

Федеральное государственное автономное
образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
Институт фундаментальной биологии и биотехнологий
Базовая кафедра защиты и современных технологий мониторинга лесов

УТВЕРЖДАЮ
Заведующий кафедрой
_____ И. Е. Ямских
« ____ » _____ 2018 г.

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Сравнительная геномика грибов комплекса *Armillaria mellea sensu lato*

Направление магистратуры «Биология» 06.04.01
Магистерская программа «Геномика и биоинформатика» 06.04.01.06

Научный руководитель	_____	д-р физ.-мат. наук	М. Г. Садовский
Выпускник	_____		В. С. Акулова
Рецензент	_____	канд. биол. наук	М. Г. Куцев

Красноярск 2018

АННОТАЦИЯ

Магистерская диссертация по теме «Сравнительная геномика грибов комплекса *Armillaria mellea sensu lato*» содержит 53 страницы текстового документа, 71 использованный источник, 9 рисунков, 13 таблиц.

СРАВНИТЕЛЬНАЯ ГЕНОМИКА, ФИЛОГЕНИЯ, СИНТЕНИЯ, ПОЛНОГЕНОМНОЕ ВЫРАВНИВАНИЕ, ТРАНСКРИПТОМИКА, РНК-СЕКВЕНИРОВАНИЕ.

Цель работы – сравнительный анализ геномов и транскриптомов грибов рода *Armillaria*.

Объектом исследования являются нуклеотидные последовательности ядерных геномов и транскриптомов грибов комплекса *Armillaria mellea sensu lato*.

Предмет исследования – эволюционная связь геномов грибов рода *Armillaria*.

Актуальность настоящего исследования обусловлена тем, что ранее ещё не производилось сравнительных исследований геномных характеристик патогенного вида грибов *Armillaria borealis* с другими представителями этого рода.

В результате работы была получена сборка транскриптома *Armillaria borealis*, проведена оценка качества сборки генома *A. borealis*, которая показала высокое содержание ортологичных генов общей группы *Basidiomycota* (95% от общего числа). Также сравнительный анализ групп сцепления генов выявил макросинтению между геномами видов *Armillaria*.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
Основная часть	7
1. Обзор литературы	7
1.1 Происхождение и эволюция грибов	7
1.2 Виды грибов рода <i>Armillaria</i>	10
1.3 Патогенность видов <i>Armillaria</i>	13
1.3.1 Разложение полисахаридов: целлюлоза.....	16
1.3.2 Разложение полисахаридов: гемицеллюлоза.....	18
1.3.3 Разложение лигнина	18
1.4 Геномы грибов	20
1.5 Синтения.....	21
1.6 Филогенетический анализ	21
2. Материалы и методы	24
2.2 Геномные входные данные.....	24
2.3 Транскриптомные входные данные.....	24
2.1 Проверка качества секвенирования.....	26
2.2 Удаление адаптеров.....	26
2.3 Сборка транскриптома	27
2.4 Филогенетический анализ	29
3. Результаты и обсуждение	30
3.1 Анализ последовательностей	30
3.1.1 Анализ данных РНК-секвенирования на содержание рРНК Ошибка! Закладка не определена.	
3.2 Транскриптомные сборки	Ошибка! Закладка не определена.
3.3 Оценка качества генома.....	Ошибка! Закладка не определена.
3.4 Филогенетический анализ	Ошибка! Закладка не определена.
3.5 Синтения видов рода <i>Armillaria</i>	Ошибка! Закладка не определена.
ЗАКЛЮЧЕНИЕ	47
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	48

ВВЕДЕНИЕ

Разрушение лесных массивов происходит повсеместно в результате антропогенных и климатических условий, а также влияния фитопатогенов — главным образом насекомых – ксилофагов, и патогенных грибов. Изменение климата повышает среднюю годовую температуру и увеличивает частоту экстремальных климатических явлений (продолжительных засух, наводнений, ураганов, штормов и т.д.). В свою очередь это может привести к тому, что последствия заболеваний, вызываемых грибами, могут стать куда более разрушительными и повсеместными [1]. В результате комбинированного воздействия экстремальных факторов среды и фитопатогенов происходит усыхание древостоев, в том числе в умеренных и бореальных лесах. Усыхание также провоцирует возгорание лесов и обширные, продолжительные лесные пожары.

Проблема сохранения лесных массивов актуальна на данный момент, так как масштабы гибели древесных растений в результате патогенного действия грибов огромны: например, в 2001 году для сосны 16,3 млн. га в Западной Канаде (патоген - *Grossmania clavigera*), 77 млн. деревьев вяза в США (патоген - *Ophiostoma ulmi*), а также многие другие [2-5].

Патогены растений отличаются от эндофитов тем, что они вызывают заболевания с выраженными симптомами. Также некоторые зооспорические и зигоспорические грибы являются патогенами растений, но большинство фитопатогенов относятся к аскомицетам и базидиомицетам. Большое количество аскомицетов и около 8000 видов базидиомицетов являются патогенами растений. Помимо патогенной роли грибов, важно помнить, что многие из них важны в естественных экосистемах и части естественного кругооборота веществ, способствуя разложению древесины [6-7].

Виды грибов рода *Armillaria* имеют как экономическое, так и экологическое значение. Они атакуют сотни древесных родов растений (например, *Abies*, *Picea*, *Pinus*, *Citrus*, *Juglans*, *Malus*) в обоих полушариях

при разных климатических условиях, а также являются одними из самых серьёзных патогенов бореальных и умеренных лесов [8].

Важность изучения грибов рода *Armillaria* заключается как в возможности использования полученных знаний в лесоохранной сфере (идентификация видов, изучение уровня патогенности), так и в промышленной биотехнологии для переработки отходов лесной и бумажно-целлюлозной промышленности (виды этого рода способны разрушать лигнин и целлюлозы). *A. borealis* обитает в Сибири, что делает его изучение особенно важным для региона и страны в целом. Кроме того грибы этого рода недостаточно изучены, что делает представленную работу важной также в плане фундаментальных исследований.

Цель настоящей работы — сравнительный анализ геномов и транскриптомов грибов рода *Armillaria*.

Для достижения указанной цели были поставлены следующие задачи:

- Освоить пайплайн по обработке данных РНК-секвенирования;
- Оценить качество данных РНК-секвенирования;
- Идентифицировать образцы с помощью филогенетического анализа;
- Собрать транскриптом *A. borealis*;
- Оценить качество сборки генома *A. borealis*.;
- Провести анализ синтении видов рода *Armillaria*.

Объектом данного исследования являются нуклеотидные последовательности ядерных геномов грибов комплекса *Armillaria mellea sensu lato*.

По материалам магистерской диссертации основные результаты работы будут опубликованы в сборнике тезисов и представлены на 11-ой международной конференции BGRS\SB 2018 – On Bioinformatics of Genome Regulation and Structure\System Biology.

Автор выражает искреннюю благодарность Н. В. Орешковой за пробоподготовку и секвенирование, Д. А. Кузьмину и В. В. Шарову за

сборку генома, Ю. А. Путинцевой, М. Г. Садовскому, И. Н. Павлову и К. В. Крутовскому за идею исследования и общее руководство на всех этапах работы. Также автор выражает благодарность всем членам лаборатории лесной геномики за участие в обсуждении работы, помощь и ценные советы, а также членам лаборатории лесных культур, микологии и фитопатологии института леса СО РАН за постановку эксперимента.

Магистерская диссертация выполнена в лаборатории лесной геномики СФУ и базовой кафедры защиты и современных технологий мониторинга лесов (зав. каф. д.б.н. И. Е. Ямских) в рамках проекта «Геномные исследования основных бореальных лесообразующих хвойных видов и их наиболее опасных патогенов в Российской Федерации», руководимого проф. К. В. Крутовским и финансируемого Правительством РФ (договор №14.У26.31.0004).

