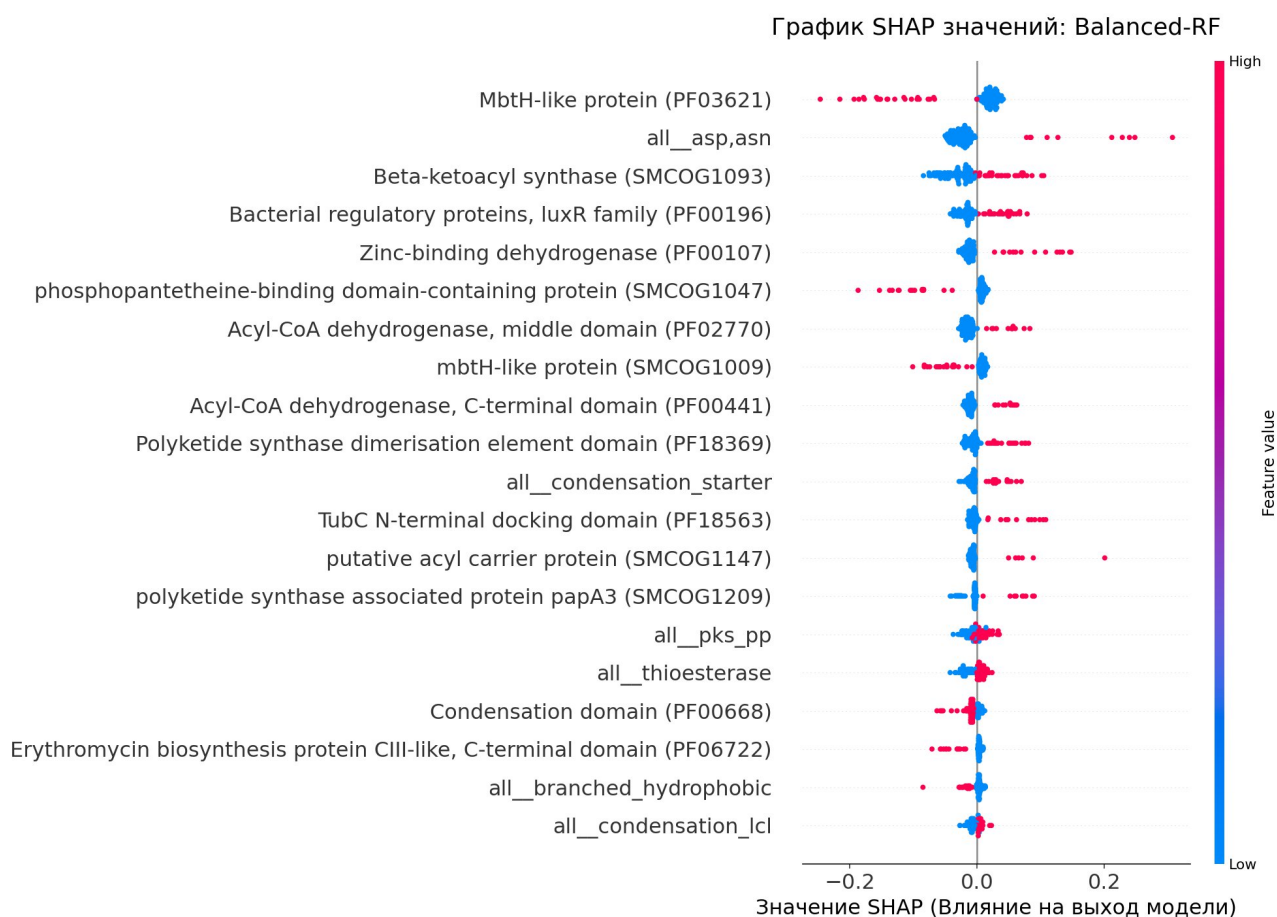


Рисунок 23. Топ 20-значений векторов Шепли (SHAP Summary Plot)

Сравнение результатов показало высокую степень согласованности обоих алгоритмов: 20 наиболее значимых признаков в значительной степени совпадают, несмотря на небольшие различия в их относительном весе, обусловленные математическими особенностями методов.

Согласно результатам анализа векторов Шепли, стартовый домен конденсации (all__condensation_starter) выступает мощнейшим положительным предиктором, что полностью согласуется с выводами предшествующего статистического анализа триад признаков. Ферменты модификации липидов ассоциируются моделью с противогрибковыми свойствами, что указывает на критическую важность липофильных хвостов липопептидной природы для проявления фунгицидной активности. Существенный положительный вклад также вносит присутствие поликетидных компонентов, включая all__ПКС_pp, кеторедуктазный домен (KR domain PF08659) и бета-кетоацилсинтазу (SMCOG1022). Это

подчеркивает высокую долю гибридных (НРПС-ПКС) или исключительно поликетидных структур среди успешных антимикотиков.

Наряду с этим обнаружена сильная ассоциация с присутствием специфических мономеров, в частности (all__asp,asn) и орнитиновых модулей (all__orn_and_hydroxy-orn_specific), что может служить надежным маркером определенного класса противогрибковых пептидов.

С другой стороны, алгоритм позволил выявить ряд доменов, присутствие которых резко снижает вероятность наличия у соединения антимикотической активности. Так, физическое присутствие MbtH-подобных белков (PF03621 и SMCOG1009) классифицируется моделью как мощнейший анти-маркер. Исторически данные белки часто ассоциированы с биосинтезом сидерофоров — железосвязывающих пептидов, что закономерно переводит содержащие их кластеры в пул соединений, лишенных фунгицидных свойств. Аналогичным образом, значительный отрицательный вклад вносит присутствие доменов эффлюксных помп и транспортных белков, таких как Drug resistance transporter, EmrB/QacA (SMCOG1005) и Major Facilitator Superfamily (PF07690). Данный факт может свидетельствовать о том, что подобные механизмы резистентности и самозащиты в большей степени характерны для бактерицидных антибиотиков или иных классов вторичных метаболитов, нежели для специализированных противогрибковых агентов.

Помимо машинного обучения для предсказания противогрибковых свойств, алгоритм AFPredictor включает в себя дополнительный шаг разбиения БКГ на подкластеры для увеличения точности работы модели и для разделения синтетаз из БКГ, содержащих более одной. Без этого, при идентификации методом BLAST, в случае если один БКГ содержит набор генов двух соединений идентифицируется только то, которое имеет в составе своей синтетазы большее число генов. Пример двойного БКГ представлен на рисунке 24.

<i>Bacillus atrophaeus</i>	фенгицин + микосуб- тилин	AG8, P1F1, ZY9
<i>Bacillus velezensis</i>	бациллибактин + ами- лоциклин	21A, 43A, 7A, FD2, Kam11, MT102, MT14, MT141, MT186, MT42, MT61, MT64, MT86, PZ2, R4.6, T5, V3.14
<i>Bacillus velezensis</i>	бацилломицин D + фенгицин	21A, 43A, 7A, FD2, Kam11, MT102, MT14, MT141, MT142, MT155, MT186, MT42, MT55, MT61, MT64, MT86, PZ2, R4.6, R5, V3.14

У штамма *Streptomyces anulatus* T5 выявлена совместная локализация путей для синтеза китацинамицинов и скилламицинов. Стоит отметить, что в паре бацилломицин – фенгицин, одна из синтетаз иногда дефектная, например, фенгицин у штаммов MT14 и MT42. Найдены несколько новых БГК, предположительно синтезирующих соединения с противогрибковыми свойствами. Результаты представлены в таблице 5.

Таблица 5. Новые противогрибковые соединения, предсказанные алгоритмом AFPredictor

Штамм	Мономерный состав	Уверенность модели
<i>Paeniacillus polymyxa</i> O1.27	fatty_acid mal Tyr D-X Ser D-Ser Ser D-Ile Ser	0.58
<i>Paeniacillus polymyxa</i> O2.11	D-Gly D-Dab D-Gly D-Ser D-Trp Ser Dab D-Dab Ile Glu	0.64
<i>Bacillus velezensis</i> R5	D-Asn Gly ccma1 ccma1 mal D-Glu	0.65
<i>Paenibacillus polymyxa</i> R3.13	D-Asn Gln Me-ohma1 mal X ohma1 X Gly ccma1 Lys X X	0.62
<i>Bacillus safensis</i> AG46	Glu Leu D-Leu Ile Asp D-Leu Ile NH2 D-Val fatty_acid Val	0.83

Полные результаты аннотации представлены в Приложении 4.

3.5. Повторный анализ штаммов с использованием алгоритма

AFPredictor

Для обоих штаммов *B. velezensis* новый алгоритм существенно расширил стандартный спектр находок antiSMASH, добавив к нему два биосинтетических кластера генов, отвечающих за синтез следующих соединений: амилоциклин - циклический бактериоцин, обладающий выраженной активностью против грамположительных бактерий, и бацилломицин D - НРПС из семейства итуринов. Кластер синтеза бацилломицина был совмещен с фенгицинсинтетазой, а амилоциклина с бациллибактинсинтетазой. Бовинимид А, которой antiSMASH стабильно находит в штаммах O1.27, O2.11, R6.14, новый алгоритм в этих конкретных позициях «не подтверждает». Это связано с тем, что AFPredictor переаннотировал данные локусы и не нашел коровых генов синтеза, однако, он ошибочно аннотировал бовинимид А у R3.13.

У штаммов *Paenibacillus* были обнаружены новые соединения и предсказан их мономерный состав, на основе которых рассчитаны масса и склонность к кислотному осаждению. Результаты представлены в таблице 6.

Таблица 6. Масса и подверженность кислотному осаждению новых предсказанных соединений.

Штамм	Номер кластера	Предсказанные мономеры	Расчетная масса (Да)	Осаждение при pH 2.0
<i>P. polymyxa</i> R4.24	14	X Val Ile	~340	Нет
<i>P. polymyxa</i> R4.5	13			
<i>P. polymyxa</i> R6.14	8	X mal	~200	Нет
<i>P. polymyxa</i> O2.11	8			
<i>P. polymyxa</i> R3.13	12			
<i>P. polymyxa</i> O1.27	8			
<i>P. polymyxa</i> O1.27	3	fatty_acid mal Tyr D-X Ser D-Ser Ser D-Ile Ser	~1050	Возможно
<i>P. polymyxa</i> R3.13	7			
<i>P. polymyxa</i> O2.11	3			

Штамм	Номер кластера	Предсказанные мономеры	Расчетная масса (Да)	Осаждение при pH 2.0
<i>P. polymyxa</i> R6.14	3			
<i>P. polymyxa</i> R4.24	8			
<i>P. polymyxa</i> R4.5	6	mal	~86	Нет
<i>P. polymyxa</i> R3.13	14			
<i>P. polymyxa</i> O2.11	12	fatty_acid D-Gly D-Dab		
<i>P. polymyxa</i> R6.14	12	D-Gly D-Ser D-Trp Ser	~1256	Возможно
<i>P. polymyxa</i> O1.27	12	Dab D-Dab Ile Glu		
<i>P. polymyxa</i> O2.11	5	fatty_acid D-Asn Gln Me-		
<i>P. polymyxa</i> O1.2	5	ohmal mal X ohmal X Gly	~1200	Возможно
7	5			
<i>P. polymyxa</i> R6.14	5	ccmal Lys X X		
<i>P. polymyxa</i> R3.13	9			

Ни одно из предсказанных новых соединений, аннотированных для штаммов *Paenibacillus*, не должны осаждаться кислотных осаждением, что может затруднить их выделение доступными нам методами; однако для крупных соединений возможно соосаждение. Пик 1257 m/z у штамма O2.11 (см. рисунок 21), потенциально является предсказанным пептидом (D-Gly D-Dab D-Gly D-Ser D-Trp Ser Dab D-Dab Ile Glu), соосажденным при осаждении фузарицидина, что может объяснять то что он встречается только у O2.11. Наличие бацилломицина D у штаммов *Bacillus* объясняет пики 1083 и 1097 m/z (см. рисунки 27 и 28). Разработанный алгоритм, таким образом, на наш взгляд, лучше справился с идентификацией соединений нерибосомальной природы, чем antiSMASH.

3.6. Подтверждение предсказанной противогрибковой активности *in vitro*

Для подтверждения эффективности предсказания алгоритма было выбрано новое соединение с предсказанным мономерным составом Glu Leu D-Leu Ile Asp D-Leu Ile NH₂ D-Val fatty_acid Val, вырабатываемое штаммом *Bacillus*

safensis AG46. Отбор был произведен по следующим критериям: высокая уверенность модели в наличии противогрибковой активности (0,83); анализ мономерного состава показал, что его можно выделить кислотным осаждением; у штамма AG46 это единственный метаболит с противогрибковой активностью и единственный осаждаемый данным методом. Для сравнения был выбран штамм *Bacillus atrophaeus* AG8, продуцент сурфактина, структурно схожего (Glu Leu D-Leu Val Asp D-Leu Ile) соединения с подтвержденной противогрибковой активностью. Сравнение зон подавления фитопатогенного гриба метаболитами AG-8 (сурфактином) и новым метаболитом из AG-46 представлено на рисунке 25.

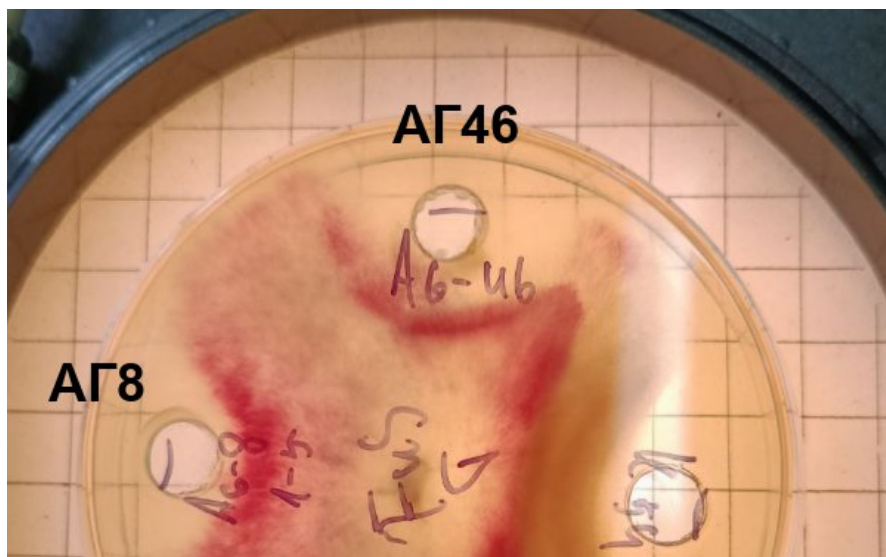


Рисунок 25. Сравнение зон подавления *Fusarium Fus1* (в центре) липопептидными фракциями штаммов AG8 и AG46.

Можно видеть, что оба штамма проявляют сопоставимую противогрибковую активность по отношению к *Fusarium sp. Fus1*.

3.7. Анализ биосинтетического потенциала штаммов-пробиотиков для аквакультуры

Штаммы серии МТ, полученные из микробиоты рыб, предварительно протестировали *in vitro* на способность подавлять рост патогенных оомицетов *Saprolegnia*. Результаты, полученные при оценке антифунгальной активности данных потенциальных пробиотиков, представлены в таблице 7.

Таблица 7. Ширина зоны отсутствия роста гриба *Saprolegnia spp.*

Штамм	MT 14	MT 42	MT 48	MT 55	MT 61	MT 64	MT 74	MT 86	MT 102	MT 141	MT 142	MT 155	MT 186
Зона от- сутствия роста, мм	5	7	7	-	7	4	-	-	5	2	4	-	-

Сравнительный анализ профилей предсказанных биосинтетических генных кластеров (БГК) выявил четкие видоспецифические различия в метаболическом потенциале исследованных штаммов. Было показано, что все проанализированные штаммы (*Bacillus velezensis* MT42, MT14 и *Bacillus subtilis* MT48, MT74) обладают базовым консервативным набором генов, детерминирующих синтез сурфактина, фенгицина, бациллена и бациллибактина.

Детальный анализ на уровне подкластеров позволил выявить четкие паттерны совместной локализации путей биосинтеза специфических соединений внутри одних и тех же локусов. В частности, алгоритм выявил устойчивые пары соединений, которые синтезируются соседствующими подкластерами: у штаммов *B. velezensis* MT42 (BGC_2) и MT14 (BGC_11) противогрибковые липопептиды фенгицин и бацилломицин D оказались локализованы в пределах единых гибридных кластеров. При этом, фенгицинсинтетаза в этих кластерах фрагментирована и фенгицин, возможно, не функционален. В то же время данные показывают, что штаммы *B. velezensis* характеризуются значительно более богатым спектром вторичных метаболитов, где на уровне отдельных подкластеров были дополнительно дифференцированы пути продукции макролактина Н и диффицидина. В качестве штаммоспецифических особенностей алгоритм обнаружил у *B. velezensis* MT42 скрытый подкластер нерибосомальных пептид-синтетаз без известных гомологов в базе данных MiBiG (AFPredictor предсказал отсутствие противогрибковых свойств), а у MT14 — подкластер со слабым уровнем сходства (30%) с путями биосинтеза ракицидина. Метаболические профили подкластеров у штаммов *B. subtilis* MT48 и MT74 оказались полностью синтетичны и ограничены исключительно консервативной базовой группой соединений. При этом, несмотря на явное преимущество *B. velezensis* в генетическом

разнообразии противогрибковых веществ, разницы в размерах зон подавления *Saprolegnia* между видами *B. velezensis* и *B. subtilis* не оказалось. Более того, хотя штаммы MT48 и MT74 имеют абсолютно одинаковый набор БГК, MT74 *Saprolegnia* не подавляет. Вероятно, это связано с разницей в регуляторных генах синтетаз.

3.8. Изучение активности штаммов-продуцентов НРП *in vivo*

Для оценки практической применимости штаммов, отобранных как потенциальные антагонисты грибов на основе анализа *in silico*, проведены опыты по изучению их потенциала как фитостимуляторов и пробиотиков для аквакультуры.

3.9. Изучение стимулирующей рост активность на растениях

В экспериментах на растениях были использованы штаммы AG46 и AG8. Была проведена оценка противогрибковой активности и способности штаммов стимулировать рост ячменя обыкновенного (*Hordeum vulgare*) в условиях повышенного инфекционного фона по фузариуму (более 5%, 1×10^5 конидий/мл). Фотографии проростков представлены на рисунках 26 и 27. Фотографии, отражающие эффекты штаммов-фитостимуляторов на дальнейших стадиях роста растений, представлены на рисунках 28 и 29.

