

*Федеральное государственное автономное образовательное
учреждение высшего образования
«ЮЖНЫЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»*

На правах рукописи



КУЛИКОВ МАКСИМ ПАВЛОВИЧ

**МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИЙ ПОДХОД К ПОИСКУ НОВЫХ
НЕРИБОСОМАЛЬНО СИНТЕЗИРУЕМЫХ ПРОТИВОГРИБКОВЫХ МЕТАБОЛИТОВ
БАКТЕРИЙ**

1.5.7. Генетика (биологические науки)

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата биологических наук

Ростов-на-Дону
2026

Работа выполнена в лаборатории экспериментального мутагенеза, лаборатории «Биотехнологический кластер» и на кафедре генетики Академии биологии и медицины им. Д.И. Ивановского Федерального государственного автономного образовательного учреждения высшего образования «Южный федеральный университет», а также в Научно-исследовательской лаборатории «Центр агробиотехнологий» в Донском государственном техническом университете (НИЛ ЦАБТ ДГТУ).

Научный руководитель:

Празднова Евгения Валерьевна,

доктор биологических наук, заведующая лабораторией «Биотехнологический кластер» Академии биологии и медицины им. Д.И. Ивановского ЮФУ (Ростов-на-Дону)

Официальные оппоненты:

Манухов Илья Владимирович, доктор биологических наук, профессор, заведующий лабораторией молекулярной генетики Московского физико-технического университета (МФТИ)

Марсова Мария Викторовна, кандидат биологических наук, старший научный сотрудник лаборатории генетики микроорганизмов Федерального государственного бюджетного учреждения науки Институт общей генетики им. Н.И. Вавилова Российской академии наук

Защита диссертации состоится **30 сентября 2026 г. в 10:00** на заседании диссертационного совета ЮФУ 801.01.23 по биологическим наукам на базе Академии биологии и медицины им. Д.И. Ивановского Южного федерального университета по адресу: 344090, г. Ростов-на-Дону, пр. Стачки 194/1, к. 712.

С диссертацией можно ознакомиться в Зональной научной библиотеке им. Ю.А. Жданова Южного федерального университета по адресу: 344090, г. Ростов-на-Дону, ул. Р. Зорге, 21Ж и на сайте Южного федерального университета <https://hub.sfedu.ru/diss/show/1356958/>

Автореферат разослан «__» _____ 2026 г.

Отзыв на автореферат в 2-х экз. (с указанием даты, полностью ФИО, учёной степени со специальностью, звания, организации, подразделения, должности, адреса, телефона, e-mail), заверенный печатью организации, просим направлять по адресу: 344090, г. Ростов-на-Дону, пр. Стачки, 194/1, к. 106, ученому секретарю диссертационного совета ЮФУ 801.01.23 Хмелевцовой Л.Е., а также в формате .pdf на e-mail: lehmelevcova@sfedu.ru.

Ученый секретарь
диссертационного совета,
кандидат биологических наук, с.н.с.

Л.Е. Хмелевцова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

Грибковые заболевания растений, животных и людей являются серьезной проблемой, в том числе из-за постепенного развития устойчивости у грибковых патогенов к лекарственным средствам (Pristov и др., 2019).

Среди фитопатогенных грибов одним из наиболее часто встречающихся родов является род *Fusarium*. Все злаковые культуры подвержены фузариозу, однако наиболее часто встречаются фузариозы пшеницы, которые могут приводить более чем 50% прямым потерям урожая и снижению качества собранного зерна (Khaytbaeva и др., 2021; Nicholson и др., 2007).

Серьезную угрозу для аквакультуры представляют патогены рода *Saprolegnia*, вызывающие массовую гибель пресноводных рыб и их икры (Engblom и др., 2023). Контроль сапролегниоза осложняется адаптацией возбудителей и строгим запретом на использование эффективных, но токсичных и канцерогенных препаратов (например, малахитового зеленого), что требует поиска новых безопасных альтернатив (Tang и др., 2024).

Ввиду возрастающей устойчивости грибковых патогенов требуются все новые и новые соединения для 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 пептидсинтетазы (НРПС) (Wang и др., 2014).

В настоящее время в большинстве случаев поиск новых противогрибковых соединений проводится путем функционального скрининга бактерий-продуцентов *in vitro*. Этот подход позволил обнаружить подавляющее большинство из ныне известных классов антибиотиков (Aminov, 2017). Однако мы полагаем, что 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 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 и др., 2022), либо наоборот, гены синтеза метаболитов известны лишь гипотетически по данным секвенирования (Agrawal и др., 2021; Blin и др., 2021; Medema и др., 2011), но их продукты не выделены и функции неизвестны, а их структура не подтверждена ни экспериментально, ни расчетно.

Разработка алгоритма для поиска генов нерибосомального синтеза и предсказания противогрибковых свойств метаболитов *in silico* 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

Достаточно хорошо изучены основные структурные элементы наиболее распространенных нерибосомальных пептид-синтетаз - домены аденилирования, пептидил-переноса, конденсации, эпимеризации и тиюэстеразный домен, изучены механизмы активации и

переноса аминокислот, а также формирования пептидных связей, установлены связи между структурой генов и производимыми пептидами (Challis и др., 2004; Dowling и др., 2016; Marahiel, 2016; Weissman, 2015). 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 (Aleti, 2016; Lee и др., 2013; Ongena и др., 2008; Sumi и др., 2015; Tajbakhsh и др., 2017; Tang и др., 2014). 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 и др., 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.

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 синтетаз.
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 пептидсинтетазы 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 пептидсинтетаз. Найден новый лихенизин-подобный противогрибковый метаболит, синтезируемый штаммом *Bacillus safensis* AG46.

