

На правах рукописи

АФОННИКОВ ДМИТРИЙ АРКАДЬЕВИЧ

**Компьютерные методы
высокопроизводительного
фенотипирования растений**

1.5.8. – математическая биология, биоинформатика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

доктора биологических наук

Новосибирск – 2023

Работа выполнена в лаборатории эволюционной биоинформатики и теоретической генетики ФГБНУ “Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук” (ИЦиГ СО РАН), г. Новосибирск, Россия

Научный консультант: академик РАН,
доктор биологических наук, профессор
Колчанов Николай Александрович

Официальные оппоненты: **Карлов Геннадий Ильич**,
доктор биологических наук, академик РАН, директор
Федерального государственного бюджетного научного
учреждения «Всероссийский научно-
исследовательский институт сельскохозяйственной
биотехнологии», г. Москва

Самсонова Мария Георгиевна,
д.б.н., профессор Высшей школы прикладной
математики и вычислительной физики, заведующий
Научно-исследовательской лабораторией
математической биологии и биоинформатики ФГАОУ
ВО "Санкт-Петербургский политехнический
университет Петр Великого", г. Санкт-Петербург

Ракутько Сергей Анатольевич,
д.т.н., главный научный сотрудник Института
агроинженерных и экологических проблем
сельскохозяйственного производства - филиала
Федерального научного агроинженерного центра
ВИМ, г.Санкт-Петербург

Ведущее учреждение: Белгородский государственный национальный
исследовательский университет (НИУ БелГУ), г.
Белгород

Защита состоится “___” _____ 2023г. на утреннем заседании
диссертационного совета **24.1.239.01** по защите диссертаций на соискание ученой
степени кандидата наук, на соискание ученой степени доктора наук в ИЦиГ СО
РАН по адресу: 630090 г.Новосибирск, пр.ак. Лаврентьева, 10. Тел/факс:
(383)3634906; Факс: (383)3331278, e-mail: **dissov@bionet.nsc.ru**

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

Автореферат разослан “___” _____ 2023 г.

Ученый секретарь диссертационного совета,
доктор биологических наук

Т.М. Хлебодарова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы. Глобальные климатические и социальные процессы оказывают всё более существенное влияние на условия ведения сельского хозяйства в разных странах. С одной стороны, в связи с увеличением населения Земли стабильно возрастает потребность в продовольствии. По данным ФАО (FAO, Food and Agriculture Organization, <http://www.fao.org/>), численность и доходы населения должны вырасти вдвое к 2050 году. Это увеличивает конкуренцию за природные (водные, земельные) и сельскохозяйственные ресурсы. С другой стороны, мировое сельское хозяйство сталкивается с такими неблагоприятными климатическими изменениями как повышение сезонных температур во всех широтах, засуха, увеличение содержания углекислого газа в атмосфере (Long, Ort, 2010). Это требует пересмотра методов и подходов в селекции. Необходимо создавать сорта сельскохозяйственных культур с более высоким потенциалом урожайности и устойчивости к абиотическим стрессам (White et al., 2012). Последнее особенно важно, так как расширение посевных площадей возможно только за счет районов, где условия возделывания отличаются от оптимальных, и где урожайность имеющихся сортов может снижаться (Long, Ort, 2010).

Важной стратегией улучшения сельскохозяйственных культур является поиск генетических образцов, которые эволюционировали в суровых климатических условиях и поэтому адаптировались к ним. Они в огромном количестве представлены в генбанках всего мира, однако большинство из этих образцов требует изучения и детального описания (Cobb et al., 2013). У растений, адаптированных к суровым условиям, необходимо охарактеризовать множество различных признаков, представляющих ценность при создании новых сортов. Только тогда они будут эффективно использоваться в селекционной работе.

Анализ данных большого масштаба актуален и при поиске генов, ответственных за важные хозяйственные признаки растений. Подавляющее большинство этих признаков, такие как урожайность, биомасса растения, сроки основных фаз роста, устойчивость к стрессу и заболеваниям контролируются полигенно. Основная идея поиска генов (локусов количественных признаков, QTL), контролирующих такие признаки, заключается в выявлении ДНК-маркеров, которые ассоциированы с уровнем экспрессии исследуемого признака (Kearsey, Farquhar, 1998; Würschum, 2012; Хлесткина, 2013). В последнее время технологии секвенирования ДНК позволяют недорого и быстро прочитывать последовательности как полных геномов индивидуальных организмов, так и отдельных его локусов (Mardis, 2008; Varshney et al., 2009). Это позволяет эффективно решать задачи идентификации геномных маркеров (Elshire et al., 2011; Miller et al., 2007; Paux et al., 2012; Нестеров и соавт, 2015) и выявления их вариаций для большого количества генотипов (Lijavetzky et al., 2007; Li et al., 2013; Cuesta-Marcos et al., 2010; Xu et al., 2010). Как результат, высокопроизводительные технологии секвенирования привели к созданию новых методов селекции, таких как маркер-контролируемый отбор, (Marker-assisted selection, MAS) (Collard, Mackill, 2008) и геномная селекция (Meuwissen et al., 2001).

Следует отметить, что в основе успешного поиска взаимосвязи между генотипом и фенотипом лежат статистические методы, эффективность и точность которых зависит от размера анализируемой выборки. Большой размер выборки позволяет уменьшить ошибку идентификации QTL, а также точность определения генетического расстояния между маркерами. Это демонстрируется, аналитическими расчетами (Frisch et al., 1999), численными экспериментами (Visscher et al., 1996), практическим опытом проведения QTL (Beavis, 1998). Фактор размера выборки также важен для методов поиска полногеномных ассоциаций (genome-wide association studies, GWAS) и геномной селекции (Meuwissen et al., 2001; Korte et al., 2012). Поэтому для идентификации локусов, отвечающих за количественные признаки, все чаще используют выборки из тысяч растений.

При проведении подобных масштабных экспериментов одним из узких мест в настоящее время является измерение фенотипических параметров растения (Furbank, Tester, 2011). Попытки решить эту задачу привели к формированию новой области науки на стыке биологии, информатики и инженерии - феномики.

Цель работы. Целью работы является создание 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.

Задачи, поставленные для достижения указанной цели, включали:

1. Разработать экспериментально-компьютерную платформу ICGPhenoPlant для фенотипирования растений, включая базы данных для хранения полученной информации о характеристиках:

- (а) опушения листьев пшеницы, картофеля и табака;
- (б) размера, формы и цвета оболочки зерен пшеницы;
- (в) размера и формы колоса пшеницы.

2. С помощью платформы исследовать разнообразие опушения листа у сортов и линий мягкой пшеницы и ее сородичей, оценить влияние генов контроля опушения листьев пшеницы на формирование этого признака и его изменчивости у растений при воздействии засухи.

3. С помощью платформы оценить разнообразие опушения листа российских сортов картофеля и изменчивость опушения листа у линий табака с модификацией гена пролиндегидрогеназы.

4. С помощью платформы провести анализ размера, формы и цвета оболочки зерен пшеницы популяции ITMI и их связи с локусами количественных признаков, длительностью хранения в генбанке и всхожестью.

5. С помощью платформы оценить характеристики формы, размера, остистости колосьев пшеницы и на этой основе построить метод предсказания их типа и индекса плотности.

Научная новизна. Экспериментально-компьютерная платформа для высокопроизводительного компьютерного фенотипирования растений ICGPhenoPlant включает впервые разработанный комплекс методов для фенотипирования опушения листьев пшеницы на основе анализа микроизображений сгибов листа программой LHDetect2. Он позволяет оценивать количественные характеристики опушения листа на изображении с высокой точностью: распределение трихомов по длинам, среднее число трихомов, среднюю длину трихомов, соотношение средней длины и количества трихомов. С помощью методов фенотипирования опушения листьев пшеницы:

(а) впервые для 47 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;

(б) впервые количественно оценено влияние на число и длину трихомов генов, контролирующих опушение листа для ряда сортов, замещенных и почти изогенных линий мягких пшениц;

(в) впервые на основе точной оценки количественных характеристик опушения листа пшеницы программой LHDetect2 показаны их изменения в ответ на засушливые условия произрастания и их связь с изменением физиологических параметров растений;

С помощью экспериментально-компьютерной платформы ICGPhenoPlant было проведено фенотипирование опушения листьев двудольных растений, картофеля и табака, и показано что:

(а) растения сибирских сортов картофеля демонстрируют разнообразие количественных характеристик опушения листьев;

(б) растения генетически-модифицированных линий табака с измененным паттерном экспрессии гена пролиндегидрогеназы демонстрируют достоверно большее количество трихомов на листе и меньшую среднюю длину трихомов.

В рамках разработки экспериментально-компьютерной платформы ICGPhenoPlant впервые было разработано приложения 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 показано, что для большинства признаков цвета зерна существует значимая статистическая линейная связь между величиной признака и годом урожая.

В рамках разработки экспериментально-компьютерной платформы ICGPhenoPlant разработано приложение WERecognizer для оценки морфометрических характеристик колосьев пшеницы, которое впервые использует модель двух четырехугольников для описания формы колоса на основе анализа его цифровых изображений. Показано, что геометрические параметры колоса позволяют предсказывать его тип и плотность с высокой точностью.

Положения, выносимые на защиту:

1. Экспериментально-компьютерная платформа ICGPhenoPlant разработанная для фенотипирования растений, включает программные продукты LHDetect, WERecognizer и SeedCounter, которые позволяют с высокой точностью на основе двумерных изображений оценивать количественные параметры опушения листьев, формы,

размеров колоса и цвета оболочки зерен растений, а также информационные ресурсы WheatPGE и SpikeDroid, обеспечивающие возможность накопления и сохранения информации о фенотипе, генотипе и месте произрастания растений в соответствующих базах данных.

2. Высокопроизводительное компьютерное фенотипирование опушения листьев пшеницы с помощью программы LHDetect2 позволяет установить достоверную статистическую связь между количественными признаками опушения листа, хромосомным составом генома, аллельным составом генов, контролирующих развитие опушения, и изменением физиологических параметров растений в процессе ответа на воддефицит.

3. Программа LHDetect2 позволяет достоверно оценить изменчивость количественных характеристик опушения листьев для генетически модифицированных линий табака с измененным паттерном экспрессии гена пролиндегидрогеназы и увеличенным содержанием пролина по сравнению с контрольным генотипом SR1.

4. Компьютерное фенотипирование зерен мягкой пшеницы урожаяв различных годов с помощью приложения SeedCounter демонстрирует достоверные изменения характера пигментации их оболочки в зависимости от срока хранения.

5. С помощью приложения SeedCounter установлено, что у мягкой пшеницы существует достоверная корреляция между характером пигментации оболочки зерновки, сроком хранения зерна и их всхожестью.

Теоретическая значимость работы. 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). Данный комплекс методов позволяет получать быстро и с высокой точностью оценивать количественные характеристики фенотипа растений, что позволяет использовать предложенные подходы для массового анализа при решении задач поиска взаимосвязи генотип–фенотип и окружающая среда–фенотип у растений.

Научно-практическая значимость работы. Разработанный комплекс экспериментально-компьютерных методов и баз данных ICGPhenoPlant доступен для исследователей в области генетики и селекции сельскохозяйственных растений в Интернет для свободного использования. Это обеспечивает широкое использование предложенных методов для решения задач в области генетики и селекции, поиска взаимосвязи между генотипом, фенотипом и окружающей средой, и, в перспективе, создания новых высокопродуктивных сортов сельскохозяйственных растений.

Апробация работы. Результаты диссертационной работы были представлены на конференциях БГРС (2012, 2014, 2016, 2018, 2020, 2022 гг.); PlantGen (2015, 2017, 2019, 2021 гг.); VII Московском международном конгрессе «Биотехнология: состояние и перспективы развития» (г. Москва, 2013); 1-м Международном семинаре «Генетика и геномика растений для продовольственной безопасности» (НГУ, 2016); конференции PhenoDays-2014 (Bon, France, 2014); XIII Новосибирском инновационно-инвестиционном форуме «Сельхозмашиностроение и переработка агропродукции»

(2018 г.); Школе-конференции по количественной генетике (ВИР, 2018 г.); Конференции «125 лет прикладной ботаники в России» (ВИР, 2019 г.); на VII съезде Вавиловского общества генетиков и селекционеров, посвященному 100-летию кафедры генетики СПбГУ (ВИР, 2019); на II Европейской школе биотехнологов (Минск, 2019); на Международной научно-практической конференции «Развитие сельского хозяйства на основе современных научных достижений и интеллектуальных цифровых технологий «Сибирь – Агробиотехнологии» («САБИТ 2019») (СибНИИРС, 2019); форуме Технопром-2019 (г. Новосибирск, 2019); на Третьем Международном форуме «Передовые цифровые и производственные технологии» (СПбПУ, 2021); конференции "В генетику - через разные отрасли науки и практики" (Центр Сириус, 2022).

Публикации. По теме диссертационной работы опубликовано 24 работы, 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. Все работы опубликованы в соавторстве.

Личный вклад автора. Данная работа была бы невозможна без участия большого количества исследователей. Автор внес решающий вклад в постановку цели и задач исследования, обоснование выбора методов для решения задач, обработку данных, интерпретацию и обобщение полученных результатов. В большинстве статей, опубликованных по тематике диссертации автор был автором для переписки.

