

На правах рукописи



Турнаев Игорь Иванович

**ИССЛЕДОВАНИЕ ЭВОЛЮЦИИ БЕЛКОВ
ТРИПТОФАН-ЗАВИСИМОГО ПУТИ БИОСИНТЕЗА
АУКСИНА У РАСТЕНИЙ**

1.5.7. –генетика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата биологических наук

Новосибирск – 2024

Работа выполнена в лаборатории эволюционной биоинформатики и теоретической генетики ФГБНУ «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук», г. Новосибирск.

Научный руководитель: **Афонников Дмитрий Аркадьевич** - доктор биологических наук, доцент, ведущий научный сотрудник, и.о.заведующего лабораторией эволюционной биоинформатики и теоретической генетики, ФГБНУ «Федеральный исследовательский центр Институт цитологии и генетики СО РАН, г. Новосибирск.

Официальные оппоненты: **Хавкин Эмиль Ефимович** - доктор биологических наук, профессор, ведущий научный сотрудник лаборатории ДНК маркеров растений ФГБНУ «Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии», г. Москва.

Букин Юрий Сергеевич - кандидат биологических наук, старший научный сотрудник лаборатории геносистематики ФГБУН «Лимнологический институт СО РАН, г. Иркутск.

Ведущая организация: ФГБУН Ботанический институт им. В.Л. Комарова РАН, г. Санкт-Петербург

Защита диссертации состоится «___»_____ 2025 года на утреннем заседании Диссертационного совета 24.1.239.01 на базе ИЦиГ СО РАН в конференц-зале Института по адресу: 630090, г. Новосибирск, проспект ак. Лаврентьева, 10, т.(383)363-49-06, e-mail: dissov@bionet.nsc.ru.

С диссертацией можно ознакомиться в библиотеке ИЦиГ СО РАН и на сайте Института <http://www.icgbio.ru>.

Автореферат разослан «___»_____ 2024 г.

Учёный секретарь
диссертационного совета,
доктор биологических наук

Хлебодарова Т.М.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность. Ауксин является важнейшим гормоном растений и участвует в регуляции таких разнообразных процессов, как гравитропизм и фототропизм, формирование архитектуры корней и побегов, морфогенез растений, развитие сосудов, деление и удлинение клеток, апикальное доминирование, старение, реакции на биотический и абиотический стрессы (Lau et al., 2009; Zhao, 2010; Finet, Jaillais, 2012; Woodward, Bartel, 2005; Teale et al., 2006; Scarpella et al., 2006; Halliday 2009; Sundberg, Ostergaard, 2009; Peris et al., 2010; Zhao, 2014). В регуляции роста и развития растений ауксин участвует совместно с другими фитогормонами такими как цитокины, этилен, абцизовая кислота, гиббереллины, но при этом считается, что он является наиболее важным фитогормоном растений (Mukherjee et al., 2022). Ауксин в настоящее время является наиболее изученным фитогормоном (Paque, Weijers, 2016). Знание механизмов биосинтеза, метаболизма и транспорта ауксина у растений крайне необходимо для понимания практически всех морфогенетических процессов, происходящих в онтогенезе растения и позволит значительно увеличить возможности контроля роста и развития растений, что представляет широкое поле деятельности для создания новых продуктивных или быстрорастущих сортов растений (Розов и др., 2013). Таким образом, изучение механизма биосинтеза ауксина является одной из важных задач для понимания биологии ауксина. Прогресс в исследованиях биосинтеза ауксина закладывает основу для улучшения важных признаков сельскохозяйственных растений, таких как ветвление или развитие цветков, позволяя нам регулировать уровни ауксина в определенных тканях/клетках. Четкое понимание биосинтеза ауксина в конечном итоге окажет значительное влияние на сельское хозяйство, а также значительно расширит наши знания в области фундаментальной биологии растений (Zhao, 2014). Знание молекулярных механизмов биосинтеза ауксина позволит в перспективе контролировать концентрацию ауксина в клетках растений, что дает возможность исследовать молекулярные механизмы регулируемых ауксином процессов развития растений (Zhao, 2014).

Основной путь синтеза ауксина (ПБА) у наземных растений включает два этапа:

- 1) синтез триптофана из хоризмата (путь биосинтеза триптофана, ПБТ);
- 2) непосредственно синтез ауксина из триптофана (путь биосинтеза ауксина, мы будем для краткости называть его ПБА(ТАА/YUCCA)).

Второй этап, ПБА(ТАА/YUCCA), состоит из двух реакций: превращение триптофана в индол-3-пировиноградную кислоту (ИПК, англ. indole-3-pyruvic acid, IPA) катализируемое ферментом ТАА и превращение ИПК в ауксин осуществляемое ферментом YUCCA (Mashiguchi et al., 2011; Won et al., 2011; Stepanova et al., 2011; Tivendale et al., 2012, Cheng et al., 2006).

ПБА(ТАА/YUCCA) наиболее исследован у *Arabidopsis thaliana*, в геноме которого присутствуют 3 гена *taa* и 11 генов *yucca*. Интересно, что растения *Arabidopsis*, по-видимому, используют разные наборы генов *yucca* для биосинтеза ауксина в побеге и в корне, а также на эмбриональной или постэмбриональных стадиях развития растения, несмотря на использование одного и того же набора генов *taa* в побеге и в корне, и на эмбриональной или постэмбриональных стадиях развития (Won et al., 2011). Таким образом, активности генов-паралогов *yucca* у растения являются специфичными к органу и стадии развития растения (Stepanova, 2008; Won et al., 2011; Chen, 2014).

Для изучения молекулярных механизмов функционирования системы биосинтеза ауксина у растений активно используются методы молекулярной эволюции (DeSmet et al., 2011; Mano, Nemoto, 2012). Они позволяют исследовать происхождение и эволюцию белков. Однако филогенетический анализ семейств генов *taa* и *yucca*, не позволил получить однозначные результаты о происхождении и дальнейшей эволюции этих белков у растений.

Юэ и др. (Yue et al., 2014) основываясь на анализе реконструкции филогении семейств белков ТАА и YUCCA у бактерий и растений предположили, что ПБА(ТАА/YUCCA) впервые появился в эволюции растений у наиболее позднего общего предка (НПОП) наземных растений в результате горизонтального переноса генов (ГПГ) от бактерий. Таким образом, ПБА(ТАА/YUCCA) может считаться инновацией первых наземных растений (Yue et al., 2014). Однако немного позже Ванг и др. (Wang et al., 2014) обнаружили гомологи *taa* и *yucca* в геноме Харофитовой водоросли *Klebsormidium nitens*. Это позволило сформулировать альтернативную гипотезу о существовании ПБА(ТАА/YUCCA) у Харофитовых водорослей (до выхода растений на сушу). В результате чего вопрос о происхождении ПБА(ТАА/YUCCA) у растений оказался дискуссионным.