Положения, выносимые на защиту

1. Субстратная специфичность доменов аденилирования нерибосомальных пептидсинтетаз 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 8 научных работ, включая 3 статьи в журналах, входящих в базы данных международных индексов научного цитирования Scopus и Web of Science и 1 статью в журналах, входящих в Перечень рецензируемых научных изданий ВАК.

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 (182 источника). 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.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

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 frigoritolerans*, *Pseudomonas allopuntida*, *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.

Выделение метаболитов и масс-спектрометрия

Для активации экспрессии генов NRPS бактериальные штаммы совместно культивировали с грибом *Fusarium graminearum* Fus1 в течение 3–5 суток при 25 °С. Собранную биомассу переносили в жидкую среду для получения инокулята (18–24 ч, 30 °С, 180 об/мин). Затем 5% инокулята вносили в модифицированную среду Лэнди и культивировали трое суток.

Выделение и очистка липопептидов

Бесклеточный супернатант отделяли центрифугированием (6000×g, 30 мин) с последующей ультрафильтрацией (Stericup, Millipore). Целевые пептиды осаждали при pH 2,0 в течение ночи при 4 °С. Осадок собирали центрифугированием (12000×g, 10 мин), промывали подкисленной водой (pH 2,0) и двукратно экстрагировали метанолом. Объединенные метанольные экстракты упаривали досуха на ротаторном испарителе при 45–50 °С, после чего взвешенный сухой остаток растворяли в 96%-ном этаноле.

Масс-спектрометрический анализ (MALDI-TOF)

Полученную смесь очищенных пептидов анализировали методом времяпролетной 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).

Выделение ДНК

Выделение геномной ДНК осуществляли методом ЦТАБ (Murray, Thompson, 1980; Doyle, Doyle, 1987) из бактериальной биомассы, выращенной в среде LB (37 °C, 24–72 ч). Клетки, предварительно отмытые в TE-bio буфере (с добавлением 2% ПВП для почвенных изолятов), подвергали ферментативному лизису смесью лизоцима (10 мг/мл) и РНКазы А при 37 °C в течение 30–60 мин, после чего последовательно обрабатывали 5 М NaCl, протеиназой К (2 мг/мл) и 10% раствором ЦТАБ (56 °C, ≥ 60 мин). Очистку лизата проводили двукратной экстракцией хлороформом, затем ДНК осаждали изопропанолом (1,5 объема), промывали 80% этанолом и растворяли в TE-НК буфере. Концентрацию и чистоту препаратов (для повышения выхода которых рекомендовался цикл замораживания–оттаивания осадка) оценивали спектрофотометрически на приборе NanoPhotometer P330 по соотношению A260/A280, а также электрофоретически в 1,5% агарозном геле.

Секвенирование

Подготовку библиотек для секвенирования Illumina проводили с помощью QIAseq FX DNA Library kit (Qiagen) под контролем распределения фрагментов на Agilent Bioanalyzer 2100, после чего секвенировали на HiSeq 2500 в режиме парных прочтений (100 п.о.); сырые данные конвертировали в FASTQ-формат (bcl-convert) и валидировали в FastQC и MultiQC (доля нуклеотидов с Q > 30 составила 85,4–88,1%). Для нанопорового секвенирования высокомолекулярную ДНК (>3 тыс. п.о.) выделяли набором MagAttract HMW DNA Kit (Qiagen), оценивали на флуориметре Qubit 3.0, очищали магнитными шариками Ampure XP и мультиплексировали с использованием наборов Rapid Barcoding (SQK-RBK114) или Native Barcoding (EXP-NBD104 + SQK-LSK109). Секвенирование осуществляли на платформе MinION с регистрацией данных в формате POD5, после чего проводили высокоточный бейсколлинг в dorado до формата FASTQ, первичный контроль качества в NanoPlot v1.42.0 и русоQC, а также фильтрацию прочтений через Filtlong v0.2.1.

Анализ геномов

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. Таксономическое определение штаммов по данным полногеномного секвенирования выполняли несколькими независимыми методами. 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 et al., 2022).

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. Полный перечень соединений представлен в Приложении 1 диссертации.

Структурно-функциональная аннотация отобранных генных кластеров проводилась с использованием программного комплекса antiSMASH версии 8.0.0. Для формирования признакового пространства базы данных из полученных 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 пептидсинтетаз (НРПС).

Биоинформатический анализ и машинное обучение

Для построения признакового пространства из результатов аннотации извлекали *pfam_id*, *smcogs_id*, домены, предсказанные мономеры и классы аминокислот. Данные векторизовали в бинарную матрицу с помощью CountVectorizer (scikit-learn), отсекая шумы фильтрацией по частоте встречаемости ($\text{min_df}=0.04$, $\text{max_df}=0.8$). Ввиду несбалансированности выборки и отсутствия размеченного негативного класса задачу классификации решали в парадигме PU-обучения (Positive-Unlabeled Learning). Протестированы алгоритмы Random Forest (со взвешиванием классов), Balanced Random Forest и Balanced Bagging Classifier (imblearn). Финальная модель построена в виде ансамбля на основе усреднения вероятностей (soft voting) двух последних классификаторов. Оптимизацию гиперпараметров проводили по кастомной метрике PU F-score. Важность признаков оценивали методом SHAP (алгоритмы TreeExplainer и PermutationExplainer), а также по снижению критерия Джини (MDI). Статистическую значимость различий в частотах признаков определяли точным критерием Фишера с поправкой Бенджамини–Хохберга на множественное тестирование ($p < 0.05$).

Пайплайн включал двухэтапное сканирование геномов платформой antiSMASH (с модулями prodigal, clusterhmmr и ss-mibig): сначала для полногеномной локализации БГК, затем — для изолированной реаннотации извлеченных подкластеров по строгим границам. Специализированный парсер агрегировал JSON-отчеты antiSMASH в единую матрицу признаков. На ее основе обученный ансамбль моделей ранжировал новые БГК (не имеющие хитов в MIBiG) по предсказанной вероятности противогрибковой активности.

