

На правах рукописи



Зорин Евгений Андреевич

**АНАЛИЗ ДИФФЕРЕНЦИАЛЬНОЙ ЭКСПРЕССИИ ГЕНОВ
ПРИ ОБРАЗОВАНИИ АЗОТФИКСИРУЮЩИХ КЛУБЕНЬКОВ И
АРБУСКУЛЯРНОЙ МИКОРИЗЫ У *PISUM SATIVUM* L.**

1.5.21. Физиология и биохимия растений

АВТОРЕФЕРАТ

диссертации на соискание учёной степени

кандидата биологических наук

Санкт-Петербург
2023

Работа выполнена в Федеральном государственном бюджетном научном учреждении «Всероссийский научно-исследовательский институт сельскохозяйственной микробиологии»

Научный руководитель

кандидат биологических наук,
Жуков Владимир Александрович

Официальные оппоненты:

Гоголев Юрий Викторович,
доктор биологических наук, профессор,
Федеральное государственное бюджетное
учреждение науки «Казанский институт
биохимии и биофизики Казанского
научного центра Российской академии
наук», заведующий лабораторией

Дорошков Алексей Владимирович,
кандидат биологических наук,
Федеральное государственное бюджетное
научное учреждение «Федеральный
исследовательский центр Институт
цитологии и генетики Сибирского
отделения Российской академии наук»,
научный сотрудник

Ведущая организация: Институт биохимии и генетики Федерального государственного бюджетного научного учреждения Уфимского федерального исследовательского центра Российской академии наук

Защита состоится «27» апреля 2023 г. в 14.00 часов на заседании диссертационного совета 24.1.002.02. на базе Федерального государственного бюджетного учреждения науки Ботанического института им. В.Л. Комарова Российской академии наук по адресу: 197022, г. Санкт-Петербург, ул. Профессора Попова, д. 2. Тел. (812) 342-54-42, факс (812) 372-54-43, Адрес электронной почты: dissovet.24100202@binran.ru

С диссертацией можно ознакомиться в библиотеке и на сайте Федерального государственного бюджетного учреждения науки Ботанического института им. В.Л. Комарова Российской академии наук <https://www.binran.ru/dissertatsionnyye-sovety/dissovet-02/>

Автореферат разослан «___» _____ 2023 г.

Ученый секретарь

диссертационного совета,
доктор биологических наук

Лянгузова Ирина Владимировна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы

Бобовые растения образуют симбиозы с клубеньковыми бактериями (КБ) и грибами арбускулярной микоризы (АМ). Формирование симбиозов повышает устойчивость растений к стрессам и способствует получению стабильного урожая даже в условиях глобального изменения климата. Изучение молекулярных основ симбиозов, образуемых бобовыми, необходимо для повышения эффективности данных симбиозов при их использовании в современном сельском хозяйстве. Поскольку Российская Федерация является одним из мировых лидеров производства гороха (FAOSTAT 2020), целесообразным представляется исследование физиологии, биохимии и молекулярной генетики симбиозов, образуемых горохом посевным (*Pisum sativum* L.).

Ранее для *P. sativum* при помощи мутационного анализа были выявлены отдельные регуляторные гены (т.н. *Sym*-гены), контролирующие основные этапы развития симбиозов. В настоящее время технологии широкомасштабного секвенирования позволяют исследовать полный набор генов, кодирующих «молекулярную машину симбиоза» и необходимых для обеспечения метаболической интеграции симбионтов (Mergaert et al., 2020; Gao et al., 2022). Среди таких генов выделяют гены нодулинов (от англ. nodule – клубенёк), специфичные для бобово-ризобияльного симбиоза (БРС), гены микоризинов, специфичные для арбускулярно-микоризного (АМ) симбиоза, а также гены симбиозинов, – гены, экспрессия которых специфична для развития как БРС, так и АМ (Küster et al., 2007). Гены нодулинов, микоризинов и симбиозинов исследованы на модельных бобовых *Medicago truncatula* Gaertn., *Lotus japonicus* (Regel.) K. Larsen и *Glycine max* (L.) Merr., в то время как симбиоз-специфичные гены гороха посевного, а также особенности их экспрессии, изучены к настоящему моменту недостаточно. Применение подходов транскриптомики (т.е. изучение всего набора экспрессирующихся генов – транскриптома) позволяет восполнить этот пробел в знаниях.

Одним из пост-транскрипционных механизмов, приводящих к увеличению транскриптомного разнообразия, является альтернативный сплайсинг (АС) – считывание нескольких изоформ транскриптов с одного гена (Chaudhary et al., 2019). Роль АС в регуляции экспрессии генов в ходе образования БРС и АМ практически не изучена (Rigo et al., 2019).

Значительная часть транскриптов, характерных для азотфиксирующих клубеньков у модельного бобового *M. truncatula*, относится к семейству генов, кодирующих регуляторные пептиды NCR (nodule-specific cysteine-rich peptides) (Mergaert et al., 2003; Nicoud et al., 2021). Под действием данных пептидов КБ дифференцируются в симбиотическую форму, называемую бактериоиды, и начинают фиксировать атмосферный азот. У гороха посевного это генное семейство не исследовано. Целесообразность изучения генов, кодирующих пептиды NCR, у гороха обусловлено их несомненной важностью

для процессов дифференцировки бактериоидов, а также формирования, функционирования и старения клубенька.

Таким образом, недостаток информации о транскриптомных изменениях в симбиотических системах гороха посевного, а также роли нодулинов, микоризинов и симбиозинов, событий АС, специфичных для симбиотических структур гороха, и особенностей экспрессии генов, кодирующих пептиды NCR, в развитии и функционировании симбиозов гороха посевного, обосновывает актуальность настоящей работы.

Цели и задачи работы. ЦЕЛЬЮ работы является анализ транскриптомных изменений в азотфиксирующих клубеньках и микоризованных корнях гороха посевного (*Pisum sativum* L.).

Для достижения поставленной цели сформулированы следующие ЗАДАЧИ:

- 1) Идентификация генов, экспрессия которых является специфичной для азотфиксирующих клубеньков или микоризованных корней *Pisum sativum* (генов нодулинов и микоризинов, соответственно);
- 2) Выявление генов симбиозинов, повышение экспрессии которых характерно одновременно для клубеньков и микоризованных корней *Pisum sativum*, их характеристика;
- 3) Анализ событий альтернативного сплайсинга, специфичных для клубеньков и микоризованных корней *Pisum sativum*;
- 4) Характеристика семейства генов, кодирующих клубенёк-специфичные пептиды NCR, и изучение разнообразия представителей данного семейства у *Pisum sativum*;
- 5) Анализ разнообразия профилей экспрессии генов, кодирующих пептиды NCR, в клубеньках *Pisum sativum* линии SGE («дикого типа») и мутантных линий с нарушением дифференцировки бактериоидов.

Научная новизна. В работе впервые выявлены гены симбиозинов гороха посевного, в том числе гены, экспрессия которых находится под контролем транскрипционных факторов (ТФ) EFD и IPD3. Показано, что азотфиксирующие клубеньки и микоризованные корни гороха демонстрируют различия профилей альтернативного сплайсинга по сравнению с интактными корнями. Выявлен ряд генов со специфичным для образующих АМ растений паттерном АС. Впервые описан полный набор генов, кодирующих пептиды NCR в геноме гороха посевного, продемонстрирована кластерная организация этих генов и их согласованная экспрессия в зависимости от возраста клубенька. На основании анализа коэкспрессии и исследования промоторных областей генов, кодирующих пептиды NCR, выявлены ТФ, предположительно регулирующие экспрессию данных генов, а также серии других генов, ассоциированных с азотфиксирующим симбиозом.

Практическая значимость. Выявленные гены симбиозинов представляются вероятными кандидатами на роль генов, определяющих эффективность симбиоза гороха с КБ и АМ, и могут быть использованы в дальнейшем при поиске маркеров эффективности симбиозов. Идентифицированные и

описанные антимикробные пептиды NCR могут служить кандидатами на роль антибиотиков и, как следствие, иметь применение в медицине.

Апробация работы. Основные результаты исследования были доложены на российских и международных конференциях: VIII Международная научно-практическая конференция «Биотехнология как инструмент сохранения биоразнообразия растительного мира (физиолого-биохимические, эмбриологические, генетические и правовые аспекты)» (2018 г., Ялта, Республика Крым, Россия), Международный конгресс «VII съезд Вавиловского общества генетиков и селекционеров, посвященный 100-летию кафедры генетики СПбГУ, и ассоциированные симпозиумы» (2019 г., Санкт-Петербург, Россия), 2-ая международная конференция «Растения и микробы: будущее биотехнологии» (2020 г., Саратов, Россия), IV школа-конференция для молодых ученых «Молекулярно-генетические и клеточные аспекты растительно-микробных взаимодействий» (2020 г., Санкт-Петербург, Россия), The 45th FEBS Virtual Congress (2021 г., Любляна, Словения).

