

**ПОЛХОВСКАЯ ЕКАТЕРИНА СЕРГЕЕВНА**

**СТРУКТУРНО-ТРАНСКРИПТОМНЫЙ АНАЛИЗ ГЕНОВ ПШЕНИЦЫ  
И ТРИТИКАЛЕ, ЭКСПРЕССИРУЮЩИХСЯ В ПРОЦЕССЕ РАЗВИТИЯ  
ЗЕРНОВКИ, С ПОМОЩЬЮ НАНОПОРОВОГО СЕКВЕНИРОВАНИЯ**

**1.5.7 Генетика (биологические науки)**

Автореферат

диссертации на соискание ученой степени

кандидата биологических наук

**Уфа – 2024**

Работа выполнена в лаборатории маркерной и геномной селекции растений ФГБНУ Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии (г. Москва, Россия).

**Научный руководитель:** Соловьев Александр Александрович, доктор биологических наук, профессор РАН

**Официальные оппоненты:** Гончаров Николай Петрович, доктор биологических наук, академик РАН, главный научный сотрудник сектора генетики пшениц ФГБНУ «Федеральный исследовательский центр Институт цитологии и генетики СО РАН» (ИЦиГ СО РАН)

Казакова Елизавета Александровна, кандидат биологических наук, старший научный сотрудник лаборатории молекулярно-клеточных основ сельскохозяйственной радиобиологии ФГБУ «Всероссийский научно-исследовательский институт радиологии и агроэкологии Национального исследовательского центра «Курчатовский институт» (НИЦ «Курчатовский институт» – ВНИИРАЭ)

**Ведущая организация:** Федеральное государственное бюджетное учреждение науки Институт общей генетики им. Н.И. Вавилова Российской академии наук (ФГБУН ИОГен РАН)

Защита диссертации состоится «27» ноября 2024 г. в «10.00» часов на заседании Диссертационного совета 24.1.218.01 при Уфимском федеральном исследовательском центре Российской академии наук по адресу: 450054, г. Уфа, пр. Октября, 71, конференц-зал ИБГ УФИЦ РАН (ком. № 406).

С диссертацией можно ознакомиться в библиотеке УФИЦ РАН по адресу 450054, г. Уфа, пр. Октября, 71 и на сайтах ВАК РФ (<https://vak.minobrnauki.gov.ru>) и УФИЦ РАН (<http://ufaras.ru>).

Автореферат разослан «\_\_\_» \_\_\_\_\_ 2024 г.

Ученый секретарь диссертационного  
совета 24.1.218.01  
доктор биологических наук, доцент

Гульназ Фаритовна Корытина

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы.** Зерновка — это важнейший орган злаковых растений, обеспечивающий их высокую ценность для питания людей. Развитие зерновки — один из ключевых и практически значимых биологических процессов, включающий разнообразные физиологические, биохимические и транскрипционные изменения (Rogers et al., 1983; Rangan et al., 2017; Wan et al., 2008). Эти изменения и соответствующие им стадии развития зерновки были хорошо охарактеризованы с помощью морфологических, гистологических и биохимических методов исследований. При этом гены, вовлеченные в развитие и формирование зерновки злаковых, динамика их транскрипции и полиморфизм среди существующих сортов остаются малоизученными. Полученные данные об экспрессирующихся генах недостаточны для отражения полной информации об изменениях транскриптома на разных этапах развития зерновки.

Следовательно, для лучшего понимания транскрипционного ландшафта зерновки требуется учитывать его высокую динамичность, а, значит, проводить исследования в различных временных точках и применять технологии секвенирования, основанные на длинных ридах, которые позволяют выявлять наиболее широкий набор экспрессируемых генов, связанных с ключевыми процессами, определяющими развитие зерновки.

Поэтому изучение молекулярно-генетических и транскрипционных особенностей развития зерновки является актуальным и крайне востребованным направлением современных исследований. Быстрая сменяемость транскрипционного ландшафта во время развития зерновки и сложность генома злаковых, многие из которых являются аллополиплоидами (например, пшеница и тритикале) и имеют большой размер генома с многочисленными повторами ДНК, существенно усложняют аннотацию и изучение генов развития зерновки. В то же время новые технологии секвенирования длинными ридами (Oxford Nanopore Technology и Pacific BioScience) позволяют улучшить точность аннотации генома и детекции новых транскриптов, включая длинные некодирующие РНК (днРНК). Поэтому использование секвенаторов третьего поколения для характеристики транскриптома и изучения структурных особенностей генов злаковых является важной задачей, решение которой позволит получить новые сведения о генетических основах развития зерновки.