Цель работы. Выявление генов пути биосинтеза ауксина у растений и биоинформатический анализ особенностей их молекулярной эволюции.

Задачи, поставленные для достижения указанной цели, включали:

1. Анализ происхождения и эволюции семейств ферментов пути биосинтеза триптофана;
2. Анализ происхождения и эволюции семейства ТАА;
3. Анализ происхождения и эволюции семейства YUCCA;
4. Оценка дупликаций генов ферментов пути биосинтеза ауксина в эволюции растений и функциональной роли этого процесса.

Научная новизна. Впервые показано, что гомолог растительных ферментов ТАА у Харофитовой водоросли *K. nitens*, kfl00051_0080, по функциональным и структурным характеристикам более близок к аллииназам, чем к

триптофанаминотрансферазам, и не принимает участия в ПБА, что впоследствии было экспериментально подтверждено независимой группой исследователей.

Впервые показано, что гомолог YUCCA у Харофитовой водоросли *K. nitens* GAQ82387.1 (kfl00109_0340), не принадлежит к семейству растительных белков YUCCA, а относится к отдельному кластеру-семейству FMO Pb; функции этого белка отличаются от функций белков семейства YUCCA. Таким образом, впервые строго показано, что ПБА(ТАА/YUCCA) появился в эволюции у общего предка наземных растений как результат горизонтального переноса генов от нерастительных таксонов (предположительно бактерий).

Впервые установлена достоверная положительная связь между увеличением количества гомологов ферментов ПБТ и сложностью организмов растений.

Положения, выносимые на защиту:

1. Белок KnТАА (NCBI: GAQ80308.1) харовой водоросли *Klebsormidium nitens*, гомологичный ферментам триптофанаминотрансфераз (ТАА) наземных растений, по функциональным и структурным характеристикам более близок к аллииназам, чем к ТАА, и не принимает участия в пути биосинтеза ауксина;
2. Белок KnFMO (NCBI: GAQ82387.1) харовой водоросли *Klebsormidium nitens*, гомологичный ферментам флавиномоноксигеназам YUCCA (YUCCA) наземных растений, относится к кластеру Pb суперсемейства флавиномоноксигеназ класса В, не принадлежит к семейству YUCCA и имеет отличные от них функции;
3. Путь биосинтеза ауксина ТАА/YUCCA появился в эволюции у общего предка наземных растений, как результат горизонтальных переносов генов, кодирующих эти ферменты, от нерастительных таксонов.

Теоретическая значимость работы. Результаты настоящего исследования позволили установить происхождение ПБА у растений. На основе биоинформатического анализа и сопоставления наших данных с результатами по анализу ферментативной активности ряда бактериальных белков суперсемейства флавин-содержащих моноксигеназ показано, что ПБА(ТАА/YUCCA) появился в эволюции у НПОП наземных растений, как результат ГПГ от нерастительных организмов.

Апробация работы. Основные результаты, полученные в рамках настоящей работы, были представлены на третьей Московской интернациональной конференции «Молекулярная филогенетика» (Москва, 2012 г.); восьмой, девятой, десятой, одиннадцатой и двенадцатой международных конференциях «Биоинформатика геномной регуляции и структурной системной биологии» (Новосибирск, 2012, 2014, 2016; 2018; 2020 гг.); IX Международном конгрессе "Биотехнология: Состояние и перспективы

развития" (Москва, 2017); международной конференции "Беляевские чтения" (Новосибирск, 2017 г.); конференции PlantGen2019 (Новосибирск, 2019).

Личный вклад автора. Все основные научные результаты были получены автором самостоятельно.

Структура и объем работы. Диссертация состоит из введения, обзора литературы, материалов и методов, результатов и обсуждений, заключения, выводов, списка литературы (382 источника), списка обозначений и сокращений. Работа изложена на 195 страницах, содержит 27 рисунков и 10 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Первая глава - обзор литературы включает семь разделов: В разделе 1.1 описан биосинтез ауксина у наземных растений и его взаимодействие с другими процессами. В разделе 1.2 описаны предполагаемые и установленные пути биосинтеза ауксина. В разделе 1.3 описаны метаболизм, транспорт ауксина и регуляция его синтеза у наземных растений. В разделе 1.4 рассмотрены эволюция ферментов ПБА наземных растений: TAA и YUCCA. В разделе 1.5 рассмотрены роль дупликации генов в эволюции живых организмов.

Вторая глава – “Материалы и методы” описывает использованные в исследовании методы. Выборки белков ферментов ПБА были сформированы с помощью поиска (программа BlastP) гомологов исходных белков в базах данных NCBI nr, Plaza 2.5, 1000 plant genomes и Green algae. Множественное выравнивание белковых последовательностей проводилось программами MAFFT и Promals. Реконструкция филогенетических деревьев выполнена программами RAxML, mrBayes и IQ-TREE. Реконструкция 3D структуры для белка kfl00051_0080 – *K. nitens* TAA осуществлена с помощью программы i-TASSER. Для анализа белковых доменов использовалась программа CD-search и информация из базы данных Pfam. Визуализация консенсусов активных сайтов в белках проведена программой WebLogo v. 2.8.2. Корреляционный анализ количества гомологов генов ферментов ПБА со сложностью организмов растений (индекс F_{БП}) проведен программой Statistica 8.0.

Глава 3 «Эволюция ферментов пути биосинтеза триптофана» Реконструированы филогенетические деревья для всех белков ферментов ПБА: ASA, ASB, PAT (TRP), PAI, IGPS, TSA и TSB. Анализ показал, что ряд генов ПБА (*asa*, *asb*, *igps*, *tsa* и *tsb*) появился у общих предков красных и зелёных водорослей в ходе вертикальной эволюции от предков цианобактерий в согласии с Юэ и др. (Yue et al., 2014).

Глава 4 «Происхождение и эволюция ферментов пути TAA/YUCCA биосинтеза ауксина у наземных растений» посвящена анализу эволюции и структур белков суперсемейств TAA и YUCCA.

Мы провели детальный анализ структуры последовательности КпТАА (kfl00051_0080/GAQ80308.1) *K. nitens* в сравнении с белками семейства ТАА наземных растений (TAR1, TAR2, TAR3 из *A. thaliana*) и аллииназами наземных растений (TAR3, TAR4 из *A. thaliana*) (рис. 1).