Основная часть

1. Обзор литературы

1.1 Происхождение и эволюция грибов

Грибы произошли приблизительно 710 - 1 060 миллионов лет назад [9-10]. К настоящему времени описано более 100 000 видов грибов, но, по оценкам, около 5 миллионов видов грибов еще не идентифицированы [11]. Грибы охватывают разнообразную группу эукариотических гетеротрофов, которые характеризуются голозойным питанием и клетками, окруженными хитиновыми стенками. Размножение грибов может быть вегетативным (соматическим) или половым. Существуют две основные формы вегетативного роста: одноклеточные «дрожжеподобные» клетки, которые делятся почкованием или делением, а также удлиненная разветвленная гифальная система, которая в совокупности составляет мицелий. Некоторые грибы имеют диморфный образ жизни, переключаясь между дрожжеподобными и гифальными формами роста.

Грибы размножаются путем образования спор, которые производятся бесполом (митотическое ядерное деление) или половым (слияние ядер и мейоз) размножением. На основании этого филогенетическая классификация «истинных грибов» выделяет четыре филогенетические группы: *Chytridiomycota*, *Zygomycota*, *Ascomycota* и *Basidiomycota*. «Классические» примеры отличительных особенностей этих филогенетических групп: 1) жгутиковые зооспоры *Chytridiomycota*, 2) несептированный мицелий *Zygomycota* и 3) септированный мицелий *Ascomycota* и *Basidiomycota*, которые различаются в зависимости от типа половой споры, которые они производят [12]. Однако недавние исследования с помощью молекулярных анализов привели к всеобъемлющей филогенетической классификации царства грибов, которая предложила четыре дополнительных филогенетических группы (*Microsporidia*, *Blastocladiomycota*, *Neocallimastigomycota* и *Nlomeromycota*) [13]. Кроме того, отделы *Ascomycota*

и *Basidiomycota* были помещены в подцарство высших грибов, основанное на образовании дикарионов во время полового размножения [14]. Время происхождения подцарства высших грибов оценивается приблизительно в 500-650 миллионов лет назад [9], и большая часть описанных видов грибов (64% and 34%) является представителями отделов *Ascomycota* и *Basidiomycota*, соответственно [15]. Отдел *Basidiomycota* поделён в свою очередь на три подотдела: *Agaricomycotina*, *Pucciniomycotina*, и *Ustilaginomycotina*, которые обычно представлены шляпочными грибами, ржавчинными и головнёвыми грибами соответственно [15]. Некоторые виды отдела *Basidiomycota* хорошо изучены и считаются модельными организмами. *Basidiomycota* включает в себя около 32 000 известных видов.

Морфологически, экологически и таксономически это очень разнообразная группа, но ее члены разделяют общую особенность — их споры полового размножения являются экзоспорами, образованными на базидиуме (и поэтому называются базидиоспоры). *Basidiomycota* включает в себя грибы-патогены растений, эктомикоризные виды, которые имеют ключевое значение для лесных экосистем, сапротрофных видов, которые могут разлагать лигнин и целлюлозу (грибы белой гнили), и наиболее заметные и часто встречающиеся шляпочные грибы [16].

Подотдел *Agaricomycotina* (традиционно известны как гименомицеты) представляет собой разнообразную группу, которая включает около 21 400 описанных видов, что, в свою очередь, составляет около 65% всех известных видов *Basidiomycota* (или около пятой части всех известных грибов). Вероятно, подотдел *Agaricomycotina* возник между 960 и 380 миллионами лет назад. В настоящее время в группу входят грибы, разрушающие деревья, сапротрофы и эктомикоризные грибы.

Ниже приведена схема классификации AFTOL для подтипа *Agaricomycotina*. Самый большой порядок - *Agaricales* (в подклассе *Agaricomycetidae*), состоит из грибов, образующих плодовые тела.

Форма и морфология зрелых плодовых тел могут вводить в заблуждение при попытке видовой идентификации грибов, поскольку они не вносят вклад в естественную классификацию. Поэтому внедрение филогенетического анализа вызывает серьезные, а иногда и неожиданные изменения в классификации этих организмов. Филогения, основанная на анализе рибосомальных последовательностей РНК, изменила классификацию *Agaricales*, а также показала, недостаточную информативность экологических признаков.

Схема классификации AFTOL для подтипа [16-17]:

- *Tremellomycetes*
 - *Cystofilobasidiales*
 - *Filobasidiales*
 - *Tremellales*
- *Dacrymycetes*
 - *Dacrymycetales*
- *Agaricomycetes*
 - *Agaricomycetidae*
 - *Agaricales*
 - *Atheliales*
 - *Boletales*
 - *Phallomycetidae*
 - *Geastrales*
 - *Gomphales*
 - *Hysterangiales*
 - *Phallales*
- *Agaricomycetes* (положение до сих пор точно не определено)
 - *Auriculariales*
 - *Cantharellales*
 - *Corticiales*
 - *Gloeophyllales*

- *Hymenochaetales*
- *Polyporales*
- *Russulales*
- *Sebacinales*
- *Thelephorales*
- *Trechisporales*

1.2 Виды грибов рода *Armillaria*

Наиболее часто применяемые концепции видов в таксономии грибов:

- Концепция морфологических видов. Исторически сложилось так, что виды определялись с помощью морфологического сходства, что, в свою очередь, создает так называемые морфологические виды. Проблема заключается в нахождении видоспецифичных признаков, которые помогут точно определить границы вида. В грибах существует небольшое количество морфологических признаков, а также они очень различны. Естественные изменения признаков, которые могут быть дополнительно подвержены влиянию изменения окружающей среды, трудно измерить. Фундамент грибной классификации был установлен Элиасом Фрисом между 1820 и 1875 годами с акцентом на морфологические особенности (такие как пластинки, поры и цвет спор).

- Концепция биологических видов. Доминирующая идея в биологии - это концепция биологических видов, в которой вид определяется как скрещивающаяся популяция, которая репродуктивно изолирована от других популяций. Идея довольно проста и очевидна, но малоприменима к грибам. Первая проблема заключается в идентификации репродуктивного барьера между популяциями, состоящими из морфологически сходных индивидуумов. Концепция биологических видов также сопряжена с трудностями с географически изолированными популяциями. Популяции, которые не скрещиваются из-за географической изоляции, будут развиваться независимо и могут

значительно расходиться, чтобы в итоге стать разными видами по другим критериям, и, тем не менее, они могут успешно скрещиваться, если объединить их искусственно (при условии, что они сохраняют общие признаки). В целом, существует слишком много серьезных ограничений на широкое применение концепции биологических видов в грибах.

- Концепции, основанные на экологии и физиологии организмов.

Представляется разумным предположение, что паразитические или симбиотические грибы имеют, по крайней мере, некоторую степень специфичности хозяина и что такой вид грибов может быть определен на основе отношений среды обитания и/или хозяина. Не менее разумными являются ожидания, что экологическая адаптация влияет на видообразование грибов и что физиологические особенности, способствующие адаптации к среде обитания и/или хозяину, могут характеризовать виды грибов. Концепция экологического и/или физиологического вида давно используется с грибами-патогенами растений. Концепция в основном отличает виды по своей экологической нише и ограничениям на их эволюцию, которые определяют их содержание и воспроизводство в этой нише. Это звучит разумно, но есть некоторые проблемы. В частности, в значительной степени неизвестно о точной физиологической/биохимической/генетической природе того, что определяет субстратную специфичность и специфичность хозяина, поэтому концепция немного лучше, чем использование морфологических признаков для определения видов. Но применение концепции на практике сильно ограничено.