Основные результаты главы 2 были получены в соавторстве с Пшеничниковой Т.А., Генаевым М.А., Дорошковым А.В., Ибрагимовой С.С., Осиповой С.В., Пермяковой М.Д., Пермяковым А.В., Колчановым Н.А. Основные результаты главы 3 получены в соавторстве с Комышевым Е.Г., Генаевым М.А., Ковалем В.С., Ефимовым В.М., Бёрнером А. Основные результаты главы 4 получены в соавторстве с Генаевым М.А., Комышевым Е.Г., Гончаровым Н.П., Кручининой Ю.В., Ковалем В.С. Основные результаты главы 5 получены в соавторстве с Комышевым Е.Г., Генаевым М.А., Ковалем В.С., Гончаровым Н.П.

Структура и объем работы. Работа состоит из Введения, 5 глав, включая Обзор литературы и 4 главы с описанием результатов работы, Заключение, Выводов, Списка публикаций по теме диссертации, Списка использованных сокращений, Списка литературы и Приложения. Объем работы составляет 391 лист. В работе 87 рисунков и 29 таблиц. Список литературы включает 438 источников.

Благодарности. Автор выражает глубокую признательность научному консультанту академику РАН Колчанову Н.А. Автор выражает благодарность сотрудникам ИЦиГ СО РАН Пшеничниковой Т.А., Генаеву М.А., Дорошкову А.В., Ибрагимовой С.С., Комышеву Е.Г., Ковалю В.С., Ефимову В.М., Шмакову Н.А., Кочетову А.В., А. Бёрнеру, Ариффу М.А.Р. за участие в работе и помощь на разных ее этапах.

ГЛАВА 1 ОБЗОР ЛИТЕРАТУРЫ

Глава 1 посвящена обзору задач и методов феномики растений. Указано, что

основными подходами является получение и анализ двумерных цифровых изображений. Описаны методы представления цифровых изображений, в том числе различные цветовые пространства (RGB, HSV, Lab, YCrCb). Описаны основные методы и алгоритмы, которые используются при решении этих задач. Даны определения основных метрик для оценки точности алгоритмов: мера Жаккара, средняя абсолютная ошибка (MAE), средняя абсолютная процентная ошибка (MAPE), коэффициент корреляции Пирсона, ошибки классификации (числа ложных положительных и отрицательных, истинных положительных и отрицательных предсказаний), матрица ошибок, точность полнота, F -мера. Описано применение этих методов к решению задач массового высокопроизводительного фенотипирования сельскохозяйственных растений. Описаны задачи, связанные с интеграцией данных по фенотипам и генотипам растений. Обоснован выбор цели работы и решаемых задач в области высокопроизводительного фенотипирования растений - нового направления биоинформатики, развитого автором в рамках настоящей диссертационной работы.

ГЛАВА 2. ВЫСОКОПРОИЗВОДИТЕЛЬНОЕ ФЕНОТИПИРОВАНИЕ ОПУШЕНИЯ ЛИСТА У РАСТЕНИЙ

Высокопроизводительное фенотипирование опушения листьев пшеницы.

В начале главы дается описание структуры и биологической роли опушения листьев у растений и мягкой пшеницы, в частности. Описан протокол для получения изображений сгибов листа пшеницы для дальнейшего анализа (рис. 1А-В).

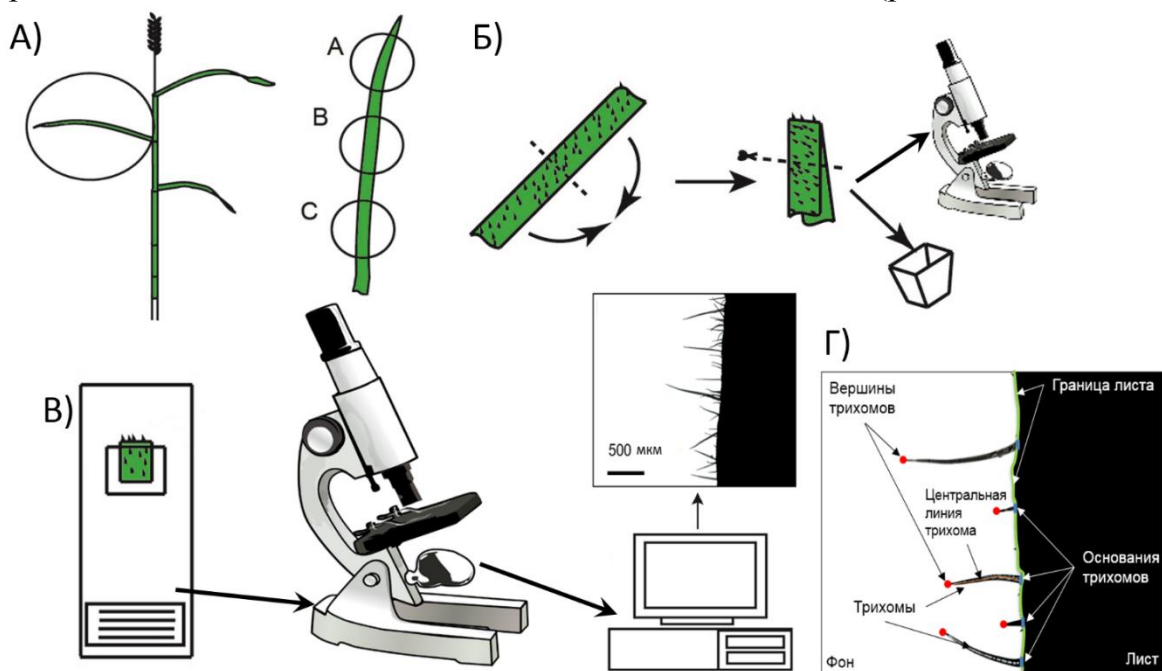


Рисунок 1. Протокол пробоподготовки образцов сгибов листа для получения микрофотографий. А) Выбор листа для фенотипирования и области поверхности листа для анализа; Б) сгиб листа и удаление остатков на концах образца; В) размещение согнутого фрагмента на предметном стекле, его фиксация липкой лентой, размещение на предметном стекле микроскопа и получение микрофотографий, для которых приведен типичный пример; Г) выделение ключевых элементов на микрофотографии сгиба листа: фон, лист, граница листа, трихомы, их основания, вершины и центральные линии.

Были разработаны алгоритмы LHDetect и LHDetect2 анализа изображений, которые состоят из нескольких шагов. LHDetect (Дорошков и соавт., 2009) включает идентификацию фона, листа и границы и оценивает число трихомов на определенном расстоянии от границы листа, преобразуя эти данные в распределение трихомов на изображении по длине. LHDetect2 включает: преобразование изображение в полутоновое; подавление шума на основе Гауссова фильтра; сегментацию по порогу интенсивности на фон и лист с трихомами; определение границы листа и фона; выделение контуров, соответствующих трихомам; идентификацию вершин трихомов и их оснований; выделение центральной линии трихомов и оценка длины для каждого трихома (рис. 1Г), включая сложные контуры, которые содержат несколько пересекающихся трихомов (Genaev et al., 2012). Параметры алгоритма LHDetect2 были подобраны на обучающей выборке 76 изображений. Точность идентификации количества трихомов на изображении и оценки их длин были верифицированы на тестовой выборке из 70 изображений (рис. 2, табл. 1).

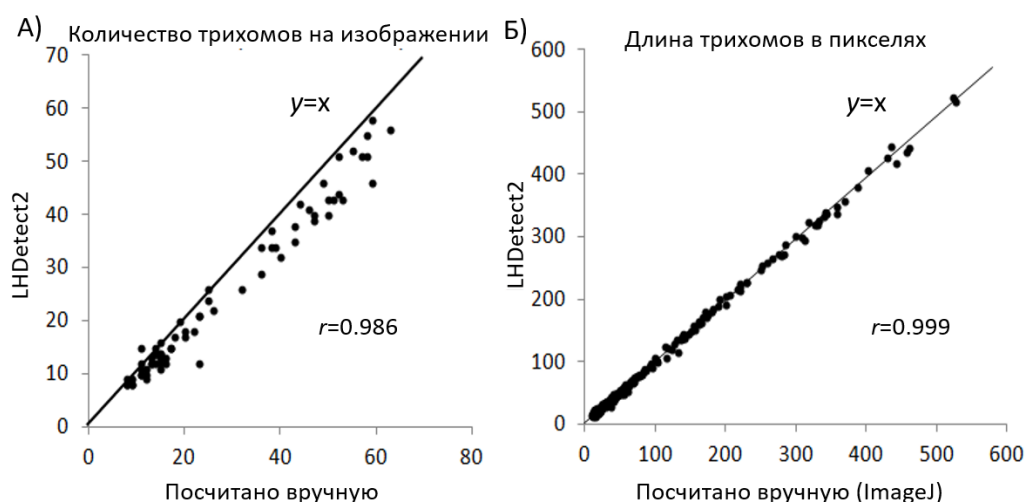


Рисунок 2. Сравнение оценок числа трихомов на изображении сгиба листа и их длины, полученных программой LHDetect2 и вручную. (А) Число трихомов, подсчитанных программой LHDetect2 (ось Y) и определенных вручную (ось X). Показана линия $y=x$ и коэффициент корреляции между двумя оценками. (В) Длины трихомов, оцененные программой LHDetect2 (ось Y) и соответствующее значение, полученное вручную с использованием пакета ImageJ (ось X). Также показана линия $y=x$ и приведен коэффициент корреляции между двумя оценками.

Таблица 1. Сравнение точности оценки числа трихомов на изображении сгиба листа, полученных методом LHDetect и LHDetect2. Приведены значения MAE и MAPE оценок числа трихомов, посчитанными автоматически в сравнении с ручной оценкой. Отдельно приведены оценки этих параметров для всей выборки из 70 тестовых изображений, включающей 44 простых (содержат мало пересекающихся трихомов) и 26 сложных изображений (много пересекающихся трихомов).

Выборка изображений	LHDetect2		LHDetect	
	MAE, шт	MAPE, %	MAE, шт	MAPE, %
Полный набор	3.33	11.65	4.54	14.22
Простые изображения	1.82	11.23	1.64	11.60
Сложные изображения	5.88	12.38	9.46	18.64

Алгоритм был реализован в программе LHDetect2, доступной по адресу <https://github.com/genaev/LHDetect2>; Интернет-версия доступна по адресу <http://wheatdb.org/lhdetect2>.

С помощью программы впервые проведен анализ характеристик опушения у 47 образцов пшениц от 12 видов родов *Triticum* и *Aegilops* (Pshenichnikova et al., 2017). Оценены распределения трихомов по длинам, средняя длина (l_m , мкм) число (n_f , шт.) трихомов на изображении, их отношение $H_{L/N}$. Проведена иерархическая кластеризация образцов на основе гистограмм распределений трихомов по длинам.

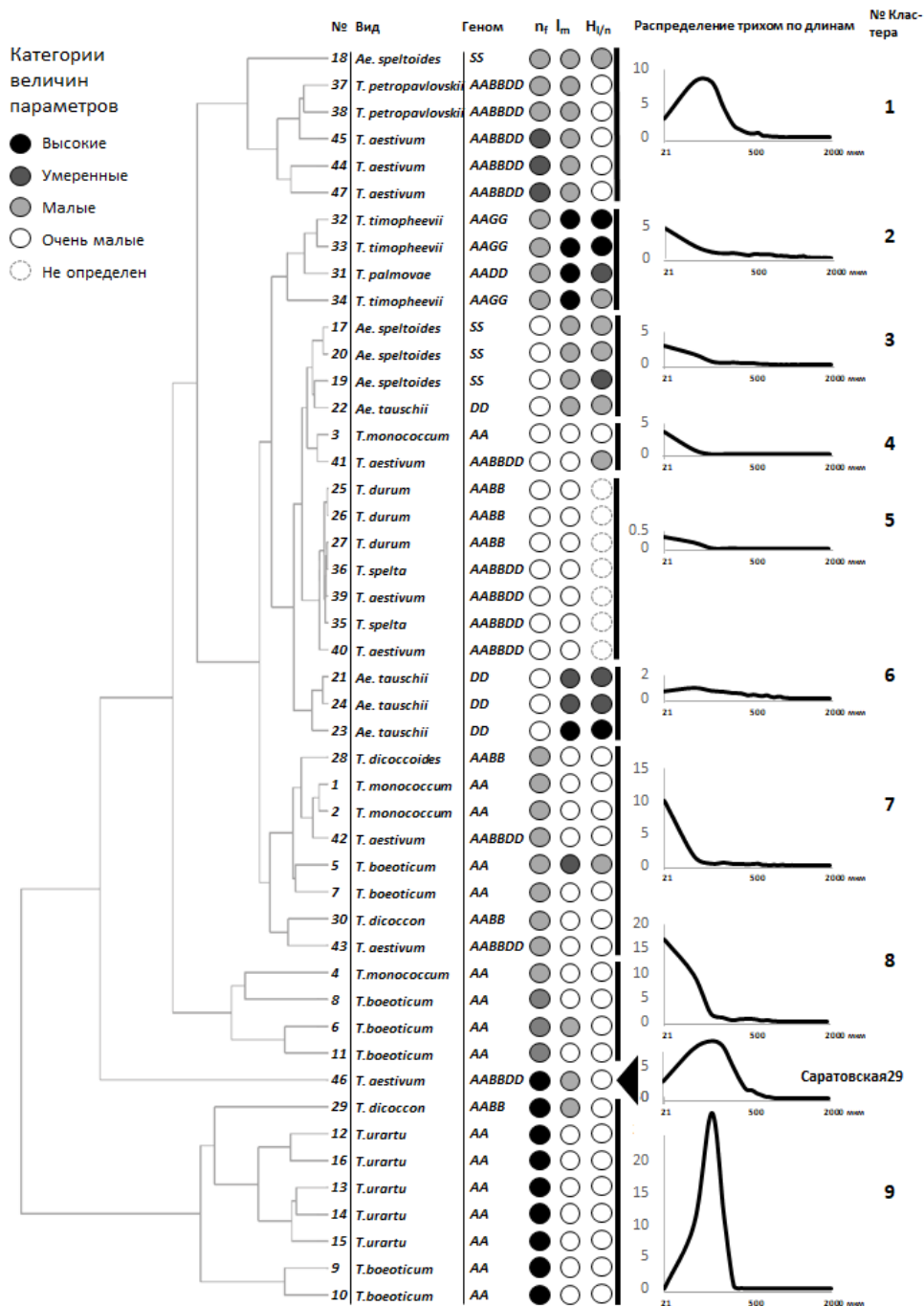


Рисунок 3. Результаты кластеризации образцов по распределению длин трихомов. Номера образцов даны в соответствии с таблицей П2 приложения к тексту диссертации, названия видов, геномная структура, качественные характеристики опушения и распределение трихомов по длинам представлено для каждого кластера. Справа указаны номера кластеров (С1-С9).