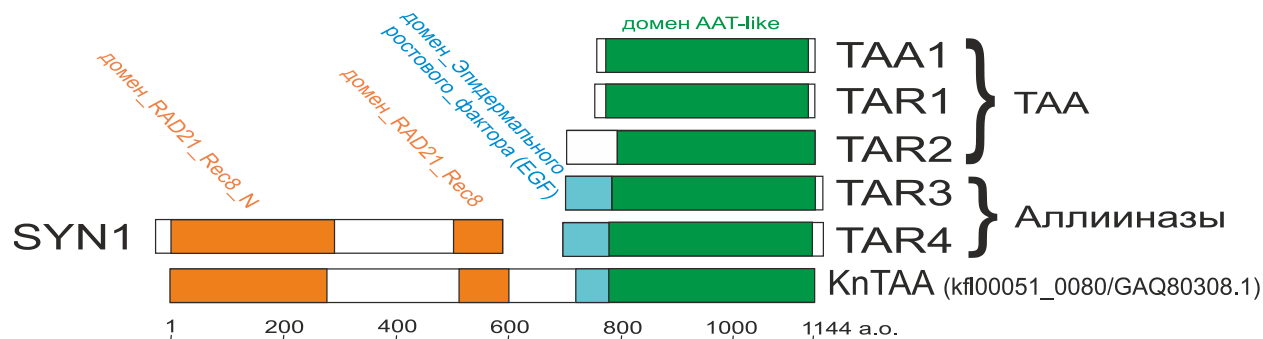


Рисунок 1. Доменный состав последовательности транскрипта КпТАА *K. nitens* по сравнению с белками, связанными с триптофан-аминотрансферазой (ТАА), и RAD21-подобным белком SYN1 из *A. thaliana*. Цветными прямоугольниками обозначены консервативные домены белковых последовательностей; белые прямоугольники соответствуют районам неконсервативных последовательностей; шкала позиций аминокислотных остатков (а.о.) в последовательности транскрипта kfl00051_0080 находится внизу рисунка.

Длина белка КпТАА, в соответствии с текущей аннотацией генома *K. nitens* (v1.0, kfl00051_0080), составила 1144 а.о. – это более чем в 2 раза превышает длину белков ТАА – 350–520 а.о.

Его последовательность включает два района: N-концевой (гомологичный мейотическому когезин-подобному белку Rad21/Rec8) и С-концевой (гомологичный ТАА белкам TAA1, TAR1 и TAR2 и аллииназам TAR3, TAR4) у *A. thaliana*. Таким образом, последовательность kfl00051_0080 представляет собой слияние когезин-подобных (Rad21/Rec8) и ферментативных (ТАА-подобных) доменов. Между этими двумя районами расположен домен связывания Эпидермального ростового фактора (EGF), характерный также для аллииназ (рис. 1, голубой прямоугольник), но отсутствующий у ТАА наземных растений.

Филогенетический анализ суперсемейства ТАА (рис. 2) показал, что белок КпТАА Харофитовой водоросли *K. nitens* расположен на дереве ближе к аллииназам, чем к ТАА наземных растений.

Проведенная нами реконструкция 3D структуры белка КпТАА (kfl00051_0080) показала, что его модель имела более высокое структурное сходство с аллииназой чеснока *Allium sativum* [PDB ID 1LK9: A; EC 4.4.1.4; Ca RMSD 0.55 Å], чем с триптофан аминотрансферазой TAA1 *A. thaliana* [PDB ID 3BWN: A; EC 2.6.1.27, 2.6.1.99; Ca RMSD 1.79 Å].

Все вышеперечисленные аргументы свидетельствуют, что белок kfl00051_0080 *K. nitens* ближе к семейству аллииназ, чем к ТАА наземных растений и не способен катализировать превращение триптофана в пировиноградную кислоту (ИПК), что свидетельствует в пользу отсутствия у Харофитовых водорослей ПБА (ТАА/YUCCA).



Рисунок 2. Филогенетическое дерево семейства ТАА, реконструированное методом максимального правдоподобия (программа IQ-TREE). Зелёным цветом ветвей обозначены белки растений и коричневым белки простейших. Шкала расстояний показана вверху рисунка слева, величины бутстреп поддержки показаны рядом с внутренними узлами дерева.

Мы провели филогенетический анализ белков класса В суперсемейства флавин-содержащих монооксигеназ, к которому относится семейство YUCCA, с помощью метода IQ-TREE (рис. 3).

Выявлено 10 групп белков: YUCCA, YUC-подобные FMO, цианобактериальные FMO, FMO типа IIa, FMO типа IIb, FMO типа IIc, NMO, FMO I типа, BVMO и флавин-содержащие монооксигеназы G класса (аутгруппа) (рис. 3).

Показано, что ближайшими к белкам YUCCA наземных растений оказались белки FMO бактерий двух групп: цианобактериальные FMO и YUC-подобные FMO. Белок KpFMO (kfl00051_0080/GAQ82387.1) харофиты *K. nitens* попадает в более удаленную от YUCCA кладу, FMO типа IIb (рис. 3).