- Концепция эволюционных/филогенетических видов. Наиболее перспективным представляется определение видов с помощью молекулярно-генетических исследований. Вид рассматривается как монофилетическая группа организмов, у которой есть общие генетические последовательности от предков. Кроме того, в концепции нет каких-либо очевидных исключений или ограничений. Особенно важным для грибов

является тот факт, что анализ может быть применен к асексуальным организмам: анаморфные и телеморфные стадии могут быть охвачены одной концепцией вида. Главное требование в применении данной концепции – использование как можно большего генетического материала при построении филогении (использование нескольких генов вместо одного дает более полную картину об эволюции организма).

Использование различных концепций видов приводит к разным результатам. Но применение концепции филогенетических видов дает преимущество определить гораздо большее количество видов, чем уже было определено с использованием морфологических, биологических или физиологических концепций видов. Также это необходимо для более точной идентификации образцов из среды.

Род *Armillaria* включает в себя больше 40 морфологических видов [8]. Большая часть видов была определена с использованием морфологической концепции видов, что привело к тому, что некоторые виды остаются неопределёнными. Использование филогенетического анализа привело к открытию новых видов, несвязанных с ранее описанными.

Виды *Armillaria* являются факультативными некротрофами: они имеют паразитическую и сапрофитную фазы. Сначала грибы поселяются на живых корнях (паразитическая фаза), затем убивают ткани корня (тем самым вызывая некротические поражения корней), а затем используют мертвые ткани в качестве источника питания (сапрофитная фаза). После смерти растения опёнок сохраняется на заражённых участках корневой системы в течение многих лет, что делает его куда более устойчивым к методам борьбы с ним. Тем не менее, виды *Armillaria*, вызывающие белую гниль являются полезными для природных экосистем (важная роль в круговороте углерода путём деградации лигнина и целлюлоз). Виды *Armillaria* значительно различаются по вирулентности, так, некоторые виды являются основной причиной гибели древесных массивов (*A. ostoyae*), в то время как другие виды колонизируют растения, ранее повреждённые разными факторами

(засуха, насекомые – вредители) [18-19]. Но данный факт не означает, что не существует различий в вирулентности внутри вида [20].

1.3 Патогенность видов *Armillaria*

Несколько видов *Armillaria* вызывают заболевания, паразитируя на деревьях и кустарниках. Такие виды относятся к чрезвычайно агрессивным патогенам. Один отдельный клон *A. ostoyae* был обнаружен в смешанном хвойном лесу в Голубых горах северо-восточного штата Орегон в США, где в итоге от деятельности гриба погибло 30% жёлтых сосен. В настоящее время это один из самых больших в мире организмов, и он охватывает 965 гектар. Максимальное расстояние между изолятами от этого 965-гектарного организма составляло примерно 3,8 км, а использование оценок скорости распространения *A. ostoyae* в хвойных лесах позволило оценить примерный возраст гриба в пределах от 1900 до 8650 лет [21].

Не так давно клон *Armillaria bulbosa* был идентифицирован как самый большой и самый старый живой организм [22], этот организм был «всего» размером в 15 гектар и возрастом 1500 лет. В Швейцарском национальном парке в Центрально-европейских Альпах организмы *A. ostoyae* составляли в размере в среднем 6,8 гектар [23]. Мицелий базидиомицет вездесущ в лесных почвах, несколько исследований показывают, что мицелий многих эктомикоризных, сапротрофных и патогенных базидиомицетов может распространяться вегетативно на значительные расстояния через лесные почвы [23-24].

Сочетание агрессивной патогенности *Armillaria* с ризоморфами (высокоорганизованные высокозащищенные структуры, которые позволяют воде и питательным веществам транспортироваться через лесную почву) помогают быстрому распространению видов *Armillaria* через древесные и кустарниковые рожи, особенно в холодных и относительно сухих условиях среды.

Грибы можно разделить в зависимости от типа питания: сапротрофы, для которых субстратом являются мертвая органика; некротрофы, которые вторгаются в живые ткани, а затем утилизируют их; биотрофы эксплуатируют организмы, которые остаются живыми. Биотрофы могут быть специфичными в выборе хозяина, но сапротрофы и некротрофы обычно имеют очень широкий диапазон мест обитания.

Представители рода *Armillaria* вызывают корневую белую гниль как у голосеменных, так и у покрытосеменных растений, в лесах, парках и даже виноградниках в более чем 500 растениях-хозяевах по всему миру [8]. Большинство видов *Armillaria* являются факультативными некротрофами, которые после колонизации и уничтожения корневого камбия переходят в сапрофитную фазу, разлагая мертвые древесные ткани хозяина. Как сапротрофы, *Armillaria spp.* представляют собой грибы белой гнили, которые могут эффективно разлагать все компоненты клеточных стенок растений, включая лигнин, целлюлозу и гемицеллюлозу.

Несмотря на огромное влияние видов *Armillaria* на лесное хозяйство, садоводство и сельское хозяйство, генетика патогенности видов этого рода плохо изучена. На данный момент уже существуют опубликованные омиксные данные о спектре ферментов, разрушающих клеточные стенки растений (PCWDE – Plant Cell Wall Degradation Enzymes), и секретирующихся белках (в частности, у *A. mellea*, *A. solidipes* и *A. ostoyae* [25-27]). Анализ геномов других патогенных базидиомицетов (*Moniliophthora* [28-29], *Heterobasidion* [30] и *Rhizoctonia* [31]) позволил идентифицировать гены, кодирующие ферменты PCWDE, а также секретируемые из клетки и эффекторные белки вторичного метаболизма в качестве предполагаемых факторов патогенности. Тем не менее, жизненный цикл и стратегия распространения представителей рода *Armillaria* предвосхищают другие эволюционные пути к патогенности, которые наряду с другими потенциальными геномными факторами (такими как мобильные генетические элементы [32]) еще не известны.

Гриб *Armillaria gallica* является агрессивным колонизатором дубовых пней. Он образует в лесной почве сеть ризоморфов, способных колонизировать большую часть дубовых деревьев в лесу. *Armillaria gallica* является условным патогеном с низким уровнем агрессивности и не в состоянии колонизировать энергично растущие деревья. Тем не менее, этот вид часто поражает деревья, ослабленные дефолиацией, насекомыми или засухой. Посевной потенциал был рассмотрен как важная особенность определения способности опенка вторгаться в хозяев. Важность посевного потенциала в основном была отмечена для агрессивных видов *Armillaria* (*A. ostoyae*) [19].

Изучение межгенных областей рибосомальной ДНК некоторых видов *Armillaria* северного полушария показало, что межгенные области *A. ostoyae*, *A. gemina* и *A. borealis* наиболее близки между собой, а также, что межгенные области *A. lutea*, *A. calvescens*, *A. cepistipes*, *A. sinapina* близки друг к другу и отличаются от *A. ostoyae*, *A. borealis* и *A. gemina* значительным числом замен [33]. Данное исследование проводилось с целью выявления филогенетических связей среди видов *Armillaria* северного полушария.

Другой вид патогенных грибов, *Phellinus sulphurascens*, относится к порядку *Hymenochaetales* и вызывает у хвойных растений ламинарную корневую гниль. Было выяснено, что кодирующая ДНК этого вида включает в себя 79 потенциальных факторов вирулентности, среди которых есть гены, связанные с вирулентностью различных патогенных видов грибов [34]. Многие из этих потенциальных факторов вирулентности также были определены как предполагаемые секретируемые белки. Данные по другим патогенным видам грибов могут помочь в изучении данного рода, используя уже известные целевые последовательности.