Личный вклад автора. Основные результаты, изложенные в диссертации, получены автором самостоятельно в лаборатории генетики растительно-микробных взаимодействий ФГБНУ ВНИИСХМ. Автор осуществлял анализ литературных данных по теме работы, проведение лабораторных исследований, обработку данных, подготовку материалов для публикации статей и докладов на конференциях. В случаях привлечения сторонних организаций (секвенирование MACE-библиотек) автор принимал непосредственное участие в подготовке проб для анализа. В постановке экспериментов участвовали сотрудники лаборатории генетики растительно-микробных взаимодействий ФГБНУ ВНИИСХМ.

Структура и объём диссертации. Диссертация изложена на 126 страницах, состоит из введения, 4 глав, заключения и выводов. Содержит 9 таблиц, 22 рисунка. Список литературы включает 162 источника, в том числе 161 на иностранных языках.

Положения, выносимые на защиту:

- 1) Азотфиксирующие клубеньки и микоризованные корни *Pisum sativum* характеризуются высоким уровнем транскриптомной специфичности: гены нодулинов составляют 48%, а микоризинов 25% от всех генов с индуцированной экспрессией в азотфиксирующих клубеньках и микоризованных корнях, соответственно.
- 2) Впервые проведённый анализ альтернативного сплайсинга в азотфиксирующих клубеньках и микоризованных корнях *Pisum sativum* показал, что в данных органах и тканях основные параметры альтернативного сплайсинга сходны, однако для отдельных мРНК характерны изоформы, специфичные для данных органов и тканей.

- 3) Представители генного семейства, кодирующего пептиды NCR у *Pisum sativum*, демонстрируют высокий уровень разнообразия. Гены, кодирующие пептиды NCR, расположены в геноме в составе кластеров, причём профили экспрессии всех генов из одного кластера сходны. Следовательно, основой молекулярной эволюции генов, кодирующих пептиды NCR, являются дубликации с последующей дивергенцией.

СОДЕРЖАНИЕ РАБОТЫ

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

В данной главе представлены результаты анализа научной литературы по физиологии и молекулярной генетике мутуалистических растительно-микробных симбиозов. Описаны генетические основы БРС и АМ. Представлен современный уровень знаний о генах нодулинов, микоризиннов и симбиозиннов и их роли в реализации консервативных генетических программ, обеспечивающих формирование и функционирование БРС и АМ. Дано описание альтернативного сплайсинга как механизма регуляции экспрессии генов, повышающего транскриптомное и протеомное разнообразие в симбиотических органах бобовых растений. Освещены современные представления о семействе генов, кодирующих короткие цистеин-богатые клубенёк-специфичные пептиды (NCR), у бобовых растений.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. Биологический материал и экспериментальные условия

В работе использован культурный сорт *P. sativum* Frisson, генетическая линия «дикого типа» SGE, а также мутантные линии, полученные на её основе: SGEFix⁻¹ (*sym40-1*) и SGEFix⁻² (*sym33-3*). Симбиотические мутанты гороха посевного SGEFix⁻¹ and SGEFix⁻² несут мутации в генах, кодирующих ТФ Sym40=PsEFD и Sym33=PsIPD3, соответственно (Tsyganov et al. 1998; Ovchinnikova et al., 2011).

Растения выращивали в климатических камерах VB 1014 (Vötsch Industrietechnik, Германия) при температуре 21 ± 1 °C и освещённости 600 ммоль фотонов м⁻² с⁻¹ с циклом день/ночь 16ч/8ч. Все эксперименты были выполнены в трёх биологических повторностях.

Для анализа клубенькообразования растения гороха инокулировали штаммом бактерий *Rhizobium leguminosarum* bv. *viciae* RCAM 1026 (Afonin et al., 2017). Штамм культивировали на твёрдой среде TY при температуре 28°C. Инокуляцию растений линий SGE, SGEFix⁻¹ и SGEFix⁻² проводили водной суспензией КБ с титром не менее 10⁶ КОЕ/мл. В экспериментах с АМ растения сорта Frisson инокулировали 300 мг высушенных корней растений сорго (*Sorghum* sp.), колонизированных *Rhizophagus irregularis* BEG144.

2.2. Молекулярно-биологические и микробиологические методы

Клубеньки, а также микоризованные и интактные корни *P. sativum* растирали в жидком азоте; выделение РНК проводили с применением реагента Тризол (Thermo Fisher Scientific, США), качество препаратов РНК оценивали в 1.5 % агарозном геле, концентрацию измеряли с помощью Qubit Fluorometer и Qubit RNA BR Assay Kit (Thermo Fisher Scientific, США). Реакцию обратной транскрипции проводили согласно протоколу производителя RevertAid Reverse Transcriptase (Thermo Fisher Scientific, США) в автоматическом амплификаторе C1000™ Thermal Cycler (Bio-Rad, США).

Подготовка библиотек для секвенирования мРНК микоризованных и интактных корней гороха сорта Frisson и секвенирование на платформе Illumina были проведены по договору с компанией Macrogen (Сеул, Южная Корея). Библиотеки для секвенирования мРНК клубеньков и корней гороха линий SGE, SGEFix⁻¹ и SGEFix⁻² были приготовлены с использованием набора 3'-MACE kit (GenXPro GmbH, Франкфурт-на-Майне, Германия) и секвенированы на приборе Illumina HiSeq 2500 по договору с компанией GenXPro GmbH (Франкфурт-на-Майне, Германия).

Существование выявленных изоформ было оценено методом ПЦР с дальнейшей визуализацией амплифицированного продукта с помощью геле-электрофореза. Уровень экспрессии изоформ был оценен методом количественной ПЦР в реальном времени.

Оценку экспрессии генов проводили при помощи метода ПЦР в реальном времени на амплификаторе CFX96 Touch Real-Time PCR Detection System. Анализ полученных данных проводили методом 2(-Delta Delta C(T)) (Livak KJ, 2001), построение графиков выполняли с применением библиотеки ggplot2, реализованной в среде программирования R.

Для проверки предсказанной антимикробной активности пептидов NCR для химически синтезированных зрелых пептидов определяли значение полной элиминирующей концентрации, минимальной ингибирующей концентрации (МИК) и минимальной бактерицидной концентрации (МБК). В качестве контроля были взяты антибиотики стрептомицин и ампициллин. Антимикробная активность исследована в отношении бактерий *Rhizobium leguminosarum*, *Sinorhizobium meliloti* и *Arthrobacter mysoarens*. Определение МИК и МИБ проводили с применением красителя резазурина (минимальную концентрацию, при которой происходило изменение цвета красителя, принимали за значение МИК, а минимальную концентрацию, при которой не наблюдалось роста бактериальной культуры, принимали за МБК).

Для оценки наличия повреждений целостности бактериальной мембраны, а также количественной оценки элиминирующей активности PsNCR478 использовали культуру *R. leguminosarum*. Для количественного анализа живых и мёртвых клеток использовали набор LIVE/DEAD® BacLight™ Bacterial Viability and Counting Kit (L34856) с последующим анализом методом проточной цитометрии.

2.3. Методы биоинформатики

Для оценки качества сырых прочтений применяли инструмент FastQC [<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>]; для удаления адаптеров, прочтений низкого качества и деконтаминации от РНК человека и бактерий применяли программу BBDuk из пакета программ BBMap. Картирование парных прочтений на референсный геном гороха посевного линии Frisson проведено в программе STAR (вер. 2.7.3a.) (Dobin and Gingeras, 2015). Для оценки экспрессии и выявления дифференциально экспрессирующихся генов (ДЭГ), картирование проводили с опцией `--quantMode GeneCounts` с последующим количественным анализом в `featureCounts` (вер. 2.0.3) (Liao et al., 2019). Для анализа альтернативного сплайсинга картирование проводили с опцией `--quantMode TranscriptomeSAM` и последующим количественным анализом в Salmon (Patro et al., 2017). ДЭГ были выявлены с помощью пакета DESeq2 (Michael Love, 2017) в среде программирования R. Анализ AC на уровне событий (типов и сайтов) был выполнен в SUPPA2 (Trincado et al., 2018), а анализ на уровне изоформ – в программах SUPPA2, DRIMSeq (Cre, 2017), IsoformSwitchAnalyseR (Vitting-Seerup, 2017). Категории Gene Ontology и KEGG были получены с использованием набора программ Trinotate (<https://github.com/Trinotate/Trinotate.github.io/blob/master/index.asciidoc>), анализ обогащения категориями GO был выполнен в пакете topGO (Adrian Alexa, 2017) (использовали тест Фишера и алгоритм “weight01”), а категориями KEGG – в пакете cluterProfiler. Визуализация данных была проведена в ggplot2 (Wickham, 2016).

