

УТВЕРЖДАЮ
И.о. проректора по науке и
инновациям НИУ «БелГУ»,
кандидат технических наук


В.В. Мишунин
« 23 » сентября 2023 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

федерального государственного автономного образовательного учреждения высшего образования «Белгородский государственный национальный исследовательский университет (НИУ «БелГУ»)» на диссертационную работу Дмитрия Аркадьевича Афонникова «Компьютерные методы высокопроизводительного фенотипирования растений», представленной к защите в диссертационном совете 24.1.239.01 на базе ФГБНУ «Федеральный исследовательский центр Институт цитологии и генетики СО РАН» на соискание ученой степени доктора биологических наук по специальности 1.5.8. Математическая биология, биоинформатика.

Актуальность исследования. Диссертационная работа Дмитрия Аркадьевича Афонникова посвящена созданию экспериментально-компьютерной платформы для высокопроизводительного компьютерного фенотипирования растений на основе анализа цифровых изображений для оценки характеристик опушения листа, размера, формы и цвета зерен, формы колоса и изучение с ее помощью взаимосвязи признаков фенотипа растений с генотипом и ответом на воздействие окружающей среды.

В последнее время технологии секвенирования ДНК позволяют недорого и быстро прочитывать последовательности как полных геномов индивидуальных организмов, так и отдельных его локусов. Как результат, высокопроизводительные технологии секвенирования привели к созданию новых методов селекции, таких как маркер-контролируемый отбор и геномная селекция. Разработка методов феномики позволяет значительно ускорить селекционный процесс и выявить взаимосвязи генома и фенотипа. Таким образом, актуальность диссертационной работы Афонникова Д.А. не вызывает сомнения.

Научная новизна исследования. Автором диссертационного исследования при помощи экспериментально-компьютерной платформы ICGPhenoPlant впервые разработан комплекс методов для фенотипирования опушения листьев пшеницы, картофеля и табака на основе анализа микроизображений сгибов листа программой LHDetect2. Впервые для образцов пшеницы и ее диких сородичей с помощью программы LHDetect2 проведен высокоточный анализ количественных характеристик опушения, выявлены их основные особенности и связь с хромосомным набором генома

пшениц. Диссертантом проведена количественная оценка влияния на число и длину трихомов генов, контролирующих опушение листа для ряда сортов, замещенных и почти изогенных линий мягких пшениц. Автором показаны изменения опушения листа пшеницы в ответ на засушливые условия произрастания и их связь с изменением физиологических параметров растений на основе точной оценки количественных характеристик программой LHDetect2.

В рамках разработки экспериментально-компьютерной платформы ICGPhenoPlant впервые было разработано:

1. Приложение SeedCounter для фенотипирования зерен злаков с помощью мобильных устройств. С его помощью впервые исследована связь количественных характеристик зерен пшеницы и сроком их хранения в генбанке и установлена значимая статистическая линейная связь между величиной признака и годом урожая;

2. Приложение WERecognizer для оценки морфометрических характеристик колосьев пшеницы, которое впервые использует модель двух четырехугольников для описания формы колоса на основе анализа его цифровых изображений.

Теоретическая значимость полученных результатов. Разработан комплекс экспериментально-компьютерных методов ICGPhenoPlant для фенотипирования важных признаков растений: опушения листа (у пшеницы, картофеля и табака), характеристик размера, формы и цвета оболочки зерен, морфометрических характеристик колоса (у пшеницы). Это позволяет использовать предложенные подходы для массового анализа при решении задач поиска взаимосвязи генотип-фенотип и окружающая среда-фенотип у растений.

Научно-практическая значимость работы. ICGPhenoPlant доступен для свободного использования в сети Интернет, что позволяет специалистам широко использовать предложенные методы для решения задач в области генетики и селекции, а в перспективе – для создания новых высокопродуктивных сортов сельскохозяйственных растений.

Структура и содержание диссертационной работы. Представленная работа оформлена по классической схеме и состоит из введения, пяти глав, включая обзор литературы и 4 главы с описанием результатов работы, заключения, выводов, списка публикаций по теме диссертаций, списка использованных сокращений, списка литературы и приложения. Список литературы содержит 438 источников. Работа представлена на 391 странице машинописного текста, содержит 87 рисунков, 29 таблиц.

Во введении раскрыта актуальность темы диссертационной работы, ее научная новизна, определены цели и задачи исследования, сформулированы положения, выносимые на защиту.