Гены, отвечающие за патогенность видов *Heterobasidion irregulare* и *H. annosum* («сестринские» фитопатогенные виды) оказались наиболее консервативными в сравнении с другими генами, такими, например, как отвечающие за сапробиотический рост, спорообразование и так далее [35].

Это доказывает, что, инвазия патогенных видов может быть достигнута без высокого уровня патогенности (так как один из этих видов патогеннее, чем другой).

В свою очередь, грибы вида *Armillaria borealis* обнаруживаются практически во всех очагах усыхания древостоев (как одиночное произрастание, так и совокупность нескольких видов комплекса), что говорит об их высоком возможном уровне патогенности [69-71].

Непреренно важную часть в патогенности грибов *Armillaria* имеют ферменты, разлагающие лигнин, целлюлозу и гемицеллюлозу (а также другие компоненты клеточных стенок растений).

1.3.1 Разложение полисахаридов: целлюлоза

Полисахариды представляют собой полимеры моносахаридов, в которых составные сахара связаны гликозидными связями. Из-за количества и разнообразия доступных сахаров и разнообразия возможностей связывания между различными атомами углерода соседних сахарных остатков существует значительное разнообразие полисахаридов. Существует соответствующее разнообразие ферментов, гидролаз или глюкозидаз, способных гидролизовать этот диапазон гликозидных связей. Ферменты, ответственные за разложение полимера (любой полимер, а не только полисахарид) могут использовать одну из двух стратегий расщепления. Они могут расщеплять случайным образом, фрагментировать молекулу полимера на несколько олигомеров (эндоферменты) или они могут крепиться на конце полимера, переваривая мономеры или димеры (экзоферменты).

Целлюлоза является наиболее распространенным органическим соединением на Земле и составляет более 50% органического углерода; ежегодно синтезируется около 1011 тонн целлюлозы. Распад целлюлозы химически прост, но осложняется его физической формой. Мягкий кислотный гидролиз целлюлозы высвобождает растворимые сахара, но не завершается; олигомеры 100-300 остатков глюкозы остаются. Фракция,

которая легко гидролизуется, называется аморфной целлюлозой, а то, что устойчиво к кислоте, называется кристаллической целлюлозой. Поскольку это влияет на химический распад, конформация и трехмерная структура целлюлозы влияет на активность целлюлитического фермента.

Целлюлолитический фермент (целлюлаза) грибов белой гнили отдела *Basidiomycota* состоит из ряда гидролитических ферментов: эндоглюканазы, экзоглюканазы и целлобиозы, которые работают синергически и, как в бактериях, так и в грибах, организованы во внеклеточный мультиферментный комплекс, называемый целлюлосомой [36]. Глюкоза является легко метаболизируемым конечным продуктом расщепления целлюлозы ферментативным гидролизом.

Целлюлосома является внеклеточной молекулярной машиной. В дополнение к каталитическим областям целлюлолитические ферменты содержат домены, не участвующие в катализе, но участвующие в связывании субстрата, образование многоферментного комплекса (так называемые «стыковочные домены») или прикрепление к поверхности [37]. Целлюлосомы эффективно деградируют кристаллические целлюлозы и связанные с ними полисахариды клеточной стенки растений и обеспечивают прикрепление к поверхности клетки. Манипулирование целлюлосомой рассматривается как перспективный способ управления бытовыми и промышленными целлюлозными отходами [38].

При выращивании на целлюлозе, грибы белой гнили, такие как *Phanerochaete chrysosporium*, продуцируют две целлобиозные оксидоредуктазы; целлобиозу: хинон-оксидоредуктаза (СВQ) и целлобиозоксидаза (СВО). Роль, первоначально отнесенная к СВQ, была связующим звеном между деградацией целлюлозы и лигнина. Целлобиозоксидаза также сокращает Fe (III) и вместе с перекисью водорода образует гидроксильные радикалы. Эти радикалы могут деградировать как лигнин, так и целлюлозу, что указывает на то, что целлобиозоксидаза играет центральную роль в деградации древесины грибами.

1.3.2 Разложение полисахаридов: гемицеллюлоза

Термин гемицеллюлоза охватывает различные полимеры с разветвленными цепями, содержащие смесь различных гексозных и пентозных сахаров, которые также могут быть замещены уроновой и уксусной кислотами. Основными гемицеллюлозами, обнаруженными в растениях, являются ксиланы, но также встречаются арабаны, галактаны, маннаны и сополимеры. Ферменты, ответственные за деградацию гемицеллюлозы, называются в соответствии с их субстратной специфичностью; например, маннаназы деградируют маннаны, ксиланазы деградируют ксиланы и т. д. Поскольку ксиланы преобладают в растительных стенках, известно больше о ксиланазах.

Ксиланазы могут быть индуцированы их субстратом, в ходе реакции грибок создает комплекс ферментов, а не один фермент. Комплекс состоит из двух эндоксилааз и β -ксилозидазы. Эндоксилаазы деградируют ксиланы до ксилобиозы и других олигосахаридов, в то время как ксилозидаза разрушает эти меньшие сахара до ксилозы. Также образуется арабиноза, показывающая, что ксиланазный комплекс способен гидролизовать точки ветвления в ксилане.

1.3.3 Разложение лигнина

Лигнины представляют собой высокомолекулярные нерастворимые растительные полимеры, которые имеют сложные и вариативные структуры. Они состоят в основном из многих метоксилированных производных бензола, особенно кониферильных, синапильных и кумарильных спиртов. Соотношение этих трех спиртов различается также между различными растениями. Как лигнины синтезируются в растениях — это вопрос, который ещё не имеет чёткого ответа. Предполагается, что монолигнолы собираются случайным образом до получения сложного полимера, который сильно сшивается в трех измерениях. Другая теория заключается в том, что лигнины имеют специфические последовательности монолигнолов, которые придают

лигнину различные функции в клеточной стенке растения и что существует лишь несколько нативных лигниновых первичных структур [39-40].

Окислительное разрушение лигнина зависит от группы ферментов:

- лигнинпероксидаза (гем (Fe) -содержащий белок), который катализирует H_2O_2 -зависимое окисление лигнина;
- марганцевая пероксидаза, которая также катализирует H_2O_2 -зависимое окисление лигнина;
- лакказы (медьсодержащий белок), которая катализирует деметилирование компонентов лигнина.

Лигнин-пероксидаза является основным лигнин-разрушающим ферментом грибов белой гнили. Существует до 15 изозимов лигнинпероксидазы, в пределах молекулярной массы от 38 000 до 43 000. Это спектр изозимов, полученных в зависимости от условий культивирования и применяемых штаммов. Тем не менее, более 10 генов лигнинпероксидазы с консервативной последовательностью были идентифицированы в *Phanerochaete chrysosporium* и выделены в три различные группы связывания. Однако лигнинпероксидазы не встречаются во всех грибах белой гнили.

Марганец-зависимые пероксидазы представляют собой еще одно семейство внеклеточных гликозилированных гем-белков, которые вырабатываются большинством грибов белой гнили. Система пероксидазы марганца создает низкомолекулярные окислители, которые диффундируют в субстрат лигнина и способны окислять остатки фенола в лигнине на некотором расстоянии от фермента.

Несмотря на то, что некоторые грибы производят лигнинолитические ферменты, гораздо более широкий диапазон грибов выделяет лакказы как внеклеточные ферменты. Это медьсодержащие оксигеназы, способные окислять некоторые фенолы и требуются для метаболизма продуктов разложения лигнина.

1.4 Геномы грибов

Возможность секвенировать и сравнивать полные геномные последовательности улучшает наше понимание многих областей биологии. Такие данные более точно показывают эволюционные отношения и указывают, как распространяются патогены и вызываются ими заболевания. Они позволяют нам приблизиться к комплексному пониманию активности живых клеток и тому, как они контролируются на молекулярном уровне.