Показано (рис. 3), что виды *T. boeoticum* и *T. urartu* характеризуются плотным опушением, *T. monococcum* характеризуется низкой плотностью трихомов. Образцы

рода *Aegilops*, носители геномов SS и DD, характеризуются умеренно длинным и длинным, но редким опушением. Аллотетраплоидные геномы характеризуются различными морфологическими признаками опушения. Опушение гексаплоидов с геномной конституцией AABBDD характеризуется умеренной длиной трихомов. При этом, в отличие от ди- и тетраплоидных видов, плотность опушения показывает высокую изменчивость от очень слабого опушения у *T. spelta* до интенсивного опушения у сортов мягкой пшеницы ‘Саратовская 29’ и ‘Хонг-Манг-май’.

На основе однофакторного дисперсионного анализа впервые установлена статистическая связь между количественными характеристиками опушения листьев и структурными особенностями генома (табл. 2).

Независимая переменная	Зависимая переменная		
	n_f	l_m	$H_{L/N}$
Кариотип	<0.005	<<0.005	<0.05
Уровень ploидности	0.13	<0.05*	<0.01
Наличие генома А	<0.05	0.30	0.96
Наличие генома В (G,S)	<0.01	0.58	0.06
Наличие генома D	<0.05	<0.005	<0.005

Таблица 2. Статистическая зависимость между структурой генома и количественных характеристик опушения (n_f , l_m , $H_{L/N}$). Приведены уровни значимости. Значимые величины выделены жирным шрифтом.

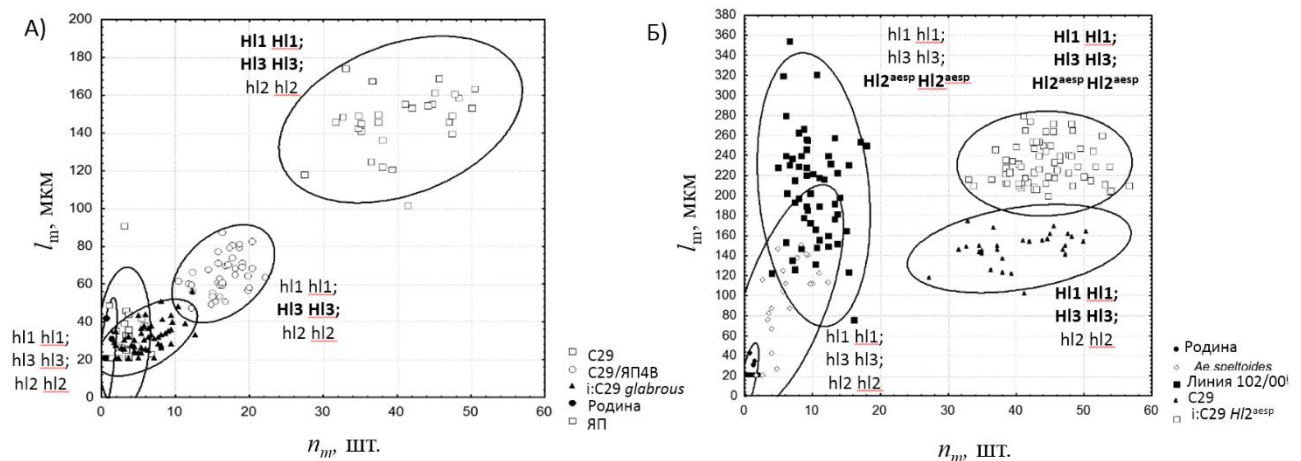


Рисунок 4. Фенотипическое проявление генов, контролирующих опушение листа пшеницы $H11$, $H12^{aesp}$, $H13$. (А) Количественная оценка опушения листьев у сортов ‘Родина’, ‘Янетцкис Пробат’ (ЯП), ‘Саратовская 29’ (С29), неопушенной линии i:C29 *glabrous* и линии, полученной на основе сорта ‘Саратовская 29’ с межвидовой заменой одной хромосомы С29/ЯП4В. Диаграмма рассеяния среднего числа (ось X) и средней длины трихомов (ось Y) на изображениях. Эллипсы определяют области доверительного интервала 95%. (Б) Количественная оценка опушения листьев в линии 102/00i линия, линии i:C29 $H12^{aesp}$ и их родительские формы сорт ‘Родина’, ‘Саратовская 29’ (С29) и *Ae. speltoides*, k-389. Обозначения те же, что и на панели А. Рядом с областями диаграммы, соответствующими генотипам приведены аллельные состояния генов.

В работе впервые было оценено фенотипическое проявление трех генов ($H11$, $H12^{aesp}$, $H13$), 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/ЯП

4В), линия моно 4В сорта С29, интрогрессированная линия 102/00ⁱ из коллекции «Арсенал» (Lapochkina, 2001) и *Ae. speltoides* Tausch) которые несут различные аллели этих генов. Результаты приведены на рис. 4 и показывают, что ген *HL2^{aesp}* в большей степени оказывает влияние на длину трихомов. Возможно его функция связана с изменением размера клеток. Гены *HL3* и *HL1* влияют одновременно и на длину трихомов и на их количество. Вероятно, они затрагивают как пути инициации трихомов на листе, так и роста клеток отдельных трихомов.

Для растений этих же сортов и линий пшеницы мы провели анализ влияния водodefицита (с 21 по 55 день выращивания растений) на количественные характеристики опушения и количество устьиц на обеих сторонах листьев, физиологические характеристики (параметры газообмена, фотосинтеза, эффективность использования воды), биомассу побега, длину и ширину листа, всего 18 признаков (Pshenichnikova et al., 2019). Результаты оценки опушения листьев (табл. 3; TN, TL - число трихомов и их средняя длина; up, bot – верхняя и нижняя стороны листа, соответственно; условия К, З - контроль, засуха), что при воздействии засухи число трихомов растет (кроме сорта ‘Родина’), а средняя длина уменьшается (за исключением сортов ‘Родина’ и ЯП).

Генотипы	Трихомы							
	TNup		TNbot		TLup		TLbot	
	К	З	К	З	К	З	К	З
С29	39.7	55.4	48.2	66.4	140.9	138.8	133.8	115.8
i:С29 <i>hl1 hl3</i> [С29]	5.3	9.4	3.3	18.1	45.0	37.5	67.8	42.7
i:С29 <i>HL2</i> [С29]	41.9	61.4	56.4	72.2	213.9	191.9	188.2	165.3
С29(ЯП 4В) [С29]	19.5	33.4	21.9	35.4	70.7	62.8	64.7	61.7
‘Родина’	0.56	0.4	0.2	0.2	18.9	11.1	11.2	17.3
102/00ⁱ [‘Родина’]	14.6	23.8	10.5	14.4	281.4	141.4	223.1	125.0
ЯП [С29]	4.0	15.5	2.7	10.3	31.4	38.8	42.9	41.8

Таблица 3.

Характеристики опушения при контрастных водных режимах. Сравнения сделаны с генотипом, указанным в квадратных скобках значимые различия выделены жирным.

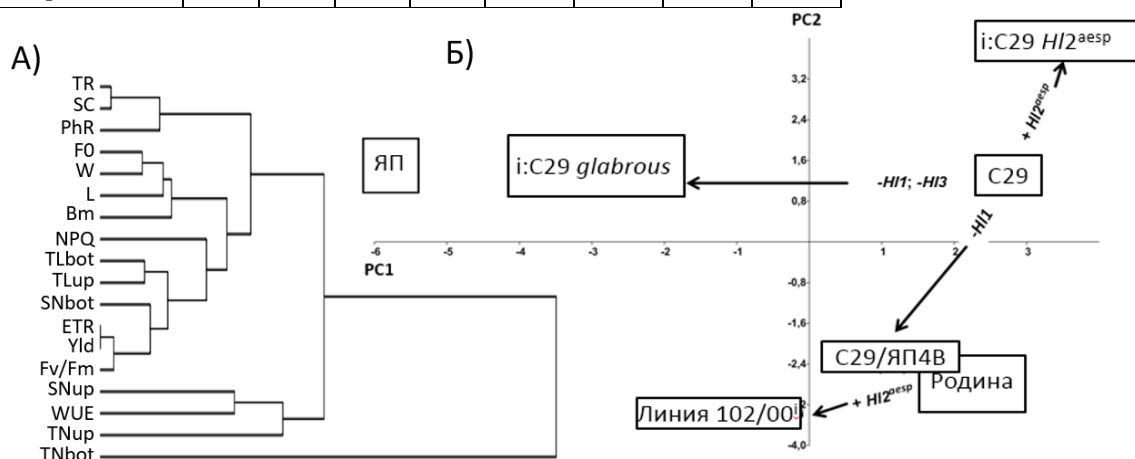


Рисунок 5. А) Кластерный анализ параметров газообмена, флуоресценции хлорофилла и морфологических признаков на основе нормализованных различий значений между нормой и засухой. Фигурные скобки обозначают кластеры, которые группируют сходную изменчивость признаков между генотипами. Б) Анализ главных компонент для генотипов пшеницы в пространстве нормализованных различий значений физиологических и морфологических признаков между нормой и засухой. Стрелками показаны изменения аллельного состава генов, контролирующего опушение.

Для изучения взаимодействия признаков и их изменчивости при изменении условий выращивания был проведён кластерный анализ, в котором в качестве данных была взята нормированная разница значений каждого признака в условиях контроля и засухи. Кластеризация производилась на основании изменения каждого признака в ответ на стресс. Дополнительно мы провели анализ главных компонент в пространстве изменения признаков при воздействии засухи. Результаты (рис. 5) показали, что длина трихомов статистически связана с показателями флуоресценции хлорофилла (NPQ, Fv/Fm, Yld, ETR) и числом устьиц на нижней стороне листа. Число трихомов на нижней и верхней стороне (TNup) и (TNbot) и число устьиц на верхней стороне листа связаны с эффективностью использования воды (WUE). Рассчитанные компоненты 1 и 2 описывали 33% и 24% вариации, соответственно. Сильно опушенный сорт С29 и практически неопушенный сорт ЯП занимали противоположные положения в пространстве первой компоненты. Замещение доминантных аллелей генов *H11* и *H13* опущения листа на рецессивные в генотипе сорта С29 вызывало смещение изучаемых характеристик в сторону неопушенного генотипа. Изменение гена *H12^{aesp}* на доминантный на фоне генотипа С29 изменило характеристики в направлении компоненты PC2. Результаты позволяют предположить, что генетически обусловленное присутствие густого опушения делает растения пшеницы преадаптированными к засухе, так как наличие пограничного слоя, обеспечиваемого наличием трихом, уменьшает транспирацию, устьичную проводимость и нетто-фотосинтез.

Высокопроизводительное фенотипирование опущения листьев картофеля.

Для этого была проведена адаптация протокола пробоподготовки фрагментов листа. Оказалось, что для исследованных сортов картофеля трихомы не имели головок, что позволяло использовать при фенотипировании алгоритм LHDetect2, предложенный для анализа трихомов пшеницы.