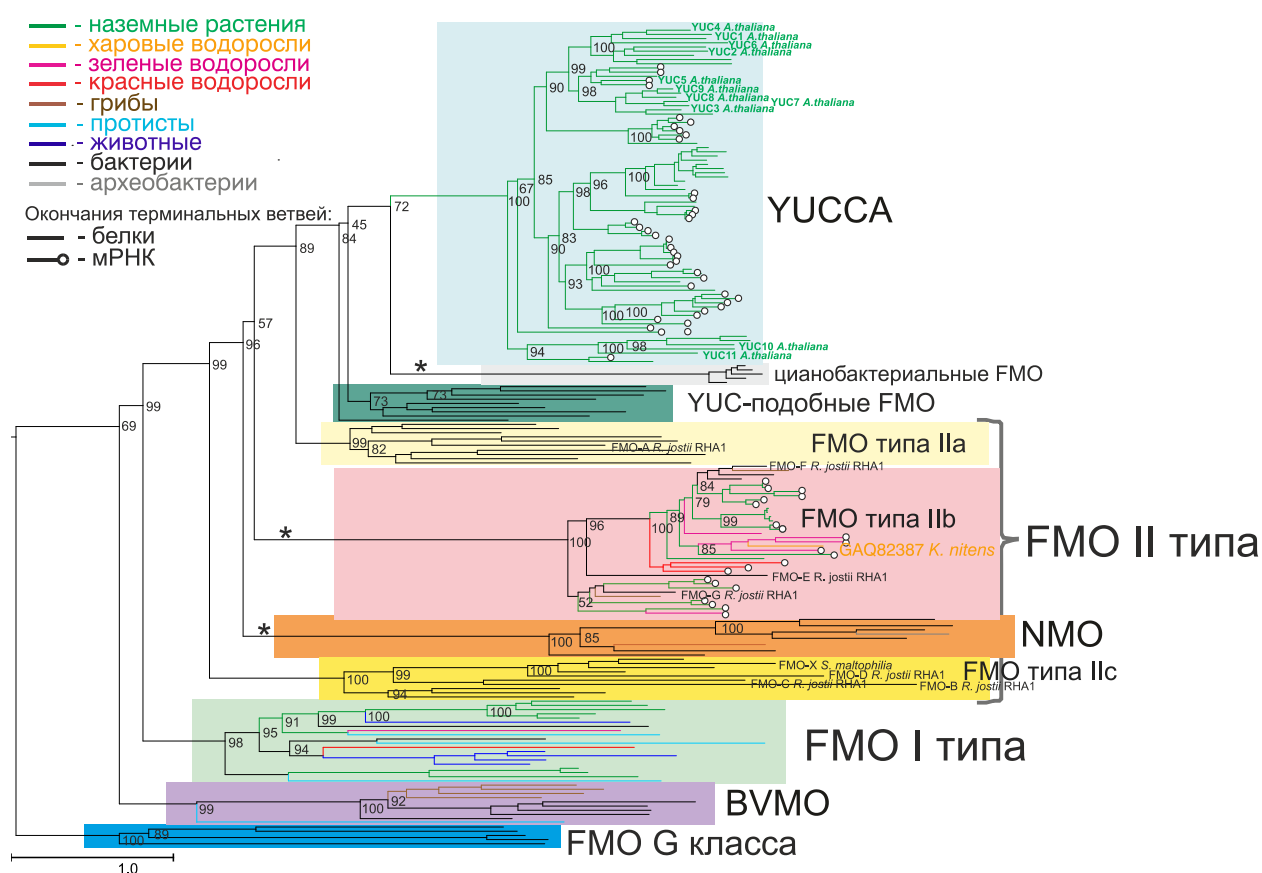


Рисунок 3. Филогения В класса флавин-содержащих монооксигеназ, включающее, как белковые последовательности, так и последовательности мРНК из транскриптомных сборок, реконструированное методом максимального правдоподобия (программы IQ-TREE). В качестве аутгруппы взяты белки G класса флавин-содержащих монооксигеназ.

* - обозначены длинные ветви, предшествующие соответствующим кластерам. Шкала расстояний показана внизу рисунка слева, величины бутстреп поддержки показаны рядом с внутренними узлами дерева.

Полученные результаты позволяют выделить белки кластера FMO типа IIb в особую группу из более обширной группы FMO II типа, так как: (а) это обособленный кластер с высоким уровнем бутстрэп-поддержки (рис. 3); (б) этот кластер единственный из FMO II типа включает белки разных царств (рис. 3); (в) этот кластер содержит последовательности ферментов FMO-E, FMO-G, FMO-F бактерии *R. jostii* *RHA1* обладающие уникальным набором каталитических свойств (Riebel et al., 2013; Riebel et al., 2014) (рис 3).

Во всех 10-ти группах белков, выделенных нами на филогенетическом дереве флавин-содержащих монооксигеназ В класса (рис. 3), был проведен сравнительный анализ белковых последовательностей по трём функциональным сайтам (ФАД-связующий, FMO и НАДФН-связующий) (рис. 4А). На рисунке 4Б показаны консенсусы этих трёх сайтов в формате LOGO для 10-ти исследованных групп последовательностей.

Для FMO сайта (рис 4Б, столбец 3) в последовательностях групп YUCCA, YUC-подобных FMO, цианобактериальных FMO, FMO типа IIa, FMO типа IIc и FMO типа I характерен консенсус FxGxxxHxxx(Y/F)(K/R) в предпоследней позиции которого преобладает тирозин (Y). В последовательностях FMO типа IIb для FMO сайта характерен консенсус: FxGxxxHxxx(H/y/f)P. Предпоследний символ этого мотива у 70% последовательностей представлен гистидином (H), однако у остальных 30% его заменяют тирозин (Y) или фенилаланин (F). Тогда как гистидин (H) на этой позиции во всех остальных группах подкласса FMO, помимо FMO типа IIb, встречается лишь у двух последовательностей – в кластере FMO I типа (рисунок 4Б, столбец 3).

Что касается мотива НАДФН-связующего сайта (рисунок 4Б, столбец 4), то в нем у всех групп последовательностей, кроме FMO типа IIb, характерны либо все три консервативных глицина (G) на 1, 3 и 5 позициях консенсуса сайта либо третий глицин может заменяться на аланин (A). Однако в белках FMO типа IIb в сайте связывания НАДФН вместо второго глицина в 75% случаев встречается N - аспарагин: хотя в других девяти исследованных группах белков эта аминокислота (N - аспарагин) не встречается ни разу (рис. 4Б, столбец 4).

Таким образом, в консенсусах FMO и НАДФН-связующего сайтов большинства белков кластера FMO типа IIb находятся аминокислоты, нехарактерные как для белков FMO типа IIa и типа IIc, так и вообще для всех остальных белков подкласса FMO.