По размеру и строению ядерного генома грибы занимают промежуточное положение между прокариотами и остальными эукариотами, в среднем размер генома грибов на 2 порядка меньше, чем у высших растений. Число хромосом колеблется от 2 до 28, у большинства видов — от 10 до 12. Размер хромосом у грибов также значительно меньше, чем у других эукариот.

В грибах размер геномов варьирует от 8,9 до 177,5 Мбр (миллион пар нуклеотидов). Средний размер генома грибов отделов *Ascomycota* и *Basidiomycota* составляет 37 и 46 Мбр, соответственно (таблица 1) [41].

Таблица 1 — Средний размер геномов, среднее число кодирующих генов и экзонов в разных отделах грибов

Отдел грибов	Средний размер генома, Мбр (миллион пар нуклеотидов)	Среднее число генов	Среднее число экзонов
<i>Ascomycota</i>	37	11 129	3
<i>Basidiomycota</i>	46	15 431	5
<i>Oomycota</i>	75	24 173	2
<i>Mucoromycota</i>	39	13 306	4

Размер транскрипта в свою очередь зависит напрямую от размера генома, а также от постановки эксперимента (разные гены и их число экспрессируется при разных условиях среды).

1.5 Синтения

Синтения — это сходство групп сцепления генов у организмов разных биологических видов. В частности, в геномах человека и мыши известно несколько десятков синтеничных групп генов [42]. Наличие феномена синтении позволяет сократить круг поиска места локализации исследуемого гена на хромосомах, ограничивая его областью известных генов, принадлежащих к конкретной синтеничной группе. Синтенный блок в свою очередь представляет из себя куда более крупными гомологичными участками между разными геномами.

Построение блока синтении происходит следующим образом: находятся гомологичные участки между геномами (это могут быть участки с хорошим выравниванием или просто гены, в нашем случае участки определённой длины). Эти участки называются якорями, которые дальше группируются в синтенные блоки по различным правилам и каждому блоку присваивается свой номер для комбинаторной интерпретации задачи поиска геномных перестроек. Один из возможных наборов правил: 1) якоря должны находиться друг от друга на расстоянии меньше, чем на N пар нуклеотидов; 2) в блоках разных организмов с одним и тем же номером лежат одни и те же якоря.

1.6 Филогенетический анализ

С помощью филогенетического анализа можно установить родственные связи между живыми организмами на основе сравнения структуры полимерных макромолекул — ДНК, РНК и белков. В данной работе используются маркерные последовательности ДНК: ITS (internal transcribed spacer) и $tef1\alpha$ (translation elongation factor 1 alpha). Результатом филогенетического анализа является построение филогенетического дерева, которое является гипотезой об эволюционных отношениях между организмами.

Наиболее часто используемые методы построения филогений на основе молекулярных данных основываются на моделях эволюции. В настоящее время самыми используемыми методами в построении филогенетических деревьев являются Байесовский подход и метод максимального правдоподобия (Maximum Likelihood, ML). Для расчета кладограммы, помимо последовательности ДНК, надо выбрать модель замены нуклеотидов, на основании которой будут рассчитываться вероятности. Также в расчет берется длина ветви или эволюционная дистанция между двумя таксонами. Во время анализа рассчитывается, какая длина ветви наиболее вероятна с точки зрения выбранной модели, вероятности всех ветвей кладограммы умножаются, и кладограмма, имеющая наибольшую вероятность, считается правильной [43-45].

Метод максимального правдоподобия стал одним из первых методов, применяющих длину ветвей и эволюционные дистанции. Существует большое разнообразие программного обеспечения для построения филогенетических деревьев на основе ML метода, как с использованием стандартных эволюционных моделей замен, так и с применением разных моделей, поправок и информационных критериев [46-51].

Байесовский подход в построении филогении использует функцию правдоподобия для создания величины, называемой апостериорной вероятностью деревьев с использованием модели эволюции, основанной на некоторых априорных вероятностях, создающей наиболее вероятное филогенетическое дерево для данных. Байесовский подход стал популярным благодаря достижениям в скорости вычислений и интеграции алгоритмов марковской цепи Монте-Карло (MCMC).

По мере того, как байесовские методы стали более популярными, MrBayes стал одним из самых используемых программных обеспечений среди молекулярных филогенетиков. В программе используется стандартный алгоритм MCMC, а также вариант MCMC, связанный с алгоритмом Метрополиса [52].

Несмотря на все плюсы двух последних методов, есть некоторые сложности. Главная слабость методов в необходимости самостоятельного подбора модели замен. Хотя уже есть программы, которые могут помочь в подборе наиболее подходящей модели [47, 49].

Подбор моделей эволюционных замен нуклеотидов характерен для каждого гена и каждой последовательности в филогенетическом анализе [53]. Модель JC69 (Jukes and Cantor, 1969) является самой простой моделью замещения. Эта модель предполагает равные частоты для всех нуклеотидов и равные скорости мутаций.

Модель K80 (Кимура, 1980) также предполагает равные частоты для всех нуклеотидов, но различает транзиции и трансверсии.

Модель F81 (Фельзенштейн, 1981) является расширением модели JC69, в которой частоты нуклеотидов могут отличаться от 0,25.

Модель HKY85 (Hasegawa, Kishino and Yano, 1985) можно рассматривать как сочетание расширений, сделанных в моделях K80 и F81. Модель различает скорость транзиций и трансверсий и позволяет использовать неравные частоты нуклеотидов.

Модель T92 (Tamura, 1992) представляет собой математический метод, разработанный для оценки количества нуклеотидных замен на сайт между двумя последовательностями ДНК, путем расширения двухпараметрического метода Кимуры (K80) до случая, когда существует смещение GC-контента.

Наиболее сложная модель – GTR (General Time-Reversible). Она использует разные частоты нуклеотидов (4 параметра), и разные частоты замен между нуклеотидами (6 параметров) (Tavare, 1986).

2. Материалы и методы

2.2 Геномные входные данные

Секвенирование и геномная сборка *Armillaria borealis* осуществлялись сотрудниками лаборатории лесной геномики СФУ под руководством проф. К. В. Крутовского. Материал для секвенирования был выделен из активного мицелия *A. borealis* из деревьев *Abies sibirica*, погибших в 2015 году. ДНК была секвенирована с использованием парноконцевых библиотек со вставкой 250 пар нуклеотидов и платформы Illumina MiSeq. *De novo* сборка генома осуществлялась сборщиком SPAdes (с использованием итерационного модуля сборки генома короткими прочтениями; значения K выбирались автоматически на основе длины прочтений и типа данных) [54]. Длина генома *A. borealis* ~ 68 Мbp. Значение N50 для контигов равняется 49 951 пар нуклеотидов (таблица 2).

Таблица 2 — Статистика сборки генома *A. borealis*

Параметр	Статистика
Количество скаффолдов	48 127
Общая длина, пар нуклеотидов	68 318 719
Максимальная длина, пар нуклеотидов	2 196 911
N50, пар нуклеотидов	49 951
N90, пар нуклеотидов	331

2.3 Транскриптомные входные данные

Для изучения транскриптомов грибов рода *Armillaria* использовались следующие данные:

- Эксперимент с двумя средами — Modified Melin-Norcrans и берёзовый сок с экстрактом хвои (Таблицы 3 и 4). Для эксперимента использовался один штамм *Armillaria borealis* в 8 повторностях (по 4 в каждой среде). Для выделения РНК использовался химический набор TruSeq RNA LT Kit. С

помощью постановки данного эксперимента планировалось получить транскриптомы из образцов с различной экспрессией генов в ответ на разное содержание среды и оценить экспрессию генов организмов в ответ на среду. Эксперимент планировался профессором И.Н. Павловым из лаборатории Института леса ФИЦ КНЦ СО РАН. Для дальнейшего анализа использовались 8 парноконцевых библиотек (8 повторностей).