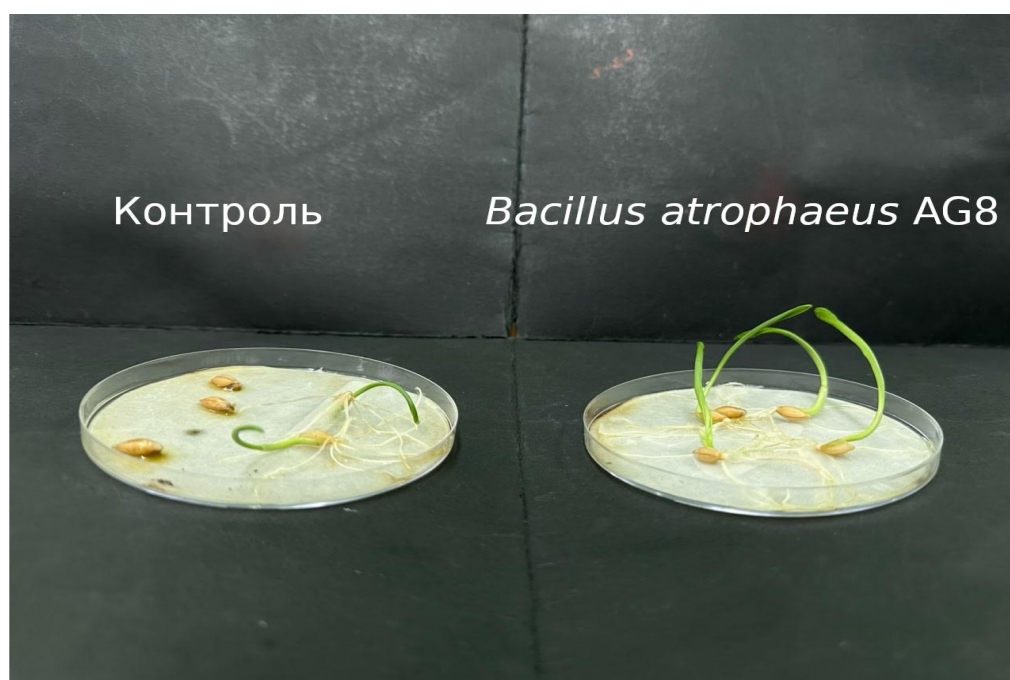


Рисунок 26. Рост семян ячменя в контроле и с обработкой штаммом AG8



Рисунок 27. Рост семян ячменя в контроле и с обработкой штаммом AG46



Рисунок 28. Рост семян ячменя в контейнерах с перлитом в контроле и с обработкой штаммом AG8



Рисунок 29. Рост семян ячменя в контейнерах с перлитом в контроле и с обработкой штаммом AG46

Анализ средней длины стеблей и корней в данном опыте показывает, что штамм 46 стимулирует прорастание семян сильнее, чем штамм 8, хотя оба штамма показывают значимый эффект (рисунки 30 и 31).

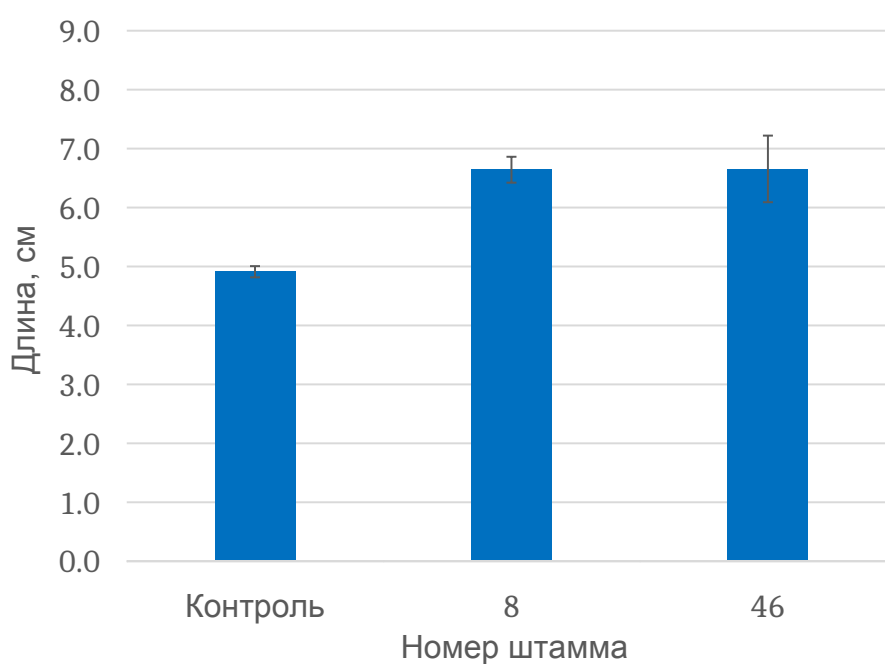


Рисунок 30. Влияние штаммов *Bacillus* на длину стеблей ячменя в чашках Петри.

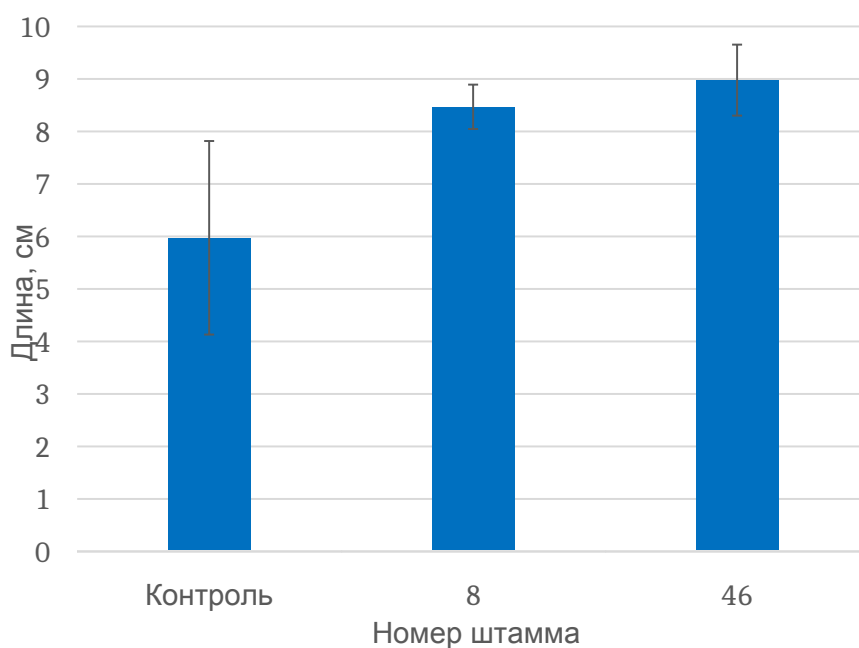


Рисунок 31. Влияние штаммов *Bacillus* на длину корней ячменя в чашках Петри.

Штамм 46 на 10 день опыта увеличивает длину стеблей на 35.5%, а штамм 8 на 35.2%, тогда как длина корней увеличивается на 50.3 и 41.8 %, соответственно. Результаты 14 дня опыта представлены на рисунках 32 и 33.

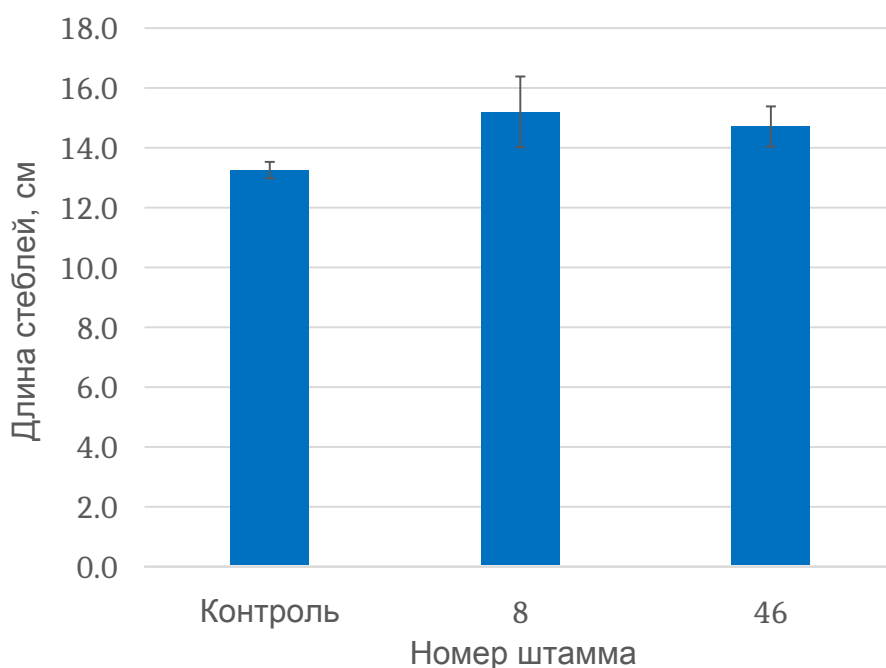


Рисунок 32. Влияние штаммов *Bacillus* на длину стеблей ячменя в контейнерах с перлитом

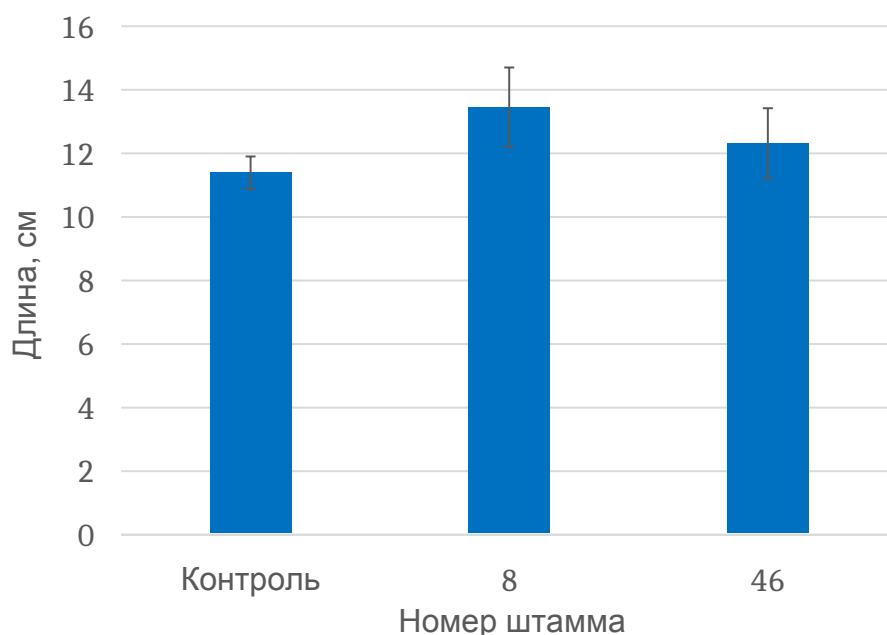


Рисунок 33. Влияние штаммов *Bacillus* на длину корней ячменя в контейнерах с перлитом

В вегетационном опыте влияние бактерий было менее выражено: на 14-й день штамм AG46 увеличил длину стеблей на 10,99%, а штамм AG8 — на 14,68%, при этом значимое увеличение длины корней (на 18,10%) было отмечено только для штамма AG8. Таким образом, фитостимуляция данными бактериями наиболее эффективна именно на ранней стадии прорастания семян. Данные о дифференциальной экспрессии генов иммунного ответа ячменя, обработанного штаммами AG46 и AG8 представлены на рисунке 34.

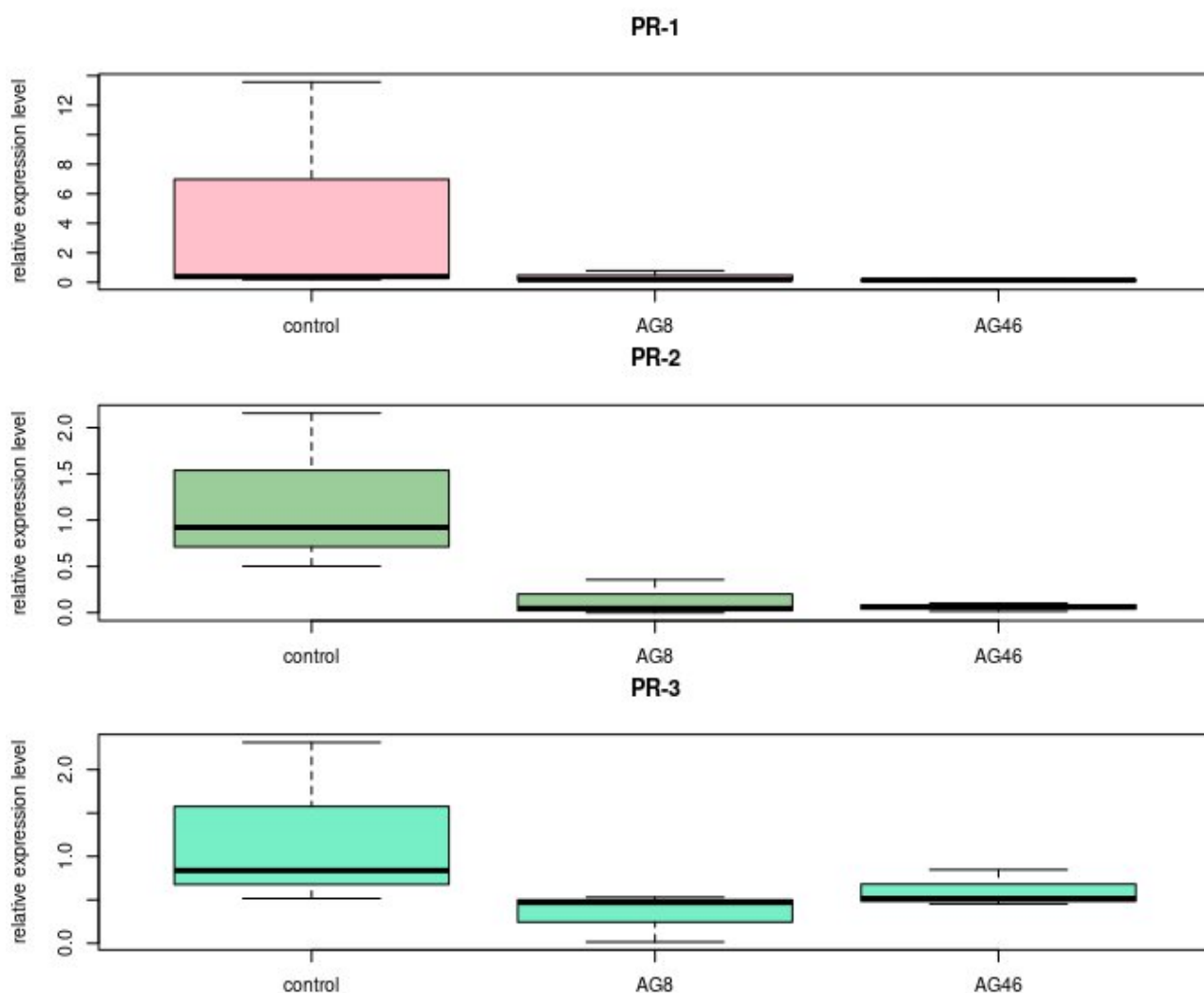


Рисунок 34. Изменение экспрессии генов PR-1, PR-2 и PR-3 ячменя.

Можно видеть, что внесение штаммов AG8 и AG46 вызвало статистически значимое снижение экспрессии генов PR-1 и PR-2 ($p < 0,05$), при этом не оказав значимого влияния на экспрессию гена PR-3. Снижение экспрессии генов PR-1 и PR-2, которые обычно активируются в рамках защитных реакций растений в ответ на воздействие грибковых патогенов, может указывать на изменение физиологического статуса ячменя (снижение необходимости борьбы с фитопатогеном). Такое изменение экспрессии генов базового иммунного ответа напрямую коррелирует с ускоренным прорастанием семян и интенсивным ростом, что обусловлено сдвигом в распределении ресурсов и перестройкой гормональной регуляции на ранних стадиях развития растения в присутствии данных штаммов (Zamioudis, Pieterse, 2012; Zhou и др., 2022; Huot и др., 2014).

3.10. Изучение пробиотической активности на объектах аквакультуры

Была проведена оценка способности штаммов серии МТ подавлять рост *Saprolegnia* – грибкового патогена рыб. Наибольшую активность против *Saprolegnia* spp. имели штаммы *B. velezensis* MT14, *B. velezensis* MT42, *B. subtilis* MT48 и *B. subtilis* MT74 (зоны подавления в тестах *in vitro* составили 5, 7, 7 и 7 мм соответственно). В рамках последующего эксперимента *in vivo* стерлядь получала корма с включением данных бактерий. Опытная группа 1 получала препарат на основе *B. velezensis* MT14 и MT42 ($5,2 \cdot 10^6$ КОЕ/г), опытная группа 2 — препарат на основе *B. subtilis* MT48 и MT74 ($6,1 \cdot 10^6$ КОЕ/г). Для оценки того, как применение различных пробиотических комплексов отражается на физиологическом состоянии и иммунном статусе рыб, мы проанализировали уровень транскрипции ключевых маркерных генов. Данные о дифференциальной экспрессии представлены на рисунке 35.

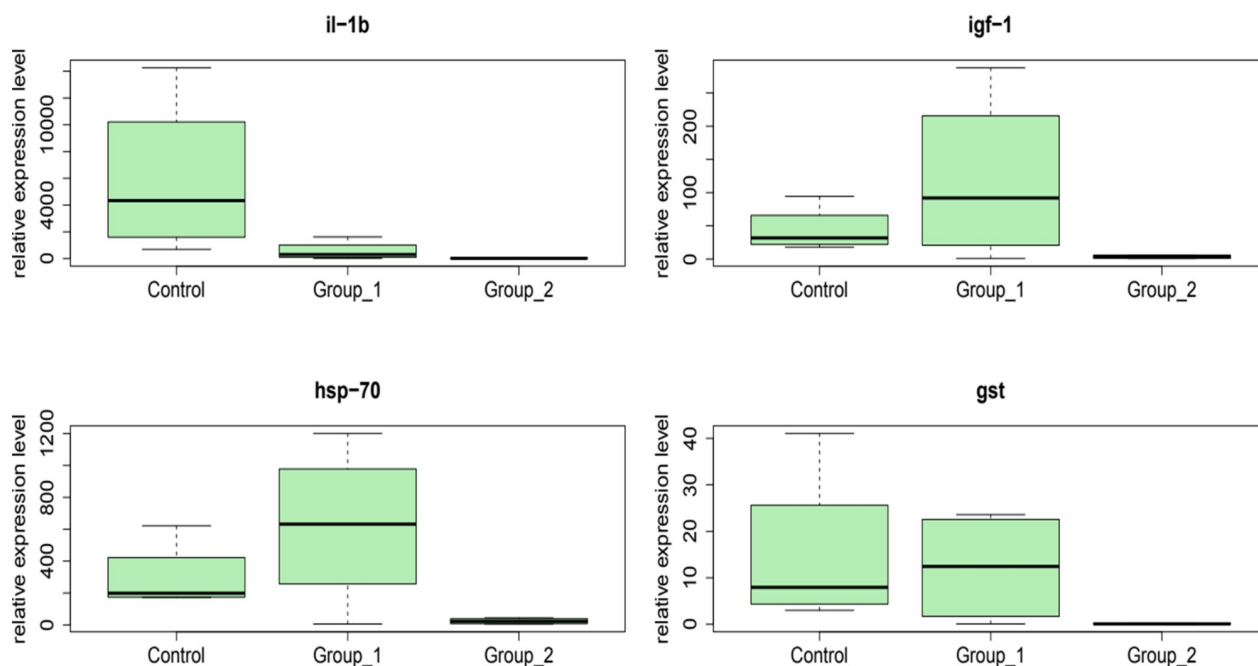


Рисунок 35. Экспрессия генов IL-1 β , gst, igf-1 и hsp70.

В группе 2 наблюдается статистически достоверное снижение экспрессии генов, IL-1 β , gst, igf-1 и hsp70 ($p < 0,05$). Группа 1 не показала статистически достоверного отличия от контроля.