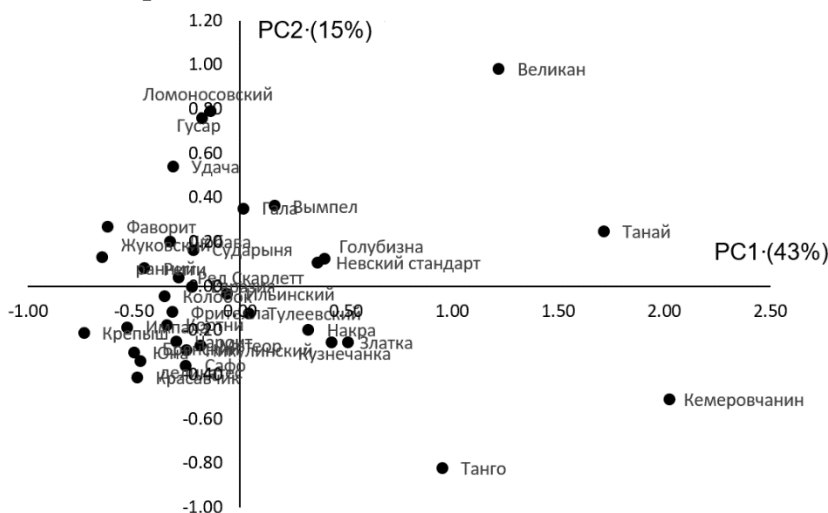


Рисунок 6. Диаграмма рассеяния 35 сортов картофеля в пространстве двух первых компонент в пространстве бинов распределения трихомов по длинам. В скобках рядом с наименованиями компонент указана доля общей дисперсии, которая приходится на компоненту.

С его помощью мы провели анализ разнообразия опущения у 35 сортов картофеля Российского происхождения. Результаты представлены на рис. 6. Анализ показал, что первая компонента характеризует увеличение числа трихомов при одновременном уменьшении их длины. Наибольшее значение первой компоненты наблюдается для

наиболее опушенных сортов ('Великан', 'Танай', 'Кемеровчанин'), которые, преимущественно, имеют умеренные и малые длины трихомов. Вторая компонента отражает преобладание трихомов умеренной длины. Предложенный метод оценки количественных характеристик опушения листьев картофеля показал высокую производительность и эффективность. С его помощью удалось охарактеризовать количественные характеристики опушения листьев и выделить сорта, имеющие наиболее интенсивное опушение ('Великан', 'Танай', 'Кемеровчанин').

Высокопроизводительное фенотипирование опушения листьев табака. В отличие от простых, несекретирующих трихомов пшеницы, трихомы листа табака характеризуются большими размерами, высокой плотностью расположения и высокой метаболической активностью. У табака два вида железистых трихомов: короткие с одноклеточной головкой и длинные трихомы с многоклеточной или одноклеточной головкой (Akers et al. 1978). Поэтому в настоящей работе нам необходимо было модифицировать алгоритм LHDetect2, предложенный ранее для автоматического анализа опушения листьев пшеницы для учета морфологии трихомов табака.

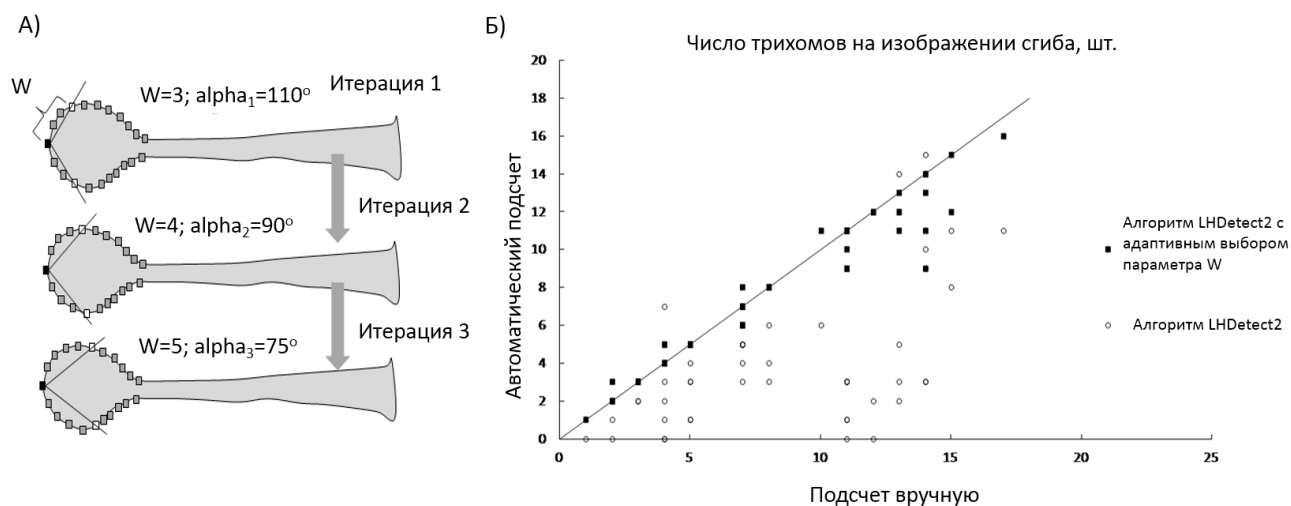


Рисунок 7. А). Модификация алгоритма поиска вершины для анализа трихомов табака. Серым контуром на изображении показан трихом на трех итерациях алгоритма. Серыми квадратами показаны пиксели границы контура трихома в области головки. Проиллюстрировано изменение угла α_i при различных значениях параметра W для пикселя контура секретирующей головки одного трихома, показанного черным цветом (кандидат на вершину). В данном примере пороговое значение $\alpha=70^\circ$. Белым цветом показаны пиксели границы, удаленные от текущего на W пикселей. Линиями показан угол, образованный тремя пикселями. Критерий $\alpha_i < \alpha$ будет выполняться на четвертой итерации алгоритма. Таким образом, пиксель, показанный черным, является вершиной трихома. Б). Сравнение точности подсчета трихомов вручную с автоматической (LHDetect2 и модифицированным, с адаптивным выбором W).

Применение алгоритма LHDetect2 продемонстрировало низкую точность определения вершин трихомов (рис. 7Б; $F=0.37$) и их числа (MAE=4.57, MAPE=52%). Причина этому – секретирующие головки сферической и овальной формы, не имеющие резкого изгиба линии границы в области вершины. Чтобы правильно выбрать вершину (пиксель головки, наиболее удаленный от основания

головки трихомы), мы провели модификацию алгоритма (рис. 7А). Алгоритм зависит от параметра W (в пикселях) – количество соседних пикселей вдоль границы контура трихома на которые удалены пиксели для определения угла $alpha$ (град.; максимальное значение угла изгиба контура для пикселей, которые считаются кандидатами на вершину трихома). В результате модификации, значения параметра W для каждого трихома в процессе поиска вершины менялись от W_{min} до W_{max} , где W_{min} и W_{max} – целые числа, параметры алгоритма. После модификации алгоритма был получен новый набор оптимальных параметров на обучающей выборке изображений, учитывающих овальную форму головки и позволяющих идентифицировать в качестве вершины пиксель границы, максимально удаленный от основания головки. 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%, рис. 7Б).

Модифицированный алгоритм позволил исследовать влияние повышенного содержания пролина в генетически-модифицированных линиях табака (Pdh5, Pdh6, Pdh8) по сравнению с исходной линией SR1 (Ibragimova et al., 2022, рис. 8).

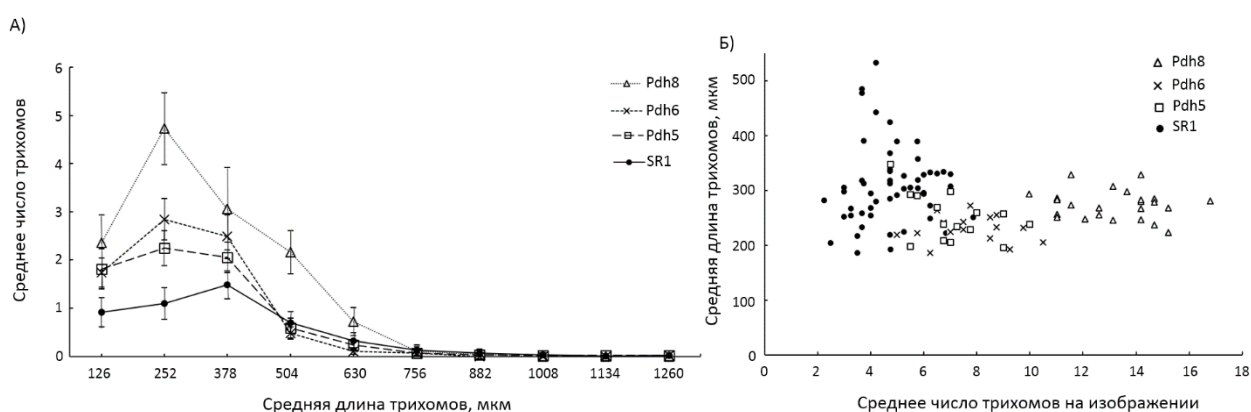


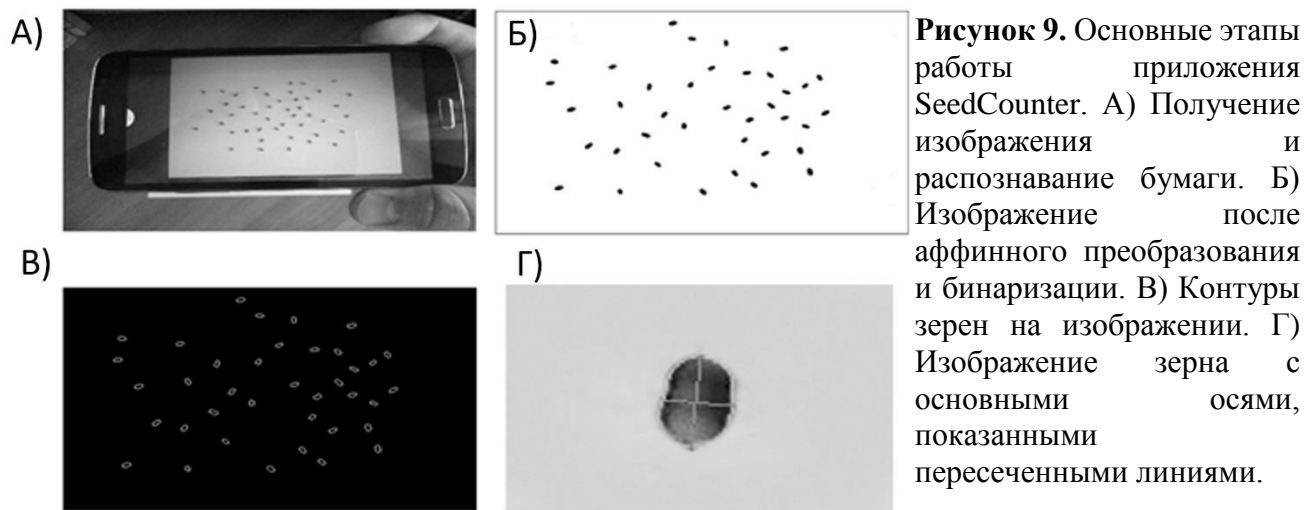
Рисунок 8. Сравнение количественных характеристик опушения листа контрольной (SR1) и трех ГМ линий табака. А) распределение трихомов по длинам. Б) распределение образцов на графике зависимости среднего числа и средней длины трихомов.

Полученные результаты демонстрируют, что ГМ линии табака с измененным паттерном экспрессии гена пролиндегидрогеназы и повышенным содержанием пролина по сравнению с контрольным генотипом имеют достоверно большее количество трихом на адаксиальной стороне листа, и при этом длина трихом достоверно меньше, чем у растений контрольной линии SR1.

ГЛАВА 3. ВЫСОКОПРОИЗВОДИТЕЛЬНОЕ ФЕНОТИПИРОВАНИЕ ЗЕРЕН ПШЕНИЦЫ

Приложение SeedCounter для морфометрии зерен пшеницы. Современные мобильные устройства (смартфоны и интернет-планшеты) содержат цифровые камеры с высоким разрешением, оснащены многоядерными процессорами с достаточной

вычислительной мощностью для обработки и анализа изображений. Эти функции позволяют пользователям получать и обрабатывать изображения там, где это необходимо. В рамках диссертации нами было создано приложение SeedCounter для оценки морфологических характеристик зерен пшеницы с помощью мобильных устройств на платформе Android (Komyshev et al., 2017). Для анализа изображений с помощью мобильных устройств необходимо получить изображение зерен, произвольно размещенных на листе белой бумаги стандартного формата (рис. 9).



Алгоритм обработки изображений для подсчета и морфометрии зерен реализован с использованием библиотеки обработки изображений OpenCV (Dawson-Howe, 2014) и состоит из нескольких шагов. Для распознавания исходное цветное изображение (рис. 9А) преобразуется в оттенки серого. Лист бумаги распознается как светлая область четырехугольной формы на темном фоне. Если форма листа бумаги на изображении отличается от прямоугольной, аффинные преобразования преобразуют ее в прямоугольную (рис. 9Б). Зерна идентифицируются после бинаризации изображения листа алгоритмом водораздела в пространстве HSV (рис. 9В). Полученные контуры аппроксимируются эллипсоидами зерен, что позволяет оценить размер главной и второстепенной осей, соответствующих длине и ширине зерна (рис. 9Г). Дополнительно программа определяет площадь области зерна на изображении. Пользователь мобильного приложения может настроить параметры обработки изображений и распознавания семян, используя опцию "Calibration" в главном меню. Данные о количестве подсчитанных семян и параметрах формы семян для каждого семени хранятся в формате XML и могут быть отображены с помощью меню "Данные о семенах".