- Всего было просеквенировано 5 образцов 2015 года, которые были выделены из активного мицелия *A. borealis* из деревьев *Abies sibirica*, погибших в 2015 году. Собирались в одном месте, дистанция между деревьями 2-10 метров.

Таблица 3 — Повторности эксперимента РНК-секвенирования 2-го запуска

Название	Повторность	Среда	Индекс Illumina
1.1n	1	MMN	19
2.2n	2	MMN	18
3.1n	3	MMN	16
4.2n	4	MMN	15
1.2b	1	Берёзовый сок с экс. хв.	14
2.2b	2	Берёзовый сок с экс. хв.	13
3.2b	3	Берёзовый сок с экс. хв.	2
4.3b	4	Берёзовый сок с экс. хв.	4

Для эксперимента с двумя средами было проведено два запуска секвенирования на платформе Illumina MiSeq.

Таблица 4 — Повторности эксперимента РНК-секвенирования 1-го запуска

Название	Повторность	Среда	Индекс Illumina
fung5	1	MMN	10
fung6	2	MMN	11
fung7	3	MMN	12
fung8	4	MMN	13
fung1	1	Берёзовый сок с экс. хв.	5
fung2	2	Берёзовый сок с экс. хв.	6
fung3	3	Берёзовый сок с экс. хв.	7
fung4	4	Берёзовый сок с экс. хв.	9

2.1 Проверка качества секвенирования

Секвенирование проводилось на геномном секвенаторе MiSeq Illumina, выходные данные которого (парноконцевые прочтения) проверялись в программе FastQC.

FastQC обеспечивает простое выполнение некоторых проверок контроля качества данных секвенирования. Программа предоставляет модульный набор анализов для быстрого представления о характеристиках используемых данных (качество секвенирования, GC-контент и т.д.).

Целесообразно проверять качество используемых последовательностей на каждом этапе обработки данных.

Для поиска контаминаций использовался BLASTn с $e\text{-value } 1 \times 10^{-3}$. Результаты показали наличие небольшого числа последовательностей (данные РНК-секвенирования), схожих с другими классами со средней точностью (~85% pident).

2.2 Удаление адаптеров



Рисунок 1 — Пайплайн для первичной обработки данных секвенирования

Прочтения очищались с помощью пайплайна, написанного в нашей лаборатории. Пайплайн (рисунок 1) использует Trimmomatic – это многопоточный инструмент командной строки, который можно использовать для обрезки данных низкого качества геномной платформы Illumina (FASTQ), а также для удаления адаптеров.

Существует два основных режима работы программы: парноконцевой и одноконцевой. Парноконцевой режим устанавливает соответствие между парами прочтений для лучшего поиска адаптеров или праймеров реакции ПЦР. Так как цель данной работы изучение транскриптома, то очистка проводилась без удаления коротких ридов.

2.3 Сборка транскриптома

Существуют разные инструменты для сборки данных РНК-секвенирования, такие как Trans-ABYSS [55], Velvet-Oases [56] и SOAPdenovo-trans (а также другие) [57].

Относительно новый алгоритм сборки транскриптома — Trinity. Он разбивает данные РНК-секвенирования на независимые графы де Брюйна, в идеале один граф на экспрессированный ген, и использует параллельные вычисления для восстановления транскриптов из этих графов, включая альтернативные сплайсированные изоформы. Trinity может использовать цепь-специфичные парноконцевые библиотеки Illumina. Несколько независимых исследований показали, что данный алгоритм очень эффективен по сравнению с альтернативными алгоритмами [58-59].

Oases и Trinity – два часто используемых пакета программного обеспечения для сборки коротких чтений в транскрипты. Оба инструмента содержат несколько сложных шагов и их трудно оценить на основе одного алгоритма. Тем не менее, было показано, что для анализа с участием известных генов Trinity является лучшим инструментом для этих данных. Однако может случиться так, что Oases лучше обнаруживает новые

транскрипты, так как количество ложноположительных среди обнаруженных транскриптов неизвестно [60-62].

Пайплайн Trinity состоит из трех последовательных модулей: Inchworm, Chrysalis и Butterfly. Вначале, все перекрывающиеся k -меры извлекаются из прочтений РНК-секвенирования. Затем, Inchworm исследует каждый уникальный k -мер в порядке убывания его частоты встречаемости и генерирует контиги, используя «жадный» алгоритм, основанный на $(k-1)$ -мерном перекрытии. Inchworm генерирует полноразмерные транскрипты для доминирующей изоформы, а также сообщает об уникальных альтернативно сплайсированных транскриптах.

Затем, Chrysalis кластеризует Inchworm контиги, используя исходные прочтения для группировки транскриптов на основе общих последовательностей прочтений и парноконцевых данных. Это кластеризует вместе области, которые могут происходить из альтернативных сплайсированных изоформ или близкородственных семейств генов. После Chrysalis создаёт граф де Брюйна для каждого кластера и распределяет прочтения среди кластеров. Такое разделение на кластеры и распределение прочтений позволяет проводить массовую параллельную обработку последующих вычислений.

Наконец, Butterfly обрабатывает отдельные графы параллельно, в конечном счете выдавая полные транскрипты для альтернативно сплайсированных изоформ и разделяя гены-паралоги. Butterfly прослеживает прочтения РНК-секвенирования через граф и определяет связанность с помощью последовательности прочтения и дополнительных данных (парноконцевые прочтения). Когда связанность не может быть установлена с помощью прочтений, Butterfly разделяет граф на несколько несвязанных подграфов и обрабатывает их по отдельности. Наконец, Butterfly связывает достоверные графы и реконструирует последовательности транскриптов [63-64].

2.4 Филогенетический анализ

Для секвенирования сборки тотального транскриптома использовали мицелий с деревьев, собранных в 2015 году. Всего было просеквенировано 5 образцов, и было необходимо идентифицировать их видовую принадлежность. Для этого был использован филогенетический анализ.

Последовательность применяемого программного обеспечения была следующая:

- 1) SeaView — графический редактор выравнивания последовательности [51]. С помощью SeaView возможно прочитывать и записывать различные форматы выравнивания, что полезно для дальнейшего анализа. Применяется алгоритм выравнивания MUSCLE (MUltiple Sequence Comparison by Log-Expectation), который является оптимальным по времени выравнивания и качеству.
- 2) PartitionFinder 2 предоставляет комбинированный выбор моделей эволюционных замен и схем программы для выравнивания ДНК последовательностей [49]. Он может использовать все информационные поправки: AIC (информационный критерий Акаике), AICc (информационный скорректированный критерий Акаике) и BIC (Байесовский информационный критерий).
- 3) Далее для построения дерева методом Байеса использовалась программа MrBayes [52].

3. Результаты и обсуждение

3.1 Анализ последовательностей

Первоначально данные по двум средам планировалось использовать для оценки дифференциальной экспрессии генов. Ожидалось, что после очистки прочтений по качеству и длине мы сможем дальше использовать данные для анализа экспрессии. Для анализа использовалось программное обеспечение CLC Genomics Workbench, для которого в качестве исходных данных используются прочтения всех повторностей и тотальный транскриптом, собранный со всеми этими прочтениями. Это программа разработана для анализа и визуализации данных секвенирования следующего поколения.

На первом этапе для визуальной проверки данных использовался метод главных компонент (PCA – Principal Component Analysis), который рассчитывает компоненты по максимальному различию. На рисунках 2 и 3 изображены результаты PCA анализа транскриптов всех повторностей и специально отобранных повторностей соответственно. В первом случае транскрипты повторностей среды MMN почти все собираются в один кластер, в то время как транскрипты повторностей среды с берёзовым соком делятся на два кластера как по первой, так и по второй компонентам.