Наблюдаемое снижение экспрессии гена IL-1 β может указывать на уменьшение воспалительных процессов у рыб, получающих пробиотики, что может быть связано с противогрибковой активностью штаммов MT48 и MT74. Ген IL-1 β кодирует интерлейкин-1 бета (IL-1 β), который действует как провоспалительный цитокин, усиливающий активацию иммунитета и аутоиммунные риски при гиперэкспрессии.

Разница в экспрессии между препаратами может быть обусловлена механизмами противогрибкового действия: у штаммов группы 2 (MT48 и MT74) БГК фенгицинсинтетазы содержит только саму фенгицинсинтазу, у группы 1 (MT14 и MT42) гомологичный БГК содержит поврежденную фенгицинсинтазу и синтазу бацилломицина D и его противогрибковая активность обеспечивается именно бацилломицином. Хотя бацилломицин D эффективно разрушает грибковые клетки за счет связывания с эргостеролом, он также способен неспецифично взаимодействовать с холестерином животных клеток (Nasir, Besson, 2012; Quentin и др., 1982). Это приводит к их повреждению, что и объясняет наблюдаемое влияние на экспрессию исследованных генов.

ЗАКЛЮЧЕНИЕ

В рамках исследования была создана предиктивная модель машинного обучения, обеспечивающая прогнозирование противогрибковой активности нерибосомальных пептидов на основе геномных данных с точностью до 81%. С ее помощью был выделен ряд новых характерных особенностей противогрибковых метаболитов, позволяющих предсказать их функциональную активность.

Так, показано, что только 38 мономеров из 51, встречающихся в структуре НРП, ассоциированы с противогрибковой активностью; в частности, аминокислоты Leu, Glu и Gln статистически значимо чаще встречаются среди противогрибковых соединений. С помощью анализа векторов Шепли (SHAP) доказано, что критически важным маркером фунгицидной активности является совместная встречаемость стартового домена конденсации с алифатическими и разветвленными гидрофобными аминокислотами, тогда как присутствие MbtH-подобных белков и транспортных доменов эффлюксных помп классифицируется как строгий предиктор отсутствия противогрибковых свойств.

Обоснована и доказана эффективность нового подхода к биоинформатическому скринингу геномов не на уровне кластеров, а на субкластерном уровне. Применение разработанного алгоритма позволило выявить на 104 биосинтетических генных кластера больше, чем при использовании традиционных методов (в частности, платформы antiSMASH), а также обнаружить скрытые паттерны совместной локализации множественных синтетаз в рамках сложносоставных кластеров. Ко-локализация путей синтеза бацилломицина D и фенгицина у *Bacillus velezensis* была успешно верифицирована масс-спектрометрическим анализом. Более того, алгоритм позволил предсказать, а эксперимент — подтвердить синтез нового противогрибкового метаболита: у штамма *Bacillus safensis* AG46 был идентифицирован новый липопептид, способный к кислотному осаждению. Тестирование данного соединения *in vitro* продемонстрировало его высокую антагонистическую активность против фитопатогена *Fusarium* Fus1.

Наряду с исследованиями *in vitro*, был подтвержден выраженный фитостимулирующий эффект бактерий-антагонистов, продуцирующих предсказанные противогрибковые вещества, *in vivo*. Так, обработка семян ячменя (*Hordeum vulgare*) штаммами AG46 и AG8 привела к статистически значимому увеличению длины корней и стеблей на стадии прорастания. На молекулярном уровне этот процесс сопровождался достоверным снижением экспрессии генов базового иммунного ответа (PR-1 и PR-2), что свидетельствует о физиологическом перераспределении ресурсов растения от защитных реакций в пользу активного вегетативного роста. Также был доказан пробиотический потенциал отобранных штаммов в аквакультуре и раскрыт механизм их взаимодействия с клетками хозяина. Применение штаммов *Bacillus subtilis* MT48 и MT74 способствовало улучшению функционального состояния стерляди (*Acipenser ruthenus*) при субкритическом уровне фоновой инфекции *Saprolegnia* spp., достоверно снизив транскрипцию маркеров воспаления и клеточного стресса IL-1 β , gst, igf-1, hsp70. Выдвинута гипотеза, что отсутствие аналогичного противовоспалительного эффекта у штаммов *Bacillus velezensis* MT14 и MT42 может быть связано с присутствием в их метаболитном профиле бацилломицина D. В отличие от фенгицина, бацилломицин D способен неспецифично связываться с холестерином мембран животных клеток, что приводит к индукции стрессового ответа.

ВЫВОДЫ

- 1) Сформирована база данных из 623 известных нерибосомальных пептидов с подтвержденной противогрибковой активностью и кластеров генов связанных с ними синтетаз.
- 2) Разработана модель машинного обучения для предсказания противогрибковых свойств у нерибосомально синтезируемых метаболитов бактерий на основе геномных данных. С использованием модели определены признаки-предикторы противогрибковой активности. Значительная часть этих признаков связана с мембранолитическим механизмом противогрибкового действия нерибосомально синтезируемых соединений.
- 3) Проведен отбор и полногеномное секвенирование 69 штаммов бактерий, проявляющих антагонистическую активность против грибов, патогенных для растений и животных, родов *Fusarium*, а также осуществлен их биохимический и генетический анализ.
- 4) Выполнена валидация разработанного алгоритма на целевой группе 69 штаммов-продуцентов, обладающих подтвержденной противогрибковой активностью *in vivo*. Показано, что тридекаптин, пенилогенптин, тетракарцин и аурантинин, могут проявлять противогрибковые свойства, ранее для них не описанные. Обнаружен новый противогрибковый лихенизин-подобный метаболит, вырабатываемый штаммом AG-46, а его активность против фитопатогенного *Fusarium* подтверждена *in vitro*.
- 5) Штамм AG-46 характеризуется как фитостимулятор в опытах с ячменем, а штаммы MT48 и MT74 показали потенциал как пробиотики для объектов аквакультуры, что подтверждает практическую значимость предсказаний, сделанных моделью машинного обучения *in silico*.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Васильченко Н.Г. Роль генов нерибосомальных пептидсинтетаз аэробных спорообразующих бактерий в развитии антагонизма против грибов рода *Fusarium* : кандидат наук / Васильченко Н.Г. – ФГАОУ ВО «Южный федеральный университет», 2023. – URL: <https://dissercat.com/content/rol-genov-neribosomalnykh-peptidsintetaz-aerobnykh-sporoobrazuyushchikh-bakterii-v-razviti> (дата обращения: 09.07.2026). – Текст : электронный.
2. Возможности биodeградации микотоксинов, образуемых грибами рода *Fusarium* / Т. Ю. Гагкаева, О. П. Гаврилова, И. Н. Никонов, Г. Ю. Лаптев. – 2016. – С. 202-238.
3. Сидорова, Т. М. БИОЛОГИЧЕСКИ АКТИВНЫЕ МЕТАБОЛИТЫ *Bacillus subtilis* И ИХ РОЛЬ В КОНТРОЛЕ ФИТОПАТОГЕННЫХ МИКРООРГАНИЗМОВ / Т. М. Сидорова, А. М. Асатунова, А. И. Хомяк // Сельскохозяйственная Биология. – 2018. – Т. 53. – № 1. – С. 29-37.
4. Соколова, Г. Д. Механизмы устойчивости к фунгицидам фитопатогенного гриба *Fusarium graminearum* / Г. Д. Соколова, А. П. Глинушкин // Микология И Фитопатология. – 2020. – Т. 54. – № 6. – С. 391-403.
5. Сравнительный геномный, транскриптомный и протеомный анализ штамма *Limosilactobacillus fermentum* U-21, перспективного для создания фармабиотика / Е. У. Полуэктова, Д. А. Мавлетова, М. В. Одорская и др. // Генетика. – 2022. – Т. 58. – № 9. – С. 1029-1041.
6. Ульцеративный дермальный некроз лососевых: обзор / Д. С. Провоторов, С. А. Мурзина, Е. П. Иешко, Н. Н. Немова (Калинина) // Журнал Общей Биологии. – 2022. – Т. 83. – Ульцеративный дермальный некроз лососевых. – № 2. – С. 138-148.
7. Щеклеина, Л. М. ВРЕДНОСТЬ ФУЗАРИОЗА КОЛОСА ЯРОВОЙ ПШЕНИЦЫ И ПОИСК УСТОЙЧИВЫХ ГЕНОТИПОВ ДЛЯ СЕЛЕКЦИИ / Л. М. Щеклеина, Т. К. Шешегова // Пермский аграрный вестник. – 2024. – № 3 (47). – С. 64-70.
8. Эффективность применения хлорида натрия для обработки икры русского осетра (*Acipenser gueldenstaedtii*) от сапролегниоза / Р. Р. Тангатарова, О. В. Пятикопова, И. Н. Бедрицкая, Ю. Н. Грозеску // Вопросы рыболовства. – 2024. – Т. 25. – № 3. – С. 163-168.
9. A strategy for the identification of patterns in the biosynthesis of nonribosomal peptides by Betaproteobacteria species / J. L. Baldim, B. L. da Silva, D. A. Chagas-Paula и др. // Scientific Reports. – 2017. – Т. 7. – С. 10400.

10. Abh, AbrB3, and Spo0A play distinct regulatory roles during polymyxin synthesis in *Paenibacillus polymyxa* SC2 / Y. Cui, D. Zhao, K. Liu и др. // *Microbiology Spectrum*. – 2023. – Т. 12. – № 1. – С. e02293-23.
11. Action of peptidolipidic antibiotics of the iturin group on erythrocytes. Effect of some lipids on hemolysis / M. J. Quentin, F. Besson, F. Peypoux, G. Michel // *Biochimica Et Biophysica Acta*. – 1982. – Т. 684. – № 2. – С. 207-211.
12. Activation of the Pacidamycin PacL Adenylation Domain by MbtH-Like Proteins / W. Zhang, J. R. Heemstra, C. T. Walsh, H. J. Imker // *Biochemistry*. – 2010. – Т. 49. – № 46. – С. 9946-9947.
13. Additive Effect of the Composition of Endophytic Bacteria *Bacillus subtilis* on Systemic Resistance of Wheat against Greenbug Aphid *Schizaphis graminum* Due to Lipopeptides / S. D. Rummyantsev, V. Y. Alekseev, A. V. Sorokan et al. // *Life*. – 2023. – Vol. 13. – № 1. – P. 214.
14. AdenPredictor: accurate prediction of the adenylation domain specificity of nonribosomal peptide biosynthetic gene clusters in microbial genomes / M. Mongia, R. Baral, A. Adduri и др. // *Bioinformatics*. – 2023. – Т. 39. – AdenPredictor. – № Suppl 1. – С. i40-i46.
15. Aleti, G. Secondary metabolite biosynthetic gene diversity - Bacillales non-ribosomal peptides and polyketides in plant-microbe interactions / G. Aleti. – 2016.
16. Alonzo, D. A. Biosynthesis of depsipeptides, or Depsi: The peptides with varied generations / D. A. Alonzo, T. M. Schmeing. – Text: electronic. – 2020. – Biosynthesis of depsipeptides, or Depsi. – URL: <https://onlinelibrary.wiley.com/doi/10.1002/pro.3979> (date accessed: 27.08.2025).
17. Aminov, R. History of antimicrobial drug discovery: Major classes and health impact / R. Aminov // *Biochemical Pharmacology*. – 2017. – Т. 133. – History of antimicrobial drug discovery. – С. 4-19.
18. An antibiotic fusaricidin: a cyclic depsipeptide from *Paenibacillus polymyxa* E681 induces systemic resistance against *Phytophthora* blight of red-pepper / S. H. Lee, Y. E. Cho, S.-H. Park et al. // *Phytoparasitica*. – 2013. – Vol. 41. – An antibiotic fusaricidin. – № 1. – P. 49-58.
19. An Integrated Pipeline and Overexpression of a Novel Efflux Transporter, YoeA, Significantly Increases Plipastatin Production in *Bacillus subtilis* / M. Wang, J. Zheng, S. Sun et al. // *Foods*. – 2024. – Vol. 13. – № 11. – P. 1785.
20. An Orphan MbtH-Like Protein Interacts with Multiple Nonribosomal Peptide Synthetases in *Myxococcus xanthus* DK1622 / K. J. Esquilín-Lebrón, T. O. Boynton, L. J. Shimkets, M. G. Thomas // *Journal of Bacteriology*. – 2018. – Т. 200. – № 21. – С. 10.1128/jb.00346-18.

21. Antifungal Activities of *Bacillus subtilis* Lipopeptides to Two *Venturia inaequalis* Strains Possessing Different Tebuconazole Sensitivity / H. Desmyttere, C. Deweer, J. Muchembled и др. // *Frontiers in Microbiology*. – 2019. – Т. 10. – С. 2327.
22. Antimicrobial peptides of the genus *Bacillus*: a new era for antibiotics / C. D. Sumi, B. W. Yang, I.-C. Yeo, Y. T. Hahm // *Canadian Journal of Microbiology*. – 2015. – Т. 61. – Antimicrobial peptides of the genus *Bacillus*. – № 2. – С. 93-103.
23. antiSMASH 6.0: improving cluster detection and comparison capabilities / K. Blin, S. Shaw, A. M. Kloosterman и др. // *Nucleic Acids Research*. – 2021. – Т. 49. – antiSMASH 6.0. – № W1. – С. W29-W35.
24. antiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation / K. Blin, S. Shaw, L. Vader и др. // *Nucleic Acids Research*. – 2025. – Т. 53. – antiSMASH 8.0. – № W1. – С. W32-W38.
25. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences / M. H. Medema, K. Blin, P. Cimermancic и др. // *Nucleic Acids Research*. – 2011. – Т. 39. – antiSMASH. – № suppl_2. – С. W339-W346.
26. Application of Psychrotolerant Antarctic Bacteria and Their Metabolites as Efficient Plant Growth Promoting Agents / M. Styczynski, G. Bieganski, P. Decewicz и др. // *Frontiers in Bioengineering and Biotechnology*. – 2022. – Т. 10. – С. 772891.
27. Assembling Single-Cell Genomes and Mini-Metagenomes From Chimeric MDA Products / S. Nurk, A. Bankevich, D. Antipov и др. // *Journal of Computational Biology*. – 2013. – Т. 20. – № 10. – С. 714-737.
28. Assembly of long, error-prone reads using repeat graphs / M. Kolmogorov, J. Yuan, Y. Lin, P. A. Pevzner // *Nature Biotechnology*. – 2019. – Vol. 37. – № 5. – P. 540-546.
29. Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes / H. Wang, D. P. Fewer, L. Holm и др. // *Proceedings of the National Academy of Sciences*. – 2014. – Т. 111. – № 25. – С. 9259-9264.
30. Bacillaenes: Decomposition Trigger Point and Biofilm Enhancement in *Bacillus* / H. Li, X. Han, Y. Dong и др. // *ACS Omega*. – 2021. – Т. 6. – Bacillaenes. – № 2. – С. 1093-1098.
31. Biocontrol and plant growth promotion potential of endophytic *Bacillus subtilis* JY-7-2L on *Aconitum carmichaelii* Debx / L. Zou, Q. Wang, R. Wu и др. // *Frontiers in Microbiology*. – 2022. – Т. 13. – С. 1059549.

32. Bioinformatic Analysis Reveals both Oversampled and Underexplored Biosynthetic Diversity in Nonribosomal Peptides / B.-S. Jian, S.-L. Chiou, C.-C. Hsu и др. // *ACS Chemical Biology*. – 2023. – Т. 18. – № 3. – С. 476-483.
33. Biosynthesis of HSAF, a Tetramic Acid-Containing Macrolactam from *Lysobacter enzymogenes* / L. Lou, G. Qian, Y. Xie и др. // *Journal of the American Chemical Society*. – 2011. – Т. 133. – № 4. – С. 643-645.
34. Biosynthetic Pathway for Mannopeptimycins, Lipoglycopeptide Antibiotics Active against Drug-Resistant Gram-Positive Pathogens / N. A. Magarvey, B. Haltli, M. He и др. // *Antimicrobial Agents and Chemotherapy*. – 2006. – Т. 50. – № 6. – С. 2167-2177.
35. Biotechnological application of *Streptomyces* for the production of clinical drugs and other bioactive molecules / F. Del Carratore, E. K. Hanko, R. Breitling, E. Takano // *Current Opinion in Biotechnology*. – 2022. – Т. 77. – С. 102762.
36. Broad-Spectrum Gramicidin S Derivatives with Potent Activity Against Multidrug-Resistant Gram-Negative ESKAPE Pathogens / J. T. Kalyvas, Y. Wang, O. Romeo et al. // *Antibiotics*. – 2025. – Vol. 14. – № 5. – P. 423.
37. Cai, X. Engineering of Specific Single-Module Nonribosomal Peptide Synthetases of the RXP Type for the Production of Defined Peptides / X. Cai, L. Zhao, H. B. Bode // *ACS Synthetic Biology*. – 2023. – Т. 12. – № 1. – С. 203-212.
38. Cane, D. E. Harnessing the Biosynthetic Code: Combinations, Permutations, and Mutations / D. E. Cane, C. T. Walsh, C. Khosla // *Science*. – 1998. – Т. 282. – Harnessing the Biosynthetic Code. – № 5386. – С. 63-68.
39. Catalytic trajectory of a dimeric nonribosomal peptide synthetase subunit with an inserted epimerase domain / J. Wang, D. Li, L. Chen et al. // *Nature Communications*. – 2022. – Vol. 13. – № 1. – P. 592.
40. Cereulide synthetase gene cluster from emetic *Bacillus cereus*: Structure and location on a mega virulence plasmid related to *Bacillus anthracis* toxin plasmid pXO1 / M. Ehling-Schulz, M. Fricker, H. Grallert и др. // *BMC Microbiology*. – 2006. – Т. 6. – Cereulide synthetase gene cluster from emetic *Bacillus cereus*. – С. 20.
41. Challis, G. L. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains / G. L. Challis, J. Ravel, C. A. Townsend // *Chemistry & Biology*. – 2000. – Т. 7. – № 3. – С. 211-224.
42. Challis, G. L. Structural aspects of non-ribosomal peptide biosynthesis / G. L. Challis, J. H. Naismith // *Current Opinion in Structural Biology*. – 2004. – Т. 14. – № 6. – С. 748-756.