Секвенирование третьего поколения позволило получить информацию о генах, вовлеченных в процесс развития зерновки, в частности, глютенинов. Это открыло широкие возможности для идентификации аллелей глютенинов,

которые по-разному влияют на хлебопекарные качества. Однако геномная организация даже хорошо изученных глютенинов остается не полностью охарактеризована ввиду того, что их аллели могут различаться структурными вариациями, расположенными как в кодирующей части гена, так и, что немаловажно, в промоторе.

Помимо сложности и полиморфности самих генов, вовлеченных в процесс развития зерновки, их изучение осложняется геномной организацией, как пшеницы, геном которой состоит из трех субгеномов и содержит в себе повторяющиеся элементы разной природы, так и тритикале, которая, являясь аллополиплоидом, объединяет в себе сложнейшие геномы пшеницы и ржи, что затрудняет аннотацию ее генома.

Именно поэтому появилась необходимость оптимизации новых методов, ранее не использованных на зерновых, таких как секвенирование целых генов. Такие подходы, помимо идентификации структурных вариаций, позволяют проводить профилирование метилирования, что является важным эпигенетическим фактором, влияющим на функцию генов.

**Степень разработанности темы.** Исследования, посвященные изучению биологии развивающейся зерновки, проводились на протяжении XX века (Villareal et al., 1990, Rogers et al., 1983, Shewry et al., 1992). В частности, были установлены основные запасные белки зерновки пшеницы, которые представлены проламинами (глиадином) и глютелинами (глютенином). Наибольшее влияние на качественные характеристики муки оказывают глютенины, которые в зависимости от молекулярной массы делятся на две группы: высокомолекулярные HMW-GS-белки (80-130 кДа) и низкомолекулярные LMW-GS-белки (30–50 кДа) (Gianibelli et al., 2001). HMW-GS представляют собой полиморфное семейство, кодируемое тремя локусами (*Glu-A1*, *Glu-B1* и *Glu-D1*) (Bietz et al., 1975, Payne et al., 1982, Branlard et al., 1985), каждый из которых включает в себя два ортологичных гена, кодирующих субъединицу *x*-типа с большей молекулярной массой и субъединицу *y*-типа с меньшей молекулярной массой (Nucia et al., 2019). Разные аллели HMW-GS влияют на хлебопекарные качества пшеницы по-разному (Payne et al., 1987). Аллели глютенинов с большим молекулярным весом, таких как *Glu-A1(1)/(2\*)*, *Glu-B1(7+8)/(7+9)/(17+18)*, считаются наиболее ценными. В связи с тем, что повышается интерес исследователей к высокомолекулярным глютенинам из-за их полиморфизма, многие работы направлены на изучение влияния глютенинов на качество хлеба с целью исключения возможной отрицательной зависимости между урожайностью и процентным содержанием белка в зерне (Filho, 1997).

Гены глютенинов, как и многие другие гены в геноме злаковых, содержат повторяющиеся элементы разной природы. Поэтому установление полной нуклеотидной последовательности таких генов долгое время оставалось трудоёмкой задачей. Разработка платформ для секвенирования длинными ридами совершила революцию в области сборки геномных последовательностей. Использование данного подхода позволило существенно улучшить сборку геномов злаковых, включая пшеницу, ячмень и рожь (Zimin et al., 2017, IWGSC et al., 2018, Li et al., 2021). Это открыло широкие возможности для поиска новых генов и их аллелей среди разных сортов. С помощью секвенирования длинными ридами были отсеквенированы ДНК разных тканей пшеницы и изучена сравнительная организация кластеров генов глиадинов, точная сборка которых практически невозможна с помощью коротких ридов (Aury et al., 2022).

Данные нанопорового секвенирования ДНК могут быть использованы для установления паттернов метилирования цитозина в геноме на совершенно новом уровне чувствительности. Данный метод позволяет изучить метилирование цитозина (meC) полноразмерных генов глютенинов с использованием недавно опубликованного алгоритма DeepSignal-plant ([https://github.com/PengNi/plant\\_5mC\\_analysis](https://github.com/PengNi/plant_5mC_analysis)), добавляющий тег MM к каждому уникальному выравниванию в bam-файле. Появление методов РНК-секвенирования стало новым этапом в изучении развивающейся зерновки (Rangan et al., 2017; Pellny et al., 2012; Pfeiffer et al., 2014; Yu et al., 2016; Guan et al., 2019). Использование транскриптомных данных с помощью секвенаторов второго поколения (секвенирование короткими ридами) показало, что транскрипционная программа значительно меняется на каждом этапе развития зерновки. Анализ транскриптома 35 генотипов пшеницы на двух этапах развития зерновки (14 dpa (days post anthesis) и 30 dpa) показал, что на 14 dpa наблюдалась активация генов, связанных с синтезом основных компонентов зерна, включая запасные белки и крахмал. Экспрессия генов снижалась к 30 dpa, но увеличивалось разнообразие генов, связанных с защитой зерновки от стрессовых факторов. Это свидетельствует о том, что гены, влияющие на хлебопекарные качества, активны на ранних стадиях развития зерновки (Rangan et al., 2017).