Тестирование антагонистической и противогрибковой активности *in vitro*

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 использовалась стерильная питательная среда.

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

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 *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.

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.

Анализ экспрессии генов

Тотальную РНК из зафиксированных в IntactRNA мышечной ткани стерляди и замороженных при -80 °С 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 АСТВ в качестве референсного гена, для растений — генов патогенез-ассоциированных белков (PR-1, PR-2, PR-3), в качестве референсного использовался ген β -тубулина. 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).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

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 диссертации. Результаты таксономической идентификации с использованием базы данных GTDB представлены в Приложении 3 диссертации.

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 на наличие известных биосинтетических кластеров генов с помощью antiSMASH. Профили двух штаммов *B. velezensis* (R4.6 и V3.14) 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). Для найденных соединений были предсказаны масса и подверженность кислотному осаждению (Таблица 1).

Таблица 1. 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

Соединение	Осаждение при pH 2.0	Масса / Диапазон масс (Да)
пелгипептин A/B/C/D	Нет	1057.6, 1071.6, 1085.6
пенилодин	Нет	2326.1
пенилогептин	Возможно	1099.6
пенилан	Нет	2509.2
пеницидин A, B	Нет	3374.5
пенибациллин	Нет	194.5
октапептин C4	Нет	1023.7
макролактин H	Нет	402.2
фузарицидин B	Да	881.5, 895.5, 909.5
фенгицин	Да	1462.8, 1476.8, 1490.8
диффицидин	Возможно	544.3
бовинимид A	Возможно	710.4
богорол A	Нет	1583.9
бацилизин	Нет	270.1
бациллибактин	Нет	882.2
бациллен	Нет	580.3
аурантинин B/C/D	Возможно	780.4, 766.4

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

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 фенгицину 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 аминокислотные группы 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. polytuxa*.

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.

Разработка улучшенного алгоритма AFPredict

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 признаков: 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 2.

Таблица 2. 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.

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); Гидроксилированные 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 (NRPS, PKS, Hybrids) представлены на рисунках 1, 2 и 3.

В целом можно обобщить, что алифатические и разветвленные гидрофобные аминокислоты на 38% чаще встречаются среди противогрибковых соединений. 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.

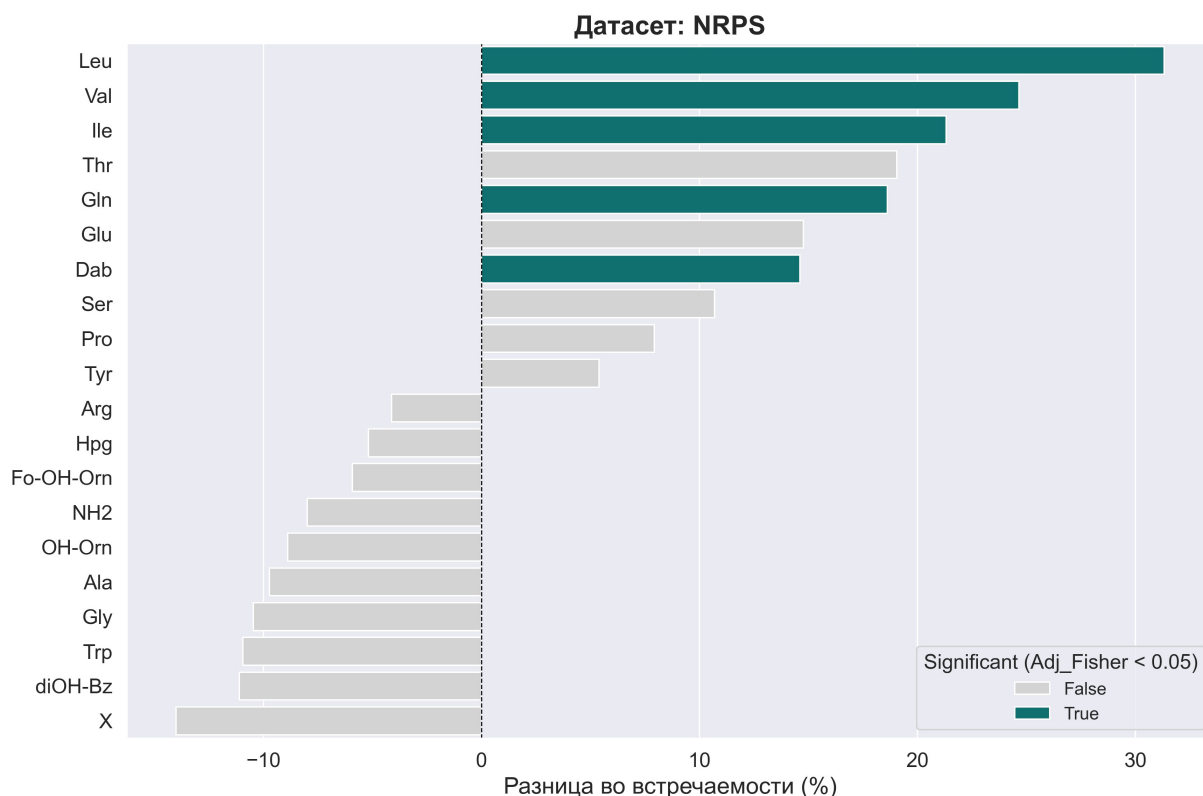


Рисунок 1. Частота встречаемости признака в зависимости от того, является соединение противогрибковым или нет.