43. Characterization of an iron-dependent regulatory protein (IdeR) of *Mycobacterium tuberculosis* as a functional homolog of the diphtheria toxin repressor (DtxR) from *Corynebacterium diphtheriae* / M. P. Schmitt, M. Predich, L. Doukhan и др. // *Infection and Immunity*. – 1995. – Т. 63. – № 11. – С. 4284-4289.
44. Chemical Logic of Peptide Branching by Iterative Nonlinear Nonribosomal Peptide Synthetases / J. Yang, A. Balutowski, M. Trivedi, T. A. Wencewicz // *Biochemistry*. – 2025. – Т. 64. – № 3. – С. 719-734.
45. Communication Breakdown: Dissecting the COM Interfaces between the Subunits of Nonribosomal Peptide Synthetases / C. D. Fage, S. Kosol, M. Jenner и др. // *ACS Catalysis*. – 2021. – Т. 11. – Communication Breakdown. – № 17. – С. 10802-10813.
46. Controlling Substrate- and Stereospecificity of Condensation Domains in Nonribosomal Peptide Synthetases / H. Peng, J. Schmiederer, X. Chen и др. // *ACS Chemical Biology*. – 2024. – Т. 19. – № 3. – С. 599-606.
47. Copper Stress Affects Iron Homeostasis by Destabilizing Iron-Sulfur Cluster Formation in *Bacillus subtilis* / S. Chillappagari, A. Seubert, H. Trip и др. // *Journal of Bacteriology*. – 2010. – Т. 192. – № 10. – С. 2512-2524.
48. Corpuz, J. C. Protein-protein interface analysis of the non-ribosomal peptide synthetase peptidyl carrier protein and enzymatic domains / J. C. Corpuz, J. O. Sanley, M. D. Burkart // *Synthetic and Systems Biotechnology*. – 2022. – Т. 7. – № 2. – С. 677-688.
49. Cyanobacterial Abundance and Microcystin Profiles in Two Southern British Lakes: The Importance of Abiotic and Biotic Interactions / D. M. Hartnell, I. J. Chapman, N. G. H. Taylor и др. // *Toxins*. – 2020. – Т. 12. – Cyanobacterial Abundance and Microcystin Profiles in Two Southern British Lakes. – № 8. – С. 503.
50. Design, Synthesis, and Biophysical Evaluation of Mechanism-Based Probes for Condensation Domains of Nonribosomal Peptide Synthetases / C. Shi, B. R. Miller, E. M. Alexander и др. // *ACS Chemical Biology*. – 2020. – Т. 15. – № 7. – С. 1813-1819.
51. Direct α -Hydroxy Acid Loading onto a Bacterial Thiotemplate Assembly Line via a Multienzyme Gateway / J. Fiedler, F. Trottmann, K. Ishida и др. // *Angewandte Chemie*. – 2024. – Т. 63. – № 30. – С. e202405165.
52. Discovery of cationic nonribosomal peptides as Gram-negative antibiotics through global genome mining / Y.-X. Li, Z. Zhong, W.-P. Zhang, P.-Y. Qian // *Nature Communications*. – 2018. – Т. 9. – С. 3273.
53. Diversity of Monomers in Nonribosomal Peptides: towards the Prediction of Origin and Biological Activity / S. Caboche, V. Leclère, M. Pupin и др. // *Journal of*

Bacteriology. – 2010. – Т. 192. – Diversity of Monomers in Nonribosomal Peptides. – № 19. – С. 5143-5150.

54. Doyle, J. J. A rapid DNA isolation procedure for small quantities of fresh leaf tissue / J. J. Doyle, J. L. Doyle. – Текст : электронный // PHYTOCHEMICAL BULLETIN. – 1987. – URL: <https://worldveg.tind.io/record/33886> (дата обращения: 12.07.2026).

55. Duban, M. Nonribosomal Peptide Synthesis Definitely Working Out of the Rules / M. Duban, S. Cociancich, V. Leclère // Microorganisms. – 2022. – Vol. 10. – № 3. – P. 577.

56. Effects of fengycin from *Bacillus subtilis* fmbJ on apoptosis and necrosis in *Rhizopus stolonifer* / Q. Tang, X. Bie, Z. Lu et al. // Journal of Microbiology. – 2014. – Vol. 52. – № 8. – P. 675-680.

57. Engineering and elucidation of the lipoinitiation process in nonribosomal peptide biosynthesis / L. Zhong, X. Diao, N. Zhang и др. // Nature Communications. – 2021. – Т. 12. – С. 296.

58. Evaluation of Bio-Friendly Formulations from Siderophore-Producing Fluorescent *Pseudomonas* as Biocontrol Agents for the Management of Soil-Borne Fungi, *Fusarium oxysporum* and *Rhizoctonia solani* / G. A. Abo-Zaid, A. S. Abdullah, N. A.-M. Soliman et al. // Agriculture. – 2023. – Vol. 13. – № 7. – P. 1418.

59. Expression of degQ gene and its effect on lipopeptide production as well as formation of secretory proteases in *Bacillus subtilis* strains / L. Lilge, M. Vahidinasab, I. Adiek et al. // MicrobiologyOpen. – 2021. – Vol. 10. – № 5. – P. e1241.

60. Foliar Application of Iron Fortified Bacteriosiderophore Improves Growth and Grain Fe Concentration in Wheat and Soybean / S. Sharma, S. Chandra, A. Kumar и др. // Indian Journal of Microbiology. – 2019. – Т. 59. – № 3. – С. 344-350.

61. Formylation Domain: An Essential Modifying Enzyme for the Nonribosomal Biosynthesis of Linear Gramicidin / G. Schoenafinger, N. Schracke, U. Linne, M. A. Marahiel // Journal of the American Chemical Society. – 2006. – Т. 128. – Formylation Domain. – № 23. – С. 7406-7407.

62. Functional Diversity and Engineering of the Adenylation Domains in Nonribosomal Peptide Synthetases / M. Zhang, Z. Peng, Z. Huang et al. // Marine Drugs. – 2024. – Vol. 22. – № 8. – P. 349.

63. Fundamentals of bio-based technologies for selective metal recovery from bio-leachates and liquid waste streams / A. Sieber, S. Spiess, W. Y. Rassy и др. // Frontiers in Bioengineering and Biotechnology. – 2024. – Т. 12. – С. 1528992.

64. Gaudelli, N. M. β -Lactam formation by a non-ribosomal peptide synthetase during antibiotic biosynthesis / N. M. Gaudelli, D. H. Long, C. A. Townsend // *Nature*. – 2015. – T. 520. – № 7547. – C. 383-387.
65. Genome-based analysis of non-ribosomal peptide synthetase and type-I polyketide synthase gene clusters in all type strains of the genus *Herbidospora* / H. Komaki, N. Ichikawa, A. Oguchi и др. // *BMC Research Notes*. – 2015. – T. 8. – № 1. – C. 548.
66. Genome-Wide Identification and Evolutionary Analyses of SrfA Operon Genes in *Bacillus* / Y. Xu, J.-Y. Wu, Q.-J. Liu, J.-Y. Xue // *Genes*. – 2023. – Vol. 14. – № 2. – P. 422.
67. Growth–Defense Tradeoffs in Plants: A Balancing Act to Optimize Fitness / B. Huot, J. Yao, B. L. Montgomery, S. Y. He // *Molecular Plant*. – 2014. – T. 7. – Growth–Defense Tradeoffs in Plants. – № 8. – C. 1267-1287.
68. GTDB-Tk v2: memory friendly classification with the genome taxonomy database / P.-A. Chaumeil, A. J. Mussig, P. Hugenholtz, D. H. Parks // *Bioinformatics*. – 2022. – T. 38. – GTDB-Tk v2. – № 23. – C. 5315-5316.
69. Gulick, A. M. Nonribosomal peptide synthetase biosynthetic clusters of ESKAPE pathogens / A. M. Gulick // *Natural Product Reports*. – 2017. – T. 34. – № 8. – C. 981-1009.
70. Hahn, M. Selective interaction between nonribosomal peptide synthetases is facilitated by short communication-mediating domains / M. Hahn, T. Stachelhaus // *Proceedings of the National Academy of Sciences of the United States of America*. – 2004. – T. 101. – № 44. – C. 15585-15590.
71. Herbicidal secondary metabolites from *Bacillus velezensis* JTB8-2 against *Orobancha aegyptiaca* / W. He, Y. Li, W. Luo et al. // *AMB Express*. – 2022. – Vol. 12. – № 1. – P. 52.
72. High-throughput reprogramming of an NRPS condensation domain / I. B. Folger, N. F. Frota, A. Pistofidis et al. // *Nature Chemical Biology*. – 2024. – Vol. 20. – № 6. – P. 761-769.
73. Hoyer, K. M. The Iterative Gramicidin S Thioesterase Catalyzes Peptide Ligation and Cyclization / K. M. Hoyer, C. Mahlert, M. A. Marahiel // *Chemistry & Biology*. – 2007. – T. 14. – № 1. – C. 13-22.
74. Hu, F. Rational strain improvement for surfactin production: enhancing the yield and generating novel structures / F. Hu, Y. Liu, S. Li // *Microbial Cell Factories*. – 2019. – T. 18. – Rational strain improvement for surfactin production. – C. 42.

75. Hur, G. H. Explorations of Catalytic Domains in Non-Ribosomal Peptide Synthetase Enzymology / G. H. Hur, C. R. Vickery, M. D. Burkart // Natural product reports. – 2012. – T. 29. – № 10. – C. 1074-1098.
76. Hutchison, M. L. Lipopeptide Phytotoxins Produced by *Pseudomonas syringae* pv. *syringae*: Comparison of the Biosurfactant and Ion Channel-Forming Activities of Syringopeptin and Syringomycin / M. L. Hutchison, D. C. Gross // Molecular Plant-Microbe Interactions®. – 1997. – T. 10. – Lipopeptide Phytotoxins Produced by *Pseudomonas syringae* pv. *syringae*. – № 3. – C. 347-354.
77. Iacovelli, R. Nonribosomal peptide synthetases and their biotechnological potential in *Penicillium rubens* / R. Iacovelli, R. A. L. Bovenberg, A. J. M. Driessen // Journal of Industrial Microbiology & Biotechnology. – 2021. – T. 48. – № 9-10. – C. kuab045.
78. Identification of Key Amino Acids in the A Domains of Polymyxin Synthetase Responsible for 2,4-Diaminobutyric Acid Adenylation in *Paenibacillus polymyxa* NBRC3020 Strain / M. Nemoto, W. Ando, T. Mano и др. // ACS Chemical Biology. – 2025. – T. 20. – № 2. – C. 321-331.
79. Impact of non-proteinogenic amino acids in the discovery and development of peptide therapeutics / Y. Ding, J. P. Ting, J. Liu et al. // Amino Acids. – 2020. – Vol. 52. – № 9. – P. 1207-1226.
80. In Tuber Biocontrol of Potato Late Blight by a Collection of Phenazine-1-Carboxylic Acid-Producing *Pseudomonas* spp. / G. Léger, A. Novinscak, A. Biessy et al. // Microorganisms. – 2021. – Vol. 9. – № 12. – P. 2525.
81. Induction of cell wall thickening by the antifungal compound dihydromaltophilin disrupts fungal growth and is mediated by sphingolipid biosynthesis / S. Li, A. M. Calvo, G. Y. Yuen и др. // The Journal of Eukaryotic Microbiology. – 2009. – T. 56. – № 2. – C. 182-187.
82. Insights into the antifungal mechanism of *Bacillus subtilis* cyclic lipopeptide iturin A mediated by potassium ion channel / S. Yang, Y. Ji, P. Xue и др. // International Journal of Biological Macromolecules. – 2024. – T. 277. – C. 134306.
83. Ishii, H. The *Bacillus subtilis* Response Regulator Gene *degU* Is Positively Regulated by CcpA and by Catabolite-Repressed Synthesis of ClpC / H. Ishii, T. Tanaka, M. Ogura // Journal of Bacteriology. – 2013. – T. 195. – № 2. – C. 193-201.
84. Karpiński, T. M. Anticancer Activity of Bacterial Proteins and Peptides / T. M. Karpiński, A. Adameczak // Pharmaceutics. – 2018. – Vol. 10. – № 2. – P. 54.
85. Kaspar, F. Bioactive Secondary Metabolites from *Bacillus subtilis*: A Comprehensive Review / F. Kaspar, P. Neubauer, M. Gimpel // Journal of Natural

Products. – 2019. – T. 82. – Bioactive Secondary Metabolites from *Bacillus subtilis*. – № 7. – C. 2038-2053.

86. Khabthani, S. In Silico/In Vitro Strategies Leading to the Discovery of New Nonribosomal Peptide and Polyketide Antibiotics Active against Human Pathogens / S. Khabthani, J.-M. Rolain, V. Merhej // *Microorganisms*. – 2021. – T. 9. – № 11. – C. 2297.

87. Kleinkauf, H. INTERRELATION BETWEEN ACTIVATION AND POLYMERIZATION IN GRAMICIDIN S BIOSYNTHESIS* / H. Kleinkauf, W. Gevers, F. Lipmann // *Proceedings of the National Academy of Sciences of the United States of America*. – 1969. – T. 62. – № 1. – C. 226-233.

88. Kohli, R. M. The thioesterase domain from a nonribosomal peptide synthetase as a cyclization catalyst for integrin binding peptides / R. M. Kohli, J. Takagi, C. T. Walsh // *Proceedings of the National Academy of Sciences of the United States of America*. – 2002. – T. 99. – № 3. – C. 1247-1252.

89. Konz, D. How do peptide synthetases generate structural diversity? / D. Konz, M. A. Marahiel // *Chemistry & Biology*. – 1999. – T. 6. – № 2. – C. R39-R48.

90. Kooner, A. Draft Genome Sequence of *Pseudomonas* sp. Strain MWU13-3659, Isolated from Commercial Cranberry Bog Soil in Massachusetts, USA / A. Kooner, S. Soby // *Microbiology Resource Announcements*. – 2022. – T. 11. – № 11. – C. e00880-22.

91. Kotowska, M. Roles of type II thioesterases and their application for secondary metabolite yield improvement / M. Kotowska, K. Pawlik // *Applied Microbiology and Biotechnology*. – 2014. – T. 98. – № 18. – C. 7735-7746.

92. Krättschmar, J. Gramicidin S biosynthesis operon containing the structural genes *grsA* and *grsB* has an open reading frame encoding a protein homologous to fatty acid thioesterases / J. Krättschmar, M. Krause, M. A. Marahiel // *Journal of Bacteriology*. – 1989. – T. 171. – № 10. – C. 5422-5429.

93. Kries, H. Biosynthetic engineering of nonribosomal peptide synthetases / H. Kries // *Journal of peptide science*. – 2016. – T. 22. – № 9. – C. 564-570.

94. Lewis, K. Platforms for antibiotic discovery / K. Lewis // *Nature Reviews. Drug Discovery*. – 2013. – T. 12. – № 5. – C. 371-387.

95. Linalool combats *Saprolegnia parasitica* infections through direct killing of microbes and modulation of host immune system / T. Tang, W. Zhong, P. Tang и др. // *eLife*. – 2025. – T. 13. – C. RP100393.

96. Lindholm-Lehto, P. C. Saprolegniosis in aquaculture and how to control it? / P. C. Lindholm-Lehto, P. Pylkkö. – Text: electronic. – 2024. – URL: <https://jukuri.luke.fi/handle/11111/9535> (date accessed: 11.07.2026).
97. Lipopeptides from *Bacillus*: unveiling biotechnological prospects—sources, properties, and diverse applications / D. Saiyam, A. Dubey, M. A. Malla, A. Kumar // *Brazilian Journal of Microbiology*. – 2024. – T. 55. – Lipopeptides from *Bacillus*. – № 1. – C. 281-295.
98. Long, D. H. Acyl Donor Stringency and Dehydroaminoacyl Intermediates in β -Lactam Formation by a Non-ribosomal Peptide Synthetase / D. H. Long, C. A. Townsend // *ACS Chemical Biology*. – 2021. – T. 16. – № 5. – C. 806-812.
99. Long, D. H. Mechanism of Integrated β -Lactam Formation by a Nonribosomal Peptide Synthetase during Antibiotic Synthesis / D. H. Long, C. A. Townsend // *Biochemistry*. – 2018. – T. 57. – № 24. – C. 3353-3358.
100. Mach, B. Separation of the biosynthesis of the antibiotic polypeptide tyrocidine from protein biosynthesis* / B. Mach, E. Reich, E. L. Tatum // *Proceedings of the National Academy of Sciences*. – 1963. – T. 50. – № 1. – C. 175-181.
101. Mansour, B. Computational Insights into Amide Bond Formation Catalyzed by the Condensation Domain of Nonribosomal Peptide Synthetases / B. Mansour, J. W. Gauld // *ACS Omega*. – 2024. – T. 9. – № 26. – C. 28556-28563.
102. Marahiel, M. A. A structural model for multimodular NRPS assembly lines / M. A. Marahiel // *Natural Product Reports*. – 2016. – Vol. 33. – № 2. – P. 136-140.
103. Markelova, N. Antimicrobial Activity of *Bacillus* Cyclic Lipopeptides and Their Role in the Host Adaptive Response to Changes in Environmental Conditions / N. Markelova, A. Chumak // *International Journal of Molecular Sciences*. – 2025. – Vol. 26. – № 1. – P. 336.
104. Martínez-Núñez, M. A. Nonribosomal peptides synthetases and their applications in industry / M. A. Martínez-Núñez, V. E. L. y. López // *Sustainable Chemical Processes*. – 2016. – T. 4. – № 1. – C. 13.
105. McErlean, M. Refining and expanding nonribosomal peptide synthetase function and mechanism / M. McErlean, J. Overbay, S. Van Lanen // *Journal of Industrial Microbiology & Biotechnology*. – 2019. – T. 46. – № 3-4. – C. 493-513.
106. Metabolically versatile psychrotolerant Antarctic bacterium *Pseudomonas* sp. ANT_H12B is an efficient producer of siderophores and accompanying metabolites (SAM) useful for agricultural purposes / M. Musialowski, Ł. Kowalewska, R. Stasiuk и др. // *Microbial Cell Factories*. – 2023. – T. 22. – C. 85.

107. MIBiG 4.0: advancing biosynthetic gene cluster curation through global collaboration / M. M. Zdouc, K. Blin, N. L. L. Louwen и др. // *Nucleic Acids Research*. – 2025. – Т. 53. – MIBiG 4.0. – № D1. – C. D678-D690.
108. Microbial community drivers of PK/NRP gene diversity in selected global soils / C. Borsetto, G. C. A. Amos, U. N. da Rocha et al. // *Microbiome*. – 2019. – Vol. 7. – № 1. – P. 78.
109. Miller, B. R. Structural Biology of Non-Ribosomal Peptide Synthetases / B. R. Miller, A. M. Gulick // *Methods in molecular biology* (Clifton, N.J.). – 2016. – Т. 1401. – C. 3-29.
110. Minimum Information about a Biosynthetic Gene cluster / M. H. Medema, R. Kottmann, P. Yilmaz et al. // *Nature Chemical Biology*. – 2015. – Vol. 11. – № 9. – P. 625-631.
111. Miras, M. A DegU-P and DegQ-Dependent Regulatory Pathway for the K-state in *Bacillus subtilis* / M. Miras, D. Dubnau // *Frontiers in Microbiology*. – 2016. – Т. 7. – C. 1868.
112. Molecular modeling of the reductase domain to elucidate the reaction mechanism of reduction of peptidyl thioester into its corresponding alcohol in non-ribosomal peptide synthetases / B. Manavalan, S. K. Murugapiran, G. Lee, S. Choi // *BMC Structural Biology*. – 2010. – Т. 10. – C. 1.
113. Moon, H. J. Biocontrol of Southern Blight Caused by *Sclerotium rolfsii* in Pepper Plants Using *Bacillus subtilis* GJ6-14 / H. J. Moon, M. K. Sang // *Research in Plant Disease*. – 2024. – Т. 30. – № 2. – C. 181-188.
114. Munkvold, G. P. Fusarium Species and Their Associated Mycotoxins / G. P. Munkvold // *Methods in Molecular Biology* (Clifton, N.J.). – 2017. – Т. 1542. – C. 51-106.
115. Murray, M. G. Rapid isolation of high molecular weight plant DNA / M. G. Murray, W. F. Thompson // *Nucleic acids research*. – 1980. – Т. 8. – № 19. – C. 4321-4325.
116. Nasir, M. N. Conformational analyses of bacillomycin D, a natural antimicrobial lipopeptide, alone or in interaction with lipid monolayers at the air-water interface / M. N. Nasir, F. Besson // *Journal of Colloid and Interface Science*. – 2012. – Т. 387. – № 1. – C. 187-193.
117. Nature's Antimicrobial Arsenal: Non-Ribosomal Peptides from PGPB for Plant Pathogen Biocontrol / A. Ranjan, V. D. Rajput, E. V. Prazdnova et al. // *Fermentation*. – 2023. – Vol. 9. – Nature's Antimicrobial Arsenal. – № 7. – P. 597.
118. NCBI prokaryotic genome annotation pipeline / T. Tatusova, M. DiCuccio, A. Badretdin и др. // *Nucleic Acids Research*. – 2016. – Т. 44. – № 14. – C. 6614-6624.