Секвенирование РНК с помощью технологий третьего поколения (секвенирование длинными ридами) позволяет детектировать транскрипцию в локусах генома, которые трудно поддаются анализу с помощью секвенирования короткими ридами. Применение технологии PacBio SMRT позволило получить информацию о транскриптом развивающейся зерновки

пшеницы и использовать полученные данные для улучшения аннотирования генома пшеницы (Dong et al., 2015). Это и другие (Liang et al., 2023) исследования продемонстрировали перспективность использования секвенирования третьего поколения для поиска новых генов растений, которые не были аннотированы ранее.

#### **Цель исследования:**

Аннотация, характеристика и структурно-функциональный анализ генов, экспрессирующихся в процессе развития зерновки тритикале и пшеницы с помощью технологии нанопорового секвенирования (Oxford Nanopore Technologies) РНК/кДНК.

#### **Задачи исследования:**

1. Адаптировать и сравнить методы целевого высокопроизводительного нанопорового секвенирования полноразмерных последовательностей генов развития зерновки для анализа вариаций полноразмерных генов HMW-GS и их промоторов в коллекции.
2. Изучить особенности транскриптома на разных стадиях развития и в разных частях зерновки с помощью нанопорового секвенирования кДНК.
3. Идентифицировать ранее неаннотированные белок-кодирующие гены и гены длинных некодирующих РНК пшеницы, экспрессирующиеся на разных стадиях развития зерновки.
4. Изучить полиморфизм ранее неаннотированных генов в коллекции яровой тритикале.

**Научная новизна полученных результатов.** Впервые проведен комплексный транскриптомный анализ с использованием секвенирования длинными ридами кДНК и РНК на ранней (10 dpa), средней (15 dpa) и поздней (20 dpa) стадиях развития зерновки тритикале. Впервые идентифицированы 7128 ранее неаннотированных генов, экспрессирующихся на разных стадиях развития зерновки тритикале.

Продемонстрировано, что Cas9-опосредованное нанопоровое секвенирование (nCATS) можно использовать в качестве потенциального инструмента для таргетного секвенирования целевых генов яровой тритикале. Определены достоинства и недостатки этого подхода. Полученная глубина секвенирования была достаточной для различных задач, включая обнаружение InDels и SNPs, гаплотипирование и профилирование метилирования. Впервые установлен профиль распределения метилирования цитозина по всей длине генов глютенинов, включая промоторную область, и показано, что данные гены имеют существенное метилирование кодирующей области.

Был впервые оптимизирован метод нанопорового секвенирования ампликонов (ONT Amplicon-Seq) для выявления аллельных вариаций в шести генах HMW-GS. Использование ONT Amplicon-Seq позволило выявить SNPs, InDels, новые аллельные варианты, а также различия как в кодирующей области генов глютенинов, так и в промоторах, которые невозможно выполнить с помощью электрофореза SDS-PAGE. Данный метод, в отличие от pCATS, позволяет проводить анализ вариабельности целевых генов у большого числа образцов, но не подходит для установления профилей метилирования целевых генов.

#### **Теоретическое и практическое значение полученных результатов.**

По результатам проведённых исследований значительно расширен список известных генов пшеницы и тритикале, экспрессирующихся на разных стадиях развития зерновки. Идентифицированные ранее неаннотированные гены используются для последующего функционального анализа, что позволит глубже понять молекулярные механизмы, вовлеченные в развитие зерновки. Идентифицированные гены могут быть использованы для поиска мишеней для маркер-опосредованной селекции и ускоренного отбора ценных генотипов с улучшенными хлебопекарными качествами. Для решения задач генотипирования в работе были оптимизированы и использованы новые методы секвенирования целых генов, а именно Cas9-опосредованное нанопоровое секвенирование и метод ONT Amplicon-Seq. Результаты, полученные в ходе данного исследования, показывают высокий потенциал применения технологии Oxford Nanopore для анализа генов HMW-GS и оценки их профиля метилирования. Основным преимуществом ONT Amplicon-Seq является его применимость в практической селекции для идентификации известных и выявления новых аллелей целевых генов в обширном генетическом разнообразии тритикале и пшеницы, а также других видов с большими и сложными геномами.