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

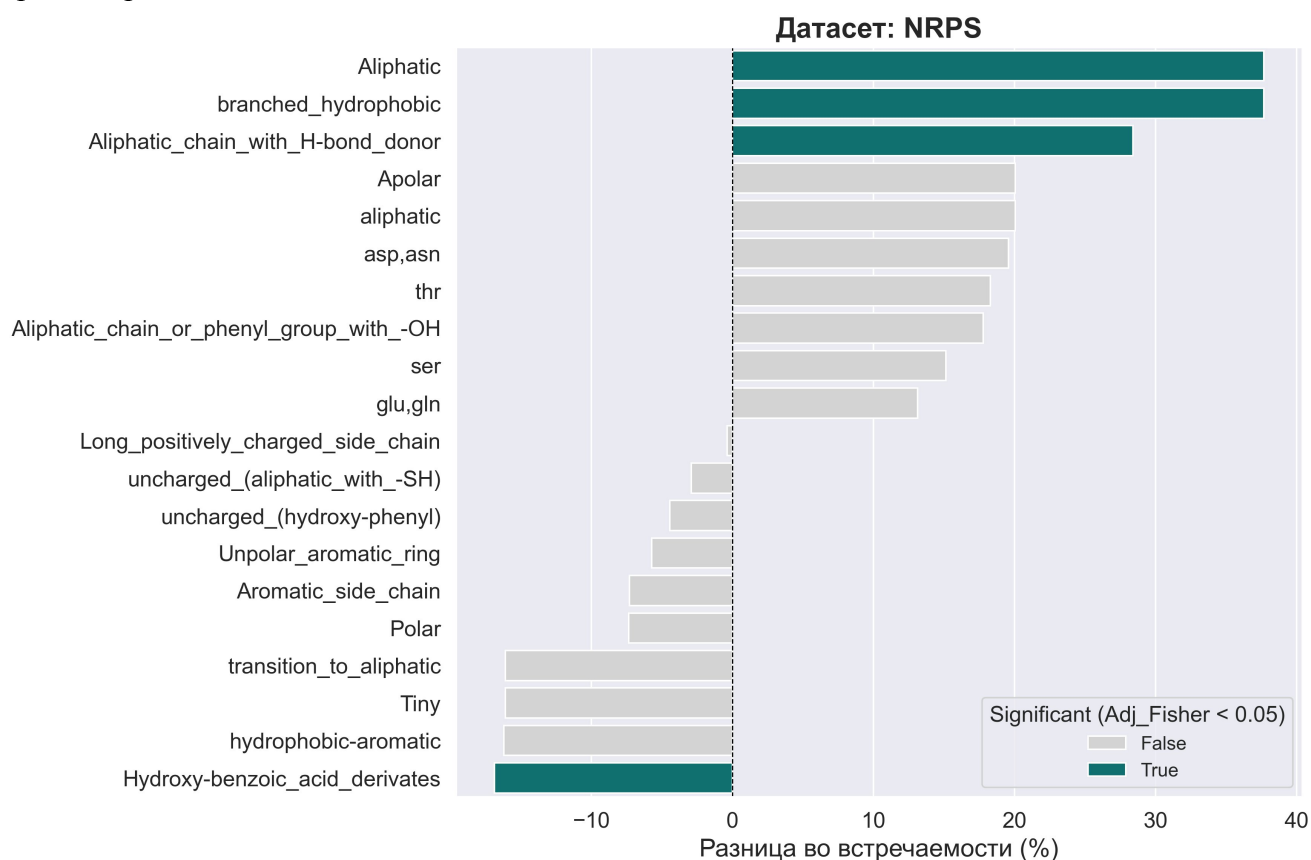


Рисунок 2. Частота встречаемости признака в зависимости от того, является соединение противогрибковым или нет.

