

## ОТЗЫВ

официального оппонента  
на диссертацию Шмакова Николая Александровича, представленную на соискание ученой степени кандидата биологических наук, на тему: «Поиск генов, ассоциированных с частичным альбинизмом и меланизмом у ячменя *Hordeum vulgare* L., на основе анализа транскриптомных данных», по специальностям 1.5.8. – «Математическая биология, биоинформатика» и 1.5.7. – «Генетика»

### Актуальность темы диссертационной работы

Механизмы контроля функционирования живых организмов, включая растения, является предметом исследования во многих странах. Такой контроль осуществляется на нескольких уровнях: организации хроматина, транскрипции, процессинга и сплайсинга пре-мРНК, трансляции, стабильности белкового продукта и модификации белков. Особый интерес представляют исследования, связанные с выяснением функциональной роли генов, формирующие метаболические пути биосинтеза растительных пигментов, которые играют значительную роль в механизмах контроля как в нормальных условиях жизнедеятельности, так и при действии стрессовых факторов.

Современные экспериментальные опіс инструменты геномного мониторинга экспрессии генов растений позволяют проследить за потоком генетической информации от генома до протеома и метаболома. Применение новых экспериментальных подходов, таких, например, RNA-Seq, прояснили многие ключевые механизмы регуляции транскрипции растений, как первого этапа экспрессии генов, и наиболее простого в методическом плане для экспериментального исполнения. Исследования транскриптомов, т.е. качественной и количественной оценки профиля экспрессии всего пула генов на уровне генома, позволили получить убедительные доказательства динамического изменения транскриптомов, как в процессах роста и развития, так и при действии разнообразных факторов окружающей среды у разных видов растений.

Несмотря на большое количество транскриптомных данных, полученных для различных видов растений, около 30% растительных генов до сих пор не имеют даже аннотации или их описания слабо детализированы. Это обусловлено, в том числе, и выбором конвейеров программ, используемых для анализа данных RNA-seq, и поиска оптимальных параметров анализа. И этот факт вносит существенные ограничения для функциональной аннотации генов растений. В частности, крайне скупы экспериментальные данные в отношении генов, ассоциированных с формированием частичного альбинизма и меланизма у колоса ячменя, важной сельскохозяйственной культуры. Это делает актуальным выявление функциональных особенностей генов, вовлеченных в метаболические пути биосинтеза пигментов с использованием максимально доступных методов анализа.

Такие исследования важны как с фундаментальной точки зрения, так и с практической. Диссертационная работа Николая Александровича Шмакова посвящена выявлению генов, ассоциированных с формированием частичного альбинизма и меланизма у колоса ячменя, на основе биоинформатического анализа транскриптомов сорта Bowman и почти изогенных линий i:BwAlm и i:BwBlp, а также выявлению функциональных особенностей этих генов и их

роли в метаболических путях биосинтеза пигментов. Принимая во внимание вышеизложенное, актуальность этой работы не вызывает сомнения.

### **Структура и содержание диссертационной работы**

Диссертационная работа состоит из введения, четырех глав, выводов, списка цитируемой литературы и приложений.

После краткого введения, в котором определены цель и задачи исследования, проведен анализ литературных источников, которые имеют непосредственное отношение к изучаемой проблеме. Обзор литературы охватывает широкий круг проблем и включает 3 раздела.

Обзор литературы каждого из разделов охватывает широкий круг вопросов, а именно: (1) в первой части дается краткое описание объекта исследования. Особое внимание уделено значению этой культуры в современном сельском хозяйстве; (2) во второй главе проанализированы литературные данные о состоянии исследований в отношении генома ячменя, и демонстрирует, что ячмень является подходящей моделью для исследования генетики злаков и особенностей генетического контроля реакции растений на изменяющиеся условия среды; (3) в третьей главе представлены современные данные о биологии пластид, аппарате транскрипции генов и аппарате синтеза белка, а также важность изучения развития пластид, синтеза хлорофилла, обратных сигналов из пластид в ядро, и генетической регуляции процессов, включая альбинизм и меланизм; (4) четвертая глава посвящена современным omics технологиям для исследования транскриптомов растений, включая характеристику платформ для секвенирования второго поколения, описание основных этапов RNA-seq, стадии биоинформатического анализа RNA-seq, подробному анализу алгоритмов и программных продуктов, используемых на каждой из этих стадий, и их вклад в прояснение ключевых биологических процессов растительной клетки.

В целом обзор литературы написан хорошим языком и касается тех проблем, которые имеют непосредственное отношение к теме диссертационной работы. Следует отметить, что все литературные данные анализируются соискателем квалифицированно и подробно, поэтому цель и задачи, поставленные автором работы, звучат вполне убедительно.