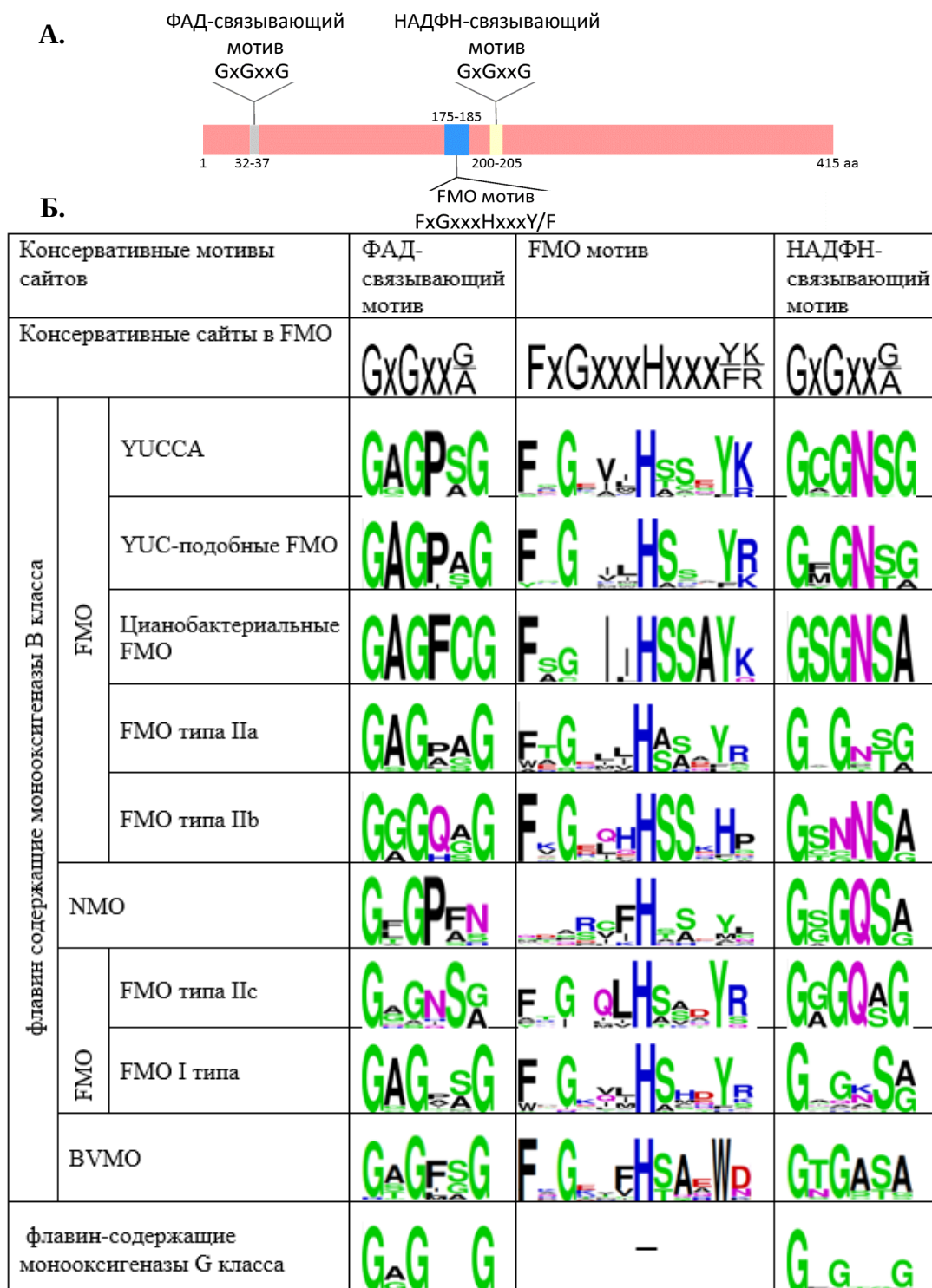


Рисунок 4 – Распределение и выравнивание ключевых функциональных сайтов у белков В класса флавин-содержащих монооксигеназ. А) Схематическое изображение локализации мотивов вдоль белковой последовательности FMO (позиции аминокислот на схеме указаны по последовательности AT4G13260 – *A. thaliana* YUCCA2). Б) В данной таблице, соответственно, в 2-4 столбцах показаны консенсусы мотивов трёх сайтов белков В класса флавин-содержащих монооксигеназ (см. рис. 4А): FMO-связывающего, FMO и НАДФН-связывающего сайтов.

Мы провели сравнительный анализ консервативных доменов трёх групп белков: (1) FMO типа IIb; (2) большой группы FMO-подобных белков (включающей FMO I типа, FMO типа IIa, FMO типа IIc, цианобактериальные FMO, YUC-подобные FMO, YUCCA) и (3) BVMO (рис. 5).

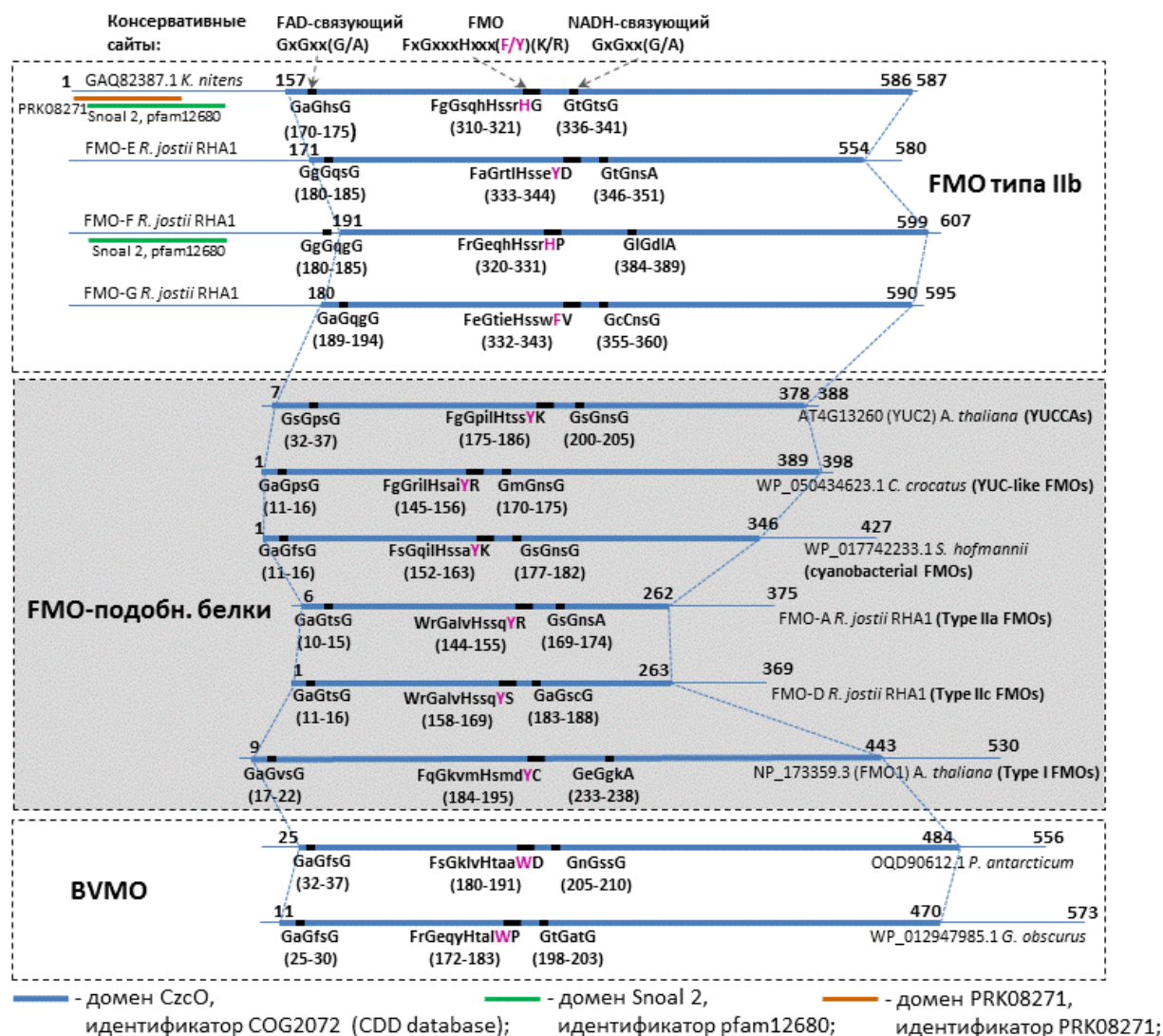


Рисунок 5. Сравнение консервативных доменов трёх групп белков: (1) FMO типа IIb; (2) FMO-подобных белков (YUCCA, YUC-подобные FMO, цианобактериальные FMO, FMO типа IIa, FMO типа IIc, FMO I типа) и (3) BVMO. Сокращения на рисунке: *C. crocatus* — бактерия *Chondromyces crocatus*; *S. hofmannii* — цианобактерия *Scytonema hofmannii*; *P. antarcticum* — гриб *Penicillium antarcticum*; *G. obscurus* — бактерия *Geodermatophilus obscurus*.