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

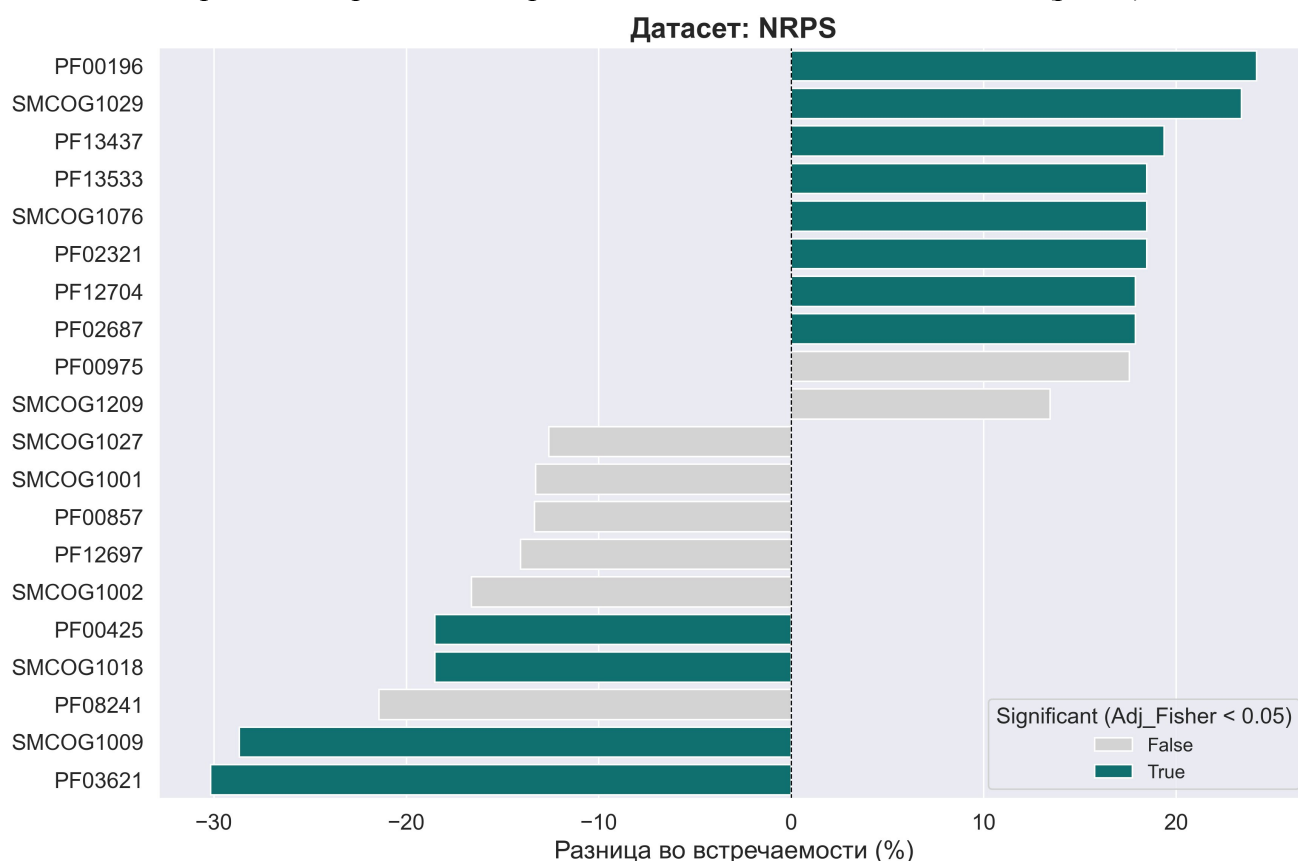


Рисунок 3. Частота встречаемости признака в зависимости от того, является соединение противогрибковым (правая часть) или нет (левая).

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, ассоциированными 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.

Машинное обучение

Было разработано несколько моделей машинного обучения, способных предсказывать противогрибковые свойства нерибосомально синтезируемых соединений (рис. 5). 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).

С целью интерпретации результатов классификации и оценки качества модели был проведен анализ векторов Шепли (SHAP Values) с использованием алгоритма TreeExplainer. Визуализация результатов в виде графиков 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. 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_pks_pp`, кеторедуктазный домен (KR domain PF08659) и бета-кетоацилсинтазу (SMCOG1022). Это подчеркивает высокую долю гибридных (NRPS-PKS) или исключительно поликетидных структур среди успешных антимикотиков.

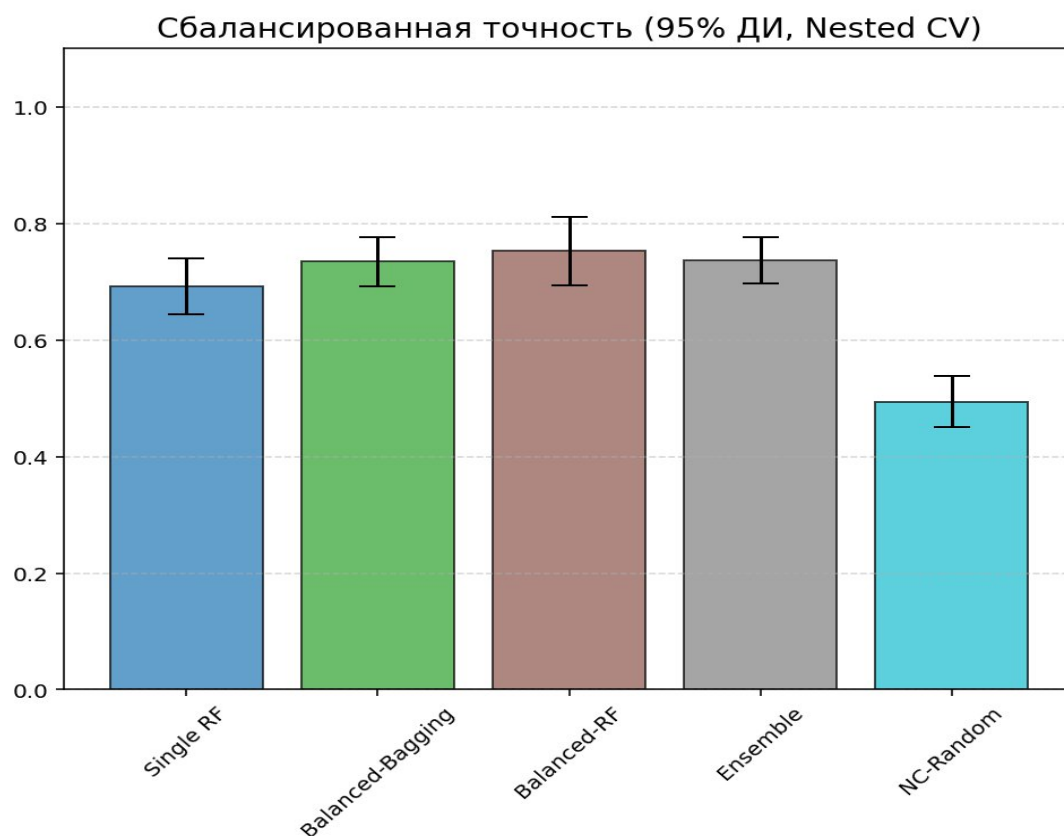


Рисунок 5. Оценка качества моделей

Наряду с этим обнаружена сильная ассоциация с присутствием специфических мономеров, в частности орнитинных 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 синтаз из БКГ, содержащих более одной. Существующий на данный момент алгоритм AntiSMASH v8.0.0 при аннотировании

руководствуется принципом «1 БГК = одна синтетаза», но этот подход устарел и не соответствует актуальным данным (Gulick A. M., 2017).