Традиционно после обзора литературы приводится описание материалов и методов исследования. В этой главе соискателем детально изложены методические особенности и приемы работы. Использован целый арсенал современных методов, применяемых в мировой практике генетических исследований растений. Следует отметить вполне удовлетворительную разрешающую способность избранных для работы методов и в ряде случаев их успешную модификацию с учетом специфики проводимых исследований.

Аналитическое рассмотрение данных, представленных в разделах "Результаты", позволяет заключить следующее: соискателем была предпринята серия экспериментов, в целом, спланированных на хорошем профессиональном уровне, которые позволили полностью решить поставленные в ходе работы задачи.

Эта глава включает два раздела.

Первоначально соискателем проведен анализ транскриптома почти изогенной линии ячменя *i:BwAlm* в сравнении с *Bowman*, эти исследования связаны с прояснением вопроса о списке генов с дифференциальной транскрипцией у ячменя сорта *Bowman* по сравнению с линией *i:BwAlm*. Для этого, Николай Александрович выбирает путь исследования, основанный на комплексном сравнительном анализе транскриптома леммы ячменя линии *i:BwAlm* и сорта *Bowman*, имеющего нормальный фенотип по окраске колоса. В сравнительном анализе диссертант использовал 36 конвейеров для транскриптома линии *i:BwAlm*. Результаты проведенного исследования в эксперименте с линией *i:BwAlm* показали высокий рейтинг конвейера, включающий программу *Bowtie* для удаления рНК, программу *Dart* для картирования и пакета *EdgeR* для оценки дифференциально экспрессируемых генов. Дополнительный анализ терминов генной онтологии и метаболических путей, позволил выявить статистически значимо обогащённые для этих генов термины. Так, было убедительно доказано, что в лемме ячменя линии *i:BwAlm* большинство дифференциально экспрессирующихся генов имеют пониженный уровень экспрессии по сравнению с сортом *Bowman*; а их функция связана с аэробным дыханием, фотодыханием и фотосинтезом, т.е. гены, которые вовлечены в метаболические пути фотосинтеза, аэробного дыхания и усвоения азота.

Поскольку диссертант убедительно продемонстрировал, что транскриптомные карты растений значительно расширяют возможности исследователя для детального анализа важных биологических процессов, соискатель ставит перед собой еще одну важную задачу. Это задача – поиск и функциональный анализ транскриптов, обнаруженных в транскриптоме сравниваемых линий, но не аннотированных ранее в геноме ячменя.

Для этого соискатель проводит *de novo* реконструкции линии *i:BwAlm* и сорта *Bowman*. В сборке транскриптома обнаружены новые контиги, а сравнительный анализ транскриптомов линии *i:BwAlm* и сорта *Bowman*, позволил выявить в линии *i:BwAlm* белок-кодирующий ген, с высоким сходством к прохибитин-1-подобному белку *Solanum pennellii*, который характеризуется высоким уровнем транскрипции в этой линии ячменя и нулевой в транскриптомах у сорта *Bowman*. Помимо этого, этот ген локализован в коротком плече хромосомы ячменя 3Н и может быть ассоциирован с проявлением альбинизма.

Второй раздел главы «Результаты» посвящен анализу транскриптома почти изогенной линии ячменя *i:BwBlp* в сравнении с *Bowman*, выбор изогенной линии ячменя *i:BwBlp* для исследований обоснован и обусловлен тем, что изогенная линия *i:BwBlp* создана на основе сорта *Bowman*, при этом контрастна с ним по окраске колоса – плотно прилегающая к перикарпу зерна цветковая чешуя (лемма) этой линии содержат меланин, что придаёт колосу чёрный цвет.

Соискателем проведён комплексный сравнительный анализ транскриптома линии ячменя *i:BwBlp* с использованием конвейера, включающий программу *Bowtie* для удаления рНК, программу *Dart* для картирования и пакета *EdgeR* для оценки дифференциально экспрессируемых генов. Дальнейший анализ терминов генной онтологии и метаболических

путей, позволил выявить статистически значимо обогащённые для этих генов термины. Так, было убедительно продемонстрировано, что гены, дифференциально экспрессирующиеся в лемме ячменя линии i:BwB/p и растениях сорта Bowman, функционально связаны с биосинтезом о-дихинонов и фенилпропаноидов, и процессом фотосинтеза. При этом отмечено, что гены, понижающие транскрипцию в линии i:BwB/p вовлечены в метаболический путь ассимиляции азота, а также цикл Кальвина-Бенсона и «шунт РБФК».