**Методология и методы исследования.** Работа выполнена с использованием современного оборудования и применением соответствующих мировому уровню исследований методов молекулярной биологии и биоинформатики. Ключевым инструментом для идентификации ранее неаннотированных генов, новых аллелей, структурных вариаций и эпигенетических характеристик является высокопроизводительная технология Oxford Nanopore.

### **Основные положения, выносимые на защиту:**

1. С помощью Cas9-опосредованного целевого нанопорового секвенирования (nCATS) выявлены структурные вариации (InDels и SNPs) полноразмерных генов высокомолекулярных глютеинов *Glu-1Ax*, *Glu-1Bx* и *Glu-1By* гексаплоидной тритикале и показан высокий уровень метилирования кодирующей области гена по сравнению с промоторной.

2. Метод нанопорового секвенирования ампликонов (ONT Amplicon-Seq) позволяет быстро и эффективно производить генотипирование полноразмерных генов в коллекциях зерновых.

3. Нанопоровое секвенирование транскриптома развивающейся зерновки тритикале идентифицировало 7128 (17,6%) ранее неаннотированных генов, экспрессирующихся на разных стадиях развития зерновки.

4. Значительная часть (более 15% или 4697) транскриптов, экспрессирующихся в процессе развития зерновки тритикале, относится к длинным некодирующим РНК, гены большинства из них (67,5% или 3174) не были аннотированы ранее.

5. Анализ полиморфизма генов длинных некодирующих выявил наличие различных аллелей среди 23 сортообразцов яровой тритикале.

**Личный вклад соискателя.** Соискателем изучены научные труды, посвященные теме исследования, разработаны и проведены экспериментальные работы, освоены методы анализа, проведена интерпретация полученных результатов и подготовка текстов и иллюстраций для публикаций и диссертационной работы.

**Степень достоверности.** Достоверность научной работы основана на обширном объеме экспериментальных исследований, проведенных на сертифицированном оборудовании с применением методов, соответствующих современным стандартам и общемировым требованиям. Эти исследования отличаются высоким уровнем чувствительности и объективности. Также значительное внимание уделено использованию программного обеспечения, необходимого для проведения биоинформатического и статистического анализа экспериментальных данных. В работе продемонстрирована воспроизводимость результатов в различных условиях экспериментов. Выводы, представленные в диссертации, тщательно обоснованы теоретически и подкреплены экспериментальными данными, что гарантирует их соответствие целям и задачам исследования.

**Апробация работы.** Экспериментальные данные, полученные в ходе выполнения диссертационных исследований, были доложены и обсуждены на международных и российских конференциях и конкурсах в том числе: 20-я

Всероссийская конференция молодых учёных, посвященная памяти академика РАСХН Г.С. Муромцева «Биотехнология в растениеводстве, животноводстве и сельскохозяйственной микробиологии» (Москва, 2020); Международная научно-практическая конференция «Научное обеспечение технологического развития и повышения конкурентоспособности в пищевой и перерабатывающей промышленности» (Москва, 2020); конкурс работ молодых ученых ГБС РАН, посвященный году науки и технологий в России (I место, Москва, 2021); VII Всероссийский конкурс научно-исследовательских работ студентов и аспирантов «Наука будущего – наука молодых» (Новосибирск, 2022); I Международная молодежная конференция «Генетические и радиационные технологии в сельском хозяйстве» (Обнинск, 2022); Международная конференция «Генетика, биотехнология, селекция, семеноводство, технологии выращивания и переработки тритикале», посвященная 110-летию со дня рождения основателя ВНИИСБ академика Н.В. Турбина» (Москва, 2022); V Вавиловская международная конференция «Тритикале – культура будущего» к 85-летию У.К. Куркиева (Санкт-Петербург, 2022); II Международная молодежная конференция «Генетические и радиационные технологии в сельском хозяйстве» (Обнинск, 2023); XXIII-я научная конференция молодых ученых «ТРИТИКАЛЕ: генетика, биотехнология, селекция и семеноводство, технологии выращивания и переработки» (Москва, 2023); VII Международная научная конференция «Генетика, геномика, биоинформатика и биотехнология растений» (PLANTGEN 2023) (Казань, 2023); VIII Съезд Вавиловского общества генетиков и селекционеров, посвященный 300-летию российской науки и высшей школы (ВОГиС, Саратов, 2024).

**Финансовая поддержка.** Исследования выполнены при частичной финансовой поддержке Министерства образования и науки Российской Федерации (№ 0431-2019-0005, № FGUM-2022-0005).