Изъято 16 страниц.

ЗАКЛЮЧЕНИЕ

В результате работы были выполнены следующие задачи:

1. Проверка данных РНК-секвенирования выявила избыточное содержание рибосомальных РНК последовательностей.
2. Филогенетический анализ позволил с высокой вероятностью определить образцы, относящиеся к виду *Armillaria borealis*.
3. Была получена сборка транскриптома *A. borealis*, где общая длина сборки и уровень представленности достаточны для того, чтобы использовать данный транскриптом в дальнейшей работе. Ранее не осуществлялось сборок транскриптомов сибирского вида опёнка.
4. Оценка качества сборки генома *A. borealis* показала высокое содержание ортологичных генов, что говорит о высоком качестве сборки.
5. Сравнительный анализ групп сцепления генов выявил макросинтению геномов между видами грибов рода *Armillaria*, что говорит о таксономической и эволюционной близости этих видов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. La Porta N. et al. Forest pathogens with higher damage potential due to climate change in Europe/ N. La Porta, P. Capretti , I. M. Thomsen , R. Kasanen , A. M. Hietala, K. Von Weissenberg // Canadian Journal of Plant Pathology. – 2008. – Т. 30. – №. 2. – С. 177-195.
2. Fisher M. C. et al. Emerging fungal threats to animal, plant and ecosystem health / Matthew C. Fisher, Daniel. A. Henk, Cheryl J. Briggs, John S. Brownstein, Lawrence C. Madoff, Sarah L. McCra, Sarah J. Gurr // Nature. – 2012. – Т. 484. – №. 7393. – С. 186.
3. Магдеев Н. Г. и др. Вредители и болезни основных лесообразующих пород в Республике Татарстан / Магдеев Н.Г., Селиховкин А.В., Мусин Х.Г., Ахматович Н.А. //Лесной вестник/Forestry bulletin. – 2013. – №. 6 (98).
4. Василюскас А., Юодвалькис А., Трейгене А. Причины массового усыхания ясеня обыкновенного в лесах Литвы // Проблемы лесной фитопатологии и микологии: материалы V международной конференции/Институт лесоведения РАН, М. – 2002. – С. 35-37.
5. Раздорожная Т. Ю., Шилкина Е. А. Изучение комплекса грибов *Armillaria mellea* sensu lato в центральных районах Красноярского края //Лесохозяйственная информация. – 2016. – №. 1.
6. Farr D. F. et al. Fungi on plants and plant products in the United States, 2nd ed. – APS press, 1989.
7. Burdon J.J. The structure of pathogen populations in natural plant communities / //Annual Review of Phytopathology. – 1993. – Т. 31. – №. 1. – С. 305-323.
8. Baumgartner K. et al. Secrets of the subterranean pathosystem of *Armillaria* / Baumgartner K., Coetzee M., Hoffmeister D. // Molecular plant pathology. – 2011. – Т. 12. – №. 6. – С. 515-534.
9. Lucking R. et al. Fungi evolved right on track // Mycologia. – 2009. – Т. 101. – №. 6. – С. 810-822.

10. Hawksworth D. L. The magnitude of fungal diversity: the 1· 5 million species estimate revisited // *Mycological research*. – 2001. – T. 105. – №. 12. – C. 1422-1432.
11. Blackwell M. The Fungi: 1, 2, 3... 5.1 million species? // *American journal of botany*. – 2011. – T. 98. – №. 3. – C. 426-438.
12. Webster J., Weber R. *Introduction to Fungi* // Cambridge University Press, New York, U.S.A. – 2007.
13. Hibbett D.S. et al. A higher-level phylogenetic classification of the Fungi // *Mycological research*. – 2007. – T. 111. – №. 5. – C. 509-547.
14. James T. Y. et al. Reconstructing the early evolution of Fungi using a six-gene phylogeny // *Nature*. – 2006. – T. 443. – №. 7113. – C. 818-822.
15. Stajich J. E. et al. *Primer--The Fungi* // *Current biology: CB*. – 2009. – T. 19. – №. 18. – C. R840-845.
16. Moore D., Robson G.D., Trinci A.P. J. *21st century guidebook to fungi*. // Cambridge University Press, New York. – 2011.
17. Blackwell M. et al. Research coordination networks: a phylogeny for kingdom Fungi (Deep Hypha) // *Mycologia*. – 2006. – T. 98. – №. 6. – C. 829-837.
18. Prospero, S. Comparison of the virulence of *Armillaria cepistipes* and *Armillaria ostoyae* on four Norway spruce provenances / S. Prospero, O. Holdenrieder, D. Rigling // *Forest Pathology* . – 2004. – T. 34. – №. 1. – C. 1-14.
19. Marcais B., Breda N. Role of an opportunistic pathogen in the decline of stressed oak trees // *Journal of Ecology*. – 2006. – T. 94. – №. 6. – C. 1214-1223.
20. Morrison D. J., Pellow K. W. Variation in virulence among isolates of *Armillaria ostoyae* // *Forest pathology*. – 2002. – T. 32. – №. 2. – C. 99-107.
21. Ferguson B. A. et al. Coarse-scale population structure of pathogenic *Armillaria* species in a mixed-conifer forest in the Blue Mountains of northeast Oregon // *Canadian Journal of Forest Research*. – 2003. – T. 33. – №. 4. – C. 612-623.

22. Smith M. L., Bruhn J. N., Anderson J. B. The fungus *Armillaria bulbosa* is among the largest and oldest living organisms // *Nature*. – 1992. – T. 356. – №. 6368. – C. 428-431.
23. Bendel M., Kienast F., Rigling D. Genetic population structure of three *Armillaria* species at the landscape scale: a case study from Swiss *Pinus mugo* forests // *Mycological research*. – 2006. – T. 110. – №. 6. – C. 705-712.
24. Cairney J. W. G. *Basidiomycete* mycelia in forest soils: dimensions, dynamics and roles in nutrient distribution // *Mycological Research*. – 2005. – T. 109. – №. 1. – C. 7-20.
25. Collins C. et al. Genomic and proteomic dissection of the ubiquitous plant pathogen, *Armillaria mellea*: toward a new infection model system // *Journal of proteome research*. – 2013. – T. 12. – №. 6. – C. 2552-2570.
26. Ross-Davis A. L. et al. Transcriptome of an *Armillaria* root disease pathogen reveals candidate genes involved in host substrate utilization at the host–pathogen interface // *Forest Pathology*. – 2013. – T. 43. – №. 6. – C. 468-477.
27. Sipos G. et al. Genome expansion and lineage-specific genetic innovations in the forest pathogenic fungi *Armillaria* // *Nature ecology & evolution*. – 2017. – T. 1. – №. 12. – C. 1931-1941.
28. Meinhardt L.W. et al. Genome and secretome analysis of the hemibiotrophic fungal pathogen, *Moniliophthora roreri*, which causes frosty pod rot disease of cacao: mechanisms of the biotrophic and necrotrophic phases // *BMC Genomics*. – 2014. – T. 15. – №. 1. – C. 164.
29. Mondego J. M. C. et al. A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao // *BMC genomics*. – 2008. – T. 9. – №. 1. – C. 548.
30. Olson Å. et al. Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen // *New Phytologist*. – 2012. – T. 194. – №. 4. – C. 1001-1013.