Поиск консервативных доменов в базе данных CDD для последовательностей белков FMO/BVMO (это белки, принадлежащие к трем группам (1) FMO типа IIb, (2) FMO-подобные белки и (3) BVMO) выявил один консервативный домен CzcO (accession COG2072), присутствовавший во всех

последовательностях. Этот домен находится в центральной части белков, занимая от 65 до 98% их длины (рис. 5).

Все последовательности FMO типа IIb имеют N-концевой район длиной около 160 а.о., который отсутствует у белков других семи кластеров: FMO I типа, FMO типа IIa, FMO типа IIc, YUC-подобных FMO, YUCCA, цианобактериальные FMO и BVMO (рис. 5). Анализ, проведенный с помощью программы CD-search (порог E-value по умолчанию 0.01) позволил обнаружить сходство данной области у последовательности GAQ82387.1 *K. nitens* с доменами: Snoal2 (SnoalL-подобный домен, идентификатор домена – pfam12680, E-value = 1.70e-04) и сигма-70 фактором РНК-полимеразы (идентификатор домена PRK08241, E-value = 9.67e-04). Известно, что домен Snoal2 характерен для суперсемейства белков NTF2-like.

Таким образом, последовательности FMO типа IIb отличаются от последовательностей YUCCA наличием дополнительного домена на их N-конце, который, вероятно, принадлежит NTF2-подобному суперсемейству.

В таблице 1 представлены результаты идентификации гомологов белков *K. nitens* (kfl00051_0080/GAQ82387.1) и *A. thaliana* YUCCA2 (AT4G13260) программой blastP в двух базах данных: базе 1000 plant genomes (1KP), содержащей транскрибируемые транскриптомные последовательности растений и базе NCBI, содержащей белковые последовательности.

Из таблицы 1 видно, что оба гомолога AT4G13260 (YUCCA *A. thaliana*) и GAQ82387.1 (KpFMO из семейства FMO типа IIb; *K. nitens*) представлены у печеночных (liverworts) и антероцеротовых (hornworts) мхов, лептоспорангиевых папоротников (leptosporangiate monilophytes). Тогда как у настоящих мхов (Bryophyta), однодольных и двудольных белки YUCCA также широко представлены, но гомологи GAQ82387.1 (FMO типа IIb) обнаруживаются только у единичных видов. При этом важно отметить, что одновременно гомологи обоих белков присутствовали у одного вида настоящих мхов, трех видов печеночных мхов, двух видов антероцеротовых мхов, 13 видов лептоспорангиевых папоротников, одного вида однодольных и двух видов двудольных. Факт наличия обоих ферментов для ряда важных таксонов от мхов до покрытосеменных предполагает, что YUCCA и FMO типа IIb – это два фермента с различными функциями важными для жизни растений.

Таким образом, результаты проведенного исследования белков суперсемейства флавин-содержащих монооксигеназ класса В показали, что семейства белков FMO типа IIb и YUCCA не являются близкородственными, что поддерживает гипотезу Юэ и др. (Yue et al., 2014) о происхождении YUCCA у ранних наземных растений.

Таблица 1. Поиск гомологов двух последовательностей – *K. nitens* GAQ82387.1 (группа FMO типа IIb) и *A. thaliana* YUCCA2 AT4G13260 (группа YUCCA) – в двух базах данных: 1000 plant genomes и NCBI.

Таксономия				Гомологи GAQ82387.1 в базе 1000 plant genomes (1KP)	Гомологи GAQ82387.1 в базе NCBI	Гомологи YUCCA в базе 1000 plant genomes (1KP)	Гомологи YUCCA в базе NCBI
Streptophyta	Streptophytina	Embryophita	Eudicots (596)	4(3)	-	434(333)	1680(114)
			Monocots (104)	1(1)	-	47(35)	444(26)
			Conifers (73)	1(1)	-	4(4)	-
			Cycadales (4)	-	-	2(2)	-
			Leptosporangiate Monilophytes (65)	90(33)	-	25(19)	-
			Eusporangiate Monilophytes (12)	1(1)	-	-	-
			Lycophytes (22)	5(4)	6(1)	-	10(1)
			Hornworts (9)	9(5)	-	3(3)	-
			Liverworts (28)	17(9)	2(1)	10(10)	6(2)
			Bryophyta (41)	1(1)	-	24(18)	6(1)
		Zygnemophyceae (5)		-	-	-	-
		Coleochaetophyceae (3)		-	-	-	-
		Charophyceae (1)		-	-	-	-
	Mesostigmatophyceae (1)		-	-	-	-	-
	Chlorokybophyceae (1)		-	-	-	-	-
	Kbsormidiophyceae (2)		2(2)	1(1)	-	-	-
Chlorophyta	Green algae* (152)		5(4)	1(1)	-	-	-
Rhodophyta	Red algae (28)		3(3)	-	-	-	-

Обозначения: В столбце 1 – приведены названия таксонов согласно классификации NCBI (в скобках – число видов каждого таксона). В столбцах 2-5 – показано число гомологичных последовательностей (в скобках – число видов, у которых обнаружены гомологи). *- база данных The Green Algal Tree of Life project (Cooper, Delwiche, 2016).

Глава 5 «Анализ числа копий генов ферментов пути биосинтеза ауксина» посвящена анализу количеств гомологов ферментов ПБА.

Нами была выявлена положительная и достоверная корреляция между числом гомологов генов ферментов биосинтеза триптофана и сложностью организмов растений.

Для анализа были взяты последовательности ферментов ПБТ у *A. thaliana*: антранилат синтазы α (ASA, идентификаторы TAIR (Lamesch, 2012) AT5G05730, AT2G29690); антранилат синтазы β (ASB; AT1G25220, AT5G57890); фосфорибозил антранилат-трансферазы (PAT(TRP); AT5G17990); фосфорибозил-антранилат-изомеразы (PAI; AT1G07780, AT5G05590, AT1G29410); индол-3-глицерол фосфат синтазы (IGPS; AT2G04400, AT5G48220); триптофан синтазы α (TSA; AT4G02610,

AT3G54640) и триптофан синтазы β (TSB; AT5G54810, AT4G27070). С помощью программы BLASTP 2.2.29+ (E-value <10e-40) проведен поиск гомологов этих ферментов в 24-х полностью секвенированных геномах растений из базы данных Plazma 2.5 (Van Bel, 2012): 5-ти геномов зелёных водорослей, 1-го мха, 1-го плауна, 4-х однодольных и 13-ти двудольных.

Для оценки сложности организмов мы использовали параметр $F_{\text{БАТ}}$ - отношение количества генов белков, ассоциированных с транскрипцией (БАТ) к общему числу генов, кодирующих белки в геноме. $F_{\text{БАТ}}$ может быть точно оценен на основе полногеномных данных и хорошо коррелирует с такой широко известной характеристикой сложности организмов, как число клеточных типов (Lang et al., 2010). Данные по доле генов БАТ, выраженной в процентах, у растений вышеуказанных таксонов взяты из статьи (Lang et al., 2010). Для оценки значимости взаимосвязи между числом гомологов и количеством генов БАТ мы использовали коэффициент корреляции Пирсона.