**Публикации.** По материалам диссертации опубликовано 14 печатных работ, в том числе 3 статьи в журналах рекомендованных ВАК при Министерстве науки и высшего образования Российской Федерации, из них 3 статьи в изданиях первого квартиля, индексируемых в международных базах WoS, Scopus, PubMed и имеется один патент на изобретение.

**Соответствие диссертационной работы паспорту научной специальности.** Диссертационная работа соответствует паспорту научной специальности 1.5.7. Генетика (биологические науки) (пункты: 2. Геномы, их структура и функция; 7. Структурная и функциональная геномика. Эволюционная геномика; 8. Эпигенетика: эпигеном/ эпипротеом/

эпитранскриптом; 10. Молекулярно-генетические механизмы основных биологических процессов; 16. Генетическая/молекулярно-генетическая биоинформатика и методы многомерного анализа; 17. Частная генетика микроорганизмов, растений и животных; 25. Прикладные аспекты генетики), охватывающей проблемы оценки полиморфизма генов глютенинов яровой тритикале, изучения изменения их статуса метилирования и его связи с экспрессией, идентификации ранее неаннотированных генов, экспрессирующихся в процессе развития зерновки.

**Объем и структура диссертации.** Диссертационная работа изложена на 145 страницах, состоит из введения, 3 разделов, заключения, выводов, списка сокращений, списка использованной литературы и 3 приложений. Текст содержит 11 таблиц, 30 рисунков. Список использованной литературы включает 271 источник, в том числе 254 – на иностранных языках.

## **ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ**

### **ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ**

В обзоре литературы приведены сведения о ботанических особенностях и селекционно-генетических аспектах качества зерна пшеницы и тритикале, обобщены данные о динамике транскриптома развивающейся зерновки и о результатах молекулярно-генетических исследований, а также представлены данные о современных методах изучения генетического разнообразия и геномного анализа.

### **ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ**

**Растительный материал.** Синтез кДНК и поли(А)-РНК для последующего нанопорового секвенирования выполняли из РНК, выделенной из развивающихся зерновок яровой тритикале линии Л8665. Для оценки полиморфизма генов были использованы проростки яровой тритикале, из которых 8 образцов зарубежной и 15 образцов отечественной селекции.

Для разработки метода одновременного секвенирования полноразмерных ампликонов глютенинов с помощью технологии The Oxford Nanopore было использовано 23 сорта яровой мягкой пшеницы. Девятнадцать образцов гибридного происхождения были получены из СИММУТ, а четыре образца были разного географического происхождения.

**Подготовка ДНК и РНК.** Выделение ДНК проводили с использованием модифицированного ЦТАБ-метода. Выделение РНК проводили методом гуанидин изотиоцианата с водонасыщенным фенолом. Синтез двуцепочечной кДНК (дц-кДНК) (Евроген Mint, ЗАО Евроген, Москва, Россия) и поли(А)-РНК (Poly(A)Purist Mag Kit, Thermo Fisher Scientific, Waltham, Massachusetts, USA) проводили в соответствии с инструкциями производителя. Во время

синтеза дц-кДНК было подобрано оптимальное количество циклов ПЦР (22 цикла) для достижения экспоненциальной фазы амплификации.

Концентрацию и качество выделенной ДНК и РНК оценивали с помощью спектрофотометра NanoDrop One UV-Vis (Thermo Fisher Scientific, Waltham, MA, USA) и Quantus Fluorometer (Promega, Madison, WA, USA), а также с помощью гель-электрофореза в агарозном геле.

**Обогащение геномной ДНК высокомолекулярными фрагментами.** Для секвенирования nCATS генов глютелинов была выделена высокомолекулярная геномная ДНК из гексаплоидной яровой тритикале линии Л8665. Обогащение геномной ДНК высокомолекулярными фрагментами проводили с помощью набора Short Read Eliminator XL Kit (Circulomics, Baltimore, MD, США).

**Дизайн олигонуклеотидов.** Все олигонуклеотиды, используемые в данном исследовании, были разработаны с помощью программного обеспечения Primer3.0 (<https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) и проверены PrimerBLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

**Подготовка библиотеки для секвенирования nCATS.** гРНК для nCATS получали путем *in vitro* транскрипции с ДНК-матриц, содержащих промотор Т7. После синтеза всех гРНК была произведена подготовка ДНК-библиотеки, включающей в себя сборку РНП-комплекса, дефосфорилирование геномной ДНК, ее расщепление путем добавления РНП-комплекса, лигирование адаптеров и очистку на магнитных шариках.