Мы оценили точность идентификации количества зерен на тестовой выборке из 40 изображений. Использовали различные мобильные устройства, разрешения камер и условия освещения. Для каждой серии оценок количества зерен мы рассчитали ошибки MAE и MAPE. MAE оказалось близкой к 1, а MAPE к 2%, а коэффициент корреляции Пирсона выше 0.99. Более детальный анализ показал, что наибольшие ошибки при подсчете количества зерен возникают, если два или более зерен на бумаге

соприкасаются и что в условиях плохого освещения алгоритм не разделяет часть соприкасающихся зерен.

Мы оценили точность определения длины и ширины зерна. Мы измерили длину и ширину 250 зерен пяти сортов пшеницы, причем каждое зерно располагалось в строгом порядке, используя микроскоп Carl Zeiss Aioscop 2 plus, оснащенный цифровой камерой с адаптером AxoCamHRc TV2/3c 0.63. Мы разместили зерна на листе бумаги в том же порядке и применили программное обеспечение SeedCounter для оценки их длины и ширины. Была проведена серия морфометрических измерений с использованием различных мобильных устройств, разрешений камер и условий освещения. Точность оценки размера зерна (суммарно длины и ширины) составляла приблизительно 0.30 мм (среднее значение для всех серий: 0.31 мм), что составляет примерно 8% от линейных размеров зерна (среднее значение для всех серий: 8.03%). Коэффициенты корреляции между контрольной длиной и ее оценкой во всех экспериментах были не ниже 0.9. Для ширины зерна этот параметр был выше 0.84.

Результаты позволили оценить влияние на точность оценок как типа устройства, так и освещения. Мобильные устройства демонстрируют наилучшие показатели при оценке размера зерна в условиях равномерного освещения двумя лампами дневного света, световой поток 2500 лк. Наименьшая точность была получена в условиях освещения от окна в помещении в пасмурный день. Двухфакторный дисперсионный анализ показал, что условия освещения существенно влияют на оценку количества зерен, а также длины и ширины зерен (табл. 4).

Тип ошибки	Условия освещения	Тип устройства
Подсчет зерен, MAE	0.004	0.365
Подсчет зерен, MAPE	0.003	0.306
Оценка размера, MAE	0.036	0.771
Оценка размера, MAPE	0.094	0.890

Таблица 4. Оценка влияния типа мобильного устройства и условий освещения на точность морфометрии зерен. Приведены *p*-значения для двухфакторного дисперсионного анализа. Жирный шрифт для значимых величин ($p < 0.05$).

Поиск локусов количественных признаков для характеристик размера, формы и цвета зерен мягкой пшеницы. Мы создали версию приложения SeedCounter для настольных ПК для модифицированного протокола съемки зерен с помощью цифровых камер (Afonnikov et al., 2022). В этом протоколе съемка производилась на листе бумаги формата A4. Зерна рассыпались по 20 штук. В область каждого кадра помещалась калибровочная палитра ColorChecker (<https://xritephoto.com/camera>), которая использовалась при обработке изображений для определения масштаба. С ее помощью также проводилась цветокоррекция изображений (Genaev et al., 2019).

Описание размера и формы зерен было расширено и включало длину (*sL*), ширину (*sW*) и проецируемую на изображении площадь (*sA*). Дополнительно оценивали округлость (*sCi*), закругленность (*sRo*), шероховатость (*sRg*), цельность (*sSo*) (Cervantes et al., 2016; Zdilla et al., 2016).

Для описания цвета зерен использовались 48 дескрипторов. Первый тип дескрипторов: 12 средних значения интенсивности компонент для пикселей зерна для

четырёх цветовых пространств (RGB, HSV, Lab, YCrCb). Следующий набор характеристик - доминантные цвета зерен. Для вычисления этих дескрипторов, все пиксели, относящиеся к изображению зерна в цветовом пространстве RGB кластеризуются по интенсивности компонент методом k -средних ($k=3$). Полученные кластеры ранжируются по доле пикселей, которые к ним относятся. Для каждого кластера определяется центроид, представляющий доминантный цвет. Для каждого такого центроида определялись значения интенсивностей цветовых компонент в 4 цветовых пространствах (всего 12).

Поиск QTL проводился в популяции Международной Инициативы по картированию пшеницы (ITMI), которая была получена путем опыления растений *T. aestivum* (сорт 'Opata 85') пыльцой синтетической гексаплоидной яровой пшеницы 'Synthetic W7984' (Arif et al., 2021). Растения каждого генотипа были выращены в сезоне 2014 года на экспериментальных полях ИРК в Гатерслебене, Германия. Материал был предоставлен проф. Андреасом Бёрнером, Институт генетики растений и исследований растениеводства Лейбница (ИРК), Гатерслебен, Германия. Выделение ДНК, генотипирование и выявление QTL было проведено д-ром Миан Абдур Рехман Арифом, Институт ядерных исследований для сельского хозяйства и биологии, Файзалабад, Пакистан, детали анализа приведены в работе (Arif et al., 2022).

Чтобы идентифицировать возможные гены-кандидаты, связанные с признаками семян, мы провели приоритизацию на основе нескольких критериев с использованием аннотации генов, предоставленной в базе данных KEGG (Kanehisa et al., 2017). Анализ включал локусы, для которых значение LOD превышало 3, и состоял из нескольких этапов. На первом этапе анализа мы определили физическую локализацию маркеров, выровняв их последовательности с последовательностью сборки генома пшеницы IWGS 2.1, в результате для каждого из выбранных нами QTL мы получили список генов пшеницы согласно аннотации. Поскольку растительные пигменты могут синтезироваться в различных тканях и органах, при определении приоритетных генов для QTL цвета семян мы дополнительно фильтровали гены по уровню экспрессии ($TPM > 1$) в семенах с использованием БД expVIP (Borrill et al., 2016).

На втором этапе анализа мы сгенерировали список ортологичных групп белков из базы данных KEGG. Так как цвет семенной оболочки определяется наличием в ней специфических растительных пигментов (Lachman et al., 2017), мы выбрали ортогруппы, участвующие в 7 путях KEGG биосинтеза этих пигментов и ряда их предшественников. Размер семян зависит от множества биологических процессов, происходящих на молекулярном уровне и выделить специфические пути, соответствующие этим процессам в KEGG было невозможно. Поэтому мы использовали три недавних обзора литературы, описывающих молекулярные процессы развития семян у арабидопсиса и риса (Chen et al., 2021; Li et al., 2019; Li et al., 2018) для получения списков генов, контролирующих размер и форму семян. Для них в БД KEGG определили соответствующие ортогруппы.

Присвоение ортологичных групп KEGG генам пшеницы по их последовательности было выполнено с использованием инструментов BlastKOALA и KofamKOALA.

Ортогруппы были присвоены генам по результатам по крайней мере одной из программ. Мы считали гены в локусах QTL приоритетными если ортогруппы к которым они принадлежали были в списках, связанных с признаками цвета и размера зерна.

Генетический анализ трех характеристик размера семян (sL , sW и sA), четырех характеристик формы (sCi , sRo , sRu и sSo) и 48 характеристик цвета выявил в общей сложности 20, 22 и 170 QTL (всего 212) (рис. 10А), соответственно, на всех хромосомах генома пшеницы, кроме хромосом 1А и 5D.

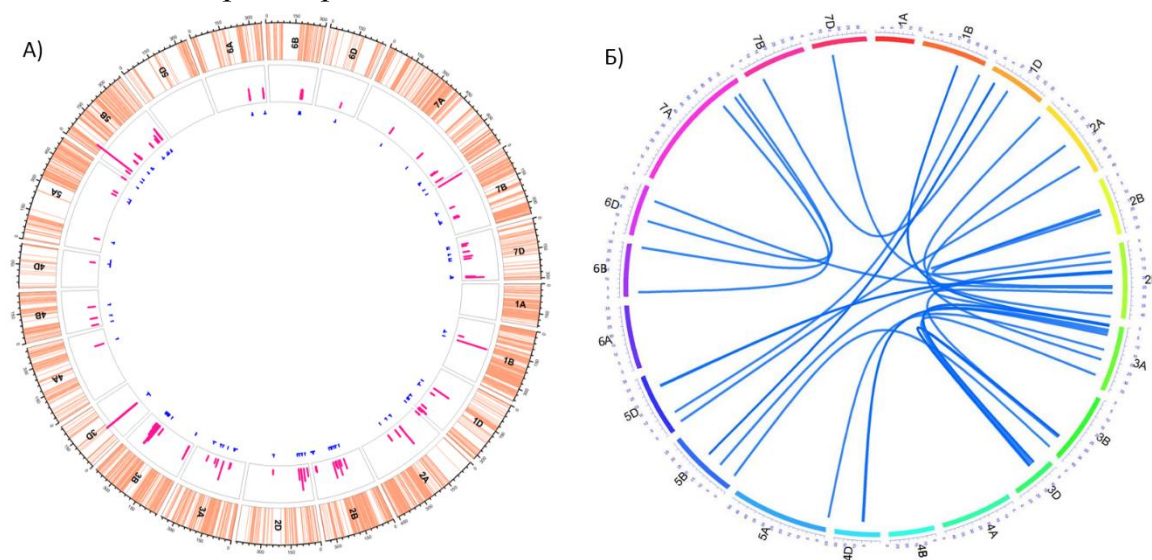


Рисунок 10. А). Распределение аддитивных QTL (синие линии во внутреннем круге) на хромосомах мягкой пшеницы. Светло-оранжевые линии на внешней дорожке указывают позиции SNP на каждой хромосоме; розовые полосы во втором круге указывают значения LOD QTLs. Координаты хромосом даны в сМ. Синие линии под окружностью трека указывают доверительный интервал QTL, а маленькие вертикальные линии указывают на положение пика QTL. Б). Сеть эпистатических пар QTL. Внешний круговой график представляет координаты генома мягкой пшеницы, расположенные на хромосомах 1A-7D по часовой стрелке.

Кроме того, мы также смогли обнаружить 34 пары эпистатических QTL, контролирующих по меньшей мере 22 признака в линиях пшеницы (рис. 10Б). Эти QTL включали все хромосомы пшеницы, кроме хромосом 1А, 4А, 4В и 6А.

Мы наблюдали заметное совпадение между местоположениями QTL для разных признаков. Например, хромосома 3В имеет локусы с множественными ассоциациями признаков: позиция 39.179 сМ (2 признака размера), позиция 298.179 сМ (7 признаков цвета), позиция 299.179 сМ (12 признаков цвета), позиция 300.179 сМ (2 признака цвета) и еще несколько подобных локусов. Чтобы оценить сходство различных анализируемых признаков в рамках подобной вырожденности, когда несколько признаков контролируются одним локусом, мы сгруппировали признаки иерархически по степени перекрытия местоположений QTL на основе индекса Очаи (рис. 11).

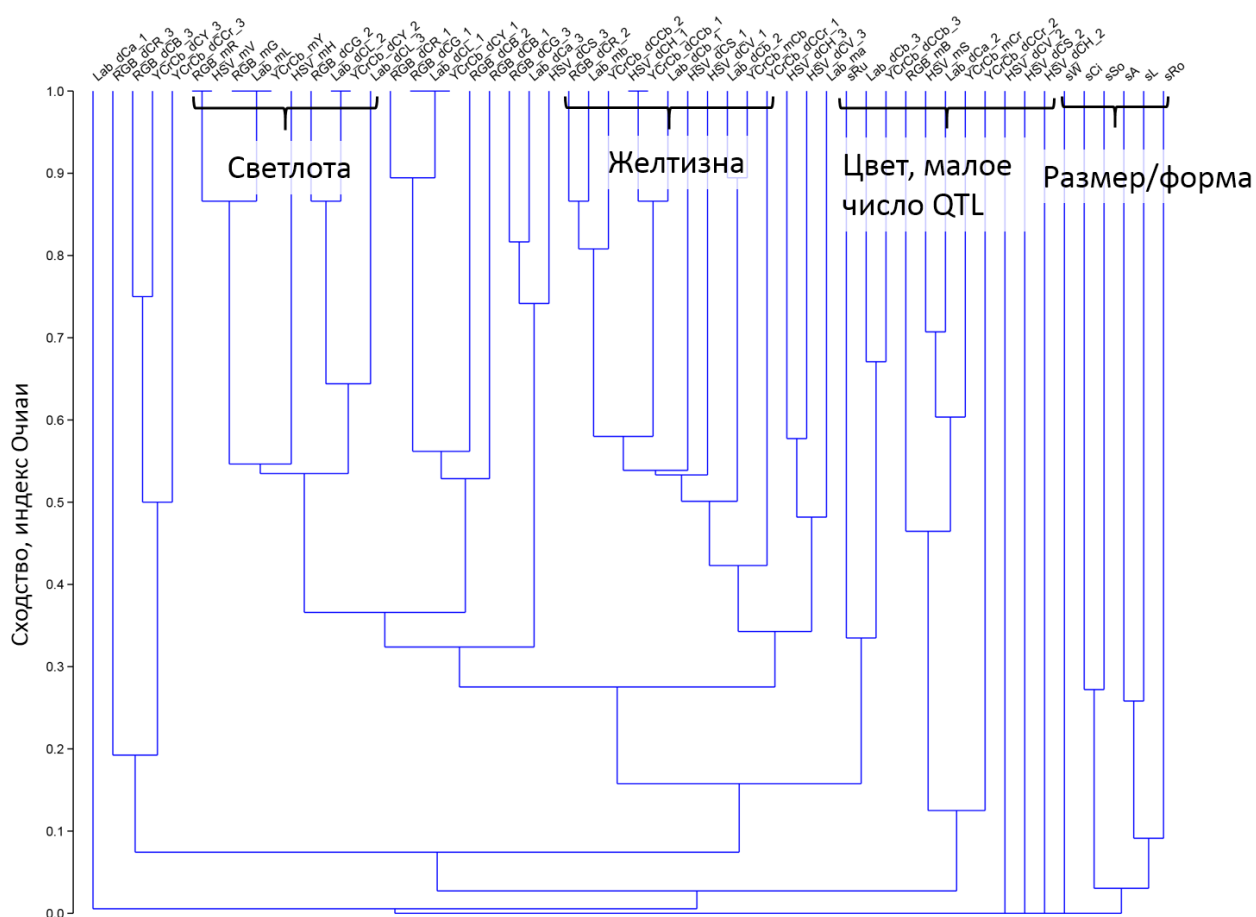


Рисунок 11. Дерево сходства для признаков семян, полученных по степени перекрытия между их местоположениями QTL. Вертикальная ось представляет меру сходства, основанную на индексе Оchiaи (ось Y). Листья соответствуют признакам семян. Группы признаков с сильным перекрытием местоположений QTL показаны фигурными скобками.