Анализ взаимосвязи между параметром $F_{\text{БАТ}}$ растений и числом гомологов ферментов синтеза триптофана, сделанный с помощью программы statistica 8.0, выявил статистически значимую положительную корреляцию для 4-х из 7 белков ферментов ПБТ: ASA ($p=0.0025$), ASB ($p=0.031$), PAI ($p=0.0033$), IGPS ($p=0.015$) (табл. 2).

Таблица 2. Зависимость числа гомологов ПБТ от параметра $F_{\text{БАТ}}$ для каждого из ферментов ПБТ

семейство белков	r	p
ASA	0.56	0.0025*
ASB	0.69	0.031*
TRP	0.29	0.27
PAI	0.79	0.0033*
IGPS	0.59	0.015*
TSA	0.43	0.094
TSB	0.38	0.15

В таблице представлены коэффициенты корреляции (r) и уровни статистической значимости (p) для этих двух параметров. * - достоверные значения ($p < 0.05$).

Также была выявлена значимая положительная корреляция между параметром $F_{\text{БАТ}}$ и числом всех гомологов ферментов пути синтеза триптофана у растений ($r=0.77$; $p=0.0005$, рис. 6).

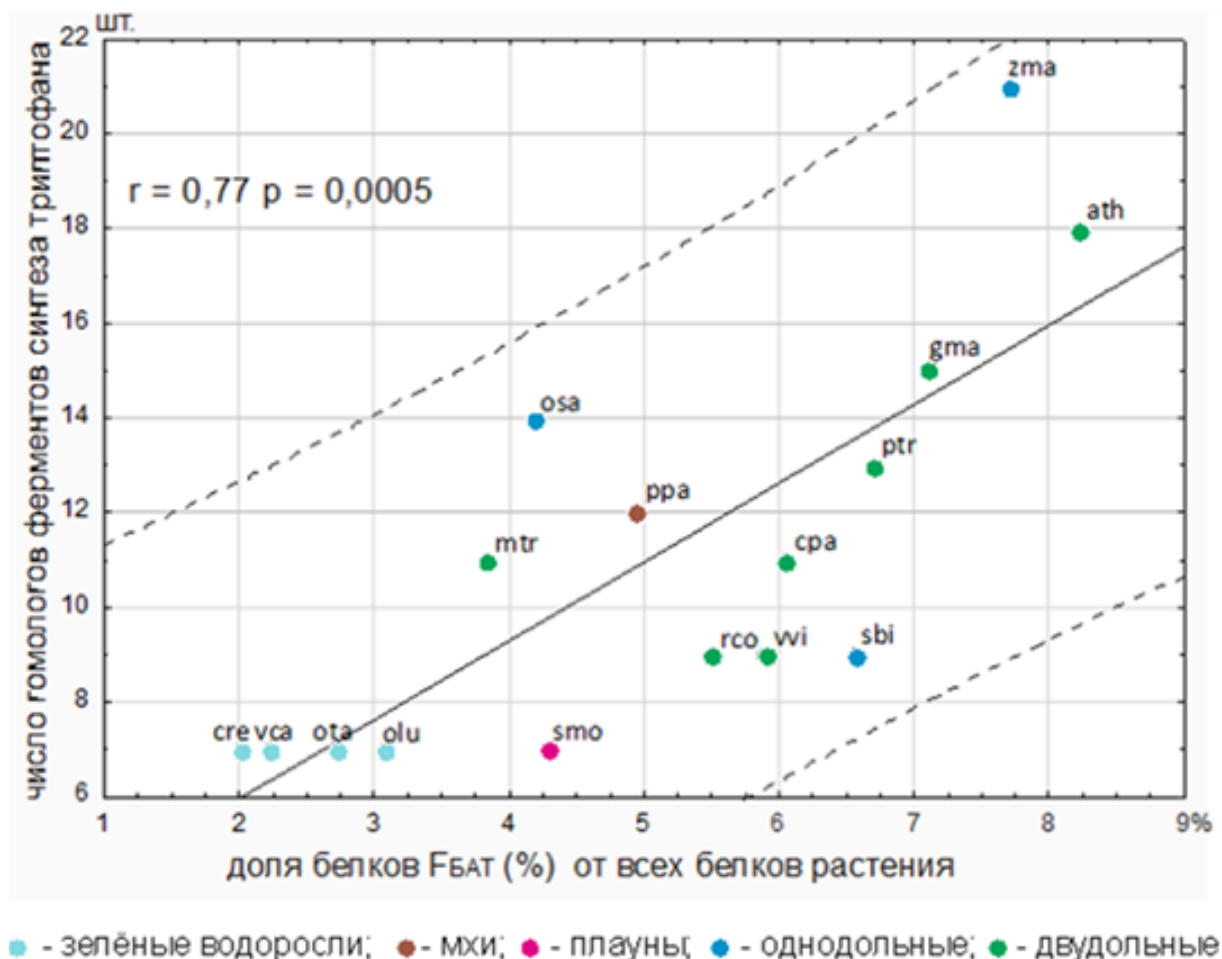


Рисунок 6. Зависимость числа гомологов ферментов пути биосинтеза триптофана (ось X) от доли генов БАТ в геномах растений, F_{BAT} . По оси Y отложено число всех гомологов ферментов ПБТ. По оси X отложено значение параметра F_{BAT} (в %). Точки соответствуют видам и окрашены в зависимости от таксона (соответствие показано внизу рисунка). Обозначение видов: **cre** - *Chlamidomonas reinhardtii*, **ota** - *Ostreococcus tauri*, **olu** - *Ostreococcus lucimarinus*, **vca** - *Volvox carteri*, **ppa** - *Physcomitrella patens*, **smo** - *Selaginella moellendorffii*, **zma** - *Zea mays*, **sbi** - *Sorghum bicolor*, **osa** - *Oryza sativa spp japonica*, **vvi** - *Vitis vinifera*, **ath** - *Arabidopsis thaliana*, **ptr** - *Populus trichocarpa*, **cpa** - *Carica papaya*, **rco** - *Ricinus communis*, **gma** - *Glycine max*, **mtr** - *Medicago truncatula*.

Таким образом, полученный результат свидетельствует о функциональной связи между уровнем дупликаций генов ферментов пути биосинтеза триптофана и усложнением организмов растений в ходе эволюции.