Идентификацию генов, кодирующих пептиды NCR, осуществляли в программе SPADA (Zhou et al., 2013), филогенетический анализ проводили в пакетах ggtree (Yu et al., 2017) и phangorn (Schliep, 2011) в среде программирования R. Направление отбора анализировали в программе RNYLIP (Yang, 2007), SNP детектировали с помощью bcftools (Danecsek et al., 2021) и vcftools (Danecsek et al., 2011). Локализацию генов в геноме анализировали с помощью chromoMap, тепловые карты экспрессии строили в пакете pheatmap, коэкспрессионный анализ проводили в пакете SEMitool (Russo et al., 2018), мишени транскрипционных факторов на основе данных коэкспрессии определяли с помощью GENIE3, сканирование промоторных областей на наличие регуляторных участков проводили в программе MEME (Bailey et al., 2009), а поиск потенциальных сайтов связывания с ТФ – в программе SEA (Bailey et al., 2009).

Предсказание антимикробной активности пептидов NCR проводили с помощью сервиса CAMP3 (<http://www.camp3.bicnirrh.res.in/prediction.php>), а 3D моделирование – с помощью программы I-TASSER (Yang and Zhang, 2015).

ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3.1. Идентификация генов нодулинов

Анализ экспрессии генов был проведён в образцах интактных корней 12-дневных растений и клубеньков (12, 21, 28 д.п.и. (дней после инокуляции)) гороха линии SGE. Среди генов со специфично повышенным уровнем экспрессии в молодых клубеньках (12 д.п.и.) бóльшая часть связана, согласно функциональному анализу KEGG, с гормон-опосредованной сигнальной трансдукцией, метаболизмом сахаров, углерода и биосинтезом аминокислот, метаболизмом азота, биосинтезом кутина и суберина и фенилпропаноидов (Рис. 1).

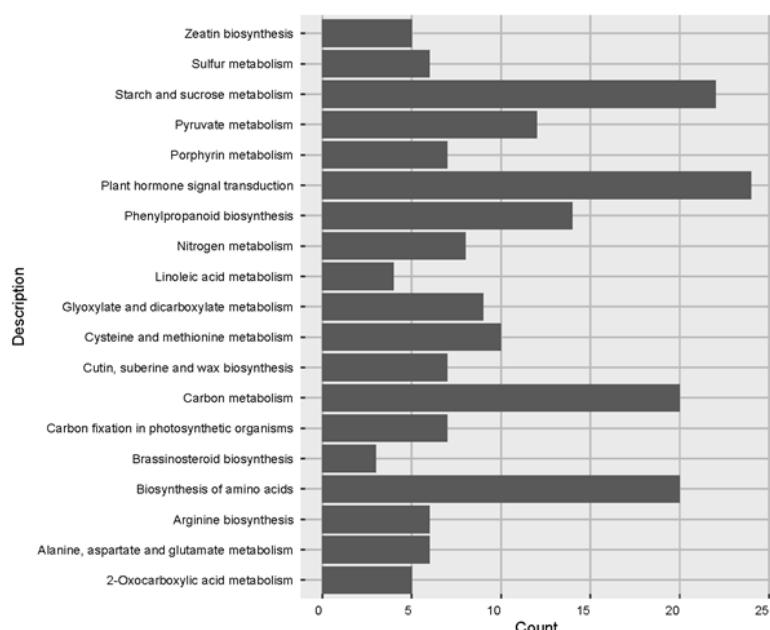


Рисунок 1. Биологические процессы, в которых принимают участие гены нодулинов в клубеньках *P. sativum* на сроке 12 д.п.и.

В клубеньках на сроке 21 д.п.и. активируются гены метаболизма глутатиона, MAPK-сигналинга и взаимодействия с микроорганизмами, по-прежнему активно работают гены метаболизма сахаров и углерода. В клубеньках на сроке 28 д.п.и. также активны гены биосинтеза глутатиона, MAPK-сигналинга и взаимодействия с микроорганизмами, и дополнительно активируются гены метаболизма цистеина и метионина.

Среди идентифицированных генов нодулинов были выявлены гены, находящиеся под контролем TF EFD и IPD3 (с пониженной экспрессией в клубеньках мутантных линий SGEFix⁻¹ и SGEFix⁻², соответственно, в сравнении с клубеньками линии «дикого типа» SGE на сроке 21 д.п.и.). Эти гены вовлечены в процессы гормон-опосредованной сигнальной трансдукции, метаболизма сахаров и азотистых соединений у SGEFix⁻¹, и в процессы гормон-опосредованной сигнальной трансдукции, метаболизма углерода, сахаров, глутатиона, цистеина, метионина и азотистых соединений у SGEFix⁻². Примечательно, что 26% и 15% из всех генов с пониженной

экспрессией у *SGEFix⁻¹* и *SGEFix⁻²*, соответственно, составляют гены, кодирующие пептиды NCR.

3.2. Идентификация генов микоризинов

На основе анализа данных секвенирования мРНК микоризованных и интактных (контрольных) корней 4-недельных растений *P. sativum* сорта Frisson были выявлены 196 генов микоризинов, вовлечённых в такие биологические процессы как ответ на стимулы (преимущественно биотического происхождения), ответ на другой организм, защитные реакции, межорганизменное взаимодействие и гормональный ответ (Рис. 2). Среди выявленных генов микоризинов обнаружены гены с ранее описанной функцией в АМ симбиозе, в частности, *MYB1*, *RAM1*, *RAM2*.

3.3. Выявление и описание генов симбиозиннов у *P. sativum*

В результате пересечения списков идентифицированных генов нодулинов и микоризинов был выявлен набор генов симбиозиннов гороха (86 генов) (Рис. 3). Выявленные гены вовлечены в такие биологические процессы как ответ на биотические и абиотические стимулы, взаимодействие между организмами, защитные реакции, симбиотические взаимодействия и реакции сверхчувствительности.

Анализ экспрессии генов симбиозиннов в клубеньках мутантов по генам *PsSym40=PsEFD* и *Sym33=PsIPD3* показал, что среди них три гена находятся под контролем ТФ EFD и 23 гена - под контролем ТФ IPD3.

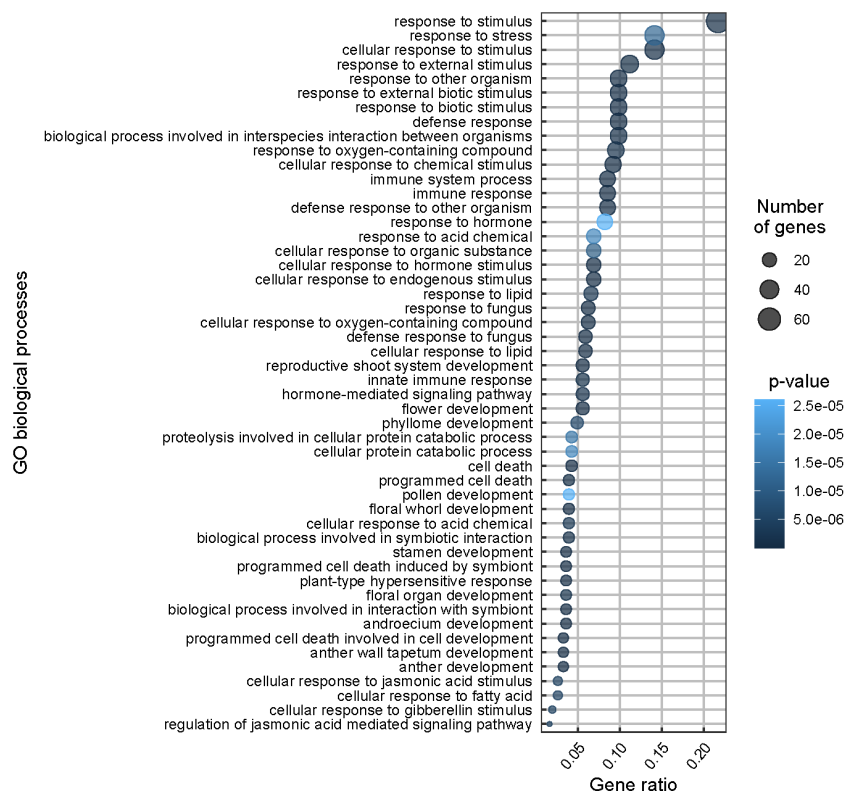


Рисунок 2. Биологические процессы, в которых принимают участие микоризины *P. sativum*