Аннотация геномов новых штаммов с помощью AFPredictor

С использованием разработанной модели AFPredictor была проведена аннотация БГК из геномов 69 штаммов из коллекции лаборатории. За счет поиска по подкластерам, было найдено 1096 соединений, на 104 больше, чем с помощью antiSMASH. Часть из них - это терпены (41) и кластеры генов рибосомально синтезируемых пептидов (23). Было показано, 54 штамма имеют хотя бы 1 БГК, содержащий более одного «скрытого» подкластера, не обнаруживаемого традиционными методами анализа. Данные показывают четкие паттерны совместной локализации (ко-локализации) путей биосинтеза специфических соединений внутри одних и тех же кластеров. Особенно ярко это выражено у штаммов *Bacillus velezensis* и *Bacillus atrophaeus*. Алгоритм выявил устойчивые пары соединений, которые синтезируются соседствующими подкластерами. (Таблица 4).

Таблица 4. Наиболее распространенные пары подкластеров внутри одного кластера

Вид бактерии	Выявленные пары соединений в одном BGC	Характерно для штаммов
<i>Bacillus atrophaeus</i>	фенгицин + микосубтилин	AG8, P1F1, ZY9
<i>Bacillus velezensis</i>	бациллибактин + амилоциклин	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 китациннамицинов 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 2 диссертации.

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 фенгедисинсинтетазой, 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 <i>P. polymyxa</i> R4.5	14 13	X Val Ile	~340	Нет
<i>P. polymyxa</i> R6.14 <i>P. polymyxa</i> O2.11 <i>P. polymyxa</i> R3.13 <i>P. polymyxa</i> O1.27	8 8 12 8	X mal	~200	Нет
<i>P. polymyxa</i> O1.27 <i>P. polymyxa</i> R3.13 <i>P. polymyxa</i> O2.11 <i>P. polymyxa</i> R6.14	3 7 3 3	fatty_acid mal Tyr D-X Ser D-Ser Ser D-Ile Ser	~1050	Возможно
<i>P. polymyxa</i> R4.24 <i>P. polymyxa</i> R4.5 <i>P. polymyxa</i> R3.13	8 6 14	mal	~86	Нет
<i>P. polymyxa</i> O2.11 <i>P. polymyxa</i> R6.14 <i>P. polymyxa</i> O1.27	12 12 12	fatty_acid D-Gly D-Dab D-Gly D-Ser D-Trp Ser Dab D-Dab Ile Glu	~1256	Возможно
<i>P. polymyxa</i> O2.11 <i>P. polymyxa</i> O1.27 <i>P. polymyxa</i> R6.14 <i>P. polymyxa</i> R3.13	5 5 5 9	fatty_acid D-Asn Gln Me-ohmal mal X ohmal X Gly ccmal Lys X X	~1200	Возможно

Ни одно из предсказанных новых соединений, аннотированных для штаммов *Paenibacillus* не должны осаждаться кислотным 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 в диссертации), 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 бацилломицин D у штаммов

Bacillus объясняет пики 1083 и 1097 m/z. Разработанный алгоритм, таким образом, на наш взгляд, лучше справился с идентификаций соединений нерибосомальной природы, чем antiSMASH.

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 6.

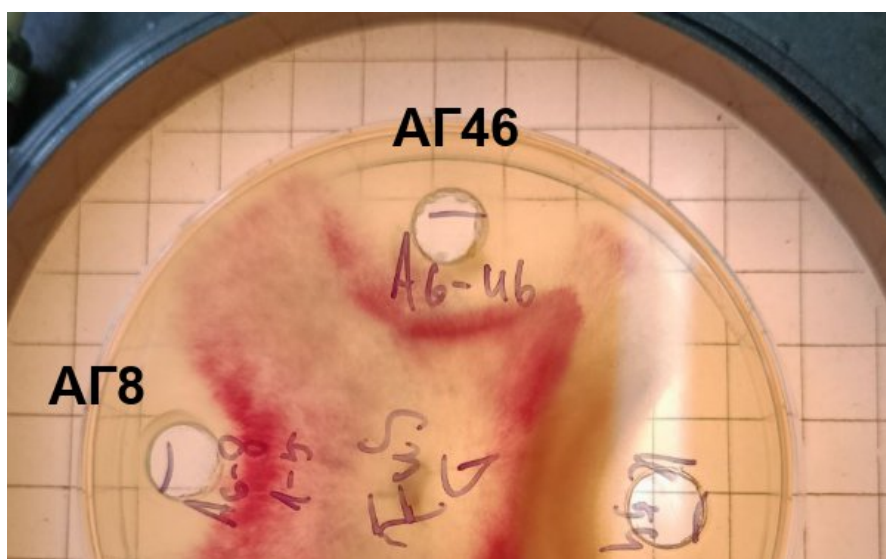


Рисунок 6. Сравнение зон подавления *Fusarium Fus1* (в центре) липопептидными фракциями штаммов AG8 и AG46.

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

Анализ биосинтетического потенциала штаммов-пробиотиков для аквакультуры

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

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

Штамм	МТ 14	МТ 42	МТ 48	МТ 55	МТ 61	МТ 64	МТ 74	МТ 86	МТ 102	МТ 141	МТ 142	МТ 155	МТ 186
Зона отсутствия 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, фенгигина, 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 фенгигин и бацилломицин D оказались локализованы в пределах единых гибридных кластеров. При этом, фенгигинсинтетаза в этих кластерах фрагментирована и фенгигин, 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 макролактинина Н и диффицидина. В качестве штаммоспецифических особенностей алгоритм обнаружил у *B. velezensis* MT42 скрытый подкластер нерибосомальных пептидсинтетаз 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.

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.

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). По результатам измерений проростков обе экспериментальные добавки статистически достоверно улучшили рост по сравнению с контролем: 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. 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 7.