ВЫВОДЫ

1. На основании сравнительного и филогенетического анализа показано, что ферменты пути биосинтеза триптофана произошли у высших растений (Streptophyta) в результате вертикальной эволюции от зелёных водорослей (Chlorophyta).
2. На основе сравнительного анализа последовательностей и структуры гомологов белка KnTAA (NCBI: GAQ80308.1) Харофитовой водоросли *K. nitens* у растительных и нерастительных таксонов впервые показано более высокое сходство KnTAA с аллииназами наземных растений, чем с белками TAA, свидетельствующее об их неспособности выполнять функцию триптофанаминотрансфераз по синтезу индолил-3-пировиноградной кислоты, что позднее было подтверждено независимым экспериментальным исследованием.
3. На основе филогенетического анализа последовательностей белков суперсемейства флавиномоноксигеназ класса В и белка KnFMO (NCBI: GAQ82387.1) *K. nitens* впервые показано, что этот белок принадлежит к новому семейству FMO IIb типа, включающему последовательности бактерий, красных, одноклеточных и многоклеточных зелёных водорослей, наземных растений и грибов и не является родственным к белкам семейства YUCCA.
4. На основе сравнительного анализа последовательностей и структур суперсемейства флавиномоноксигеназ класса В, YUCCA и белка KnFMO *K. nitens* впервые были показаны различия в их последовательностях и структурах, свидетельствующие о неспособности KnFMO выполнять функцию по синтезу ауксина (индолил-3-уксусной кислоты), что предполагает происхождение пути биосинтеза ауксина у наземных растений в результате горизонтального переноса предковых генов от организмов нерастительных таксонов к общему предку наземных растений.
5. Количество копий генов в геноме у ферментов пути биосинтеза триптофана (как для числа всех копий генов ферментов этого пути, так и для некоторых отдельных ферментов ASA/ASB, PAI и IGPS) достоверно коррелирует со сложностью организмов растений.

СПИСОК РАБОТ ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

Основные статьи:

1. **Turnaev, I.I.** The Number of Homologs of Some Enzymes Involved in Tryptophan Biosynthesis Is Correlated to the Proportion of Proteins Associated with Transcription in Plants. / И.И. Турнаев, И.Р. Акбердин, В.В. Суслов, Д.А. Афони́ков // Russian Journal of Genetics: Applied Research. - **2015**. - Т. 5 № 3. - С. 308-312. **IF. Wos 0.6**.
2. **Turnaev, I.I.** Plant auxin biosynthesis did not originate in charophytes. / I.I. Turnaev, K.V. Gunbin, D.A. Afonnikov // Trends Plant Sci. - **2015**. - Т. 20. № 8. - С. 463–465. **IF. 22.5**.
3. **Turnaev, I.I.** The Phylogeny of Class B Flavoprotein Monooxygenases and the Origin of the YUCCA Protein Family. / I.I. Turnaev, K.V. Gunbin, V.V. Suslov, I.R. Akberdin, N.A. Kolchanov, D.A. Afonnikov // Plants. - **2020**. - Т. 9. № 9. - 1092. **IF 4.0**.

Тезисы конференций:

4. **Turnaev, I.I.** Molecular evolution of proteins belonging to auxin biosynthesis gene network in plants. / I.I. Turnaev, Akberdin I.R., Mironova V.V., Omelyanchuk N.A., Afonnikov D.A. В сб.: Abstracts of the eighth international conference on bioinformatics of genome regulation and structure/systems biology (BGRS-SB – 2012). // - Novosibirsk: IC&G Press. - **2012**. - С. 320.
5. **Turnaev, I.I.** Molecular evolution of proteins belonging to auxin biosynthesis gene network in plants. В сб.: Molecular Phylogenetics: Contributions to the 3rd Moscow International Conference "Molecular Phylogenetics", Moscow, Russia, July 31 - August 4, 2012). Compiled by A. Troitsky, L. Rusin, and V. Aleoshin. // I.I. Turnaev, I.R. Akberdin, V.V. Mironova, N.A. Omelyanchuk, D.A. Afonnikov. - M. - 2012. - С. 89.
6. **Turnaev, I.I.** Increasing the number of paralogs for enzymes involved in tryptophan biosynthesis during the evolution of land plants. В сб.: Abstracts of the Ninth International Conference on Bioinformatics of Genome Regulation and Structure/Systems Biology (BGRS/SB, June 23-28, 2014). // I.I. Turnaev, K.V. Gunbin, I.R. Akberdin, V.V. Mironova, N.A. Omelyanchuk, D.A. Afonnikov. - Novosibirsk: Publishing House SB RAS. - 2014. - С. 163.
7. **Turnaev, I.I.** Molecular evolution of yucca protein family. В сб.: The tenth international conference on bioinformatics of genome regulation and structure / systems biology (BGRS-SB – 2016). // I.I. Turnaev, V.V. Suslov, K.V. Gunbin, D.A. Afonnikov. – Novosibirsk. - **2016**. - С. 324.

8. **Турнаев, И.И.** Молекулярная эволюция белков семейства YUCCA, В книге: Беляевские чтения. Тезисы докладов Международной конференции, посвященной 100-летию со дня рождения академика АН СССР Д.К. Беляева. // И.И. Турнаев, В.В. Суслов, К.В. Гунбин, Д.А. Афонников. - Новосибирск: ИЦиГ СО РАН. - **2017**. - С. 181.
9. **Турнаев, И.И.** Анализ эволюции пути биосинтеза ауксина у наземных растений, В сборнике: БИОТЕХНОЛОГИЯ: СОСТОЯНИЕ И ПЕРСПЕКТИВЫ РАЗВИТИЯ. материалы IX международного конгресса. // И.И. Турнаев, К.В. Гунбин, В.В. Суслов, И.Р. Акбердин, Н.А. Колчанов, Д.А. Афонников. - Новосибирск: ИЦиГ СО РАН. - **2017**. - С. 394–395.
10. **Turnaev, I.** Discordant evolution of YUCCA family proteins demonstrated along sequence. В книге: Bioinformatics of Genome Regulation and Structure\Systems Biology (BGRS\SB-2018). The Eleventh International Conference. // I. Turnaev, K. Gunbin, V. Suslov, D. Afonnikov. - Новосибирск: ИЦиГ СО РАН. - **2018**. - С. 192.
11. **Turnaev, I.I.**, Suslov V.V., Afonnikov D.A., Gunbin K.V., Phylogenetic relationships between fmo classes and the origin of YUCCA, В книге: Plant Genetics, Genomics, Bioinformatics, and Biotechnology (PlantGen2019). Abstracts. Eds. A.V. Kochetov, E.A. Salina; Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences. // I.I. Turnaev, V.V. Suslov, D.A. Afonnikov, K.V. Gunbin. - Новосибирск: ИЦиГ СО РАН. - **2019**. - С. 217.
12. **Turnaev, I.** FMO superfamily protein phylogeny and the origin of yucca family. В сб.: Bioinformatics of genome regulation and structure/systems biology (BGRS/SB-2020). The Twelfth International Multiconference Abstracts. // I. Turnaev, V. Suslov, K. Gunbin, D. Afonnikov. - Новосибирск: ИЦиГ СО РАН. - **2020**. - С. 261–262.