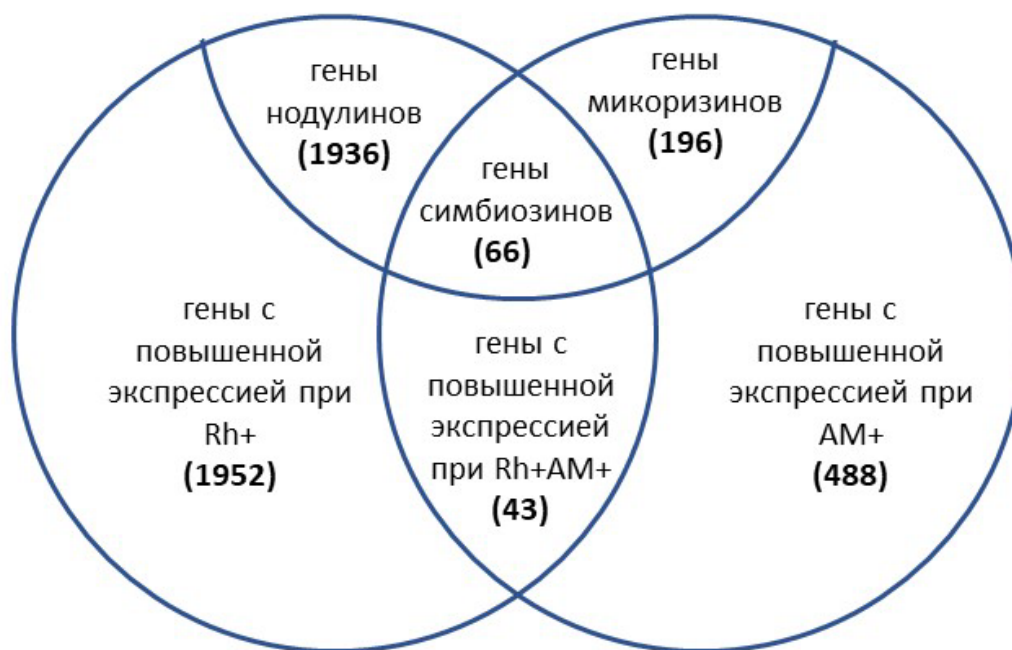


Рисунок 3. Диаграмма Венна, демонстрирующая сходства и различия транскриптомов клубеньков и микоризованных корней *P. sativum*

3.4. Анализ альтернативного сплайсинга в клубеньках и микоризованных корнях гороха посевного

В ходе анализа данных РНК-секвенирования установлено, что соотношения событий АС в клубеньках и в микоризованных корнях гороха не различаются.

Гены, для которых были выявлены изоформы, специфичные для клубеньков, участвуют в защитных реакциях растения, метаболизме нуклеотидов и росте организма. Изоформы, специфичные для микоризованных корней, характерны для генов, связанных с ответом на стресс, гипоксию и белковым автофосфорилированием. Изоформ, общих для клубеньков и микоризованных корней гороха, но отсутствующих в интактных корнях, не выявлено.

Методом количественной ПЦР в реальном времени подтверждено существование специфичных изоформ для транскриптов генов *PsWRKY40* и *PsPR-10* в азотфиксирующих клубеньках.

С помощью компьютерного анализа для 32 генов идентифицированы изоформы, специфичные для микоризованных корней гороха. Из этого списка для пяти генов известна функциональная аннотация (Табл. 1). Среди выявленных генов, ген *ESP3* вовлечён в контроль сплайсинга, что указывает на существование регуляторных механизмов, действующих по типу обратной связи и осуществляющих тонкую настройку реализации генетического экспрессии генов за счёт посттранскрипционных модификаций.

Таблица 1. Гены с известной функциональной аннотацией, для которых описаны изоформы, специфичные для микоризованных корней *P. sativum*

Ген гороха	Гомолог у <i>M. truncatula</i>	Функция
<i>API5</i>	<i>MTR_4g087935</i>	белок-ингибитор апоптоза
<i>AFC3</i>	<i>MTR_3g109390</i>	серин/треониновая протеинкиназа
<i>ESP3</i>	<i>MTR_6g061840</i>	АТФ-зависимая РНК хеликаза DEAN1 (фактор сплайсинга пре-мРНК)
<i>NUS1</i>	<i>MTR_5g007260</i>	дигидролицил дифосфат синтаза
<i>FTSH4</i>	<i>MTR_5g075360</i>	АТФ-зависимая Zn-металлопептидаза

Среди 32 выявленных генов гороха с изоформами, специфичными для АМ корней, для 23 были найдены ортологи у *M. truncatula*, для которых существуют такие же изоформы. В то же время, у *A. thaliana*, модельного растения, не способного формировать АМ, сходные изоформы описаны лишь для 11 ортологичных генов. Предполагается, что образование изоформ транскриптов, сходных у *P. sativum* и *M. truncatula*, указывает на общие молекулярные механизмы, характерные для растений, формирующих АМ, и утраченные у растений, не образующих АМ.

3.5. Поиск генов *P. sativum*, кодирующих пептиды NCR, и анализ их разнообразия и геномной организации.

У бобовых растений, относящихся к группе IRLC (inverted repeat-lacking clade), в том числе, у гороха посевного, в клубеньках экспрессируются гены, кодирующие пептиды NCR (nodule-specific cysteine-rich peptides). Пептиды NCR – это дефензин-подобные молекулы длиной 30-60 аминокислот, содержащие 4 либо 6 цистеиновых остатков в консервативных положениях (Рис. 4). Пептиды NCR способствуют дифференцировке клубеньковых бактерий в симбиотическую форму – бактериоды. Генное семейство, кодирующее пептиды NCR, наиболее полно описано у *M. truncatula*.

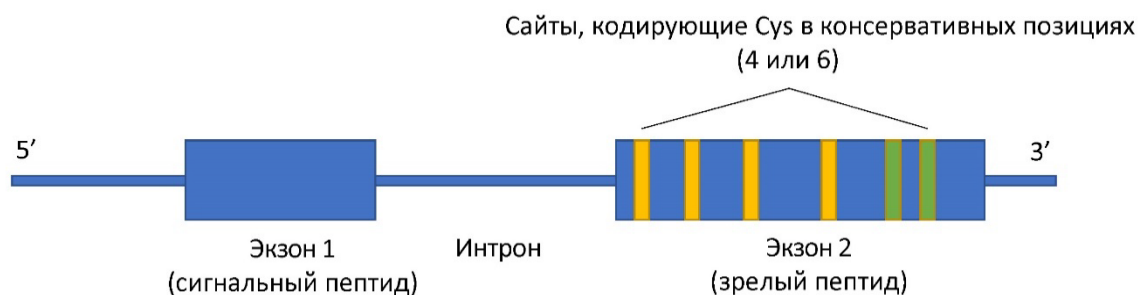


Рисунок 4. Схематичное изображение структуры генов, кодирующих пептиды NCR.

Для идентификации генов гороха, кодирующих пептиды NCR, известные последовательности пептидов NCR *M. truncatula* были использованы как модель для поиска в специализированной программе SPADA. В геноме гороха сорта Frisson (JANEYU000000000 в базе NCBI) поиск выявил 653 последовательности, среди которых 360 были охарактеризованы как кодирующие канонические пептиды NCR (длина >30 аминокислот, 4 или 6 цистеинов в составе, возможность формировать дисульфидные связи, полноразмерный сигнальный пептид). Все указанные 360 генов экспрессируются в клубеньках гороха и кодируют пептиды с 4 или 6 цистеиновыми остатками в составе (Рис. 4).

Последовательности генов пептидов NCR у гороха очень разнообразны и отличаются не только друг от друга (сходство менее 85% на уровне зрелого пептида), но и от гомологичных генов пептидов NCR *M. truncatula* (сходство менее 71% на уровне зрелого пептида). Только одна пара генов гороха и люцерны может рассматриваться как ортологичная (*PsNCR47 – MtNCR312* [*Ps – Pisum sativum*, *Mt – Medicago truncatula*]) со сходством 70.9%. В геномах других бобовых группы IRLC (*Cicer arietinum* L., *Trifolium pratense* L. и *Vicia faba* L.) ортологи генов пептидов NCR гороха не обнаружены, что подтверждает идею независимой эволюции представителей этого семейства генов у различных эволюционных линий бобовых растений.