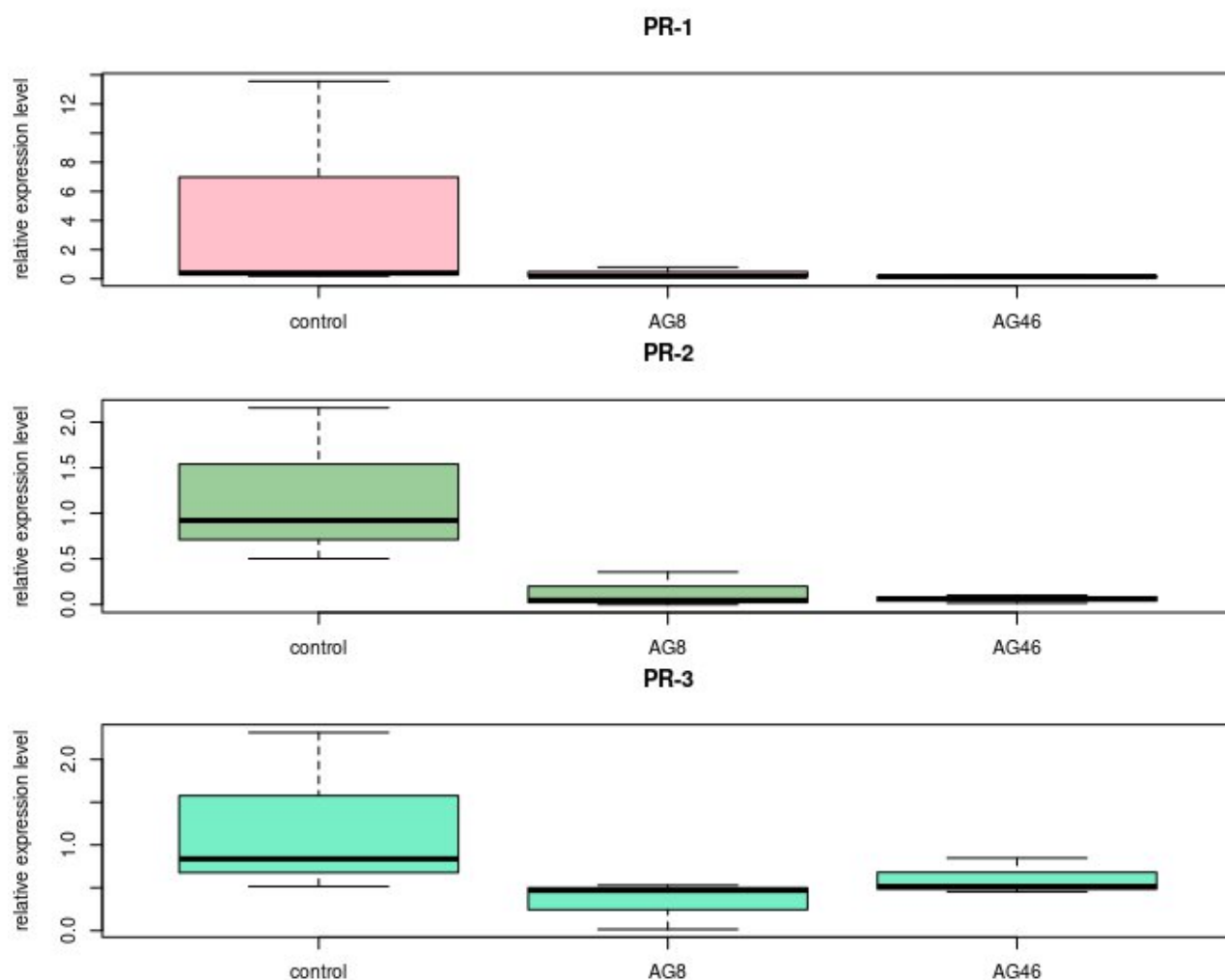


Рисунок 7. Изменение экспрессии генов PR-1, PR-2 и PR-3 ячменя.

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 (Rojas et al., 2022; Constantin et al., 2019; Zamioudis and Pieterse, 2012; Huot et al., 2014).

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 8.