119. Negative Interactions Balance Growth and Defense in Plants Confronted with Herbivores or Pathogens / H. Zhou, J. Hua, J. Zhang, S. Luo // *Journal of Agricultural and Food Chemistry*. – 2022. – T. 70. – № 40. – C. 12723-12732.
120. Negative regulation of bleomycins biosynthesis by ArsR/SmtB family repressor BlmR in *Streptomyces verticillus* / H. Chen, J. Wang, J. Cui et al. // *Applied Microbiology and Biotechnology*. – 2019. – Vol. 103. – № 16. – P. 6629-6644.
121. Nerpa 2: probabilistic linking of biosynthetic gene clusters to nonribosomal peptides. Nerpa 2 / I. Olkhovskii, A. Kushnareva, A. Tagirdzhanov, A. Gurevich. – bioRxiv, 2026. – URL: <https://www.biorxiv.org/content/10.1101/2024.11.19.624380v3> (date accessed: 12.07.2026). – Text : electronic.
122. Nerpa: A Tool for Discovering Biosynthetic Gene Clusters of Bacterial Nonribosomal Peptides / O. Kunyavskaya, A. M. Tagirdzhanov, A. M. Caraballo-Rodríguez et al. // *Metabolites*. – 2021. – Vol. 11. – Nerpa. – № 10. – P. 693.
123. NMR resonance assignments for a docking domain pair with an attached thiolation domain from the PAX peptide-producing NRPS from *Xenorhabdus cabanillasii* / J. Watzel, S. Sarawi, E. Duchardt-Ferner et al. // *Biomolecular NMR Assignments*. – 2021. – Vol. 15. – № 1. – P. 229-234.
124. Nonribosomal antibacterial peptides isolated from *Streptomyces agglomeratus* 5-1-3 in the Qinghai-Tibet Plateau / K. Jiang, X. Chen, W. Zhang и др. // *Microbial Cell Factories*. – 2023. – T. 22. – № 1. – C. 5.
125. Nonribosomal Peptide Synthetases Involved in the Production of Medically Relevant Natural Products / E. A. Felnagle, E. E. Jackson, Y. A. Chan и др. // *Molecular pharmaceutics*. – 2008. – T. 5. – № 2. – C. 191-211.
126. Nonribosomal Peptides from Marine Microbes and Their Antimicrobial and Anticancer Potential / S. Agrawal, D. Acharya, A. Adholeya и др. // *Frontiers in Pharmacology*. – 2017. – T. 8. – C. 828.
127. Nonribosomal Peptides Produced by Minimal and Engineered Synthetases with Terminal Reductase Domains / A. Tietze, Y. Shi, M. Kronenwerth, H. B. Bode // *Chembiochem*. – 2020. – T. 21. – № 19. – C. 2750-2754.
128. NRPS Substrate Promiscuity Diversifies the Xenematides / J. M. Crawford, C. Portmann, R. Kontnik и др. // *Organic Letters*. – 2011. – T. 13. – № 19. – C. 5144-5147.
129. NRSPredictor2—a web server for predicting NRPS adenylation domain specificity / M. Röttig, M. H. Medema, K. Blin и др. // *Nucleic Acids Research*. – 2011. – T. 39. – № suppl_2. – C. W362-W367.

130. Ogura, M. Regulation of the Response Regulator Gene *degU* through the Binding of SinR/SlrR and Exclusion of SinR/SlrR by DegU in *Bacillus subtilis* / M. Ogura, H. Yoshikawa, T. Chibazakura // *Journal of Bacteriology*. – 2014. – T. 196. – № 4. – C. 873-881.
131. Ongena, M. *Bacillus* lipopeptides: versatile weapons for plant disease biocontrol / M. Ongena, P. Jacques // *Trends in Microbiology*. – 2008. – T. 16. – *Bacillus* lipopeptides. – № 3. – C. 115-125.
132. Overinduction of CYP51A Gene After Exposure to Azole Antifungals Provides a First Clue to Resistance Mechanism in *Fusarium solani* Species Complex / M. D'Agostino, T. Lemmet, C. Dufay и др. // *Microbial Drug Resistance*. – 2018. – T. 24. – № 6. – C. 768-773.
133. Overview of ribosomal and non-ribosomal antimicrobial peptides produced by Gram positive bacteria / M. Tajbakhsh, A. Karimi, F. Fallah, M. M. Akhavan // *Cellular and Molecular Biology (Noisy-Le-Grand, France)*. – 2017. – T. 63. – № 10. – C. 20-32.
134. Parallel pathways of repression and antirepression governing the transition to stationary phase in *Bacillus subtilis* / A. V. Banse, A. Chastanet, L. Rahn-Lee и др. // *Proceedings of the National Academy of Sciences of the United States of America*. – 2008. – T. 105. – № 40. – C. 15547-15552.
135. Piazza, F. Mutational Analysis and Membrane Topology of ComP, a Quorum-Sensing Histidine Kinase of *Bacillus subtilis* Controlling Competence Development / F. Piazza, P. Tortosa, D. Dubnau // *Journal of Bacteriology*. – 1999. – T. 181. – № 15. – C. 4540-4548.
136. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement / B. J. Walker, T. Abeel, T. Shea et al. // *PLOS ONE*. – 2014. – Vol. 9. – Pilon. – № 11. – P. e112963.
137. Principles of the distribution of wheat with fusariois disease, its effect on yield, harm and disease measures in the case of Uzbekistan / N. Khaytbayeva, M. Sattorova, U. Rakhimov, N. Tillyakhodjaeva // *E3S Web of Conferences*. – 2021. – Vol. 258. – P. 04043.
138. PRISM 3: expanded prediction of natural product chemical structures from microbial genomes / M. A. Skinnider, N. J. Merwin, C. W. Johnston, N. A. Magarvey // *Nucleic Acids Research*. – 2017. – T. 45. – PRISM 3. – № W1. – C. W49-W54.
139. Prospection for potential new non-ribosomal peptide gene clusters in *Bacillus* genus isolated from fermented foods and soil through genome mining / B. Waongo, L. Ndayishimiye, F. Tapsoba и др. // *Frontiers in Microbiology*. – 2025. – T. 16. – C. 1515483.

140. *Pseudomonas* spp. in biological plant protection and growth promotion / E. Nerek, B. Sokołowska, E. Nerek, B. Sokołowska // AIMS Environmental Science. – 2022. – Vol. 9. – № 4. – P. 493-504.
141. QUASt: quality assessment tool for genome assemblies / A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler // Bioinformatics. – 2013. – T. 29. – QUASt. – № 8. – C. 1072-1075.
142. Rational Design of a Bimodular Model System for the Investigation of Heterocyclization in Nonribosomal Peptide Biosynthesis / T. Duerfahrt, K. Eppelmann, R. Müller, M. A. Marahiel // Chemistry & Biology. – 2004. – T. 11. – № 2. – C. 261-271.
143. Recent advances in engineering nonribosomal peptide assembly lines / M. Winn, J. K. Fyans, Y. Zhuo, J. Micklefield // Natural Product Reports. – 2016. – Vol. 33. – № 2. – P. 317-347.
144. Recent Trends in Therapeutic Drug Monitoring of Peptide Antibiotics / Z. Kilianova, I. Cizmarova, M. Spaglova, J. Piestansky // Journal of Separation Science. – 2024. – T. 47. – № 19. – C. e202400583.
145. Reduction of Mycotoxigenic Fungi Growth and Their Mycotoxin Production by *Bacillus subtilis* QST 713 / T. Bertuzzi, G. Leni, G. Bulla, P. Giorni // Toxins. – 2022. – T. 14. – № 11. – C. 797.
146. Regulation of Nonribosomal Peptide Synthesis as a Mechanism of Antifungal Activity of Probiotics Based on the Bacteria Genera *Bacillus* and *Paenibacillus* / N. Vasilchenko, M. Kulikov, V. Stacenko et al. // XIV International Scientific Conference "INTERAGROMASH 2021" : Lecture Notes in Networks and Systems / A. Beskopylny, M. Shamtsyan eds. . – Cham : Springer International Publishing, 2022. – P. 111-121.
147. Roggiani, M. ComA, a phosphorylated response regulator protein of *Bacillus subtilis*, binds to the promoter region of *srfA*. / M. Roggiani, D. Dubnau // Journal of Bacteriology. – 1993. – T. 175. – № 10. – C. 3182-3187.
148. Schneider, T. L. Oxidase Domains in Epothilone and Bleomycin Biosynthesis: Thiazoline to Thiazole Oxidation during Chain Elongation / T. L. Schneider, B. Shen, C. T. Walsh // Biochemistry. – 2003. – T. 42. – Oxidase Domains in Epothilone and Bleomycin Biosynthesis. – № 32. – C. 9722-9730.
149. Scholz-Schroeder, B. K. The *sypA*, *sypB*, and *sypC* Synthetase Genes Encode Twenty-Two Modules Involved in the Nonribosomal Peptide Synthesis of Syringopeptin by *Pseudomonas syringae* pv. *syringae* B301D / B. K. Scholz-Schroeder, J. D. Soule, D. C. Gross // Molecular Plant-Microbe Interactions®. – 2003. – T. 16. – № 4. – C. 271-280.

150. Sharma, V. Streptomyces: host for refactoring of diverse bioactive secondary metabolites / V. Sharma, R. Kaur, R. Salwan // 3 Biotech. – 2021. – T. 11. – Streptomyces. – № 7. – C. 340.
151. Specific Phylotypes of *Saprolegnia parasitica* Associated with Atlantic Salmon Freshwater Aquaculture / K. V. Shreves, M. Saraiva, T. Ruba et al. // Journal of Fungi. – 2024. – Vol. 10. – № 1. – P. 57.
152. Stachelhaus, T. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases / T. Stachelhaus, H. D. Mootz, M. A. Marahiel // Chemistry & Biology. – 1999. – T. 6. – № 8. – C. 493-505.
153. Stanišić, A. Adenylation Domains in Nonribosomal Peptide Engineering / A. Stanišić, H. Kries // Chembiochem: A European Journal of Chemical Biology. – 2019. – T. 20. – № 11. – C. 1347-1356.
154. Staunton, J. Polyketide biosynthesis: a millennium review / J. Staunton, K. J. Weissman // Natural Product Reports. – 2001. – T. 18. – Polyketide biosynthesis. – № 4. – C. 380-416.
155. Structural advances toward understanding the catalytic activity and conformational dynamics of modular nonribosomal peptide synthetases / K. D. Patel, M. R. MacDonald, S. F. Ahmed и др. // Natural Product Reports. – 2023. – T. 40. – № 9. – C. 1550-1582.
156. Structural and functional investigation of the intermolecular interaction between NRPS adenylation and carrier protein domains / J. A. Sundlov, C. Shi, D. J. Wilson и др. // Chemistry & biology. – 2012. – T. 19. – № 2. – C. 188-198.
157. Structural Basis of the Interaction of MbtH-like Proteins, Putative Regulators of Nonribosomal Peptide Biosynthesis, with Adenylating Enzymes * / D. A. Herbst, B. Boll, G. Zocher и др. // Journal of Biological Chemistry. – 2013. – T. 288. – № 3. – C. 1991-2003.
158. Structural elements of an NRPS cyclization domain and its intermodule docking domain / D. P. Dowling, Y. Kung, A. K. Croft и др. // Proceedings of the National Academy of Sciences. – 2016. – T. 113. – № 44. – C. 12432-12437.
159. Structural insights into the substrate-bound condensation domains of non-ribosomal peptide synthetase AmbB / M.-J. Chu Yuan Kee, S. R. Bharath, S. Wee et al. // Scientific Reports. – 2022. – Vol. 12. – № 1. – P. 5353.
160. Structure and Function of a Dehydrating Condensation Domain in Nonribosomal Peptide Biosynthesis / J. B. Patteson, C. M. Fortinez, A. T. Putz и др. // Journal of the American Chemical Society. – 2022. – T. 144. – № 31. – C. 14057-14070.

161. Structure Elucidation and Activity of Kolossin A, the D-/L-Pentadecapeptide Product of a Giant Nonribosomal Peptide Synthetase / H. B. Bode, A. O. Brachmann, K. B. Jadhav и др. // *Angewandte Chemie (International Ed. in English)*. – 2015. – T. 54. – № 35. – С. 10352-10355.
162. Structures of a non-ribosomal peptide synthetase condensation domain suggest the basis of substrate selectivity / T. Izoré, Y. T. Candace Ho, J. A. Kaczmariski et al. // *Nature Communications*. – 2021. – Vol. 12. – № 1. – P. 2511.
163. Süssmuth, R. D. Nonribosomal Peptide Synthesis—Principles and Prospects / R. D. Süssmuth, A. Mainz // *Angewandte Chemie International Edition*. – 2017. – Vol. 56. – № 14. – P. 3770-3821.
164. Taj-Aldeen, S. J. Reduced Multidrug Susceptibility Profile Is a Common Feature of Opportunistic Fusarium Species: Fusarium Multi-Drug Resistant Pattern / S. J. Taj-Aldeen // *Journal of Fungi*. – 2017. – Vol. 3. – Reduced Multidrug Susceptibility Profile Is a Common Feature of Opportunistic Fusarium Species. – № 2. – P. 18.
165. The Fusarium Head Blight Pathosystem / P. Nicholson, N. Gosman, R. Draeger et al. // *Wheat Production in Stressed Environments* / H. T. Buck et al. eds. . – Dordrecht : Springer Netherlands, 2007. – P. 23-36.
166. The Iturin and Fengycin Families of Lipopeptides Are Key Factors in Antagonism of *Bacillus subtilis* Toward *Podosphaera fusca* / D. Romero, A. de Vicente, R. H. Rakotoaly и др. // *Molecular Plant-Microbe Interactions®*. – 2007. – T. 20. – № 4. – С. 430-440.
167. The Linear Pentadecapeptide Gramicidin Is Assembled by Four Multimodular Nonribosomal Peptide Synthetases That Comprise 16 Modules with 56 Catalytic Domains * / N. Kessler, H. Schuhmann, S. Morneweg и др. // *Journal of Biological Chemistry*. – 2004. – T. 279. – № 9. – С. 7413-7419.
168. The Zinc-Responsive Regulator Zur Controls Expression of the Coelibactin Gene Cluster in *Streptomyces coelicolor* / D. Kallifidas, B. Pascoe, G. A. Owen и др. // *Journal of Bacteriology*. – 2010. – T. 192. – № 2. – С. 608-611.
169. Timofeeva, A. M. Bacterial Siderophores: Classification, Biosynthesis, Perspectives of Use in Agriculture / A. M. Timofeeva, M. R. Galyamova, S. E. Sedykh // *Plants*. – 2022. – T. 11. – Bacterial Siderophores. – № 22. – С. 3065.
170. Transcription Factor Spo0A Regulates the Biosynthesis of Difficidin in *Bacillus amyloliquefaciens* / N. Liu, H. Sun, Z. Tang и др. // *Microbiology Spectrum*. – 2023. – T. 11. – № 4. – С. e01044-23.

171. Troxell, B. Transcriptional regulation by Ferric Uptake Regulator (Fur) in pathogenic bacteria / B. Troxell, H. M. Hassan // *Frontiers in Cellular and Infection Microbiology*. – 2013. – T. 3. – C. 59.
172. Type II Non-ribosomal Peptide Synthetase Proteins: Structure, Mechanism, and Protein-Protein Interactions / M. J. Jaremko, T. D. Davis, J. C. Corpuz, M. D. Burkart // *Natural product reports*. – 2020. – T. 37. – Type II Non-ribosomal Peptide Synthetase Proteins. – № 3. – C. 355-379.
173. Walker, A. S. A Machine Learning Bioinformatics Method to Predict Biological Activity from Biosynthetic Gene Clusters / A. S. Walker, J. Clardy // *Journal of Chemical Information and Modeling*. – 2021. – T. 61. – № 6. – C. 2560-2571.
174. Walsh, C. T. Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds / C. T. Walsh, R. V. O'Brien, C. Khosla // *Angewandte Chemie (International Ed. in English)*. – 2013. – T. 52. – № 28. – C. 7098-7124.
175. Weckermann, R. Complete nucleotide sequence of the *tycA* gene coding the tyrocidine synthetase 1 from *Bacillus brevis*. / R. Weckermann, R. Fürbass, M. A. Marahiel // *Nucleic Acids Research*. – 1988. – T. 16. – № 24. – C. 11841.
176. Weissman, K. J. The structural biology of biosynthetic megaenzymes / K. J. Weissman // *Nature Chemical Biology*. – 2015. – Vol. 11. – № 9. – P. 660-670.
177. Williams, G. J. Engineering polyketide synthases and nonribosomal peptide synthetases / G. J. Williams // *Current Opinion in Structural Biology*. – 2013. – T. 23. – № 4. – C. 603-612.
178. X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis / K. Haslinger, M. Peschke, C. Brieke et al. // *Nature*. – 2015. – Vol. 521. – № 7550. – P. 105-109.
179. Yasumura, A. Involvement of Nitrogen Regulation in *Bacillus subtilis* *degU* Expression / A. Yasumura, S. Abe, T. Tanaka // *Journal of Bacteriology*. – 2008. – T. 190. – № 15. – C. 5162-5171.
180. Zamioudis, C. Modulation of Host Immunity by Beneficial Microbes / C. Zamioudis, C. M. J. Pieterse // *Molecular Plant-Microbe Interactions®*. – 2012. – T. 25. – № 2. – C. 139-150.
181. Zhi, Y. Genome and transcriptome analysis of surfactin biosynthesis in *Bacillus amyloliquefaciens* MT45 / Y. Zhi, Q. Wu, Y. Xu // *Scientific Reports*. – 2017. – T. 7. – C. 40976.
182. Zinc-Responsive Regulator Zur Regulates Zinc Homeostasis, Secondary Metabolism, and Morphological Differentiation in *Streptomyces avermitilis* / M. Lyu,