Аминокислотные последовательности сигнальных пептидов менее полиморфны, чем зрелая часть пептидов NCR гороха, как показано в результате анализа dN/dS статистик отдельно для сигнальной и зрелой части пептида. Зрелая часть пептидов NCR эволюционирует в соответствии с нейтральной моделью эволюции (dN/dS близко к 1), тогда как сигнальная часть подвержена стабилизирующему отбору (dN/dS < 1).

Гены пептидов NCR расположены в геноме в составе кластеров (Рис. 5А). Сходство последовательностей внутри каждого кластера значительно выше, чем сходство последовательностей из различных кластеров (Рис. 5Б), а гены, локализованные в одном кластере, имеют близкий уровень экспрессии в клубеньках (Рис. 5В). Эти факты свидетельствуют о том, что разнообразие генов, кодирующих пептиды NCR, является результатом дупликационных событий, затрагивающих не только кодирующие части генов, но и их промоторные области.

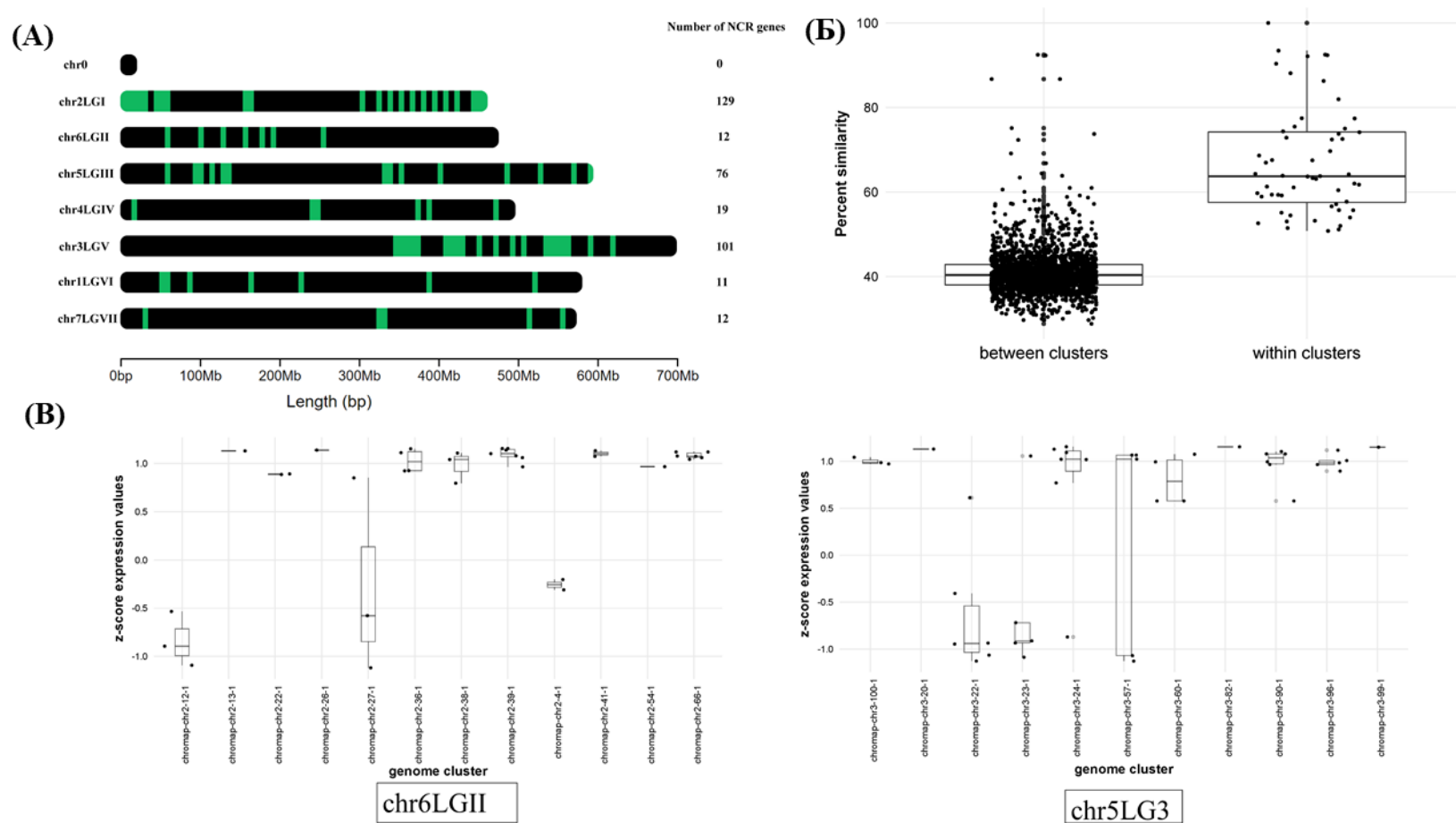


Рисунок 5. Геномная локализация генов, кодирующих пептиды NCR.

(A) Локализация генов пептидов NCR (зелёные вертикальные блоки) в геноме *P. sativum*. (Б) Сравнение усреднённого процента сходства последовательностей генов пептидов NCR внутри и между кластерами на геноме. (В) Оценка уровня экспрессии генов пептидов NCR на сроке 12 д.п.и. внутри и между кластерами в геноме *P. sativum*. Уровень экспрессии в кластерах представлен в виде трансформированного в z-score логарифмированного значения CPM (Counts per million reads). Хромосомы с более выраженным эффектом отобраны для визуализации.

3.6 Предсказание антимикробных свойств пептидов NCR и их экспериментальная проверка

Для аминокислотных последовательностей зрелых пептидов NCR гороха были рассчитаны их физико-химические параметры. Поскольку известно, что антимикробная активность пептидов определяется их зарядом и значением коэффициента гидропатии (индекса Бомана), для проверки антимикробной активности в опытах *in vitro* были и выбраны два пептида NCR – PsNCR276 с максимальным зарядом (+10) и PsNCR478 с максимальным индексом Бомана (3.55). Антимикробная активность химически синтезированных пептидов PsNCR276 и PsNCR478 была подтверждена в микробиологических опытах (Табл. 2).

Таблица 2. Значения МИК и МБК для пептидов NCR и антибиотиков

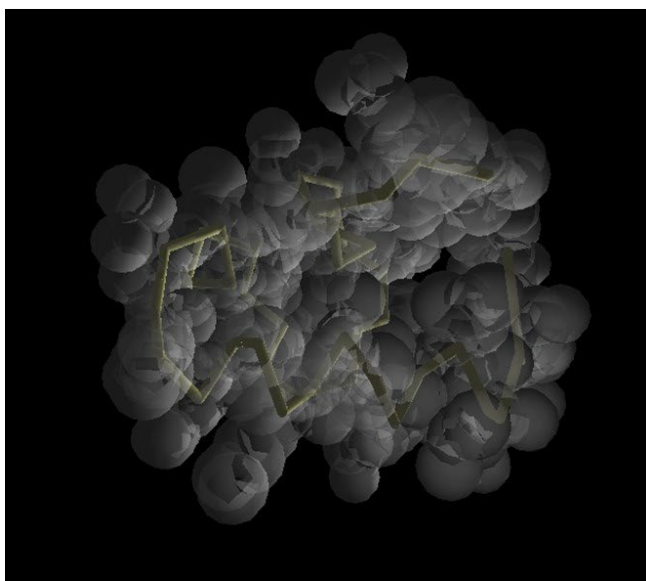
Пептид NCR/ антибиотик	Бактерия	МИК μМ	МБК μМ
PsNCR276	<i>R. leguminosarum</i>	128	>128
PsNCR478	<i>R. leguminosarum</i>	128	>128
Ампициллин	<i>R. leguminosarum</i>	32	>128
Стрептомицин	<i>R. leguminosarum</i>	16	128
PsNCR276	<i>S. meliloti</i>	32	>128
PsNCR478	<i>S. meliloti</i>	32	>128
Ампициллин	<i>S. meliloti</i>	8	>128
Стрептомицин	<i>S. meliloti</i>	>128	>128
PsNCR276	<i>A. mysorens</i>	4	128
PsNCR478	<i>A. mysorens</i>	2	32
Ампициллин	<i>A. mysorens</i>	2	64
Стрептомицин	<i>A. mysorens</i>	<0.1	0.5

Поскольку PsNCR478 показал более высокую антимикробную активность, было дополнительно оценено его действие на экспрессию генов в клетках клубеньковых бактерий *R. leguminosarum* (симбиотический партнёр гороха) и *S. meliloti* (симбиотический партнёр люцерны слабоусечённой). При обработке PsNCR478 в течение 10 минут наблюдалось увеличение уровня экспрессии гена *rpoC* у обоих штаммов бактерий с последующим снижением до контрольного уровня или ниже после 30 минут инкубации, что свидетельствует о включении общих систем защиты бактерий (активации транскрипции и синтеза белков) в ответ на обработку антимикробным агентом.