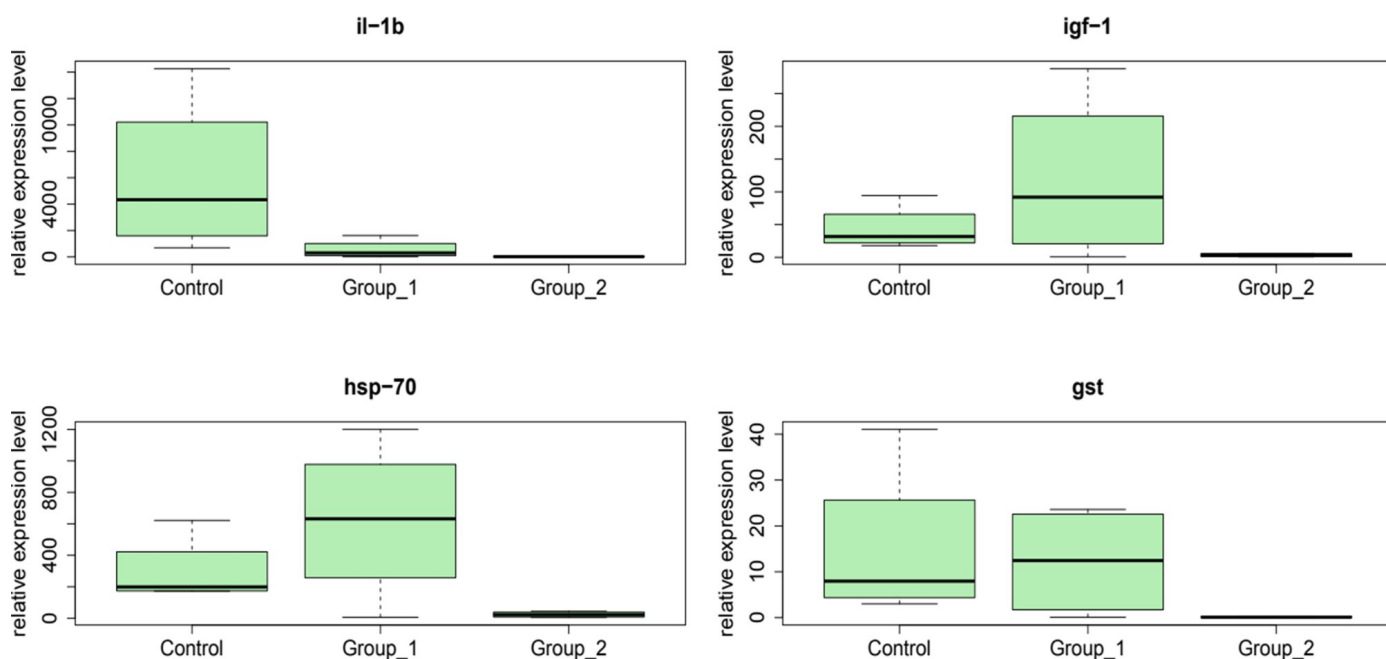


Рисунок 8. Экспрессия генов IL-1 β , *gst*, *igf-1* и *hsp70*.

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 МТ48 и МТ74. 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 (МТ48 и МТ74) БГК фенгицинсинтетаза содержит только саму фенгицинсинтетазу, у группы 1 (МТ14 и МТ42) гомологичный БГК содержит поврежденную фенгицинсинтетазу и синтетазу 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. Это приводит к их повреждению, что и объясняет наблюдаемое влияние на экспрессию исследованных генов.

ЗАКЛЮЧЕНИЕ

В рамках исследования была создана предиктивная модель машинного обучения (Balanced Random Forest), обеспечивающая прогнозирование противогрибковой активности нерибосомальных пептидов на основе геномных данных с точностью 76,4%. 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 аурантицин, 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 РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи, опубликованные в журналах, входящих в базы данных международных индексов научного цитирования Scopus, Web of Science, RSCI:

1. The Potential of Non-Ribosomal Peptide Engineering for Creating New Antimicrobial Complexes / E. V. Prazdnova, M. P. Kulikov, L. E. Khmelevtsova // *Molecules*. – 2026. – Vol. 31. – № 4. – P. 683. (Категория 1)
2. Evaluation of Effects and Genomic Analysis of Aquaculture Probiotic Strains With Antioxidant and Enzymatic Activity / E. V. Prazdnova, M. S. Mazanko, M. P. Kulikov et al. // *Molecular Nutrition & Food Research*. – 2026. – Vol. 70. – № 1. – P. e70277. (Категория 1)
3. In Silico and In Vitro Analysis of Antifungal Compounds of Ten New Bacillus and Paenibacillus Strains / N. Vasilchenko, M. Kulikov, Y. Brislavsky et al. // *Applied Microbiology*. – 2025. – Vol. 5. – № 4. – P. 120. (Категория 1)

Статьи, опубликованные в журналах, входящих в Перечень рецензируемых научных изданий ЮФУ и ВАК:

1. Куликов М. П., Молекулярно-генетические основы нерибосомального синтеза пептидов у прокариот // «Живые и биокосные системы». – 2026. – № 55; URL: <https://jbks.ru/archive/issue-55/article-13>; DOI: 10.18522/2308-9709-2026-55-13 (Категория 3)

Свидетельства о государственной регистрации

1. Свидетельство о государственной регистрации базы данных № 2025624504 Российская Федерация. База данных генетических и метаболических характеристик перспективных для борьбы с болезнью подсолнечника штаммов *Bacillus velezensis* PZ2 и *Bacillus velezensis* PZ4 : № 2025624186 : заявл. 10.10.2025 : опубл. 16.10.2025, Бюл. № 10 / М. П. Куликов, Е. В. Празднова, Н. Г. Дуплий, В. А. Чохели ; правообладатель общество с ограниченной ответственностью "БИОКОД-АГРО".
2. Свидетельство о государственной регистрации базы данных № 2024625944 Российская Федерация. База данных генетических маркеров для идентификации некультивируемых бактерий : № 2024625361 : заявл. 19.11.2024 : опубл. 12.12.2024, Бюл. № 12 / М. П. Куликов, Д. Б. Куликова, К. А. Демин [и др.] ; правообладатель федеральное государственное автономное образовательное учреждение высшего образования «Южный федеральный университет».

Публикации в сборниках трудов конференций

1. Effect of growth medium composition on the efficiency of non-ribosomal synthesis in bacteria of the genus *Bacillus* / V. Vibe, M. Kulikov, E. Prazdnova [et al.] // *BIO Web of Conferences* : [XVII International Scientific and Practical Conference "State and Development Prospects of Agribusiness" (INTERAGROMASH 2024), Rostov-on-Don, Russia, May 22-25, 2024]. – 2024. – Vol. 113. – P. 02020. – DOI 10.1051/bioconf/202411302020.
2. 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", February 24-26, 2021, Rostov-on-Don, Russia : Precision Agriculture and Agricultural Machinery Industry : [in 2 volume]. Vol. 1 / A. Beskopylny, M. Shamtsyan (eds). – Springer Publishing, 2022. – P. 111-121. – (Lecture Notes in Networks and Systems, vol. 246). – DOI 10.1007/978-3-030-81619-3_12. – URL: https://doi.org/10.1007/978-3-030-81619-3_12 (data accesses 14.07.2026)