**Подготовка библиотеки для секвенирования ампликонов (ONT Amplicon-Seq).** Получение ампликонов для ONT Amplicon-Seq проводилось с использованием ДНК-полимеразы Encyclo (Evrogen, Москва, Россия) и полимеразы Biolabmix LR-HS (Biolabmix, Новосибирск, Россия) в соответствии с инструкциями производителя. В процессе амплификации использовались нефосфорилированные праймеры. Фосфорилирование ампликонов было проведено с использованием полинуклеотидкиназы Т4 (PNK). Полученные ампликоны были пулированы и баркодированы с последующим секвенированием на ячейке SQK-LSK109. Бейсколлинг был выполнен с помощью Guppy v6.3.8 (Oxford Nanopore Technologies, Oxford, United Kingdom).

**Подготовка библиотеки для секвенирования кДНК.** Синтезированную дц-кДНК очищали с помощью 1,8× шариков Agencourt AMPure XP (Beckman Coulter, Pasadena, CA, USA) в соответствии с инструкциями производителя. Каждый образец подвергался баркодированию

с использованием набора Native Barcoding Kit (Oxford Nanopore Technologies, Oxford, UK).

**Секвенирование библиотек.** Секвенирование проводили с помощью проточной ячейки MinION R9.4.1. Процесс секвенирования контролировали программным обеспечением MinKnow (Oxford Nanopore Technologies, Oxford, UK). Бейсколлинг был выполнен с помощью Guppy (Oxford Nanopore Technologies, Oxford, UK).

**Анализ данных секвенирования nCATS.** Полученные риды выравнивали на полноразмерные последовательности ВАС с помощью программы minimap2 (Li, 2018). Для обнаружения SNP использовали программу Nanocaller (Ahsan et al., 2021). Для реконструкции последовательностей аллелей *Glu-1B* применяли команду bcftools (Danecek et al., 2017). Ранжирование ридов по гаплотипам было выполнено с помощью WhatsHap (Patterson et al., 2015). Идентификация метилирования с использованием необработанных данных ридов с помощью нанопора была выполнена с помощью программного обеспечения DeepSignal-plant (Ni et al., 2021).

**Анализ данных секвенирования ONT Amplicon-Seq.** Полученные риды после ONT Amplicon-seq были выровнены на референсные последовательности генов HMW-GS из NCBI с использованием minimap2 (Li, 2018). Полученные файлы BAM были визуализированы с помощью JBrowse2.

**Анализ данных секвенирования кДНК.** Для анализа данных, полученных после секвенирования дц-кДНК и поли(А)-РНК, объединялись последовательности генома *Triticum aestivum* с последовательностями генома *Secale cereale* в один fasta-файл. Сборка и оценка количества расшифровок на основе количества выровненных ридов были выполнены с использованием программы StringTie2. Данные фильтровали по уровню экспрессии (TPM > 5), а затем пересекали (bedtools intersect и StringTie2 merge) координаты экспрессирующихся локусов с аннотацией пшеницы IWGSC54 (ABD) (IWGSC, 2018) и ржи Lo7 (R) (Rabanus-Wallace et al., 2021), чтобы идентифицировать аннотированные и неаннотированные транскрипты.

### ГЛАВА 3. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

#### 1. Cas9-опосредованное нанопоровое секвенирование (nCATS) и определение ДНК метилирования генов глютенинов

Гены глютенинов на >80% состоят из коротких и разнородных повторяющихся мотивов, что делает их секвенирование короткими ридами крайне проблематичным. Поэтому для секвенирования нами был апробирован метод nCATS. Для применения этого подхода были выбраны

последовательности генов глютеинов *Glu-1Ax*, *Glu-1Bx* и *Glu-1By*. Для nCATS были подобраны 12 гРНК на промоторную и кодирующую области генов. Сайт посадки гРНК на кодирующую область находился за пределами повторяющегося региона. Ожидаемая длина последовательности для целевых областей составляла 3,4 тыс.п.н., 5,1 тыс.п.н. и 3,6 тыс.п.н. для *Glu-1Ax*, *Glu-1Bx* и *Glu-1By*, соответственно.