Результаты показывают, что характеристики размера/формы (правая часть дерева) отделены от цветковых признаков (за исключением шероховатости, sRu). Некоторые признаки с небольшим числом QTL (от 1 до 2) также отделены от других признаков. Значительное количество признаков, связанных с желтизной, образуют большую группу. Наконец, признаки, связанные со светлотой (Lab_mL, HSV_mV, YCrCb_mY), попадают в один и тот же кластер, и их QTL сильно перекрываются. С одной стороны, это отражает вырожденность оцениваемых признаков, что указывает на их избыточность. С другой стороны, расположение нескольких локусов, связанных с цветовыми признаками, в одном и том же регионе может указывать на более надежную идентификацию ассоциации локуса с конкретным признаком.

Мы идентифицировали 2787 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. Для 29 генов из локусов, связанных с размером семян, мы обнаружили совпадение в списке ортогрупп, полученном на основе анализа литературных данных. 11 генов, идентифицированных таким образом, принадлежат регуляторным белкам (факторы транскрипции EREBP,

HD-ZIP, MYBP; локусы на хромосомах 3A, 2B, 2D, 7D). Они могут быть вовлечены в регуляцию роста и развития семян. 6 генов относятся к факторам инициации трансляции (ELF2C; локусы на хромосомах 2B, 7D). 5 генов относятся к ферментам, связанным с процессами убиквитинирования (локусы на хромосомах 2D). 4 гена обладают хитиназной активностью (локус на хромосоме 7D). 2 гена обладают цитокиндегидрогеназной активностью (хромосома 7D).

Для генов из локусов, связанных с цветом семенной оболочки, 14 обнаружили совпадение с ортогруппами метаболических путей базы данных KEGG, связанных с биосинтезом пигментов. Восемь генов были вовлечены в биосинтез фенилпропаноидов (локусы на хромосомах 3A, 3B, 6A и 6B), два гена были вовлечены в путь биосинтеза каротиноидов (локусы на хромосомах 2A и 6A), по одному гену каждый был вовлечен в биосинтез флавонов и флавонолов, биосинтез флавоноидов, метаболизм триптофана и биосинтез терпеноидов.

Изучение взаимосвязи между характеристиками размера, формы и цвета зерен мягкой пшеницы, временем их хранения и всхожестью. Мы изучали 44 образца из популяции рекомбинантных инбредных линий (РИЛ) картирующей популяции мягкой пшеницы (*Triticum aestivum* L.) ITMI, описанной выше. Растения каждого генотипа были выращены в 2003, 2004, 2009 и 2014 гг. Для 19 генотипов проф. А. Бёрнером были получены данные по всхожести зерен в 2003, 2004, 2009 и 2014 г. Эти данные позволяют исследовать влияние длительности хранения зерен в генбанке на их характеристики.

Для оценки влияния длительности хранения зерен на их характеристики мы проводили дисперсионный анализ (независимая переменная – год урожая). Для выявления трендов между годом урожая и значением признака использовали коэффициент корреляции Пирсона. Поскольку годы урожая (2003, 2004, 2009 и 2014) разделены неравномерными временными интервалами, мы применили три варианта кодировки независимой переменной. Первый вариант кодировки (Year01) приписывал для годов 2003 и 2004 значение 0, а для 2009 и 2014 года значение 1. Второй вариант кодировки (YearRank) приписывал разным годам ранг в порядке возрастания (соответственно, 1 для 2003, 2 для 2004, 3 для 2009 и 4 для 2014). Третий вариант кодировки (Year) использовал непосредственно значение года урожая. Оценку значимости корреляций проводили на основе рандомизационных тестов (бутстреп и перестановочного).

Дисперсионный анализ оценивал влияние двух факторов (генотип и год урожая) независимо, для значений p -value мы делали поправку Бонферрони на число анализируемых признаков (55), p_{corr} . Анализ показал, что генотип оказывает значимое влияние на подавляющее большинство анализируемых признаков, для 53 из 55 признаков за исключением sRu (rugosity) и HSV_dCH_2. Для второго фактора, года урожая, число таких признаков составило 34, а для 21 из них $p > p_{\text{corr}}$. К признакам размера/форм зерен, для которых влияние года урожая не значимо, относятся sA (площадь), sRu (шероховатость), sW (ширина). К признакам цвета, для которых влияние года урожая не значимо, относятся компоненты разных цветовых шкал в основном для

2 и 3 доминантных цветов.

Для 45 из 55 признаков зависимость между YearRank и признаками зерен является значимой как для бутстреп, так и для перестановочного теста. Это также наблюдается и для других кодировок года урожая. Большинство характеристик цвета зерен, 44 из 48, демонстрируют значимую связь с годом урожая. Среди признаков размера/формы только один, шероховатость, демонстрирует значимую линейную связь с годом урожая. Наибольшие по абсолютному значению отрицательные значения коэффициентов корреляции (все меньше или равные -0.6) наблюдаются для признаков, связанных с компонентой a пространства Lab, которая характеризует красный оттенок цвета. Это означает, что зерна ранних годов урожая имеют более красный оттенок оболочки, по сравнению с более поздними годами урожая. Несколько меньшие по модулю отрицательные значения коэффициентов корреляции наблюдаются для признаков компоненты Cr шкалы YCrCb, которая является красной цветоразностной компонентой этого пространства и также связана с красным оттенком. Наибольший по модулю положительный коэффициент корреляции (0.49) наблюдается для признака HSV_mH, который отражает среднее значение цветовой компоненты насыщенность шкалы HSV.

На рис. 12 приведены интервальные диаграммы распределений значений признаков Lab_ma, HSV_mv (значимая положительная связь с годом урожая) HSV_mH (значимая отрицательная связь с годом урожая) и sL (нет значимых корреляций с годом урожая) для всех генотипов в зависимости от года урожая.

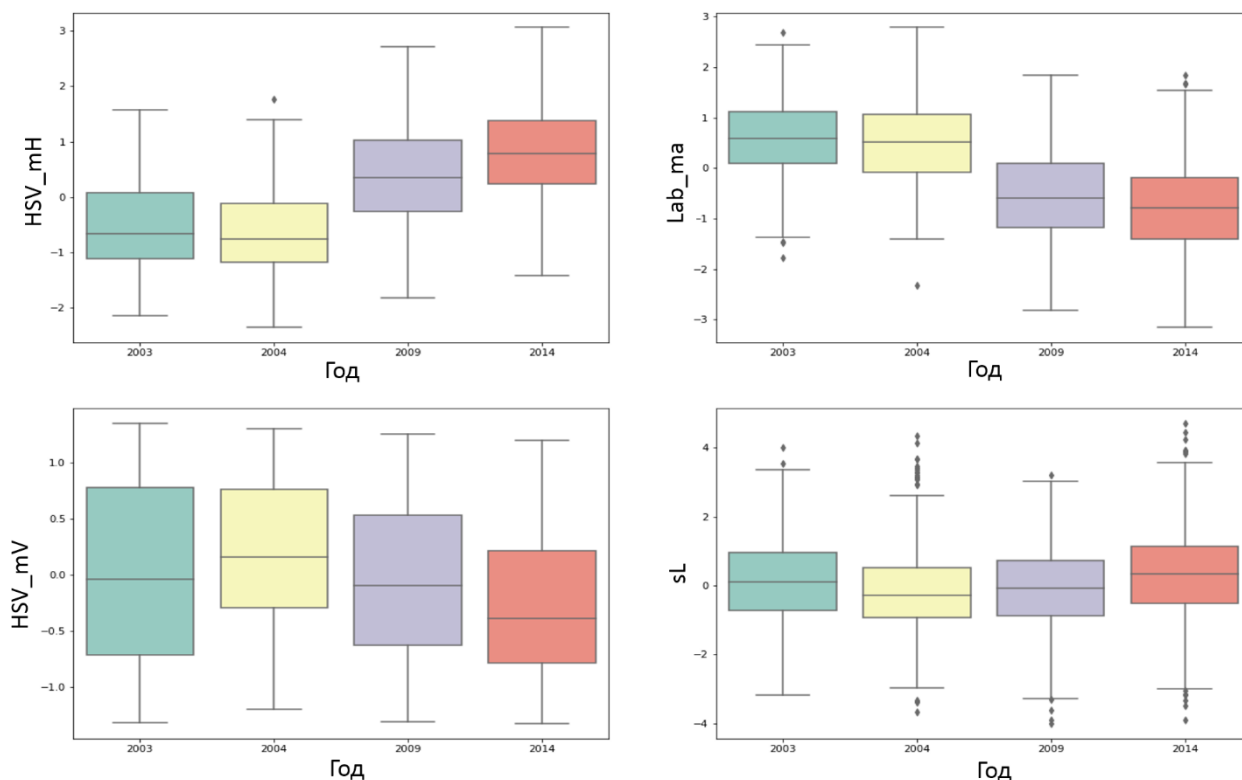


Рисунок 12. Интервальные диаграммы распределений нормализованных параметров HSV_mH, Lab_ma, HSV_mv и sL (оси Y) в образцах зерен, собранных в сезоны разных лет (ось X).

Для того, чтобы исследовать зависимости между всхожестью зерен и их признаками мы вначале с помощью метода ANOVA оценили их зависимость от двух

факторов: генотипа и года урожая. Оказалось, что для генотипа такая связь оказалась не значимой ($p\text{-value}=0.98$), а для года урожая – значимой ($p\text{-value}=1.2 \cdot 10^{-15}$). Значимая положительная связь была обнаружена только между всхожестью и средним значением hue (шкала HSV). Отрицательные значимые коэффициенты корреляции были выявлены для ряда признаков, которые связаны с красным оттенком зерна (рис. 13).

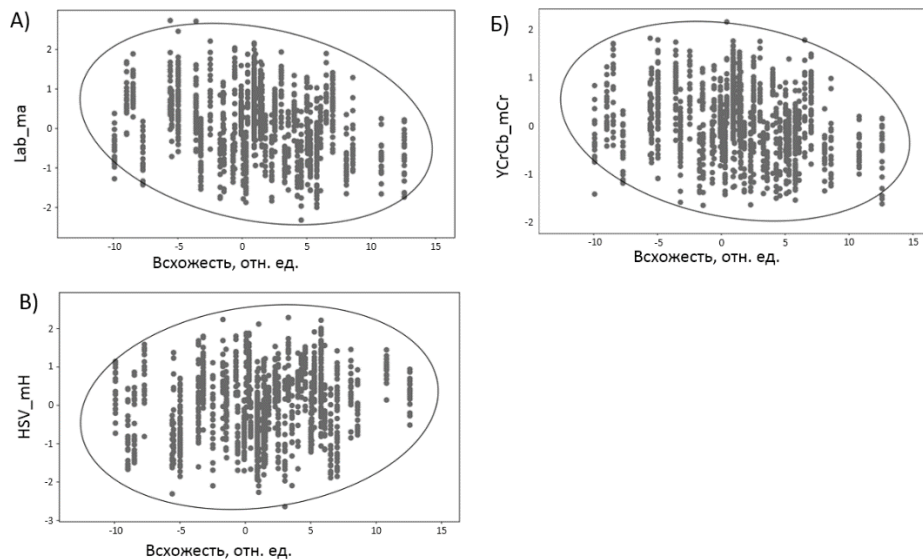


Рисунок 13. Диаграммы рассеяния для значений всхожести (оси X) и признаков цвета зерен (оси Y): (A) Lab_ma, (B) YCrCb_mCr, (C) HSV_mH. Показаны эллипсоиды рассеяния при уровне значимости 95%.

ГЛАВА 4. ВЫСОКОПРОИЗВОДИТЕЛЬНОЕ ФЕНОТИПИРОВАНИЕ КОЛОСА ПШЕНИЦЫ

Протокол и алгоритм анализа изображений для фенотипирования колосьев.