Совокупность полученных результатов позволили Николаю Александровичу Шмакову сделать общие заключения: (1) формирование частичного дефицита хлорофилла в колосе ячменя (*Hordeum vulgare* L.) мутантной линии i:BwA/m сопровождается понижением уровня экспрессии генов фотосинтеза, аэробного дыхания и усвоения азота, а также активацией в клетках оболочки зерновки гена, локализованного в коротком плече хромосомы 3Н и кодирующего белок с доменом прохибитина и (2) формирование меланиновой окраски колоса ячменя в линии i:BwB/p связано с повышением экспрессии генов в перикарпе зерновки и цветковой чешуе, участвующих в биосинтезе о-дихинонов и фенилпропаноидов.

В четвертой главе диссертантом представлено обсуждение примененных биоинформатических подходов обработки данных RNA-seq, полученных результатов транскриптомного анализа линий i:BwA/m и i:BwB/p, включая результаты функционального анализа дифференциально экспрессирующихся генов и *de novo* реконструированных транскриптомов.

#### **Степень новизны результатов научных исследований.**

Соискателем впервые проведен сравнительный транскриптомный анализ почти изогенных линий ячменя, контрастных по окраске колоса. Выявлены закономерности функционального профиля дифференциально экспрессируемых генов, а именно, продемонстрировано, что формирование частичного дефицита хлорофилла в колосе ячменя мутантной линии i:BwA/m сопровождается понижением уровня экспрессии генов фотосинтеза, аэробного дыхания и усвоения азота, а также активацией в клетках оболочки зерновки гена, локализованного в коротком плече хромосомы 3Н и кодирующего белок с доменом прохибитина, а формирование меланиновой окраски колоса ячменя в линии i:BwB/p связано с повышением экспрессии генов в перикарпе зерновки и цветковой чешуе, участвующих в биосинтезе о-дихинонов и фенилпропаноидов.

Приоритетными можно назвать и результаты по сравнительному анализу нескольких конвейеров биоинформатической обработки библиотек RNA-seq с последующим выбором наиболее оптимального конвейера. Использование нескольких сборщиков транскриптома *de novo* и последующее объединение полученных сборок в одну общую гибридную сборку повышает точность и чувствительность в определении последовательностей транскриптов.

#### **Научная и практическая значимость результатов**

Диссертационная работа Николая Александровича Шмакова совмещает в себе и фундаментальность, и практическую значимость. Полученные соискателем результаты важны

для развития фундаментальных представлений о молекулярных механизмах регуляции транскрипции генов и функционального анализа генов растений. С практической точки данная работа интересна тем, что полученные данные могут быть востребованы при решении вопросов повышения устойчивости культурных растений к неблагоприятным факторам внешней среды, а также при выведении новых высокопродуктивных линий сельскохозяйственных культур и разработке акклимационных моделей взаимодействия природы и растительного мира.

#### **Обоснованность и вероятность заключительных выводов**

Использование для исследований классических и современных биоинформатических, молекулярно-биологических и генетических методов, а также методов анализа экспериментального материала подтверждают обоснованность и достоверность экспериментальных результатов, представленных в диссертационной работе Николая Александровича, а также выносимых на защиту положений и выводов.

#### **Полнота опубликованности положений и результатов диссертации**

Основные положения и результаты исследований по диссертации Николая Александровича Шмакова опубликованы в 4 статьях в престижных зарубежных изданиях, рекомендованных ВАК, доложены на международных и российских научных форумах. Рукопись автореферата соответствует содержанию рассматриваемой диссертации, результатам и положениям, выносимым на защиту.

#### **Вопросы, замечания и комментарии к диссертационной работе**

При аналитическом рассмотрении представленных в диссертационной работе материалов возникло ряд вопросов:

1. На основании каких соображений или данных выбрана платформа секвенирования Ion Torrent PGM, в частности для анализа дифференциальной экспрессии генов, поскольку данная платформа характеризуется сравнительно низкой производительностью и низкой точностью в гомополимерных участках?
2. Предложенные соискателем вычислительные конвейеры с использованием программ биоинформатического анализа транскриптомных данных RNA-seq для реконструкции и анализа последовательностей транскриптов будут также эффективны при анализе транскриптомов других видов растений или транскриптомных данных, полученных на других платформах секвенирования?
3. Диссертантом не указаны сведения о размещении полученных первичных данных (Accession ID) в открытых базах данных (например, SRA - Sequence Read Archive (SRA) data), а только приводится описание. Вопрос: размещены ли первичные данные в открытых базах данных?