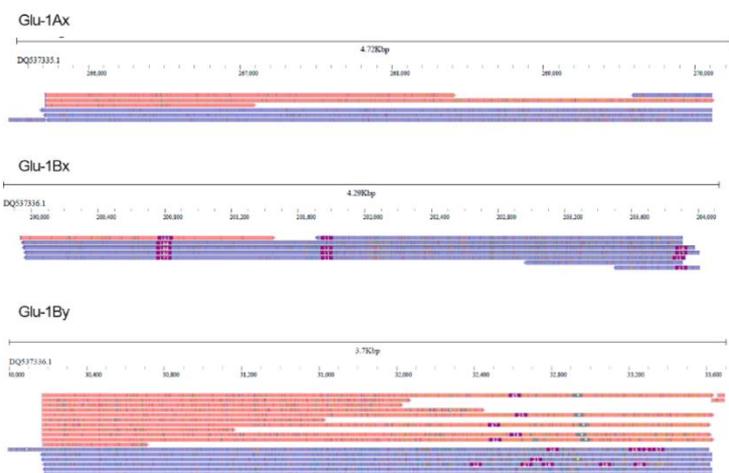


Рисунок 1 – Выравнивание ридов nCATS на референсные последовательности генов *Glu-1Ax*, *Glu-1Bx*, *Glu-1By*

путем поиска сходства прочитанных последовательностей с референсными генами глютеинов и обнаружено 7, 8 и 17 ридов для локусов *Glu-1Ax*, *Glu-1Bx* и *Glu-1By* (рисунок 1), соответственно (около 0,03% всех ридов были целевыми). Несмотря на общее низкое количество целевых ридов, расчетная степень обогащения для трех целевых генов колебалась от ~200× до ~645×.

Большая часть ридов охватывала всю целевую область, предоставляя полезные данные для выявления структурных вариаций генов глютеинов тритикале линии Л8665. В результате были обнаружены две инсерции в гене

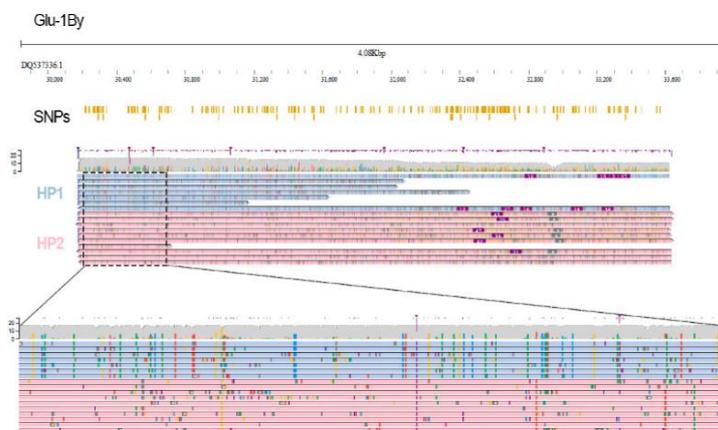


Рисунок 2- Гаплотипирование гена *Glu-1By* на основе SNP

На основании двух запусков секвенирования методом nCATS на проточной ячейке MinION, используя смесь из 12 гРНК, было получено 120681 высококачественных ридов (Qscore > 8, N50 = 3,1 Kb) при общей длине ридов около ~547 Mb, что соответствует 0,02× покрытию генома тритикале. Было подсчитано количество целевых ридов

*Glu-1Bx*: одна длиной ~180 п.н. в промоторной области и вторая - 12 п.н. в кодирующей области. Результаты ПЦР и секвенирования по Сэнгеру подтвердили наличие вставок в вариантах *Glu-1Bx*.

Далее была проведена идентификация SNP для гена *Glu-1By*, имеющего наибольшую глубину

секвенирования (~15×). В общей сложности выявлено 222 SNPs, различающих вариант гена *Glu-1By* яровой тритикале линии Л8665 от аналогичного гена у мягкой пшеницы сорта Chinese Spring. Примененный метод WhatsHap позволил провести гаплотипирование на основе идентифицированных SNP, выявив, что аллели гаплотипов HP1 и HP2 (рисунок 2) группируются с известными генами *Glu-1By* пшеницы. Однако выявленные аллели не полностью соответствуют известным, что, вероятно, указывает на новый вариант гена *Glu-1By*.

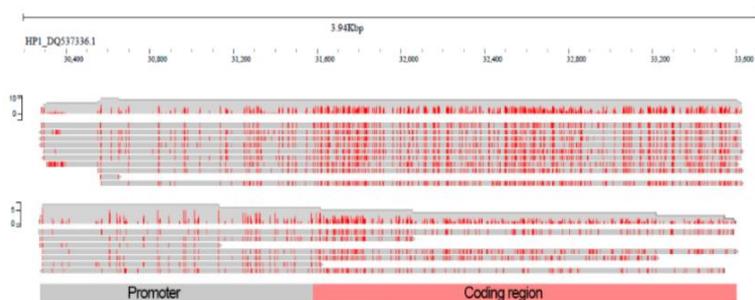


Рисунок 3 - Профиль метилирования гена *Glu-1By*, красным отмечены участки метилирования