Обзор литературы состоит из четырех основных частей, посвященных описанию задач и методов феномики растений. Указано, что основными подходами являются получение и анализ двумерных цифровых изображений. Описаны основные методы и алгоритмы их применения для решения задач массового высокопроизводительного фенотипирования сельскохозяйственных растений – нового направления биоинформатики, развитого автором в рамках настоящей диссертационной работы.

Во второй главе дано описание структуры и биологической роли опушения листьев у растений и мягкой пшеницы, в частности. Были разработаны алгоритмы LHDetect и LHDetect2 анализа изображений. С помощью программы впервые проведен анализ характеристик опушения у 47 образцов пшениц 12 видов родов *Triticum* и *Aegilops*.

В работе впервые было оценено фенотипическое проявление трех генов (H11, H12^{asp}, H13), контролирующих опушение листа у пшеницы на основе высокопроизводительного фенотипирования замещенных и почти изогенных линий мягкой пшеницы (сорта 'Родина', 'Янетцкис Пробат' (ЯП), 'Саратовская 29' (С29), межсортовая замещенная линия 'Саратовская 29'/'Янетцкис Пробат' 4В (С29/ЯП4В), линия моно 4В сорта С29, интрогрессированная линия 102/00ⁱ из коллекции «Арсенал» и *Ae. speltoides* Tausch), которые несут различные аллели этих генов. Для растений этих же сортов и линий пшеницы автор провел анализ влияния воддефицита (с 21 по 55 день выращивания растений) на количественные характеристики опушения и количество устьиц на обеих сторонах листьев, физиологические характеристики.

В результате проведения кластерного анализа установлено, что длина трихомов статистически связана с показателями флуоресценции хлорофилла (NPQ, Fv/Fm, Yld, ETR) и числом устьиц на нижней стороне листа.

Исследование разнообразия опушения у 35 сортов картофеля российского происхождения позволило установить, что наибольшее значение первой компоненты (увеличение числа трихомов при одновременном уменьшении их длины) наблюдается для наиболее опушенных сортов, которые, преимущественно, имеют умеренные и малые длины трихомов.

Применение алгоритма LHDetect2 для фенотипирования опушения листьев табака первоначально продемонстрировало низкую точность определения вершин трихомов ($F = 0,37$) и их числа (MAE=4.57, MAPE=52%). После модификации алгоритма был получен новый набор оптимальных параметров на обучающей выборке изображений. При этом точность оценки идентификации трихомов на выборке тестовых изображений возросла до $F = 0,90$, а ошибки определения их числа снизились (MAE=0,78, MAPE=7%).

Третья глава посвящена высокопроизводительному фенотипированию зерен пшеницы. В рамках диссертации автором было создано приложение

SeedCounter для оценки морфологических характеристик зерен пшеницы с помощью мобильных устройств на платформе Android. Оценивалась точность определения длины и ширины 250 зерен пяти сортов пшеницы, при помощи микроскопа Carl Zeiss Aioscop 2 plus, оснащенного цифровой камерой с адаптером AxoCamHRc TV2/3c 0.63.

Для идентификации возможных генов-кандидатов, связанных с признаками семян, была проведена приоритизация на основе нескольких критериев с использованием аннотации генов, предоставленной в базе данных KEGG. Автор идентифицировал 2 787 уникальных генов, локализованных в участках, ограниченных маркерами QTL с $LOD > 3$. Из них 1422 гена были связаны с размером/формой семян и 1365 генов связаны с цветом семян. После фильтрации по уровню экспрессии осталось 823 гена, связанных с цветом семян.

Четвертая глава исследования посвящена фенотипированию колоса пшеницы. Автором было предложено два протокола для получения изображения колоса: «на прищепке» и «на столе». Обработка изображений осуществлялась с использованием пакета OpenCV. Форму колоса описывали дополнительно несколькими моделями. На основе аннотированной экспертами выборки колосьев 249 растений методами машинного обучения с использованием характеристик колосьев, были отобраны по 12 признаков для протоколов «на столе» (одна проекция) и «на прищепке» (4 проекции) (общие из них 14 признаков). Установлено, что признаками с высокой значимостью для обоих протоколов оказываются длина и индекс округлости колоса. Точность лучшего метода для 4 проекций составила 0,85 (F/случ. лес), для одной проекции 0,78 (F/лог. регрессия). Использование информации о четырех проекциях увеличивают точность предсказания формы на 7% и уменьшают среднюю абсолютную ошибку при предсказании индекса плотности колоса на 1,28, что соответствует 4,25%.