4. Соискатель указывает, что в ходе исследования «предложен метод оптимизации вычислительного конвейера для биоинформатического анализа экспериментов RNA-seq, повышающий точность оценки дифференциальной активности генов, который основан на использовании данных независимой верификации изменения экспрессии генов с помощью ОТ-ПЦР». Поскольку данные анализа экспериментов RNA-seq в большинстве исследований, как правило, сопровождаются данными количественной ПЦР в реальном времени для определенного набора генов, возникает вопрос: что подразумевает Николай Александрович под независимой верификацией изменения экспрессии генов с помощью ОТ-ПЦР?

Это осталось в работе без пояснения или обсуждения.

По разделам диссертационной работы Николая Александровича Шмакова имеется ряд замечаний и пожеланий, которые могут быть учтены в дальнейших работах соискателя.

Ко всем разделам диссертационной работы:

- в тексте имеются некоторые стилистические погрешности неточности и неудачные выражения, и не вполне профессиональное использование некоторых терминов и обозначений. Например, диссертант указывает, что «В сборке транскриптома было обнаружено 943 новых контига, содержащих по одной ОРС. Из них 578 **имеют гомологию** к геному линии Wowman. Белковые продукты остальных новых контигов **имеют гомологию** с 97 известными последовательностями из базы NCBI Protein». Следует отметить, что сравниваемые белки либо гомологи, либо не гомологи, а оценка того, факта, что белки являются гомологами либо не являются, более профессионально приводить оценку гомологов в терминах «идентичность», «сходство».

- частая подмена термина «транскрипция» на термин «экспрессия». Например, при описании в разделах «Экспрессия генов пластома», «Экспрессия генов района Alm», «Экспрессия генов района Vlp» и т.д. Транскрипция – это, безусловно, важный этап экспрессии генов, и ее эффективность вносит значительный вклад, однако транскрипция не единственный биологический процесс, определяющий преобразование наследственной информации от гена в функциональный продукт, прежде всего белок. Экспериментально доказано, что регуляция экспрессии проходит и на уровне трансляции.

Все замечания к работе исчерпываются выше названными, большинство из которых, видимо, следует отнести к разряду досадных неточностей в оформлении работы. Высказанные замечания не носят принципиального характера, не затрагивают сути научных выводов, сделанных диссертантом, и не умаляют значения представленной работы, выполненной, в целом, на высоком научном и методическом уровне, и оставляющей, в целом, хорошее впечатление. Следует еще раз отметить правильность выбранной стратегии исследования и высокую квалификацию исполнения, что положительно характеризует самого исследователя.

## ЗАКЛЮЧЕНИЕ

Диссертация на тему «Поиск генов, ассоциированных с частичным альбинизмом и меланизмом у ячменя *Hordeum vulgare* L., на основе анализа транскриптомных данных» по актуальности, новизне, теоретической и практической значимости соответствует критериям пп. 9-14 «Положения о порядке присуждения ученых степеней», утвержденного постановлением Правительства РФ от 24 сентября 2013 г. № 842, и представляет собой завершённую научно-квалификационную работу, а ее автор, Николай Александрович Шмаков заслуживает присуждения ученой степени кандидата биологических наук по специальностям 1.5.8. – «Математическая биология, биоинформатика» и 1.5.7. – «Генетика».

Официальный оппонент:


Доктор биологических наук,

Руководитель лаборатории функциональной геномики

Федерального государственного бюджетного учреждения науки

Институт физиологии растений им К.А. Тимирязева

Российской академии наук,

 Голденкова-Павлова Ирина Васильевна

«05» февраля 2024 года

Контактные данные:

тел. +7 (499) 678-53-56; E-mail: [irengold58@gmail.com](mailto:irengold58@gmail.com); Специальность, по которой официальным оппонентом защищена диссертация: 03.01.07 – генетика

Адрес места работы:

127276 Российская Федерация, г. Москва, ул. Ботаническая, дом 35, Федеральное государственное бюджетное учреждение науки Институт физиологии растений им. К.А. Тимирязева Российской академии наук, группа функциональной геномики

Подпись сотрудника Федерального государственного бюджетного учреждения науки Института физиологии растений им. К.А. Тимирязева Российской академии наук Ирины Васильевны Голденковой-Павловой удостоверяю:

Заместитель директора по научной работе

Федерального государственного бюджетного учреждения науки

Института физиологии растений им К.А. Тимирязева

Российской академии наук,

к.б.н.



Александр Валерьевич Карташов

«05» февраля 2024 года