31. Hane J. K. et al. Genome sequencing and comparative genomics of the broad host-range pathogen *Rhizoctonia solani* AG8 // PLoS genetics. – 2014. – T. 10. – №. 5. – C. e1004281.
32. Raffaele S., Kamoun S. Genome evolution in filamentous plant pathogens: why bigger can be better // Nature Reviews Microbiology. – 2012. – T. 10. – №. 6. – C. 417-430.
33. Anderson J. B., Stasovski E. Molecular phylogeny of northern hemisphere species of *Armillaria* // Mycologia. – 1992. – C. 505-516.
34. Williams H. L. et al. Gene expression profiling of candidate virulence factors in the laminated root rot pathogen *Phellinus sulphurascens* // BMC genomics. – 2014. – T. 15. – №. 1. – C. 603.
35. Sillo F. et al. Comparative genomics of sibling fungal pathogenic taxa identifies adaptive evolution without divergence in pathogenicity genes or genomic structure // Genome biology and evolution. – 2015. – T. 7. – №. 12. – C. 3190-3206.
36. Bégum P., Lemaire M. The cellulosome: an exocellular, multiprotein complex specialized in cellulose degradation // Critical reviews in biochemistry and molecular biology. – 1996. – T. 31. – №. 3. – C. 201-236.
37. Shoham Y., Lamed R., Bayer E. A. The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides // Trends in microbiology. – 1999. – T. 7. – №. 7. – C. 275-281.
38. Bayer E. A., Lamed R., Himmel M. E. The potential of cellulases and cellulosomes for cellulosic waste management // Current opinion in Biotechnology. – 2007. – T. 18. – №. 3. – C. 237-245.
39. Humphreys J. M., Chapple C. Rewriting the lignin roadmap // Current opinion in plant biology. – 2002. – T. 5. – №. 3. – C. 224-229.
40. Davin L. B., Lewis N. G. Lignin primary structures and dirigent sites // Current opinion in biotechnology. – 2005. – T. 16. – №. 4. – C. 407-415.
41. Mohanta T. K., Bae H. The diversity of fungal genome // Biological Procedures Online. – 2015. – T. 17. – №. 1. – C. 8.


42. Kent W. J. et al. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes // *Proceedings of the National Academy of Sciences*. – 2003. – T. 100. – №. 20. – C. 11484-11489.
43. Baum D. A., Smith S. D. *Tree thinking: an introduction to phylogenetic biology*. – Greenwood Village, CO : Roberts, 2013.
44. Wiley E. O., Lieberman B. S. *Phylogenetics: theory and practice of phylogenetic systematics*. Second edition // Hoboken, NJ: Wiley-Blackwell, 2011.
45. Yang Z., Rannala B. *Molecular phylogenetics: principles and practice* // *Nature Reviews Genetics*. – 2012. – T. 13. – №. 5. – C. 303-314.
46. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies // *Bioinformatics*. – 2014. – T. 30. – №. 9. – C. 1312-1313.
47. Darriba D. et al. jModelTest 2: more models, new heuristics and parallel computing // *Nature methods*. – 2012. – T. 9. – №. 8. – C. 772.
48. Nguyen L. T. et al. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies // *Molecular biology and evolution*. – 2014. – T. 32. – №. 1. – C. 268-274.
49. Lanfear R. et al. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses // *Molecular Biology and Evolution*. – 2016. – T. 34. – №. 3. – C. 772-773.
50. Guindon S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0 // *Systematic biology*. – 2010. – T. 59. – №. 3. – C. 307-321.
51. Gouy M., Guindon S., Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building // *Molecular biology and evolution*. – 2009. – T. 27. – №. 2. – C. 221-224.
52. Ronquist F. et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space // *Systematic biology*. – 2012. – T. 61. – №. 3. – C. 539-542.

53. Liò P., Goldman N. Models of molecular evolution and phylogeny // Genome research. – 1998. – T. 8. – №. 12. – C. 1233-1244.
54. Bankevich A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing // Journal of computational biology. – 2012. – T. 19. – №. 5. – C. 455-477.
55. Robertson G. et al. *De novo* assembly and analysis of RNA-seq data // Nature Methods. – 2010. T. 7. – № 11. – C. 909-912.
56. Schulz M. H. et al. Oases: robust *de novo* RNA-seq assembly across the dynamic range of expression levels // Bioinformatics. – 2012. – T. 28. – №. 8. – C. 1086-1092.
57. Xie Y. et al. SOAPdenovo-Trans: *de novo* transcriptome assembly with short RNA-Seq reads // Bioinformatics. – 2014. – T. 30. – №. 12. – C. 1660-1666.
58. Duan J. et al. Optimizing *de novo* common wheat transcriptome assembly using short-read RNA-Seq data // BMC genomics. – 2012. – T. 13. – №. 1. – C. 392.
59. Zhao Q. Y. et al. Optimizing *de novo* transcriptome assembly from short-read RNA-Seq data: a comparative study // BMC bioinformatics. – BioMed Central, 2011. – T. 12. – №. 14. – C. S2.
60. Schulz M. H. et al. Oases: robust *de novo* RNA-seq assembly across the dynamic range of expression levels // Bioinformatics. – 2012. – T. 28. – №. 8. – C. 1086-1092.
61. Haas B. J. et al. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis // Nature protocols. – 2013. – T. 8. – №. 8. – C. 1494-1512.
62. Rana S. B. et al. Comparison of *de novo* transcriptome assemblers and k-mer strategies using the killifish, *Fundulus heteroclitus* // PloS one. – 2016. – T. 11. – №. 4. – C. e0153104.

63. Haas B. J. et al. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis // Nature protocols. – 2013. – Т. 8. – №. 8. – С. 1494.
64. Grabherr M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome // Nature biotechnology. – 2011. – Т. 29. – №. 7. – С. 644-652.
65. Kopylova E., Noé L., Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data // Bioinformatics. – 2012. – Т. 28. – №. 24. – С. 3211-3217.
66. Kim D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions // Genome biology. – 2013. – Т. 14. – №. 4. – С. R36.
67. Waterhouse R.M. et al. BUSCO applications from quality assessments to gene prediction and phylogenomics // Molecular biology and evolution. – 2017.
68. Kurtz S. et al. Versatile and open software for comparing large genomes // Genome biology. – 2004. – Т. 5. – №. 2. – С. R12.
69. Павлов И. Н. и др. Образование и затухание очагов куртинного усыхания сосны обыкновенной в результате воздействия *Armillaria borealis* Marxm. & Korhonen (Сообщение 1. Эдафические закономерности) // Хвойные бореальной зоны. – 2012. – Т. 29. – №. 3-4.
70. Павлов И. Н., Барабанова О. А., Корхонен К. Патогенные свойства *Heterobasidion annosum* (fr.) Bref s. str. и *Armillaria borealis* Marxm. & Korhonen в различных лесорастительных условиях сосняков Минусинской впадины // Иммунопатология, аллергология, инфектология. – 2010. – №. 1. – С. 123-124.
71. Павлов И. Н. Биотические и абиотические факторы усыхания хвойных лесов Сибири и Дальнего Востока // Сибирский экологический журнал. – 2015. – Т. 22. – №. 4. – С. 537-554.

Федеральное государственное автономное
образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
Институт фундаментальной биологии и биотехнологий
Базовая кафедра защиты и современных технологий мониторинга лесов

УТВЕРЖДАЮ
Заведующий кафедрой

 И. Е. Ямских

« 22 » июня 2018 г.


МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Сравнительная геномика грибов комплекса *Armillaria mellea sensu lato*

Направление магистратуры «Биология» 06.04.01

Магистерская программа «Геномика и биоинформатика» 06.04.01.06

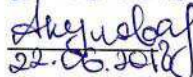
Научный руководитель


22.06.2018

д-р физ.-мат. наук

М. Г. Садовский

Выпускник


22.06.2018

В. С. Акулова

Рецензент


19.06.2018

канд. биол. наук

М. Г. Куцев

Красноярск 2018