Моделирование 3D-структуры PsNCR478 (рис. 6), выполненное в программе I-TASSER, показало, что данный пептид, вероятнее всего, не способен к прямому взаимодействию с ДНК, однако способен связывать внутриклеточные катионы (кальций, магний и цинк) за счёт сайта связывания лигандов на С-конце пептида (глутаминовая кислота в положении 38, лейцин

в положении 39 и аспаргин в положении 41). Известно, что липополисахариды внешней мембраны грамотрицательных бактерий связаны с дивалентными катионами (Mg^{2+} , Ca^{2+}), что необходимо для поддержания целостности бактериальной мембраны (Vaara, 1992), и некоторые поликатионные антибиотики группы полимиксина способны взаимодействовать с данными катионами и тем самым нарушать целостность мембраны (Vaara, 1992). Предполагается, что для PsNCR478 данное взаимодействие является основным способом проявления антимикробной активности.

(A)



(Б)

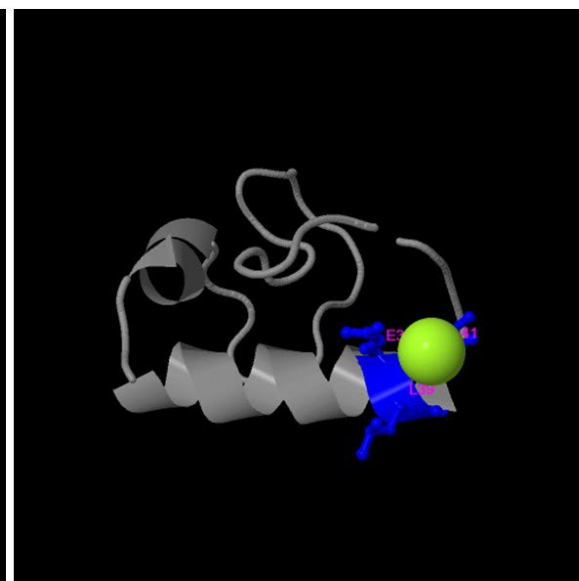


Рисунок 6. (А) 3D структура пептида PsNCR478, полученная методом, реализованным в I-TASSER. (Б) Вероятный сайт связывания лиганда на 3D-структуре молекулы пептида PsNCR478.

3.7. Экспрессионные профили генов пептидов NCR в клубеньках *P. sativum* линии SGE

Для 360 описанных генов пептидов NCR был проведён анализ профилей экспрессии в клубеньках гороха посевного линии SGE на сроках 12, 21 и 28 д.п.и. с использованием данных РНК-секвенирования. Все гены, кодирующие пептиды NCR, были разделены на пять групп в соответствии с профилем их экспрессии. Самые многочисленные группы включают в себя гены с максимумом экспрессии на сроке 12 д.п.и. (259 генов, «ранние») и 28 д.п.и. (58 генов, «поздние») (Рис. 7).

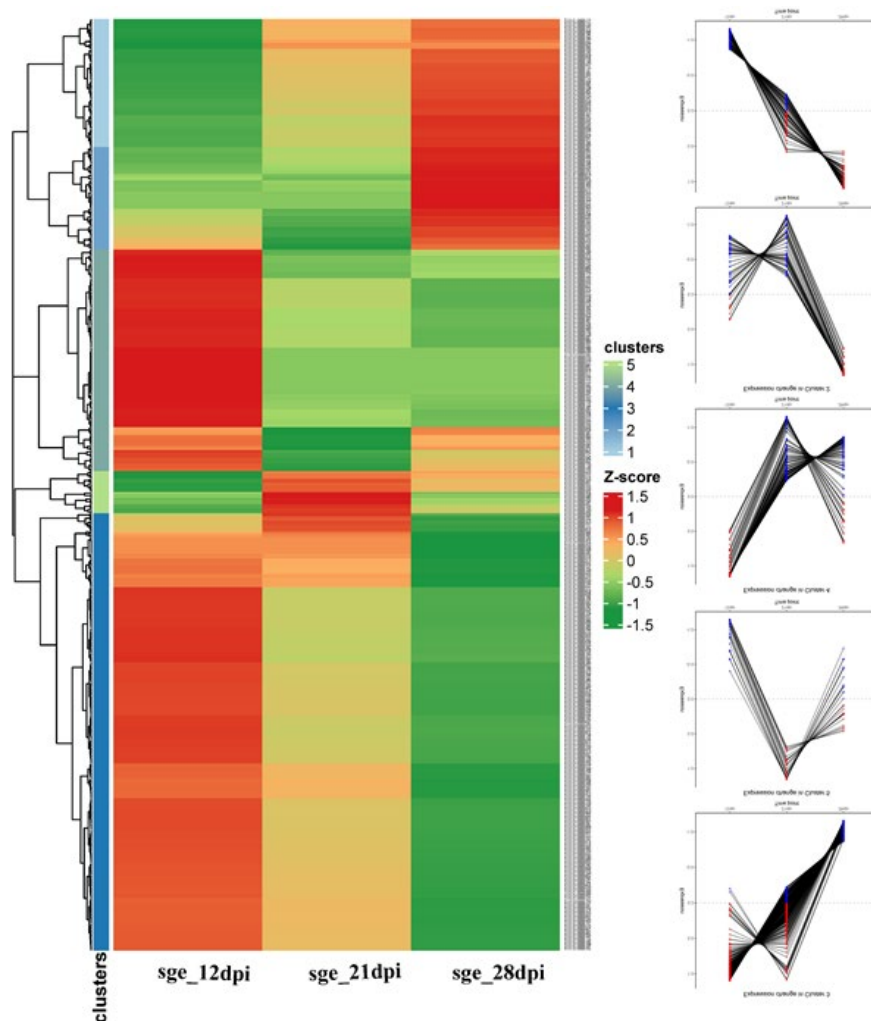


Рисунок 7. Кластерный анализ генов пептидов NCR *P. sativum* линии SGE на основе изменений уровня их экспрессии в клубеньках разного возраста (12, 21 и 28 д.п.и.)

3.8. Анализ экспрессии генов пептидов NCR в клубеньках мутантных линий SGEFix⁻¹ (*sym40-1*) и SGEFix⁻² (*sym33-3*)

В клубеньках линии SGEFix⁻² (*sym33-3*) с мутацией в гене ТФ IPD3, в которых не происходит терминальной дифференцировки бактериоидов, на сроке 21 д.п.и. экспрессия практически всех генов NCR (358 из 360) подавлена по сравнению с клубеньками «дикого типа» линии SGE (Рис. 8А). В клубеньках линии SGEFix⁻¹ (*sym40-1*) с мутацией в гене ТФ EFD, в свою очередь, снижена экспрессия только 189 генов (Рис. 8Б), а уровень экспрессии остальных генов NCR находится на уровне «дикого типа». Большинство из генов NCR со сниженной экспрессией в клубеньках SGEFix⁻¹ относятся к группе «поздних» генов NCR (с максимумом экспрессии на 28 д.п.и. в клубеньках «дикого типа» линии SGE), а гены с экспрессией на уровне

«дикого типа» относятся к группе «ранних» генов NCR (максимум экспрессии на 12 д.п.и. в клубеньках «дикого типа» линии SGE) (Рис. 8Б).

3.9. Анализ коэкспрессии генов, кодирующих пептиды NCR, и их регуляторных элементов

Для установления потенциальных механизмов регуляции экспрессии генов пептидов NCR была проведена кластеризация генов, коэкспрессирующихся вместе с ними в образцах транскриптома клубеньков линии SGE на сроках 12, 21 и 28 д.п.и. Были выявлены три модуля коэкспрессии, содержащих гены, коэкспрессирующиеся с «ранними» и «поздними» генами пептидов NCR. Среди этих генов были выявлены пять генов транскрипционных факторов, потенциально регулирующих экспрессию генов NCR: WRKY40, NAC969, RITF1, PTI5, ERF053 (Рис. 8В). Обнаружено, что ТФ NAC969 регулирует экспрессию только «ранних» генов NCR, а другие выявленные ТФ регулируют экспрессию «ранних» и «поздних» генов NCR.