Y. Cheng, Y. Dai и др. // *Applied and Environmental Microbiology*. – 2022. – Т. 88.
– № 7. – С. e00278-22.

Приложение 1

Таблица П1. Соединения и их активность

Соединение	Противогрибковая активность	Соединение	Противогрибковая активность	Соединение	Противогрибковая активность
abyssomicin	0*	poaeamide	1	lipstatin	0
aurafuron	1	nannocystin	1	lysobactin	0
aureothin	1	retimycin	0	mannopeptimycin	0
avilamycin	0	salinamide	0	mirubactin	0
bafilomycin	1	microsclerodermins_2	1	naphthyridinomycin	0
borrelidin	0	aresnopolyketides	0	nocardicin	0
chlorothricin	0	rakicidin	0	nodularin	0
dawenol	0	rakicidin_2	0	paenibacterin	0
erythromycin	0	rakicidin_d	0	paenibactin	0
galbonolides	1	BE_43547A1	0	pelgipeptin	1
kedarcidin	0	BE_43547A1_2	0	plipastatin	1
lactimidomycin	1	bananamide	0	polymyxin_2	0
neoaureothin	1	poaeamide_b	0	porothramycin	0
neocarzinostatin	0	JBIR_100	1	pseudomonine	0
nystatin	1	myxochromide_d	0	putisolvin	1
pactamycin	1	myxochromide_d_2	0	pyochelin	0
salinomycin	0	myxochromide_d_3	0	pyoverdine	0

soraphen	1	myxochromide_d_4	0	quinocarcin	0
spirangien	1	myxochromide_b	0	quinomycin	0
sporolide	0	myxochromide_a	0	rhodochelin	0
stambomycin	0	myxochromide_c	0	ristocetin	0
stigmatellin	1	myxochromide_c_2	0	safracin	0
tautomycetin	1	myxochromide_a_2	0	saframycin_a	0
tautomycin	1	myxochromide_a_3	0	scabichelin	0
tiacumicin	0	myxochromide_a_4	0	sibiromycin	0
9methylstreptim idone	1	myxochromide_s	0	streptothricin	0
difficidin	0	myxochromide_s_2	0	SW_163	0
migrastatin	1	myxochromide_s_3	0	syringafactin	0
elansolid	0	myxochromide_d_5	0	syringolin_a	0
etnangien	0	argimycin	0	syringomycin	0
macrolactin	1	ikarugamycin	1	syringopeptin	0
mupirocin	0	obafluorin	0	teicoplanin_2	0
sorangicin	0	warkmycin	0	thaxtomin	0
thailandamide	0	cervnimycin	0	tilivaline	0
anatoxin	0	actinonin	0	tomaymycin	0
actinorhodin	0	matylstatins	0	tridecaptin	0
alnumycin	0	aculeximycin	1	triostin_a	0
aranciamycin	0	ambruticin	1	tyrocidine	0

arenimycin	0	amphotericin	1	vanchrobactin	0
auricin	0	anglomycin	0	vancomycin	0
benastatin	0	ansamitocin	0	vicibactin	0
chromomycin	1	apoptolidin	0	viomycin	0
cinerubin	0	calcimycin	1	vulnibactin	0
colabomycin	0	calicheamicin	1	WAP_8294A2	0
elloramycin	0	candicidin	1	yersiniabactin	0
enterocin	0	chalcomycin	0	napyradiomycin_2	0
arthrofactin	0	concanamycin_a	0	diazepinomicin	0
aeruginosin	0	cremimycin	0	paromomycin	1
enterocin	0	curacin_a	0	auricin_deoxysugar_moi eties	0
flustatin	0	dihydrochalcomycin	0	albonoursin	0
fredericamycin	1	E_492_E_975	0	eicosapentaenoic_acid_ like_compound	0
granaticin	0	E_837	1	heterocyst_glycolipids	0
griseorhondin	0	ebelactone	1	A_90289	0
arylomycin	0	ECO_2301	1	nikkomycin	1
landomycin	0	erythromycin_2	0	flaviolin	0
lomaiviticin	0	FD_891	0	vibriobactin	0
mithramycin	0	filipin	1	A_500359	0
nogalamycin	0	fostriecin	0	pacidamycin	0

raelomycin	0	FR_008	1	ansatrienin	1
resistomycin	1	geldanamycin	1	azinomycin_b	0
SF2575	0	geldanamycin_2	1	bactobolin	1
steffimycin	0	geldanamycin_3	1	caerulomycin_a	1
tetarimycin	0	gephyronic_acid	1	crocacin	1
tetracenomycin	0	halstocacosanolide	0	cylindrospermopsin_2	0
cetoniacytone	0	herbimycin	0	cylindrospermopsin_3	0
red_pigment	0	hygrocin	1	cylindrospermopsin_4	0
A40926	0	incednine	0	cystothiazole_a	1
A47934	0	jerangolid	1	delftibactin	0
aeruginoside	0	kijanamicin	0	dkxanthene	1
myxoprincomide	0	laidlomycin	0	epothilone	0
orfamide	1	lankamycin	0	epothilone_2	0
orfamide_a	1	lasalocid	0	epothilone_3	0
ristomycin	0	marineosin	0	epothilone_4	0
serobactins	0	megalomicin	0	FK520	1
sessilin	1	meilingmycin	0	FR901464	0
spumigin	0	midecamycin	0	glidobactin	1
stenothricin	0	monensin	0	guadinomine	0
surfactin	1	mycinamicin	0	heat_stable_antifungal_factor	1

taromycin	0	mycolactone	0	jawsamycin	1
teicoplanin	0	nanchangmycin	0	locillomycin	1
thanamycin	1	natamycin	1	melithiazol_2	1
thiocoraline	0	nemadectin	0	meridamycin	0
tolaasin	1	neocarzilin	0	meridamycin_2	0
valinomycin	1	niddamycin	0	meridamycin_3	0
griseoviridin	1	nystain_like_polyene	1	microcystin_3	0
WLIP	1	oligomycin	1	micropeptin_k139	0
xantholysin	1	pactamycin_2	1	muraymycin	0
phenalinolactone	0	phoslactomycin_b	1	mycobactin	0
lactimidomycin	1	piericidin_a1_2	1	myxalamid	1
hassallidins	1	pimaricin	1	nocobactin_na	0
methyldapomycin	0	pladienolide	0	oocydin_a	1
eponemycin	0	pyoluteorin	1	oocydin_a_2	1
cupriachelin	0	pyoluteorin_2	1	paenilamicin	1
hormaomycin	0	quartromicin	0	polyoxypeptin	0
luminmycin	0	reveromycin	1	pyridomycin	0
fimsbactin	0	rifamycin	0	rapamycin	1
cichofactin	1	rifamycin_2	0	sanglifehrin_a	0
anabaenopeptin	0	rimocidin	1	sgr_ptms	0

leupyrrin	1	RK_682	0	syringolin_a_2	0
massetolide	1	rubradirin	0	tubulysin	0
lyngbyatoxin	0	salinilactam	0	yersiniabactin_2	0
cereulide	0	salinomycin_2	0	zwittermycin_a	1
epoxomicin	0	salinosporamide_a	0	dynemicin	0
cycloheximide	1	spinosad	0	tln-05220	0
grincamycin	0	tautomycetin	1	undecylprodigiosin	1
cephamycin	0	tetrocarcin_a	0	herboxidiene	0
cyclomarin	0	tetronasin	1	kosinostatin	1
calcium_dependent_antibiotic	0	tetronomycin	1	cosmomycin_d	0
clorobiocin	0	tylactone	0	alkyl_o_dihydrogeranyl_methoxyhydroquinones	0
coumermycin	0	vicenistatin	0	furaquinocin_a	0
vpe	0	basiliskamides	1	teleocidin_b	0
chloramphenicol	0	tartrolon	0	albicidin	1
promysalin	0	A_74528	0	corallopyronin	0
endophenazines	1	aclacinomycin	0	disorazole	0
esmeraldin	0	aclacinomycin_2	0	enacyloxin	0
napsamycin	0	azicemicin	0	FR901464_2	0
pristinamycin	0	BE_7585A	0	iturin	1

amicetin	0	bryostatin	0	lankacidin	0
ajudazol	1	chelocardin	0	mycosubtilin	1
althiomycin	0	chortetracycline	0	myxovirescin_2	0
andrimid	0	chrysomycin	0	oxazolomycin	1
antimycin	1	cosmomycin	0	spliceostatin	0
barbamide	0	curamycin	0	thiomarinol	0
bleomycin	0	dactylocycline	0	virginiamycin	0
burkholdac	0	daunorubicin	0	himastatin	0
C_1027	0	doxorubicin	0	puwainaphycins	0
calyculin	0	erdacin	0	jagaricin	1
chondramide	0	FD_594	0	luminmide	0
chondrochloren	1	frenolicin	1	pheganomycin	0
cinnabaramide	1	glivocarcin	1	lobophorin_2	0
colibactin	0	granaticin_2	0	misakinolide	0
collismycin	1	griseusin	1	taiwanchelin	0
cryptophycin	1	hatomarubigin	0	colistin	0
curacin	0	hedamycin	0	thiolutin	1
cylindrospermo psin	0	jadomycin	0	hitachimycin	1
didemnin	0	kiamycin	0	nocardiopsin	0
FK228	1	kiamycin_2	0	desotamide	0
frontalamide	1	lactomycin_z	0	frankiamicin	0

hectochlorin	1	lactomycin	0	akaeolide	0
jmaicamide	0	lysolipin	0	lorenic_acid_a	0
lipomycin	0	medermycin	0	landepoxcin	0
lobophorin	0	mithramycin_2	0	clarepoxcin	0
lymphostin	0	naptocyclinone	0	versipelostatin	0
lymphostin_2	0	oxytetracycline	0	teixobactin	0
maduropeptin	0	pederin_2	0	albachelin	0
melithiazol	1	pradimicin	1	nataxazole	0
micacocidin	0	prodigiosin	1	marformycins	0
microcystin	0	R1128	0	splenocin	0
microcystin_2	0	ravidomycin	1	leucanicidin	0
microsclerodermins	1	rubromycin	0	feglymycin	0
myxochromide	0	Sch47554_47555	0	antalid	0
myxothiazol	1	simocyclinone_2	0	thiotetroamide	0
myxoviresein	0	spore_pigment	0	thiolactomycin	0
nostopeptolide	0	spore_pigment_2	0	chaxamycin	0
nostophycin	0	xantholipin	1	maklamicin	0
pellasoren	0	2_4_diacetylphloroglucinol	1	cyclizidine	0
pyralomicin	0	viguiepinol	0	streptazone_e	0
salinosporamide	0	2_amino_4_methoxy_tr	0	WS9326	0

		ans_3_butenic_acid			
spiruchostatin	1	A_503083	0	4_Z_annamycin	0
streptolydigin	0	A54145	0	chlorotolil	1
tallysomylin	0	acinetobactin	0	anthracimycin	0
thuggacin	0	actinomycin	1	anthracimycin_2	0
tirandamycin	0	amychelin	0	lobosamide	0
xenocoumactin	0	anabaenopeptin_2	0	simocyclinone_4	0
zeamine	0	anthramycin	1	naringenin	0
zorbamycin	0	aureusimine	0	viscosin	0
polyketomycin	0	bacillibactin_2	0	thuggacin_2	0
cylindrocyclophane	0	bacitracin	0	aurachin	1
kendomycin	0	balhimycin	0	tubulysin_2	0
chivosazol	1	capreomycin	0	myxochelin	0
kirromycin	0	chloroeremomycin	0	heronamide	0
nosperin	0	coelibactin	0	phormidolide	0
simocyclinone	0	coelichelin	0	cahuitamycin	0
liposidomycin	0	complestatin	0	tu_3010	0
napyradiomycin	0	congoicidine	0	thiolactomycin_2	0
endophenazine	1	coronatin	0	U_68204	0
merochlorin	0	corpeptin	0	U_68204_2	0
bacillaene	0	cyanopeptin	0	carbamidocyclophane	0

bacillomycin	1	cystomanamides	0	PM100117	0
diaphorin	0	daptomycin	0	ralsolamycin	0
fengycin	1	echinomycin	1	zunyimycin_a	0
kalimantacin	0	echosides	0	haliamide	0
leinamycin	0	enduracidin	0	tambromycin	0
malleilactone	0	enterobactin	0	metatricycloene	0
onnamide	0	entolysin	0	lipopeptide_8D11	0
pederin	0	equibactin	0	hexaricin	0
psymberin	0	erythrochelin	0	uncialamycin	0
rhizopodins	0	exochelin	0	tiancimycin	0
rhizoxin	0	FK506	1	BD_12	1
thailanstatin	0	friulimicin	0	brasilinolide	1
divergolide	0	fusaricidin_2	1	svaricin	0
xenotetrapeptide	0	fuscachelin	0	macrolactin_2	0
taxllaid	0	GE81112	0	saprolmycin_e	0
distamycin	1	glycopeptidolipid	0	JBIR_76	0
brabantamide	1	glycopeptidolipid_1	0	cichozeptin	0
fusaricidin	1	glycopeptidolipid_2	0	lugdunin	0
polymyxin	0	glycopeptidolipid_3	0	phenalamide	1
marinopyrrole	0	gobichelin	0	aldgamycin	0
nonadecene	0	gramicidin	0	lidamycin	0
heptadecene	0	griseobactin	0	telomycin	0

curacin_2	0	heterobactin	0	dutomycin	0
livipeptin	0	holomycin	1	cystobactamide	0
piericidin_a1	1	indigoidine	0	griselimycin	0
chlorizidine_a	0	JBIR_34	0	althiomycin_2	0
mensacaricin	0	koranimine	0	nunapeptin	1
UK_68_597	0	kutznerides	1	bryostatin_2	0
rishirilide_b	0	laspartomycin	0	curacin_3	0
bacillibactin	0	lichenysin	0		

*1 – Наличие противогрибковой активности, 0 – отсутствие противогрибковой активности.

Приложение 2

Таблица П2 – Качество сборки геномов

Штамм	Полнота сборки %	Степень контаминации	N50 контигов	Размер генома	Общее количество кодирующих последовательностей	Всего контигов
B_atrophaeus_AG8	100.0	0.09	4164149	4302183	4186	2
B_atrophaeus_P1F1	100.0	0.03	4092400	4092400	3961	1
B_halotolerans_15A	100.0	0.02	4340782	4340782	4385	1
B_halotolerans_31A	100.0	0.07	4255552	4255552	4262	1
B_halotolerans_35A	100.0	0.09	4108934	4108934	4108	1
B_halotolerans_40A	100.0	0.0	4142210	4142210	4076	1
B_halotolerans_41A	100.0	0.11	4141287	4141287	4176	1
B_halotolerans_42A	100.0	0.03	4140596	4140596	4190	1
B_licheniformis_CL33	100.0	0.11	264201	4303445	4460	69
B_licheniformis_CL56	100.0	0.09	1037589	4278533	4458	73
B_safensis_AG46	100.0	0.05	3795137	3795137	3850	1
B_safensis_CL35	100.0	0.03	304565	3763322	3823	38
B_subtilis_50_1B	100.0	0.01	4105350	4105350	4063	1
B_subtilis_58B	100.0	0.02	4131922	4131922	4081	1
B_subtilis_60B	100.0	0.01	4105212	4105212	4062	1
B_subtilis_65AB	100.0	0.0	4179472	4179472	4183	1
B_subtilis_65BB	100.0	0.0	4179460	4179460	4181	1

B_subtilis_8A	100.0	0.07	4269882	4269882	4301	1
B_subtilis_KATMIRA	100.0	0.39	4149357	4149357	4253	1
B_subtilis_KamC12	100.0	0.02	2254312	4036221	4136	6
B_subtilis_MT48	100.0	0.03	4143098	4143098	4132	1
B_subtilis_MT74	100.0	0.05	4143087	4143087	4143	1
B_subtilis_R1	100.0	0.03	114355	4162873	4286	127
B_velezensis_21A	100.0	0.04	3940087	3940087	3763	1
B_velezensis_43A	100.0	0.03	3929400	3940217	3760	2
B_velezensis_7A	100.0	0.1	4178067	4178067	4132	1
B_velezensis_B1895	100.0	0.99	3910193	4062636	4021	3
B_velezensis_MT102	100.0	0.17	339475	4246725	4170	30
B_velezensis_MT155	100.0	0.09	3920487	3999080	3821	4
B_velezensis_MT186	100.0	0.08	4110085	4140244	4065	4
B_velezensis_MT55	100.0	0.02	3912365	3994220	3816	4
B_velezensis_MT61	100.0	0.09	2689985	4107921	3930	9
B_velezensis_MT64	100.0	0.1	4062846	4106989	3959	5
B_velezensis_MT86	100.0	0.12	4077349	4102637	4100	5
B_velezensis_PZ2	100.0	0.18	4129726	4129726	3978	1
B_velezensis_R4	100.0	0.19	437506	4064860	3980	57
B_velezensis_R4.6	100.0	0.13	3965290	4171958	4108	4
B_velezensis_R5	100.0	0.04	79957	4004650	3939	119
B_velezensis_T5	100.0	0.06	4236852	4288643	4314	5

B_velezensis_V3.14	100.0	0.13	3965290	4171958	4108	4
Cytobacillus_horneckiae_ Dr3_4	100.0	0.53	73854	5047342	5070	185
P_alloputida_A10	100.0	0.31	6097026	6097026	5483	1
P_polymyxa_O1.27	100.0	0.16	5745826	5745826	4939	1
P_polymyxa_O2.11	100.0	0.16	5745826	5745826	4939	1
P_polymyxa_R3.13	100.0	0.45	5730779	5743678	5024	2
P_polymyxa_R4.24	100.0	0.35	5760149	5760149	5103	1
P_polymyxa_R4.5	100.0	0.37	5712992	5712992	5073	1
P_polymyxa_R5.31	100.0	0.16	5866872	5871740	5304	2
P_polymyxa_R6.14	100.0	0.16	5745826	5745826	4939	1
R_opacus_Nah7IF	100.0	1.88	7308160	8035834	7428	4
Sutcliffiella_sp_Dr1_2	100.0	1.65	210282	5017272	4896	96
B_velezensis_MT14	97.35	2.19	1035922	4080572	4593	6
B_halotolerans_3F23	98.17	0.52	135438	4012235	4109	49
B_velezensis_MT42	99.91	2.32	4080892	4115985	4453	3
B_velezensis_MT142	99.94	1.19	4305438	4434607	4836	3
B_velezensis_MT141	99.97	2.1	4303216	4432692	4937	3
P_polymyxa_K1.14	99.98	0.23	5746779	5746779	5048	1

Приложение 3

Таблица ПЗ. Таксономическая идентификация штаммов с использованием базы данных GTDB