В пятой главе представлены базы данных для хранения фенотипических данных, созданные в ходе диссертационной работы:

1. WheatPGE позволяет обеспечить информационную поддержку для селекционно-генетических экспериментов у пшеницы, включая массовое фенотипирование растений, удобное для дальнейшей статистической обработки.

2. SpikeDroidDB. Ее использование способствует структурированному хранению цифровых изображений колоса и предоставляет пользователю гибкую систему для доступа к данным.

Разработанные методы в совокупности сформировали экспериментально-компьютерную платформу ICGPhenoPlant для высокопроизводительного компьютерного фенотипирования растений на основе анализа цифровых изображений.

Выводы включают 5 пунктов и полностью соответствуют задачам исследования.

Несмотря на высокий уровень представленного для рецензии диссертационного исследования в нем отмечен ряд недочетов. Изложенные ниже замечания и вопросы носят дискуссионный характер и не влияют на общую высокую оценку диссертационной работы Дмитрия Аркадьевича Афонникова:

1. Некоторые результаты, представленные в новизне диссертационной работы автора (стр. 9), как, например, «... растения сибирских сортов картофеля демонстрируют разнообразие количественных характеристик опушения листьев» требуют расшифровки. В противном случае они выглядят заранее предсказуемыми и тривиальными, поскольку любые группы сортов растений демонстрируют разнообразие количественных характеристик.

2. Четвертое и пятое положение, выносимое на защиту, имеют значительное сходство и несколько дублируют содержание друг друга. Они вполне могут быть объединены в единое положение.

3. Автор отмечает (стр. 19.), что «...искусственные условия не могут полностью воспроизвести разнообразие абиотических факторов, которые действуют на растения в естественных условиях. Поэтому одним из актуальных направлений феномики является изучение растений в полевых условиях». Однако в диссертации также (например, стр. 98) указывается, что в ряде экспериментов растения для анализа выращивались в гидропонной теплице, что несколько противоречит представленному выше тезису. При обосновании такого подхода стоило бы сделать акцент на том, что анализ растений проводился по генетически детерминированным, константным признакам, которые, очевидно, останутся в тех же рамках и в условиях культивирования растений в открытом грунте.

В качестве пожелания необходимо отметить, что несмотря на обстоятельный и качественный подход к разработке новых алгоритмов оценки степени опушенности листа, представленный в диссертационном исследовании соискателем, оригинальность и перспективность использованных подходов и результатов была бы подчеркнута значительно сильнее в случае, если бы они сравнивались с ранее разработанными, описанными в диссертации, программами, которые уже применяются в настоящее время для обработки изображений.

Заключение. Диссертационная работа Афонникова Дмитрия Аркадьевича «Компьютерные методы высокопроизводительного фенотипирования растений», представленная на соискание ученой степени доктора биологических наук по специальности 1.5.8. Математическая биология, биоинформатика, является самостоятельным законченным научным исследованием, обладающим высоким уровнем научной новизны.

Результаты исследования опубликованы в 24 работах, в том числе из них 12 в зарубежных и российских журналах с импакт-фактором, 7 в российских журналах, индексируемых в системе WoS без импакт-фактора, 2 журнала, индексируемых в системе SCOPUS и три авторских свидетельства. Материалы диссертации были доложены на конференциях различного уровня.

Рассматриваемая диссертационная работа по поставленным задачам, оригинальности подходов, актуальности и научной новизне полученных результатов полностью соответствует требованиям п. 9-14 «Положения о присуждении ученых степеней», утвержденном Постановлением Правительства РФ от 24.09.2013 г. № 842, а ее автор – Афонников Дмитрий Аркадьевич полностью заслуживает присуждения ученой степени доктора биологических наук по специальности 1.5.8. Математическая биология, биоинформатика.

Заключение принято на заседании кафедры биологии Института фармации, химии и биологии НИУ «БелГУ», протокол № 1 от 06 сентября 2023 г. федерального государственного автономного образовательного учреждения высшего образования «Белгородский государственный национальный исследовательский университет».

Присутствовало на заседании 25 человек. Результаты голосования: «за» - 25, «против» - 0, «воздержались» - 0.

Директор Научно-образовательного центра
«Ботанический сад (НИУ «БелГУ»)),
доктор биологических наук,
старший научный сотрудник

Тохтарь Валерий Константинович

ФГАО ВО «Белгородский государственный национальный
исследовательский университет» (НИУ «БелГУ»)),
308015, г. Белгород, ул. Победы 85;
Тел: (4722) 30-11-00
tokhtar@bsu.edu.ru

Личную подпись удостоверяю Ведущий специалист по кадрам департамента управления персоналом	
	« 23 » 09 20 23