Нами были предложены два протокола для получения изображения колоса. Протокол 1 («на прищепке»). Колос располагается вертикально перед синим фоном. Колос и цветовая шкала ColorChecker держатся на прищепках, которые помещаются на штатив. С помощью данного метода можно производить съемку колоса в 4 проекциях. Протокол 2 («на столе»). Колос и цветовая шкала лежат на стеклянном столике, который располагается на столе с поверхностью синего цвета (фон). Фотокамера фиксируется на стойке типа "журавль" над стеклом. С помощью данного метода можно производить съемку колоса в двух проекциях.

Обработка изображений осуществлялась с использованием пакета OpenCV. На первом этапе анализа мы выполняли идентификацию цветовой шкалы на изображении. На следующем этапе предварительной обработки мы сглаживали изображение при помощи фильтра Гаусса с ядром 3×3 для удаления шумов, удаляли элементы изображения штатива и прищепки (для протокола «на прищепке»). Бинаризацию изображения проводили после преобразования изображения в цветовое пространство HSV. Для идентификации остей использовался алгоритм, состоящий из двух этапов. На первом этапе идентифицировали пиксели остова остей. Поскольку толщина остей существенно меньше, чем толщина тела колоса, для получения остова остей

использовали алгоритм частичной скелетизации. В процессе эрозии последовательность удаляемых пикселей контура запоминалась. На основе этого можно было определить пиксели, удаление которых приводило к формированию остова. Эти пиксели отмечались как пиксели остей. Для подбора параметров алгоритмов использовали обучающую выборку из 93 изображений. После выделения областей остей и колоса рассчитываются характеристики остистости колоса: площадь области остей (количество формирующих ее пикселей, умноженное на площадь одного пикселя); количество остей; суммарная длина остей L_a , которая определялась как сумма длин пикселей, формирующих скелетизированный остов области остей; средняя длина остей.

После удаления остей выполняли сглаживание контура колоса на основе алгоритма вычисления эллиптических дескрипторов Фурье. Осевую линию колоса аппроксимировали ломаной линией, построенной на основе сегментов контура “тела” колоса. Далее производится процедура выпрямления контура колоса так, что эта линия проходит вертикально (верхний конец соответствует вершине колоса, нижний – основанию колоса).

Характеристики формы колоса определялись на основе анализа выпрямленного контура и включали несколько категорий. К первой категории относятся интегральные характеристики формы: длина колоса, периметр контура тела колоса, площадь тела колоса, отношение площади колоса к квадрату его длины. Также оценивали округлость, закругленность, шероховатость, целостность (как это описано для зерен).

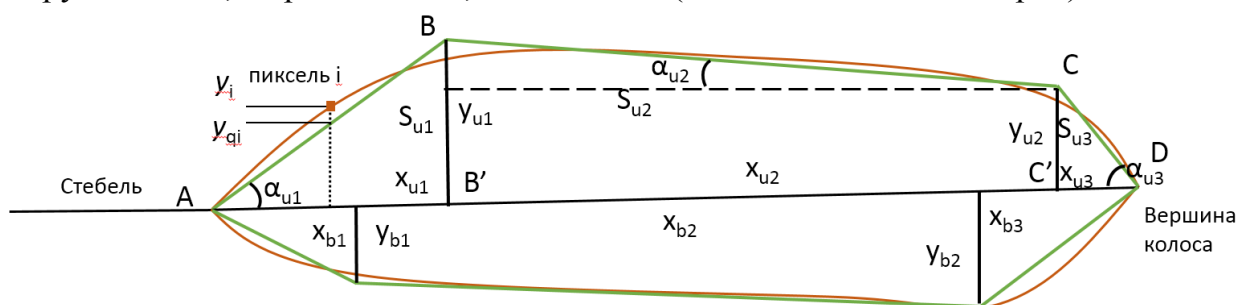


Рисунок 14. Представление формы колоса в виде двух четырехугольников. Черной горизонтальной линией показана осевая линия колоса. Контур колоса показан коричневой линией. Четырехугольники, аппроксимирующие контур колоса показаны зелеными линиями. На рисунке для верхнего четырехугольника показаны основные параметры, характеризующие его геометрию.

Форму колоса описывали дополнительно несколькими моделями. Первая модель (сечений) для описания формы колоса представляла собой набор сечений, которые определялись перпендикулярами к его осевой линии, отложенными с шагом в 1/21 длины колоса. Вторая – модель четырёхугольников. Контур колоса, расположенного горизонтально, можно представить в виде двух четырехугольников (рис. 14), верхнего и нижнего, с общим основанием – осевой линией колоса. Геометрия каждого из двух четырехугольников определяется четырьмя независимыми параметрами. Параметры верхнего и нижнего четырехугольника подбирались таким образом, чтобы минимизировать квадратичное отклонение по перпендикуляру от осевой линии между

пикселями контура и ребрами четырехугольника. Минимизация проводилась с помощью алгоритма Левенберга-Марквардта (Ranganathan, 2004).

Третья модель – радиальная. Параметры этой модели вычисляются следующим образом: от центра масс контура колоса проводятся 360 лучей, начиная от направления главной оси контура с шагом в 1 градус, до ближайшей точки линии контура к центру масс вдоль луча. Длины полученных таким образом 360 отрезков представляют радиальную модель.

Для оценки точности алгоритма сегментации изображения на фон и колос использовался индекс Жаккара J . На тестовой выборке изображений после подбора оптимальных параметров алгоритма было достигнуто среднее значение точности распознавания тела колоса $J_b = 0.925$ и остей $J_a = 0.660$. На обучающей точность распознавания тела колоса $J_b = 0.932$ и остей $J_a = 0.634$. Использование цветокоррекции не оказало существенного влияния на точность распознавания колоса и остей.

Мы провели оценку влияния различных факторов, связанных с протоколом получения изображения и его обработки на точность определения остей: масштаб съемки (количество пикселей на единицу фотографируемой площади); тип протокола («на столе» или «на прищепке»); проекция колоса (лицевая или боковая сторона колоса) для протокола «на прищепке». Результаты, представленные на рисунке, показали, что масштаб изображения значимо влияет на точность распознавания как тела колоса, так и остей. Тип протокола оказывают существенное влияние на точность определения тела колоса, но не остей.

Примеры изображений колосьев формы трех различных типов (компактный, нормальный и спельтоидный) приведены на рис. 15.



Рисунок 15. Изображения колосьев формы различных типов: компактный (А), нормальный (Б), спельтоидный (В). Каждая панель содержит три фрагмента: на левом изображено исходное изображение колоса, в центре – сегментированное на тело (зеленый цвет), ости (малиновый цвет) и линии сечений (синим). Справа приведено изображение модели четырехугольников для соответствующего колоса с указанием основных параметров модели.

Из рис. 15 видно, что использование моделей четырехугольников демонстрирует различия в параметрах этой модели для колосьев трех типов.

На основе аннотированной экспертами выборки колосьев 249 растений мы оценили точность предсказания типа колосьев и их плотности D различными методами машинного обучения с использованием параметров - характеристик колосьев,

описанных выше (параметры 3-х моделей + общие характеристики).

Вначале использовали отбор наиболее значимых признаков по их вкладу в предсказание плотности колоса. Ранжирование признаков по значимости осуществлялось по усредненному значению результатов 8 различных подходов, предлагаемых в библиотеке *sklearn*: (1) коэффициенты линейной регрессии; (2) коэффициенты регрессии с регуляризацией лассо; (3) коэффициенты регрессии с регуляризацией ridge; (4) Отбор признаков по критерию стабильности (Meinshausen, Bühlmann, 2010); (5) ранг признака, оцененный путем рекурсивного удаления признаков с наименьшей значимостью (RFE) (Guyon et al., 2002); (6) Ранжирование признаков на основе среднего уменьшения энтропии признака, которая вычисляется при построении деревьев решений в модели случайного леса; (7) Ранжирование признаков на основе вычисления корреляции между регрессором и целевым признаком; (8) ранжирование на основании значения максимального информационного коэффициента (MIC) (Reshef et al, 2011). Были отобраны по 12 признаков для протоколов «на столе» (одна проекция) и «на прищепке» (4 проекции) (общие из них 14 признаков). Признаками с высокой значимостью для обоих протоколов оказываются длина и индекс округлости колоса.

Ошибка MAE предсказания индекса плотности колоса D на тестовой выборке для одной проекции равна 4.61, для четырех проекций 3.33. Коэффициент корреляции Пирсона между предсказанными значениями и значениями определенные экспертом для тестовой выборки равен 0.51 для одной проекции и 0.74 для четырех проекций.

Метрика точности/метод предсказания	Протокол «на прищепке»	Протокол «на столе»
<i>F</i> /лог. регрессия	0.82	0.78
<i>F</i> /лог.регрессия+индекс плотности	0.83	0.77
<i>F</i> /случ. лес	0.85	0.72
<i>F</i> /случ.лес+индекс плотности	0.84	0.76

Таблица 5. Оценки качества предсказания формы колоса для одной (протокол «на столе») и четырех (протокол «на прищепке») проекций.

Провели автоматическую классификацию колосьев на три типа (компактный, нормальный и спельтоидный) с помощью нескольких методов: логистической регрессии, случайного леса, логистической регрессии и дополнительно предсказанный индекс плотности D, случайный лес и дополнительно предсказанный индекс плотности D. Результаты оценки точности классификации на тестовой выборке представлены в табл. 5. 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%.

ГЛАВА 5. БАЗЫ ДАННЫХ ДЛЯ ОПИСАНИЯ ФЕНОТИПА ПШЕНИЦЫ

База данных WheatPGE для описания генотипа, фенотипа и окружающей среды у пшеницы. Одной из актуальных задач феномики растений является создание систем хранения фенотипических данных и баз данных (БД), которые будут способствовать ускорению процесса сбора фенотипических данных.

В рамках диссертационной работы нами была создана БД WheatPGE, которая позволяет обеспечить информационную поддержку для селекционно-генетических экспериментов у пшеницы, включая массовое фенотипирование растений. БД обеспечивает интерфейс для сбора этих данных в процессе селекционно-генетического эксперимента (рис. 16). Структура базы обеспечивает интеграцию этих данных, поиск информации и ее извлечение в формате, удобном для дальнейшей статистической обработки. Это позволяет экспериментатору затрачивать меньше времени на сбор данных, обеспечивает их сохранность, доступ к данным в любой момент времени.

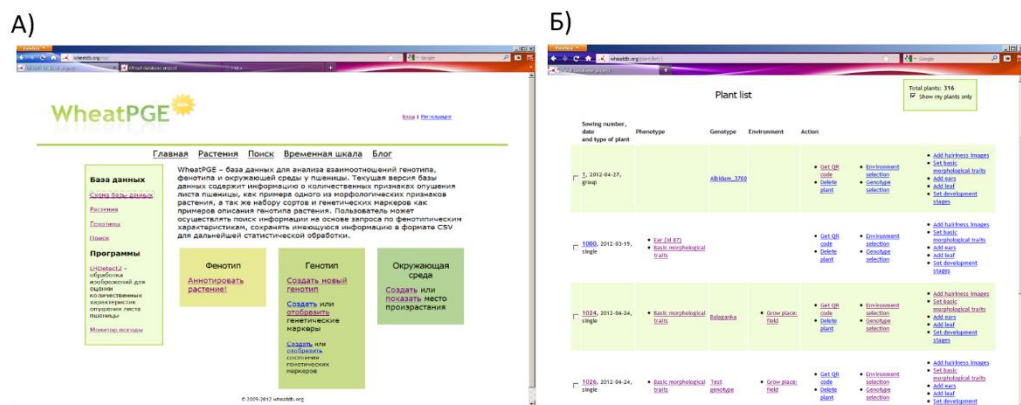


Рисунок 16. Интерфейс БД WheatPGE. А) Главная страница доступа к БД; Б) Визуализация списка растений, созданного пользователем.

Всего в текущей версии БД содержит 32 таблицы и 55 отношений между ними. Генотип растения описывается девятью таблицами. Они включают описание сорта растения или линии. Фенотип растения описывается 15 таблицами: таблица, описывающая базовые признаки растения (длина стебля, число колосьев, общую урожайность); таблицы, описывающие структуру урожая (главный и вторичные колосья); таблицы, описывающие опушение листа; таблица, описывающая длительность стадий развития растения. Место произрастания описано двумя таблицами, основная информация в которых содержит название, широту, долготу, тип климата, климатические характеристики (среднегодовую температуру, среднегодовую влажность, средние температуры января и июля).

Блок информации, связанный с проведением эксперимента, содержит две таблицы, описывающих событие и их список для растения. Событие описывает проведение измерений над фенотипом, агротехнические мероприятия, или изменение фазы развития растений. Блок информации, описывающий данные пользователя, содержит три таблицы с информацией, позволяющей управлять уровнем доступа пользователя к БД WheatPGE.

Интерфейс основан на использовании методологии MVC (Model-View-Controller) для разработки программного обеспечения. Мы использовали реализацию MVC на платформе Catalyst. Для представления объектов и отношений между ними в базе данных использована технология реляционной системы управления базой данных (СУБД) наравне с технологиями объектно-реляционного отображения (ORM – object relation mapping). В качестве СУБД использована MySQL 5, развернутая на сервере баз данных под управлением CentOS Linux.