Штамм	Таксономическая классификация	Ближайший референсный геном	ANI %
K3F3	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Achromobacter;s__	N/A	N/A
Kush2	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Achromobacter;s__	N/A	N/A
Acinetobacter_4_E6.2	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter;s__Acinetobacter baumannii	GCF_009759685.1	97.82
AG8	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus atrophaeus	GCF_001584335.1	98.53
PIF1	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus atrophaeus	GCF_001584335.1	98.57
ZY9	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus atrophaeus	GCF_001584335.1	98.53
15A	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus halotolerans	GCF_001517105.1	97.86
31A	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus halotolerans	GCF_001517105.1	98.0
35A	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus halotolerans	GCF_001517105.1	97.96
3F23	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus halotolerans	GCF_001517105.1	97.68
40A	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus halotolerans	GCF_001517105.1	97.93
41A	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus halotolerans	GCF_001517105.1	97.96
42A	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus halotolerans	GCF_001517105.1	97.96
CL33	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus	GCF_000011645.1	99.71

	licheniformis		
CL56	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus licheniformis	GCF_000011645.1	99.84
C76.2	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus pumilus	GCF_900186955.1	99.09
AG46	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus safensis	GCF_000691165.1	96.33
CL35	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus safensis	GCF_000691165.1	96.18
50_1B	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.8
58B	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.8
60B	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.8
65AB	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.75
65BB	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.75
8A	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.57
KATMIRA	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.74
Kam3	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.49
KamC12	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.75
MT48	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.51
MT74	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.47
R1	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus subtilis	GCF_000009045.1	98.84
21A	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus	GCF_001461825.1	98.22

	velezensis		
43A	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.21
7A	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.16
B1895	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	97.28
FD2	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.15
Kam11	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	99.03
MT102	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.02
MT14	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.74
MT141	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.75
MT142	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.75
MT155	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.22
MT186	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.9
MT42	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.75
MT55	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.22
MT61	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	99.03
MT64	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	97.59
MT86	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.8
PZ2	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus	GCF_001461825.1	98.17

	velezensis		
R4	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.27
R4.6	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	97.43
R5	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	97.77
Bacillus_T5	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	98.57
V3.14	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__Bacillus velezensis	GCF_001461825.1	97.43
5_S.P.2	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Enterobacterales;f__Enterobacteriaceae;g__Citrobacter_A;s__Citrobacter_A sedlakii	GCF_000759835.1	99.09
Dr3_4	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales_B;f__DSM-18226;g__Cytobacillus;s__Cytobacillus horneckiae	GCF_002835735.1	99.4
K1.14	d__Bacteria;p__Bacillota;c__Bacilli;o__Paenibacillales;f__Paenibacillaceae;g__Paenibacillus;s__Paenibacillus polymyxa_B	GCF_001719045.1	95.9
O1.27	d__Bacteria;p__Bacillota;c__Bacilli;o__Paenibacillales;f__Paenibacillaceae;g__Paenibacillus;s__Paenibacillus polymyxa	GCF_015710975.1	98.33
O2.11	d__Bacteria;p__Bacillota;c__Bacilli;o__Paenibacillales;f__Paenibacillaceae;g__Paenibacillus;s__Paenibacillus polymyxa	GCF_015710975.1	98.33
R3.13	d__Bacteria;p__Bacillota;c__Bacilli;o__Paenibacillales;f__Paenibacillaceae;g__Paenibacillus;s__Paenibacillus polymyxa	GCF_015710975.1	98.43
R4.24	d__Bacteria;p__Bacillota;c__Bacilli;o__Paenibacillales;f__Paenibacillaceae;g__Paenibacillus;s__Paenibacillus polymyxa_B	GCF_001719045.1	98.33
R4.5	d__Bacteria;p__Bacillota;c__Bacilli;o__Paenibacillales;f__Paenibacillaceae;g__Paenibacillus;s__Paenibacillus polymyxa_B	GCF_001719045.1	98.28
R5.31	d__Bacteria;p__Bacillota;c__Bacilli;o__Paenibacillales;f__Paenibacillaceae;g__Paenibacillus;s__Paenibacillus polymyxa_B	GCF_001719045.1	95.76
R6.14	d__Bacteria;p__Bacillota;c__Bacilli;o__Paenibacillales;f__Paenibacillaceae;g__Paenibacillus;s__Paenibacillus polymyxa	GCF_015710975.1	98.33
BB	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales_B;f__DSM-1321;g__Peribacillus;s__Peribacillus frigoritolerans	GCF_039526365.1	98.6
A10	d__Bacteria;p__Pseudomonadota;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseu-	GCF_021283545.2	97.2

	domonadaceae;g Pseudomonas E;s Pseudomonas E alloputida		
Nah7IF	d__Bacteria;p__Actinomycetota;c__Actinomycetes;o__Mycobacteriales;f__Mycobacteriaceae;g Rhodococcus F;s Rhodococcus F opacus	GCF_012396235.1	98.24
Staphylococcus_4_E6.2	d__Bacteria;p__Bacillota;c__Bacilli;o__Staphylococcales;f__Staphylococcaceae;g__Staphylococcus;s__Staphylococcus saprophyticus	GCF_000010125.1	98.91
Streptomyces_T5	d__Bacteria;p__Actinomycetota;c__Actinomycetes;o__Streptomycetales;f__Streptomyetaceae;g Streptomyces;s Streptomyces anulatus	GCF_014650675.1	97.29
Dr1_2	d__Bacteria;p__Bacillota;c__Bacilli;o__Bacillales;f__Bacillaceae_I;g Sutcliffiella_A;s__	N/A	N/A

Приложение 4

Таблица П4. Соединения, для которых предсказана противогрибковая активность на основе геномов отсеквенированных штаммов

Штамм	Тип вещества	Референсное соединение	Процент сходства	Предсказанные мономеры	Вероятность противогрибковой активности
Bacillus_pumilis_C76.2	НРПС, Т1ПКС	zwittermicin A	18.0	Asn Asn mal Glu mal mal mal	82,6%
Rhodococcus_opacus_Nah7IF	НРПС	WS9326B/WS9326A/WS9326G/WS9326F	7.0	Thr X X X Leu Thr X X	59,6%
Paeniacyllus_polymyxa_R6.14	НРПС	tridecaptin	100.0	Val Dab Gly Ser Trp Ser Dab Dab Trp Glu Val Ile Ala	64,1%
Paeniacyllus_polymyxa_R5.31	НРПС	tridecaptin	100.0	Val Dab Gly Ser Trp Ser Dab Dab Trp Glu Val Ile Ala	61,4%
Paeniacyllus_polymyxa_R4.24	НРПС	tridecaptin	100.0	Val Dab Gly Ser Trp Ser Dab Dab Phe Glu Val Ile Ala	60,2%
Paeniacyllus_polymyxa_R3.13	НРПС	tridecaptin	100.0	Val Dab Gly Ser Trp Ser Dab Dab Trp Glu Val Ile Ala	64,1%
Paeniacyllus_polymyxa_O2.11	НРПС	tridecaptin	100.0	Val Dab Gly Ser Trp Ser Dab Dab Trp Glu Val Ile Ala	64,1%
Paeniacyllus_polymyxa_O1.27	НРПС	tridecaptin	100.0	Val Dab Gly Ser Trp Ser Dab Dab Trp Glu Val Ile Ala	64,1%
Paeniacyllus_polymyxa_K1.14	НРПС	tridecaptin	100.0	Val Dab Gly Ser Trp Ser Dab Dab Phe Glu Val Ile Ala	59,7%
Rhodococcus_opacus_Nah7IF	Т1ПКС	tetrocarcin A	11.0	mmal mal	51,4%
Bacillus_velezensis_V3.14	НРПС	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	91,9%
Bacillus_velezensis_T5	НРПС	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	86,6%
Bacillus_velezensis_R4.6	НРПС	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	91,9%
Bacillus_velezensis_PZ2	НРПС	surfactin	91.0	Leu Val Asp Leu Glu Leu Leu	87,6%
Bacillus_velezensis_MT86	НРПС	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	86,6%
Bacillus_velezensis_MT64	НРПС	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	91,9%

Bacillus_velezensis_MT61	HPIC	surfactin	78.0	Glu Leu Leu Val Asp Leu Leu	86,6%
Bacillus_velezensis_MT55	HPIC	surfactin	91.0	Glu Leu Leu Val Asp Leu Leu	87,6%
Bacillus_velezensis_MT186	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	86,6%
Bacillus_velezensis_MT155	HPIC	surfactin	91.0	Glu Leu Leu Val Asp Leu Leu	87,6%
Bacillus_velezensis_MT102	HPIC	surfactin	91.0	Glu Leu Leu Val Asp Leu Leu	87,6%
Bacillus_velezensis_Kam11	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	86,6%
Bacillus_velezensis_FD2	HPIC	surfactin	91.0	Glu Leu Leu Val Asp Leu Leu	87,6%
Bacillus_velezensis_7A	HPIC	surfactin	91.0	Glu Leu Leu Val Asp Leu Leu	87,6%
Bacillus_velezensis_43A	HPIC	surfactin	91.0	Glu Leu Leu Val Asp Leu Leu	87,6%
Bacillus_velezensis_21A	HPIC	surfactin	91.0	Glu Leu Leu Val Asp Leu Leu	87,6%
Bacillus_subtilis_MT74	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	84,4%
Bacillus_subtilis_MT48	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	84,4%
Bacillus_subtilis_KATMIRA	HPIC	surfactin	82.0	Leu Val Asp Leu Glu Leu Leu	84,4%
Bacillus_subtilis_KamC12	HPIC	surfactin	73.0	Glu Leu Leu Val Asn Leu Leu	84,4%
Bacillus_subtilis_Kam3	HPIC	surfactin	82.0	Leu Val Asp Leu Glu Leu Leu	84,4%
Bacillus_subtilis_8A	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	84,4%
Bacillus_subtilis_65BB	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	84,4%
Bacillus_subtilis_65AB	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	84,4%
Bacillus_subtilis_60B	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	84,4%
Bacillus_subtilis_58B	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	84,4%
Bacillus_subtilis_50_1B	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	84,4%
Bacillus_halotolerans_42A	HPIC	surfactin	86.0	Glu Leu Leu Val Asp Leu Leu	85,5%
Bacillus_halotolerans_41A	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	85,0%
Bacillus_halotolerans_40A	HPIC	surfactin	86.0	Glu Leu Leu Val Asp Leu Leu	85,0%
Bacillus_halotolerans_3F23	HPIC	surfactin	86.0	Glu Leu Leu Val Asp Leu Leu	87,3%

Bacillus_halotolerans_35A	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	83,9%
Bacillus_halotolerans_31A	HPIC	surfactin	82.0	Glu Leu Leu Val Asp Leu Leu	85,0%
Bacillus_halotolerans_15A	HPIC	surfactin	86.0	Glu Leu Leu Val Asp Leu Leu	85,0%
Bacillus_atrophaeus_ZY9	HPIC	surfactin	86.0	Ile Val Asp Leu Glu Leu Leu	81,0%
Bacillus_atrophaeus_P1F1	HPIC	surfactin	86.0	Ile Val Asp Leu Glu Leu Leu	81,0%
Bacillus_atrophaeus_AG8	HPIC	surfactin	86.0	Glu Leu Leu Val Asp Leu Ile	81,0%
Streptomyces_anulatus_T5	HPIC, HPIC-like, T1PKC	skyllamycin D/skylamycin E	38.0	mal Fo-OH-Orn X diOH-Bz X	51,3%
Streptomyces_anulatus_T5	T1PKC, HPIC-like, HPIC	skyllamycin D/skylamycin E	40.0	mal Fo-OH-Orn X diOH-Bz X	52,3%
Rhodococcus_opacus_Nah7IF	HPIC	rimosamide	14.0	X Thr Leu X Thr Val X Leu Ser Leu Ser Thr X X	55,9%
Streptomyces_anulatus_T5	HPIC	qinichelins	33.0	Ser OH-Orn Gly Val X Thr His Gly Gly Thr Ser	62,1%
Paeniacyllus_polymyxa_R6.14	HPIC	polymyxin	100.0	Dab Thr Dab Dab Dab Thr Leu Leu Dab Dab	80,5%
Paeniacyllus_polymyxa_R5.31	HPIC	polymyxin	100.0	Dab Thr Dab Dab Dab Thr Leu Thr Dab Dab	74,4%
Paeniacyllus_polymyxa_R4.5	HPIC	polymyxin	100.0	Dab Thr Dab Dab Dab Thr Phe Leu Dab Dab	80,8%
Paeniacyllus_polymyxa_R3.13	HPIC	polymyxin	100.0	Dab Thr Dab Dab Dab Thr Leu Leu Dab Dab	73,4%
Paeniacyllus_polymyxa_O2.11	HPIC	polymyxin	100.0	Dab Thr Dab Dab Dab Thr Leu Leu Dab Dab	80,5%
Paeniacyllus_polymyxa_O1.27	HPIC	polymyxin	100.0	Dab Thr Dab Dab Dab Thr Leu Leu Dab Dab	80,5%
Bacillus_velezensis_T5	HPIC	plipastatin	84.0	X Ile Tyr Glu	86,2%
Bacillus_velezensis_Kam11	HPIC	plipastatin	46.0	Glu	56,7%
Bacillus_velezensis_B1895	HPIC	plipastatin	61.0	X Orn Tyr Thr	61,7%
Pseudomonas_alloputida	NRP-metallophore, HPIC	Pf-5 pyoverdine	18.0	Gly Ser OH-Orn Asp Dab Asp Orn NH2	57,2%

A10				Glu Tyr Dab	
Paeniacyllus_alloputida_A10	NRP-metallophore, HPPIC	Pf-5 pyoverdine	18.0	Gly Ser OH-Orn Asp Dab Asp Orn NH2 Glu Tyr Dab	57,2%
Paeniacyllus_polymyxa_K1.14	transAT-PIKC, HPPIC	Pelgipeptin A/Pelgipeptin B/Pelgipeptin C/Pelgipeptin D	25.0	mal fatty_acid Gly Gly Dab Phe Pro X	61,2%
Paeniacyllus_polymyxa_R5.31	transAT-PIKC, HPPIC-like, HPPIC	paenilipoheptin	57.0	mal fatty_acid Ser Dab Trp	50,7%
Paeniacyllus_polymyxa_R4.5	HPPIC	paenilipoheptin	7.0	Gly Dab Gly Ser Trp Ser Dab Dab Ile Glu	51,7%
Paeniacyllus_polymyxa_R4.24	HPPIC	paenilipoheptin	7.0	Gly Dab Gly Ser Trp Ser Dab Dab Ile Glu	62,1%
Paeniacyllus_polymyxa_R4.24	transAT-PIKC, HPPIC-like, HPPIC	paenilipoheptin	76.0	mal fatty_acid Ser Dab Trp Ala Phe Tyr Glu	52,5%
Paeniacyllus_polymyxa_R4.24	HPPIC, HPPIC-like, transAT-PIKC	paenilipoheptin	53.0	mal fatty_acid Ser Dab Trp Ala Phe Tyr Glu	50,8%
Paeniacyllus_polymyxa_R4.24	transAT-PIKC, HPPIC-like, HPPIC	paenilipoheptin	65.0	mal fatty_acid Ser Dab Trp	52,3%
Paeniacyllus_polymyxa_K1.14	transAT-PIKC, HPPIC-like, HPPIC	paenilipoheptin	65.0	mal fatty_acid Ser Dab Trp	53,5%
Paeniacyllus_polymyxa_K1.14	HPPIC	paenilipoheptin	11.0	Ala Dab Gly Ser Trp Ser Dab Dab Ile Glu	59,1%
Paeniacyllus_polymyxa_K1.14	HPPIC, HPPIC-like, transAT-PIKC	paenilipoheptin	53.0	mal fatty_acid Ser Dab Trp Ala Phe Tyr Glu	56,1%
Paeniacyllus_polymyxa_K1.14	transAT-PIKC, HPPIC-like, HPPIC	paenilipoheptin	76.0	mal fatty_acid Ser Dab Trp Ala Phe Tyr Glu	56,8%
Rhodococcus_opacus_Nah7IF	HPPIC	olimycin A/olimycin B	5.0	X X X Thr Val X	75,3%
Paeniacyllus_polymyxa_R4.24	HPPIC	octapeptin C4	35.0	Dab Thr Dab Dab Dab Thr Phe Leu Dab Dab	80,7%
Rhodococcus_opacus_Nah7IF	HPPIC	No hit	0.0	X	53,1%
Rhodococcus_opacus_Nah7IF	HPPIC	No hit	0.0	X	53,0%
Paeniacyllus_polymyxa_R6.14	HPPIC, transAT-PIKC	No hit	0.0	Asn Gln X X Gly mal Lys X X	57,7%
Paeniacyllus_polymyxa_R6.14	transAT-PIKC, HPPIC	No hit	0.0	mal fatty_acid Tyr X Ser Ser Ser Ile Ser	53,2%
Paeniacyllus_polymyxa	HPPIC	No hit	0.0	Gly Dab Gly Ser Trp Ser Dab Dab Ile	64,6%

R6.14				Glu	
Paeniaceillus_polymyxa_R3.13	HPIC, transAT-PIKC	No hit	0.0	Asn Gln X X Gly mal Lys X X	59,6%
Paeniaceillus_polymyxa_R3.13	transAT-PIKC, HPIC	No hit	0.0	mal fatty_acid Tyr X Ser Ser Ser Ile Ser	53,2%
Paeniaceillus_polymyxa_O2.11	HPIC	No hit	0.0	Gly Dab Gly Ser Trp Ser Dab Dab Ile Glu	64,6%
Paeniaceillus_polymyxa_O2.11	HPIC, transAT-PIKC	No hit	0.0	Asn Gln X X Gly mal Lys X X	57,7%
Paeniaceillus_polymyxa_O2.11	transAT-PIKC, HPIC	No hit	0.0	mal fatty_acid Tyr X Ser Ser Ser Ile Ser	53,2%
Paeniaceillus_polymyxa_O1.27	HPIC, transAT-PIKC	No hit	0.0	Asn Gln X X Gly mal Lys X X	57,7%
Paeniaceillus_polymyxa_O1.27	HPIC	No hit	0.0	Gly Dab Gly Ser Trp Ser Dab Dab Ile Glu	64,6%
Paeniaceillus_polymyxa_O1.27	transAT-PIKC, HPIC	No hit	0.0	mal fatty_acid Tyr X Ser Ser Ser Ile Ser	53,2%
Bacillus_velezensis_R5	HPIC, transAT-PIKC	No hit	0.0	Asn Gly mal Glu	64,5%
Bacillus_velezensis_MT142	transAT-PIKC, HPIC-like, HPIC, betalactone	mycosubtilin	80.0	Tyr Asn Ser Asn fatty_acid mal X	63,4%
Bacillus_velezensis_MT141	transAT-PIKC, betalactone, HPIC-like, HPIC	mycosubtilin	80.0	X mal fatty_acid Asn Tyr Asn Ser	63,4%
Bacillus_velezensis_MT14	transAT-PIKC, betalactone, HPIC-like, HPIC	mycosubtilin	80.0	X mal fatty_acid Asn X Asn Ser	61,5%
Bacillus_velezensis_MT14	HPIC, HPIC-like, betalactone, transAT-PIKC	mycosubtilin	100.0	Glu Tyr X mal fatty_acid Asn X Asn Ser Glu Ser Thr	64,0%
Bacillus_atrophaeus_ZY9	transAT-PIKC, HPIC	mycosubtilin	40.0	mal fatty_acid Asn Tyr Asn Gln Pro Ser Asn	74,9%
Bacillus_atrophaeus_P1F1	transAT-PIKC, HPIC	mycosubtilin	40.0	mal fatty_acid Asn Tyr Asn Gln Pro Ser Asn	74,9%
Bacillus_atrophaeus_AG8	HPIC, transAT-PIKC	mycosubtilin	40.0	Ser Asn Tyr Asn Gln Pro fatty_acid Asn mal	74,8%
Bacillus_velezensis_V3.14	transAT-PIKC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_T5	transAT-PIKC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_R5	transAT-PIKC	macrolactin H	90.0	mal	67,7%
Bacillus_velezensis_R4.6	transAT-PIKC	macrolactin H	100.0	mal	64,3%

Bacillus_velezensis_R4	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_PZ2	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_MT86	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_MT64	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_MT61	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_MT55	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_MT42	transAT-ПКC	macrolactin H	100.0	mal	66,0%
Bacillus_velezensis_MT186	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_MT155	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_MT142	transAT-ПКC	macrolactin H	100.0	mal	66,0%
Bacillus_velezensis_MT141	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_MT14	transAT-ПКC	macrolactin H	100.0	mal	66,0%
Bacillus_velezensis_MT102	transAT-ПКC	macrolactin H	100.0	mal	63,1%
Bacillus_velezensis_Kam11	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_FD2	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_B1895	transAT-ПКC	macrolactin H	100.0	mal	64,3%
Bacillus_velezensis_7A	transAT-ПКC	macrolactin H	100.0	mal	63,1%
Bacillus_velezensis_43A	transAT-ПКC	macrolactin H	100.0	mal	63,1%
Bacillus_velezensis_21A	transAT-ПКC	macrolactin H	100.0	mal	63,1%
Bacillus_atrophaeus_AG8	T3ПКC, transAT-ПКC	macrolactin H	40.0	pk mal	51,3%
Bacillus_velezensis_PZ2	HPIC, HPIC-like, transAT-ПКC-like	locillomycin/locillomycin B/locillomycin C	100.0	Tyr Val Asp Val Asn Asp Gly Gln Asp Gly Thr fatty_acid	73,1%
Bacillus_velezensis_PZ2	HPIC-like, transAT-ПКC-like, HPIC	locillomycin/locillomycin B/locillomycin C	50.0	Gln Asp Gly Thr fatty_acid	57,8%