Аналогичный анализ данных экспрессии генов в клубеньках мутантных линий SGEFix⁻¹ (*sym40-1*) и SGEFix⁻² (*sym33-3*) показал, что ТФ ERN1, вовлечённый в ранние этапы клубенькообразования, и другие ТФ, связанные с развитием и функционированием клубенька (ERF.C.3, ERF34 и BBM1), могут являться регуляторами экспрессии генов пептидов NCR и влиять на проявление мутантного фенотипа у линий SGEFix⁻¹ и SGEFix⁻².

Кроме того, был проведён анализ регуляторных участков (1000 п.о., предшествующие 5'-концу гена) генов, кодирующих пептиды NCR, в результате чего установлено, что регуляторные области «ранних» и «поздних» генов NCR различны. Примечательно, что в промоторных участках ряда генов-симбиозиннов были выявлены те же сайты связывания ТФ, что и у генов NCR. К этим генам относятся, в частности, гены, кодирующие нодулин-13, субъединицу NF-YB1, DELLA белок LA, Nodulin-26-like intrinsic protein (NIP), ранний нодулин-5 и рецептор-подобный белок CLAVATA2.

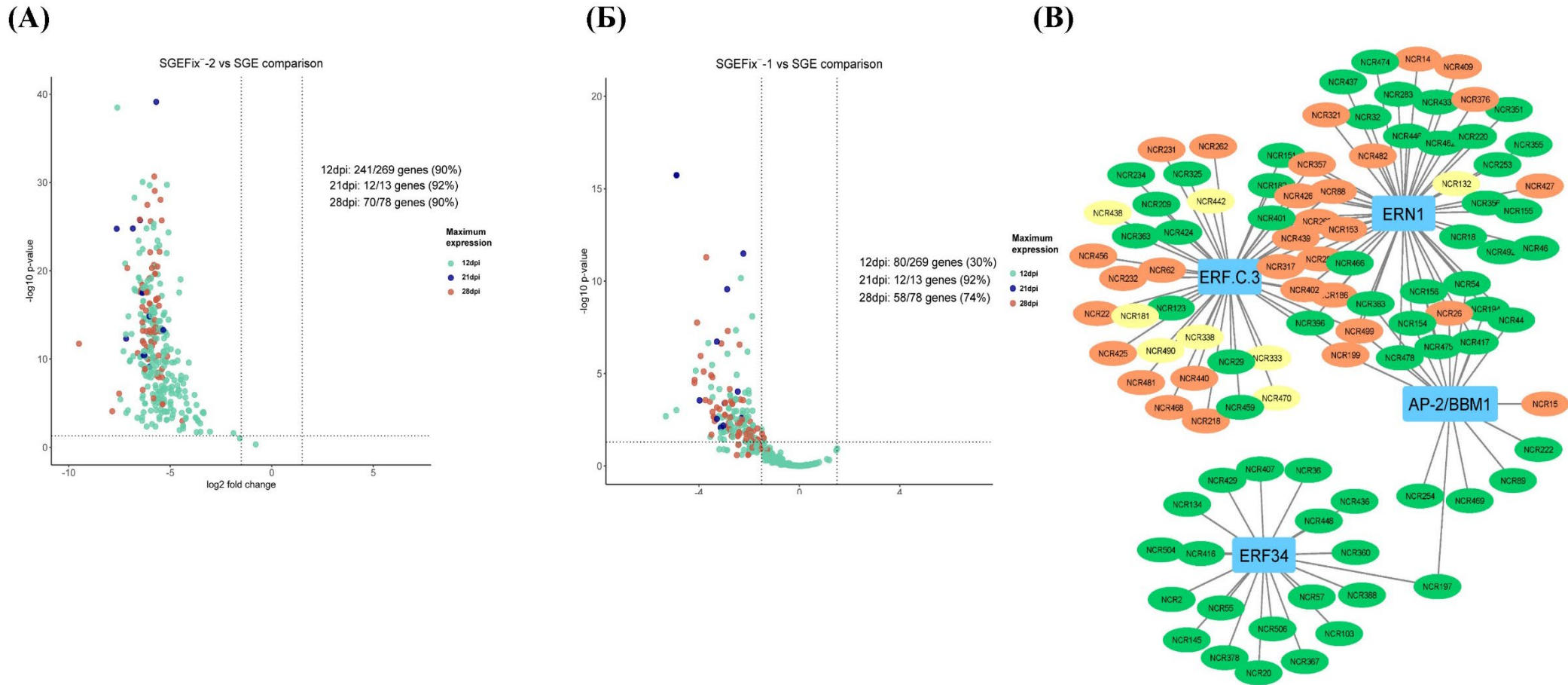


Рисунок 8. Анализ экспрессии и коэкспрессии генов NCR у мутантных линий *P. sativum*.

Распределение групп «ранних» и «поздних» среди дифференциально экспрессирующихся генов NCR у мутанта SGEFix⁻¹ (*sym40-1*) (А) и у мутанта SGEFix⁻² (*sym33-3*) (Б). Регуляторная сеть генов NCR, построенная на данных экспрессии генов в клубеньках «дикого типа» и мутантных линий SGEFix⁻¹ и SGEFix⁻² (В). Гены NCR с максимумом экспрессии на 12 д.п.и. выделены зелёным цветом. Гены NCR с максимумом экспрессии на 21 д.п.и. выделены жёлтым цветом. Гены NCR с максимумом экспрессии на 28 и 42 д.п.и. выделены красным цветом.

ЗАКЛЮЧЕНИЕ

Изучение полного набора генов, экспрессирующихся в азотфиксирующих клубеньках и микоризованных корнях *P. sativum*, показало, что при развитии БРС и АМ индуцируется экспрессия 66 генов, кодирующих симбиозины. Среди этих генов выявлены известные Sym-гены *IPD3* и *CERBERUS*, необходимые для развития обоих типов симбиозов (Roy et al., 2020). Анализ транскриптома клубеньков мутантов по генам *PsSym40=PsEFD* и *Sym33=PsIPD3* показал, что среди 66 генов симбиозинов, три гена находятся под контролем ТФ EFD и 23 гена - под контролем ТФ IPD3.

Были продемонстрированы события альтернативного сплайсинга при развитии азотфиксирующего симбиоза и арбускулярной микоризы у гороха, при этом общие параметры АС (средняя длина интрона, выбор использования различных сплайс-сайтов и распределение типов АС) в клубеньках и микоризованных корнях гороха были сходными. Среди 32 генов, демонстрирующих характерные для микоризованных корней изоформы, для 12 генов описаны изоформы, специфичные для видов, способных образовывать АМ-симбиоз (*P. sativum* и *M. truncatula*), для 11 – изоформы, общие для двудольных растений, и для 9 – изоформы, уникальные для *P. sativum*.

Среди идентифицированных 360 генов, кодирующих пептиды NCR у гороха посевного, бóльшая часть (259) имеет максимальный уровень экспрессии в клубеньках на сроках 12 и 21 д.п.и. («ранние» гены NCR), что соответствует времени дифференцировки свободноживущих бактерий в бактериоиды, а часть генов NCR (58) имеет повышенную экспрессию в точке 28 или 42 д.п.и. («поздние» гены NCR), что соответствует периоду, когда дифференцировка бактериоидов уже завершена. Также показано, что гены NCR в геноме гороха располагаются в кластерах, причём профили экспрессии генов в пределах одного кластера имеют значительное сходство между собой.

Промоторные участки генов NCR гороха имеют сайты связывания с ТФ NLP7, членом семейства NIN-подобных белков, которые накапливаются в ядре в ответ на нитрат и регулируют экспрессию генов, кодирующих пептиды CLE – негативные регуляторы клубенькообразования (Nishida et al., 2018). Результаты диссертационной работы указывают на то, что NLP7 может также регулировать экспрессию ранних генов пептидов NCR. Другой сайт связывания в промоторах генов NCR является мишенью ТФ ауксинового ответа ARF2, который вместе с ARF3 и ARF4 вовлечён в органогенез клубенька посредством регуляции ауксин-зависимых программ развития. Промоторные участки генов NCR гороха содержат консервативные мотивы, присутствующие также в промоторах других симбиотических генов, в том числе генов, связанных с гормональным сигналингом (*nodulin-13* и *gibberellin signaling DELLA protein*) и авторегуляцией клубенькообразования (рецептор-подобный белок *CLAVATA2*). Это сходство промоторных участков может

являться молекулярно-генетической основой связи между системой авторегуляции клубенькообразования и экспрессией генов пептидов NCR.