База данных SpikeDroid DB. Для сбора, хранения и анализа информации о

фенотипических характеристиках колоса пшеницы нами была разработана компьютерная информационная система SpikeDroidDB. Эта БД позволяет структурированно хранить цифровые изображения колоса, производить их аннотацию и предоставляет пользователю гибкую систему для доступа к данным. Модель объектов и отношений базы данных SpikeDroidDB показана на рис. 17.

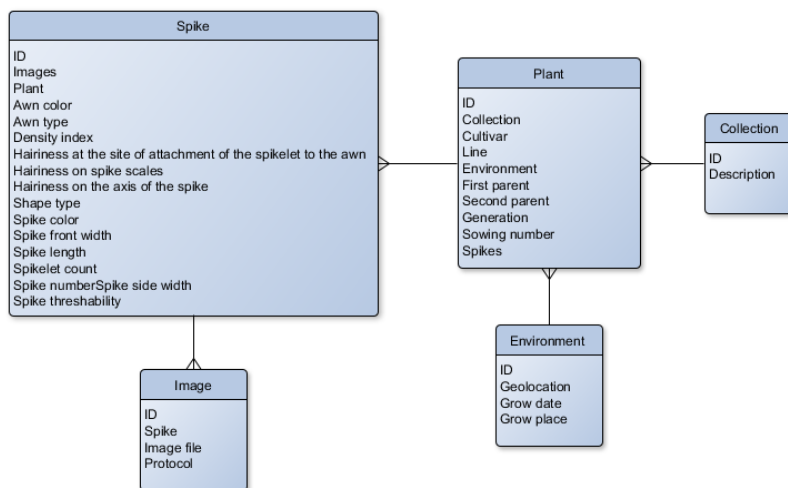


Рисунок 17. Модель объектов и отношений базы данных SpikeDroidDB.

Таблица Plant (Растение) включает описание сорта (образца) растения или линии, ссылки на родительские растения, поколение и посевной номер растения. Растение связано с таблицами Collection (Коллекция), Environment (место произрастания) и Spike (описание колоса). Для коллекции указывается её держатель и аффилиация (описание). Место произрастания содержит геоданные, дату вегетации, фенологические данные и данные об условиях выращивания растения. Фенотип колоса описывается набором из 14 признаков. Для признаков колоса пшеницы, содержащийся в нашей базе данных была выполнена привязку к признакам онтологии сельскохозяйственных растений портала CropOntology.org, в тех случаях где это представлялось возможным. Система SpikeDroidDB была разработана на основе системы управления контентом (CMS – content manager system) Drupal 8. В качестве хранилища данных используется реляционная база данных под управлением MySQL СУБД, которая развернута на сервере баз данных под управлением CentOS Linux.

Текущая версия базы данных содержит коллекцию колосьев гибридов F2 от скрещивания австралийского сорта мягкой пшеницы ‘Triple Dirk’ на образец KU506 китайской пшеницы *Triticum yunnanense*. Коллекция включает в себя 103 растения и 230 фотографий колосьев.

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

ВЫВОДЫ

1. Разработана экспериментально-компьютерная платформа ICGPhenoPlant для фенотипирования растений, позволяющая с высокой точностью на основе двумерных изображений:

- (а) с помощью приложения LHDetect2 оценивать количественные характеристики опушения листьев пшеницы, картофеля и табака;
- (б) с помощью приложения SeedCounter оценивать 3 признака размера, 4 признака формы и 48 признаков цвета оболочки зерен пшеницы;
- (в) с помощью приложения WERecognizer оценивать характеристики размера и формы колоса пшеницы;
- (г) сохранять информацию о фенотипе, генотипе и месте произрастания растений в базах данных WheatPGE и SpikeDroid.

2. Использование приложения LHDetect2 для анализа опушения пшеницы позволило:

- (а) впервые изучить опушение у 47 образцов пшеницы и ее диких сородичей; выделить 9 характерных типов опушения; установить, что для гексаплоидных пшениц наблюдается положительная корреляция между числом трихомов и их длиной; установить, что характеристики опушения листьев пшеницы значимо ассоциированы с хромосомным набором генома;
- (б) для генов, контролирующих опушение листа пшеницы впервые показано, что ген H12aesp в большей степени влияет на длину трихом, а гены H11 и H13 влияют одновременно и на длину трихомов и на их количество;
- (в) впервые на основе точной оценки количественных характеристик опушения листа пшеницы программой LHDetect2 показана связь между изменчивостью опушения листа пшеницы и ответом растения на водodefицит: число трихомов на обеих сторонах листа увеличивается, а их средняя длина уменьшается; изменение длины трихомов связано с нефотохимическим тушением и числом устьиц на нижней стороне листа; изменение числа трихомов на нижней и верхней стороне листа связано с эффективностью использования воды.

3. Использование приложения LHDetect2 для анализа опушения листьев двудольных растений показало, что:

- (а) опушение листьев для 35 сортов картофеля демонстрирует широкое разнообразие по количеству трихомов и их длине; наиболее интенсивное опушение демонстрируют сорта Великан, Танай, Кемеровчанин;
- (б) линии табака с измененным паттерном экспрессии гена пролиндегидрогеназы PDH и повышенным содержанием пролина по сравнению с контрольным генотипом SR1 имеют достоверно большее количество трихомов на верхней стороне листа, при этом средняя длина трихомов достоверно меньше, чем у растений контрольной линии.

5. Использование приложения SeedCounter для анализа зерен пшеницы растений популяции ITMI позволило:

- (а) выявить 212 аддитивных и 34 пары эпистатических QTL для признаков зерен; идентифицировать 29 генов-кандидатов для локусов, ассоциированных с

размером семян, и 14 для локусов, ассоциированных с цветом оболочки;
(б) показать, что для зерен мягкой пшеницы, хранившихся в условиях генбанка пигментация оболочки значительно изменяется в зависимости от срока хранения;
(г) выявить значимую отрицательную корреляцию между интенсивностью красной компоненты цвета оболочки и всхожестью зерен.

6. Использование приложения WERecognizer на основе количественных характеристик размера и формы колоса пшеницы позволяет с высокой точностью оценить тип колоса и индекс его плотности.

СПИСОК ОСНОВНЫХ ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Arif, M.A.R.; Komyshev, E.G.; Genaev, M.A.; Koval, V.S.; Shmakov, N.A.; Börner, A.; **Afonnikov, D.A.** (2022) QTL Analysis for Bread Wheat Seed Size, Shape and Color Characteristics Estimated by Digital Image Processing. *Plants*, 11, 2105.
2. Ibragimova, S., Genaev, M., Kochetov, A., **Afonnikov, D.** (2022). Variability of leaf pubescence characteristics in transgenic tobacco lines with partial proline dehydrogenase gene suppression. *Biologia plantarum*, 66, 24-28.
3. **Afonnikov D.A.**, Komyshev E.G., Efimov, V.M., Genaev M.A., Koval V.S., Gierke P.U., Börner A. (2022) Relationship between the Characteristics of Bread Wheat Grains, Storage Time and Germination. *Plants* 2022, 11, 35.
4. Комышев, Е. Г., Генаев, М. А., **Афонников, Д. А.** (2020). Анализ цветковых и текстурных характеристик зерен злаков на цифровых изображениях. *Вавиловский журнал генетики и селекции*, 24(4), 340-347.
5. Иванова К.А., Комышев Е.Г., Генаев М.А., Егорова А.А., Колошина К.А., Чалая Н.А., **Афонников Д.А.**, Кочетов А.В., Рогозина Е.В., С.В. Герасимова (2019). Оценка количественных характеристик клубнеобразования дикого картофеля на основе анализа изображений клубней с использованием компьютерного приложения SeedCounter. *Вавиловский журнал генетики и селекции*, 19(6), 1-8.
6. Genaev M.A., Komyshev E.G., Smirnov N.V., Kruchinina Y.V., Goncharov N.P. and **Afonnikov D.A.** (2019) Morphometry of the Wheat Spike by Analyzing 2D Images. (2019) *Agronomy*, 9(7), 390.
7. Pshenichnikova, T.A., Doroshkov, A.V., Osipova, S.V., Permyakov, A.V., Permyakova, M.D., Efimov, V.M., **Afonnikov, D.A.** (2019) Quantitative characteristics of pubescence in wheat (*Triticum aestivum* L.) are associated with photosynthetic parameters under conditions of normal and limited water supply. *Planta*, 249 (3), 839-847.
8. Дорошков А.В., **Афонников Д.А.** (2018) Опушение листа у картофеля *Solanum tuberosum*: морфология, функциональная роль и методы исследования. (2018) *Вавиловский журнал генетики и селекции*. 22(1), 46-53. DOI 10.18699/VJ18.327.
9. Генаев М.А., Комышев Е.Г., Фу Хао, Коваль В.С., Гончаров Н.П., **Афонников Д.А.** (2018) SpikeDroidDB – информационная система для аннотации морфометрических характеристик колоса пшеницы. *Вавиловский журнал генетики и селекции*. 22(1):132-140.
10. Komyshev E.G., Genaev M.A., **Afonnikov D.A.** (2017) Evaluation of the SeedCounter, a

- mobile application for grain phenotyping. *Front. Plant Sci.*, 7:1990.
11. Pshenichnikova T.A., Doroshkov A.V., Simonov A.V., **Afonnikov D.A.**, Börner A. (2017) Diversity of leaf pubescence in bread wheat and relative species. *Genet. Res. and Crop Evol.*, 64, 1761-1773.
 12. Doroshkov, A. V., **Afonnikov, D. A.**, Dobrovolskaya, O. B., Pshenichnikova, T. A. (2016). Interactions between leaf pubescence genes in bread wheat as assessed by high throughput phenotyping. *Euphytica*, 207, 491-500.
 13. Дорошков А.В., Симонов А.В., Сафонова А.Д., **Афонников Д.А.**, Лихенко Н.Е., Колчанов Н.А. (2016) Оценка количественных характеристик опушения листьев картофеля с использованием анализа цифровых микроизображений. *Достижения науки и техники АПК*. Т.30, С. 12-14.
 14. **Афонников Д.А.**, Генаев М.А., Дорошков А.В., Комышев Е.Г., Пшеничникова Т.А. (2016) Методы высокопроизводительного фенотипирования растений для массовых селекционно-генетических экспериментов. *Генетика*. Т. 52. № 7. С. 788-803.
 15. Дорошков А.В., Генаев М.А., **Афонников Д.А.** Протокол анализа количественных характеристик опушения листа картофеля. *Вавиловский журнал генетики и селекции*. 2016;20(6):863-868.
 16. Дорошков А.В., **Афонников Д.А.**, Пшеничникова Т.А. (2014) Генетический анализ опушения листа у изогенных линий мягкой пшеницы сорта Новосибирская 67. *Генетика*, Т. 50, N. 2, С. 172–180.
 17. Genaev M.A., Doroshkov A.V., Pshenichnikova T.A., Kolchanov N.A., **Afonnikov D.A.** (2012) Extraction of quantitative characteristics describing wheat leaf pubescence with a novel image-processing technique. *Planta*, 236, 1943–1954.
 18. Генаев М.А., Дорошков А.В., Пшеничникова Т.А., Морозова Е.В., Симонов А.В., **Афонников Д.А.** (2012) Информационная поддержка селекционно-генетического эксперимента у пшеницы в системе WheatPGE. *Математическая биология и биоинформатика*. 2012. Т. 7. № 2. С. 410–424.
 19. Генаев, М.А., Дорошков А.В., Морозова Е.В., Пшеничникова Т.А., **Афонников Д.А.** (2011) Компьютерная система WheatPGE для анализа взаимосвязи фенотип–генотип–окружающая среда у пшеницы. *Вавиловский журнал генетики и селекции*. 15, 784-793.
 20. Дорошков А.В., Пшеничникова Т.А., **Афонников Д.А.** (2011) Анализ особенностей морфологии и наследования опушения листа пшеницы *Triticum aestivum* L. с помощью компьютерных методов фенотипирования. *Генетика*, Т. 47, № 6, с. 839–843.
 21. Дорошков А.В., Арсенина С.И., Пшеничникова Т.А., **Афонников Д.А.** (2009) Применение компьютерной обработки изображений микрофотографий листа для анализа опушения пшеницы *Triticum aestivum* L. *Информационный вестник ВОГиС*. 13, 218-226.
 22. Комышев Е. Г., Генаев М. А., **Афонников Д. А.** Программа для оценки количественных характеристик колоса пшеницы (ОКХК) / The program for the extraction the quantitative characteristics of the wheat spike (WERecognizer) Свидетельство о регистрации программы для ЭВМ N 2019666362 от 09.12.2019, дата

приоритета 27.11.2019.

23. Генаев М.А., Дорошков А.В., **Афонников Д.А.** (2012) Программа автоматического определения количественных характеристик опущения листа (ОЛДетект2). // Свидетельство о регистрации программы для ЭВМ N 2012620897 от 21.08.2012 от 5.03.2012, дата приоритета 21.08.2012.
24. Генаев М.А., Дорошков А.В., **Афонников Д.А.** (2010) База данных для анализа взаимосвязей генотип-фенотип-окружающая среда у пшеницы (ВитПГЕ). // Свидетельство о регистрации базы данных N 2010620602 от 23.10.2010, дата приоритета 30.06.2010