Bacillus_velezensis_PZ2	HPIC, HPIC-like, transAT-PIKC-like	locillomycin/locillomycin B/locillomycin C	100.0	Tyr Val Asp Val Asn Asp Gly Gln Asp Gly Thr fatty_acid	73,1%
Bacillus_velezensis_MT64	HPIC, transAT-PIKC	locillomycin/locillomycin B/locillomycin C	35.0	Val Phe Asp Asn Gly mal Glu	51,7%
Bacillus_safensis_CL35	HPIC	lichenysin	71.0	fatty_acid Val NH2 Val Ile Ile Asp Leu Glu Leu Leu	76,0%
Bacillus_safensis_AG46	HPIC	lichenysin	92.0	Glu Leu Leu Ile Asp Leu Ile NH2 Val fatty_acid Val	83,3%
Bacillus_pumilis_C76.2	HPIC	lichenysin	85.0	Glu Leu Leu Val Asp Leu Ile Acetyl-CoA Val fatty_acid Val	83,4%
Bacillus_licheniformis_CL56	HPIC	lichenysin	50.0	Gln Leu Leu	59,5%
Bacillus_licheniformis_CL33	HPIC	lichenysin	50.0	Gln Leu	63,0%
Rhodococcus_opacus_Nah7IF	HPIC	Legonmycin A/Legonmycin B/Legonindolizine A/Legonindolizine B	8.0	X X Leu X Thr X Leu Arg Leu Ser X X X	64,8%
Paeniacyllus_polymyxa_R5.31	HPIC, T1PIKC	laterocidine	5.0	pk X Thr Ser Orn Ser Ala Phe Gln Glu	53,0%
Paeniacyllus_polymyxa_R3.13	HPIC, T1PIKC	laterocidine	5.0	pk X Thr Ser Orn Ser Ala Phe Gln Glu	56,6%
Paeniacyllus_polymyxa_K1.14	HPIC, T1PIKC, prodigiosin	laterocidine	5.0	pk X Thr Ser Orn Ser Ala Phe Gln Glu	55,8%
Bacillus_velezensis_MT42	transAT-PIKC, betalactone, HPIC-like, HPIC	iturin	88.0	X mal fatty_acid Asn Tyr Asn Ser	66,0%
Rhodococcus_opacus_Nah7IF	HPIC	glycopeptidolipid	7.0	Thr X Leu Arg Leu Ser X X X	65,9%
Paeniacyllus_polymyxa_R6.14	HPIC	fusaricidin B	100.0	X Thr Val X Thr Asn X	81,8%
Paeniacyllus_polymyxa_R5.31	HPIC	fusaricidin B	100.0	X Thr Val X Thr Asn X	81,8%
Paeniacyllus_polymyxa_R4.5	HPIC	fusaricidin B	100.0	X Thr Val X Thr Asn X	82,4%
Paeniacyllus_polymyxa_R4.24	HPIC	fusaricidin B	100.0	X Thr Val X Thr Asn X	82,4%
Paeniacyllus_polymyxa_R3.13	HPIC	fusaricidin B	100.0	X Thr Val X Thr Asn X	82,8%
Paeniacyllus_polymyxa	HPIC	fusaricidin B	100.0	X Thr Val X Thr Asn X	81,8%

O2.11					
Paeniacyllus_polymyxa_O1.27	HPIC	fusaricidin B	100.0	X Thr Val X Thr Asn X	81,8%
Paeniacyllus_polymyxa_K1.14	HPIC	fusaricidin B	100.0	X Thr Val X Thr Asn X	82,4%
Bacillus_velezensis_V3.14	HPIC, betalactone	fengycin	93.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	84,2%
Bacillus_velezensis_V3.14	HPIC, transAT-PIKC, betalactone	fengycin	100.0	Asn Ser Tyr Asn Gln Pro fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	67,2%
Bacillus_velezensis_T5	HPIC, transAT-PIKC, betalactone	fengycin	93.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X Ile Tyr Glu	62,4%
Bacillus_velezensis_R5	HPIC, betalactone	fengycin	80.0	X Ile Pro Gln Tyr X	66,8%
Bacillus_velezensis_R5	HPIC, transAT-PIKC, betalactone	fengycin	86.0	Asn Ser Tyr Asn Gln Pro fatty_acid Asn mal X Ile Pro Gln Tyr X	70,0%
Bacillus_velezensis_R4.6	HPIC, betalactone	fengycin	93.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	84,2%
Bacillus_velezensis_R4.6	HPIC, transAT-PIKC, betalactone	fengycin	100.0	Asn Ser Tyr Asn Gln Pro fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	67,2%
Bacillus_velezensis_R4	HPIC, betalactone, transAT-PIKC	fengycin	80.0	Tyr Ile X mal fatty_acid Asn Tyr Asn Ser Glu Ser Thr	61,5%
Bacillus_velezensis_R4	HPIC, betalactone	fengycin	80.0	Tyr Ile X	58,8%
Bacillus_velezensis_PZ2	HPIC, betalactone	fengycin	93.0	Glu Orn Tyr Thr Glu X Pro Gln Tyr Ile X	80,6%
Bacillus_velezensis_MT86	HPIC, transAT-PIKC, betalactone	fengycin	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X Ile Tyr Glu	63,8%
Bacillus_velezensis_MT86	HPIC, betalactone	fengycin	86.0	X Ile Tyr Glu	77,4%
Bacillus_velezensis_MT64	HPIC, transAT-PIKC, betalactone	fengycin	93.0	Asn Ser Tyr Asn Gln Pro fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	65,1%
Bacillus_velezensis_MT64	HPIC, betalactone	fengycin	86.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	71,2%
Bacillus_velezensis_MT61	HPIC, transAT-PIKC, betalactone	fengycin	100.0	Ser Thr Tyr Asn Pro Glu fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	66,5%
Bacillus_velezensis_MT	HPIC, betalactone	fengycin	93.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu	80,6%

61				Orn	
Bacillus_velezensis_MT55	HPIC, transAT-IPKC, betalactone	fengycin	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	66,5%
Bacillus_velezensis_MT55	HPIC, betalactone	fengycin	93.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	80,6%
Bacillus_velezensis_MT42	HPIC, betalactone	fengycin	86.0	Glu Tyr Ile X	77,9%
Bacillus_velezensis_MT186	HPIC, betalactone	fengycin	100.0	X Ile Tyr Glu	77,4%
Bacillus_velezensis_MT186	HPIC, transAT-IPKC, betalactone	fengycin	93.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X Ile Tyr Glu	63,3%
Bacillus_velezensis_MT155	HPIC, betalactone	fengycin	93.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	80,6%
Bacillus_velezensis_MT155	HPIC, transAT-IPKC, betalactone	fengycin	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	66,5%
Bacillus_velezensis_MT142	HPIC, transAT-IPKC, HPIC-like, betalactone	fengycin	100.0	Thr Ser Tyr Asn Ser Glu Asn fatty_acid mal X Ile Tyr Glu	64,7%
Bacillus_velezensis_MT142	HPIC, betalactone	fengycin	93.0	X Ile Tyr Glu	77,4%
Bacillus_velezensis_MT141	HPIC, betalactone	fengycin	86.0	Glu Tyr X X	77,9%
Bacillus_velezensis_MT141	HPIC, betalactone, transAT-IPKC, HPIC-like	fengycin	93.0	Glu Tyr Ile X mal fatty_acid Asn Tyr Asn Ser Glu Ser Thr	62,5%
Bacillus_velezensis_MT14	HPIC, HPIC-like, betalactone	fengycin	93.0	Glu Tyr X	75,9%
Bacillus_velezensis_MT102	HPIC, betalactone	fengycin	93.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	80,6%
Bacillus_velezensis_MT102	HPIC, transAT-IPKC, betalactone	fengycin	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	68,5%
Bacillus_velezensis_Kam11	HPIC, betalactone	fengycin	86.0	X Ile Pro Gln Tyr Glu X Tyr Thr Orn	71,2%
Bacillus_velezensis_Kam11	HPIC, transAT-IPKC, betalactone	fengycin	93.0	Ser Thr Tyr Asn Pro Glu fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Orn	64,9%
Bacillus_velezensis_FD2	HPIC, HPIC-like,	fengycin	93.0	X Ile Pro Gln Tyr Glu X Thr Tyr Orn	81,0%

	betalactone			Glu	
Bacillus_velezensis_FD2	HPIC-like, HPIC, betalactone	fengycin	80.0	X Pro Gln Tyr Glu X	66,3%
Bacillus_velezensis_FD2	HPIC, transAT-HPIC, HPIC-like, betalactone	fengycin	100.0	Ser Thr Tyr Asn Ser Glu Asn fatty_acid mal X Ile Pro Gln Tyr Glu X Thr Tyr Orn Glu	66,6%
Bacillus_velezensis_B1895	HPIC, betalactone, transAT-HPIC	fengycin	86.0	Tyr Ile X mal fatty_acid Asn Tyr Asn Gln Pro Asn Ser	64,1%
Bacillus_velezensis_B1895	HPIC, betalactone	fengycin	86.0	Tyr Ile X	59,3%
Bacillus_velezensis_7A	HPIC, transAT-HPIC, betalactone	fengycin	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	66,8%
Bacillus_velezensis_7A	HPIC, betalactone	fengycin	93.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	79,0%
Bacillus_velezensis_43A	HPIC, transAT-HPIC, betalactone	fengycin	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	66,8%
Bacillus_velezensis_43A	HPIC, betalactone	fengycin	93.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	79,0%
Bacillus_velezensis_21A	HPIC, betalactone	fengycin	93.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	79,0%
Bacillus_velezensis_21A	HPIC, transAT-HPIC, betalactone	fengycin	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	66,8%
Bacillus_subtilis_MT74	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu Val Tyr Thr Glu Orn	83,4%
Bacillus_subtilis_MT48	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu Val Tyr Thr Glu Orn	83,4%
Bacillus_subtilis_KATMIRA	HPIC, betalactone	fengycin	100.0	Glu Orn Tyr Thr Glu Val Pro Glu Tyr Ile X	83,4%
Bacillus_subtilis_KamC12	HPIC, betalactone	fengycin	93.0	Glu Tyr Ile X	77,4%
Bacillus_subtilis_Kam3	HPIC, betalactone	fengycin	100.0	Glu Orn Tyr Thr Glu Val Pro Glu Tyr Ile X	83,4%
Bacillus_subtilis_8A	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu Val Tyr Thr Glu Orn	83,4%
Bacillus_subtilis_65BB	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu Val Tyr Thr Glu	83,4%

				Orn	
Bacillus_subtilis_65AB	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu Val Tyr Thr Glu Orn	83,4%
Bacillus_subtilis_60B	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu Val Tyr Thr Glu Orn	83,4%
Bacillus_subtilis_58B	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu Val Tyr Thr Glu Orn	83,4%
Bacillus_subtilis_50_1B	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu Val Tyr Thr Glu Orn	83,4%
Bacillus_halotolerans_42A	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu X Tyr Thr Glu Orn	83,0%
Bacillus_halotolerans_41A	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu X Tyr Thr Glu Orn	83,0%
Bacillus_halotolerans_40A	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu X Tyr Thr Glu Orn	82,3%
Bacillus_halotolerans_3F23	HPIC, betalactone	fengycin	100.0	Glu Orn Tyr Thr Glu X Pro Glu Tyr Ile X	83,0%
Bacillus_halotolerans_35A	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu X Tyr Thr Glu Orn	82,3%
Bacillus_halotolerans_31A	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu X Tyr Thr Glu Orn	83,0%
Bacillus_halotolerans_15A	HPIC, betalactone	fengycin	100.0	X Ile Pro Glu Tyr Glu X Tyr Thr Glu Orn	83,0%
Bacillus_atrophaeus_ZY9	transAT-PIKC, HPIC, betalactone	fengycin	100.0	mal fatty_acid Asn Tyr Asn Gln Pro Ser Asn Glu Orn Tyr Thr Glu X Pro Gln Tyr Ile X	75,5%
Bacillus_atrophaeus_ZY9	HPIC, betalactone	fengycin	100.0	Glu Orn Tyr Thr Glu X Pro Gln Tyr Ile X	85,5%
Bacillus_atrophaeus_P1F1	HPIC, betalactone	fengycin	100.0	Glu Orn Tyr Thr Glu X Pro Gln Tyr Ile X	85,5%
Bacillus_atrophaeus_P1F1	transAT-PIKC, HPIC, betalactone	fengycin	100.0	mal fatty_acid Asn Tyr Asn Gln Pro Ser Asn Glu Orn Tyr Thr Glu X Pro Gln Tyr Ile X	75,5%
Bacillus_atrophaeus_AG8	HPIC, betalactone	fengycin	100.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn	83,0%
Bacillus_atrophaeus_AG8	HPIC, transAT-PIKC, betalactone	fengycin	100.0	X Ile Pro Gln Tyr Glu X Tyr Thr Glu Orn Ser Asn Tyr Asn Gln Pro fatty_acid Asn mal	71,7%

Bacillus_velezensis_R5	transAT-PIKC	difficidin	46.0	X mal	62,6%
Bacillus_velezensis_R4	transAT-PIKC	difficidin	46.0	X mal	57,7%
Bacillus_velezensis_MT102	transAT-PIKC	difficidin	60.0	mal X	59,5%
Streptomyces_anulatus_T5	transAT-PIKC, HPIC-like, lanthipeptide-class-iv	cycloheximide	66.0	X mal	78,7%
Streptomyces_anulatus_T5	HPIC-like, lanthipeptide-class-iv, transAT-PIKC	cycloheximide	66.0	X mal	63,6%
Rhodococcus_opacus_Nah7IF	HPIC	corynecin III/corynecin I/corynecin II	46.0	X X Ser Val X X Ser X Ser X Gln X	58,9%
Bacillus_velezensis_V3.14	HPIC, transAT-PIKC, betalactone	bacillomycin D	100.0	Asn Ser Tyr Asn Gln Pro fatty_acid Asn mal X	63,8%
Bacillus_velezensis_T5	HPIC, transAT-PIKC, betalactone	bacillomycin D	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X	63,1%
Bacillus_velezensis_R5	HPIC, transAT-PIKC, betalactone	bacillomycin D	100.0	Asn Ser Tyr Asn Gln Pro fatty_acid Asn mal X	72,4%
Bacillus_velezensis_R4.6	HPIC, transAT-PIKC, betalactone	bacillomycin D	100.0	Asn Ser Tyr Asn Gln Pro fatty_acid Asn mal X	63,8%
Bacillus_velezensis_R4	transAT-PIKC, betalactone, HPIC	bacillomycin D	100.0	X mal fatty_acid Asn Tyr Asn Ser Glu Ser Thr	62,8%
Bacillus_velezensis_PZ2	transAT-PIKC, betalactone, HPIC	bacillomycin D	100.0	X mal fatty_acid Asn Tyr Asn Ser Glu Ser Thr	62,8%
Bacillus_velezensis_MT86	HPIC, transAT-PIKC, betalactone	bacillomycin D	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X	65,4%
Bacillus_velezensis_MT64	HPIC, transAT-PIKC, betalactone	bacillomycin D	100.0	Asn Ser Tyr Asn Gln Pro fatty_acid Asn mal X	63,8%
Bacillus_velezensis_MT61	HPIC, transAT-PIKC, betalactone	bacillomycin D	100.0	Ser Thr Tyr Asn Pro Glu fatty_acid Asn mal X	66,0%
Bacillus_velezensis_MT55	HPIC, transAT-PIKC, betalactone	bacillomycin D	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X	63,1%
Bacillus_velezensis_MT42	HPIC, betalactone, transAT-PIKC, HPIC-like	bacillomycin D	100.0	Glu Tyr Ile X mal fatty_acid Asn Tyr Asn Ser Glu Ser Thr	65,9%
Bacillus_velezensis_MT42	HPIC, HPIC-like, transAT-PIKC	bacillomycin D	100.0	mal fatty_acid Asn Tyr Asn Ser Glu Ser Thr	74,3%
Bacillus_velezensis_MT186	HPIC, transAT-PIKC, betalactone	bacillomycin D	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X	62,8%
Bacillus_velezensis_MT155	HPIC, transAT-PIKC, betalactone	bacillomycin D	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X	63,1%

Bacillus_velezensis_MT142	HPIC, transAT-PIKC	HPIC-like,	bacillomycin D	100.0	Thr Ser Tyr Asn Ser Glu Asn fatty_acid mal	67,7%
Bacillus_velezensis_MT141	HPIC, transAT-PIKC	HPIC-like,	bacillomycin D	100.0	mal fatty_acid Asn Tyr Asn Ser Glu Ser Thr	66,9%
Bacillus_velezensis_MT14	HPIC, transAT-PIKC	HPIC-like,	bacillomycin D	100.0	mal fatty_acid Asn X Asn Ser Glu Ser Thr	69,9%
Bacillus_velezensis_MT102	HPIC, betalactone	transAT-PIKC,	bacillomycin D	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X	65,8%
Bacillus_velezensis_Kam11	HPIC, betalactone	transAT-PIKC,	bacillomycin D	100.0	Ser Thr Tyr Asn Pro Glu fatty_acid Asn mal X	66,0%
Bacillus_velezensis_FD2	HPIC, betalactone	transAT-PIKC,	bacillomycin D	100.0	Ser Thr Tyr Asn Ser Glu Asn fatty_acid mal X	61,9%
Bacillus_velezensis_B1895	transAT-PIKC, HPIC	betalactone,	bacillomycin D	100.0	X mal fatty_acid Asn Tyr Asn Gln Pro Asn Ser	63,8%
Bacillus_velezensis_7A	HPIC, betalactone	transAT-PIKC,	bacillomycin D	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X	62,8%
Bacillus_velezensis_43A	HPIC, betalactone	transAT-PIKC,	bacillomycin D	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X	62,8%
Bacillus_velezensis_21A	HPIC, betalactone	transAT-PIKC,	bacillomycin D	100.0	Ser Thr Tyr Asn Ser Glu fatty_acid Asn mal X	62,8%
Bacillus_atrophaeus_AG8	transAT-PIKC, PKC-like, T3PKC	T1PKC,	aurantinin B/aurantinin C/aurantinin D	21.0	pk pk mal	53,4%
Bacillus_atrophaeus_AG8	PKC-like, T1PKC, transAT-PIKC, T3PKC		aurantinin B/aurantinin C/aurantinin D	21.0	pk pk mal	53,4%
Streptomyces_anulatus_T5	HPIC		A-500359 A/A-500359 B	8.0	Gln Gly Pro mal X Pro	57,5%
Streptomyces_anulatus_T5	HPIC, T1PKC		10-epi-HSAF/10-epi-3-deOH-HSAF/10-epi-maltophilin/10-epi-xanthobaccin C/10-epi-hydroxymaltophilin/10-epi-FI-2	100.0	pk Orn	77,7%