Экспрессия генов NCR не индуцируется в клубеньках мутантной линии SGEFix⁻² (*sym33-3*), что согласуется с фенотипом мутантных клубеньков, в которых не происходит дифференцировки бактериоидов, в то время как в клубеньках мутанта SGEFix⁻¹ (*sym40-1*) подавлена лишь экспрессия «поздних» генов NCR что также согласуется с фенотипом мутантных клубеньков, в которых дифференцировка бактериоидов не завершена.

Компьютерное предсказание свойств пептидов NCR гороха посевного позволило идентифицировать пептиды с максимальной вероятной антимикробной активностью – PsNCR276 и PsNCR478. В экспериментах *in vitro* PsNCR478 подтвердил свою антимикробную активность, что открывает возможность для его использования в качестве антибиотика.

В целом, выявление новых генетических детерминант, обуславливающих транскриптомное разнообразие симбиотических структур гороха посевного, создаёт основу для дальнейших исследований формирования и функционирования азотфиксирующего и арбускулярно-микоризного симбиозов.

ВЫВОДЫ

1. Среди генов гороха с профилями экспрессии, специфичными для бобово-ризобияльного и арбускулярно-микоризного симбиозов, 66 генов является общими для двух типов симбиозов. Из них под контролем ТФ IPD3 находится 13 генов, а под контролем EFD – 3. При этом транскриптомные изменения в азотфиксирующих клубеньках и микоризованных корнях демонстрируют высокую степень специфичности.
2. Клубеньки и микоризованные корни гороха характеризуются высокой степенью сходства таких общих параметров альтернативного сплайсинга как средняя длина интрона, выбор использования различных сплайс-сайтов и распределение типов альтернативного сплайсинга.
3. Из 32 генов, подвергающихся альтернативному сплайсингу в микоризованных корнях гороха, для 12 генов описаны изоформы, специфичные для видов, способных образовывать арбускулярно-микоризный симбиоз, для 11 – изоформы, общие для двудольных растений, и для 9 – изоформы, уникальные для гороха посевного.
4. Из идентифицированных 360 генов, кодирующих пептиды NCR у гороха посевного, большая часть (259 генов) имеет максимальный уровень экспрессии в клубеньках на сроках 12 и 21 день после инокуляции, что соответствует времени дифференцировки свободноживущих бактерий в бактериоиды, а остальные (58 генов) имеют повышенную экспрессию на сроке 28 дней после инокуляции, что соответствует периоду, когда дифференцировка бактериоидов завершена.

5. Гены, кодирующие пептиды NCR, в геноме гороха располагаются в кластерах, и профили экспрессии генов в пределах одного кластера имеют более высокое сходство, чем профили экспрессии генов из разных кластеров.
6. Экспрессия генов, кодирующих пептиды NCR, не индуцируется в клубеньках мутантной линии SGEFix⁻² (*sym33-3*), что согласуется с фенотипом мутантных клубеньков, в которых выход бактерий из инфекционных капель происходит лишь в отдельных клетках, при этом не происходит дифференцировки бактериоидов, а экспрессия «поздних» генов NCR подавлена в клубеньках мутанта SGEFix⁻¹ (*sym40-1*), что также согласуется с аномальной дифференцировкой в них бактериоидов.
7. По результатам компьютерного предсказания, регуляция экспрессии генов, кодирующих пептиды NCR, может быть опосредована транскрипционными факторами ERN1, ERF.C.3, ERF34 и BBM1.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в рецензируемых научных изданиях, рекомендуемых ВАК РФ:

1. **Zorin E.A.**, Kliukova M.S., Afonin A.M., Gribchenko E.S., Gordon M.L., Sulima A.S., Zhernakov A.I., Kulaeva O.A., Romanyuk D.A., Kusakin P.G., Tsyganova A.V., Tsyganov V.E., Tikhonovich I.A., Zhukov V.A. A variable gene family encoding nodule-specific cysteine-rich peptides in pea (*Pisum sativum* L.) // **Frontiers in Plant Science**. 2022. 13:884726. doi: 10.3389/fpls.2022.884726.
2. **Зорин Е. А.**, Жуков В. А. Симбиотические гены бобовых растений, выявляемые путём транскриптомного анализа // **Биомика**. 2022. Т. 14. №3. С. 285-294. doi: 10.31301/2221-6197.bmcs.2022-26
3. **Zorin E.A.**, Afonin A.M., Kulaeva O.A., Gribchenko E.S., Shtark O.Y., Zhukov V.A. Transcriptome analysis of alternative splicing events induced by arbuscular mycorrhizal fungi (*Rhizophagus irregularis*) in pea (*Pisum sativum* L.) roots // **Plants**. 2020. V. 9. P. 1700. <https://doi.org/10.3390/plants9121700>.
4. **Зорин Е.А.**, Кулаева О.А., Афонин А.М., Жуков В.А., Тихонович И.А. Анализ событий альтернативного сплайсинга в кончиках корней и клубеньках *Pisum sativum* L // **Экологическая генетика**. 2019. Т. 17. №1. С. 53-63. doi: 10.17816/ecogen17153-63.
5. **Зорин Е.А.**, Клюкова М.С., Кулаева О.А., Афонин А.М., Тихонович И.А., Жуков В.А. Идентификация последовательностей, кодирующих NCR-пептиды и дефензины, в метасборке транскриптома азотфиксирующих клубеньков гороха посевного (*Pisum sativum* L.) // **Экологическая генетика**. 2019. Т. 17. №3. С. 39-46. doi: 10.17816/ecogen17339-46.

Публикации в других научных изданиях:

1. Zhukov V., **Zorin E.**, Afonin A., Zhernakov A., Romanyuk D., Kulaeva O., Sulima A., Shtark O., Tikhonovich I. (2021). Molecular genetic bases of high symbiotic performance of pea (*Pisum sativum* L.). 14th European Nitrogen Fixation Conference. September–October 2021, Aarhus, Denmark. P. 53.
2. Gribchenko E.S., Afonin A.M., Kulaeva O.A., **Zorin E.A.**, Shtark O.Yu., Zhukov V.A. The study of transcriptomes of symbiotic tissue of pea using the third-generation sequencing technology Oxford Nanopore. 2nd International Conference “Plants and Microbes: the Future of Biotechnology”. October 5-9 2020. Saratov, Russia. P. 93.
3. **Zorin E.A.**, Kulaeva O.A., Shtark O. Y., Zhukov V.A. Regulation of development of arbuscular-mycorrhizal symbiosis by alternative splicing in garden pea (*Pisum sativum* L.). 2nd International Conference “Plants and Microbes: the Future of Biotechnology”, October 5-9 2020. Saratov, Russia. P.291.
4. **Зорин Е.А.**, Ключкова М.С., Афонин А.М., Тихонович И.А., Жуков В.А. Поиск представителей семейства NCR-пептидов гороха посевного (*Pisum sativum* L.), обладающих антимикробной активностью. / В книге: VII Съезд Вавиловского общества генетиков и селекционеров, посвященный 100-летию кафедры генетики СПбГУ, и ассоциированные симпозиумы. Сборник тезисов Международного Конгресса. 18-22 июня 2019. Санкт-Петербург, Россия, С. 951
5. **Зорин Е.А.**, Ключкова М.С., Афонин А.М., Жуков В.А., Тихонович И.А. Идентификация последовательностей NCR-пептидов в собранном de novo «метатранскриптом» клубеньков *Pisum sativum* L. / В книге: Биотехнология как инструмент сохранения биоразнообразия растительного мира (физиолого-биохимические, эмбриологические, генетические и правовые аспекты) Материалы VIII Международной научно-практической конференции. 1-5 октября 2018. Ялта, Россия. С. 214.
6. **Зорин Е.А.**, Афонин А.М., Кулаева О.А., Тихонович И.А., Жуков В.А. Альтернативный сплайсинг в азотфиксирующих клубеньках гороха посевного. Биология - наука XXI века. Материалы 22-й Международной Пущинской школы-конференции молодых ученых. 23-27 апреля 2018. Пущино, Россия. С. 121.

Благодарности

Автор выражает признательность и благодарность своему научному руководителю В.А. Жукову за предоставление интересной темы диссертации, за обсуждение результатов, конструктивную критику, помощь и поддержку. Автор выражает благодарность своим коллегам, сотрудникам лаб. генетики растительно-микробных взаимодействий ФГБНУ ВНИИСХМ Штарк О.Ю., Ахтемовой Г.А., Кулаевой О.А., Романюк Д.А., Ключковой М.С., Сулиме А.С., Жернакову А.И., Дворяниновой Л.Е., Кичигиной Н.Е., Гордону М.Л., а также Афонину А.М. за помощь в рабочем процессе и дружескую атмосферу.