Метод nCATS позволил впервые получить информацию о метилировании ДНК полноразмерных генов глютеинов. Профиль метилирования для трех генов был схожим: низко метилированная промоторная область и сильно метилированная кодирующая последовательность (рисунок 3). Повышенное метилирование кодирующей последовательности ("gene body methylation" (gbM)) характерно для экспрессирующихся консервативных генов. Однако гены глютеинов не соответствуют этим характеристикам, поскольку они имеют специфические для эндосперма паттерны экспрессии и демонстрируют высокую вариабельность, что подтверждается SNP-анализом аллелей гена *Glu-1By*. Тем не менее, анализ профиля метилирования ДНК генов с высокой экспрессией в эндосперме риса (Zemach et al., 2010) показал, что эти гены часто имеют gbM, что коррелирует с полученными результатами для тритикале.

## 2. Разработка метода нанопорового секвенирования полноразмерных генов глютеинов.

Одновременное секвенирование генов глютеинов и их промоторов интересно тем, что полученные данные могут быть использованы для изучения вариабельности генов и их регуляторных последовательностей. Используя три последовательности генов глютеинов, было продемонстрировано, что nCATS является подходящим инструментом для анализа вариабельности целевых генов только одного сорта.

Альтернативным методом секвенирования является подход, основанный на ПЦР-амплификации. Этот метод становится особенно привлекательным при оценке вариаций, как в кодирующей части генов глютеинов, так и в промоторных областях в коллекциях растений. В данном исследовании были

разработаны праймеры для амплификации шести субъединиц глютенинов (*Ax/y*, *Bx/y*, *Dx/y*). Длина последовательностей составляла от 2,6 Кб для *Glu-1Ax* до 5,0 Кб для *Glu-1Dx*. Для успешного синтеза ПЦР-продукта условия амплификации индивидуально оптимизировали, учитывая сложные структуры генов глютенинов. Результатом амплификации стали 138 ПЦР-продуктов (6 генов глютенинов для 23 образцов мягкой пшеницы) (рисунок 4). Фосфорилированные ПЦР-ампликоны были баркодированы и объединены в одну пробирку для нанопорового секвенирования. Анализ ридов показал, что глубина покрытия 50× – 150× была получена уже через один час секвенирования. В общей сложности было получено ~215000 ридов со значением N50 около 3,2 тыс.п.н.

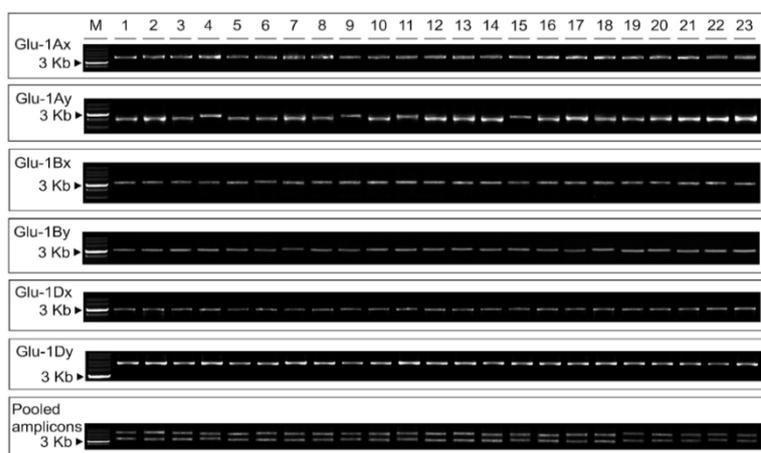


Рисунок 4 - Результаты амплификации генов HMW-GS (*Glu-1Ax*, *Glu-1Ay*, *Glu-1Bx*, *Glu-1By*, *Glu-1Dx*, *Glu-1Dy*) в коллекции мягкой пшеницы. Представленные сорта: MC№23, MC№35, MC№59, MC №65, MC№66, MC№67, MC№70, MC№79, MC№147, MC№150, MC№151, MC№152, MC№153, MC№178, MC№187, MC№215, MC№217, MC№220, MC№223, Glenlea, Bombona, Agatha, Kanyuk. ДНК-маркер Sky-High (Biolabmix, Novosibirsk, Russia) (M).

Для идентификации аллельных вариантов каждого из шести генов полученные риды были картированы на референсные последовательности генов глютенинов. Затем были идентифицированы SNPs и InDels, а также присвоены соответствующие идентификаторы аллелей HMW-GS, если они были ранее известны и депонированы в GenBank. Анализ вариации *Glu-1Ax* в 23 образцах пшеницы выявил аллели *Glu-Ax2\** (14 образцов) и *Ax-null* (2

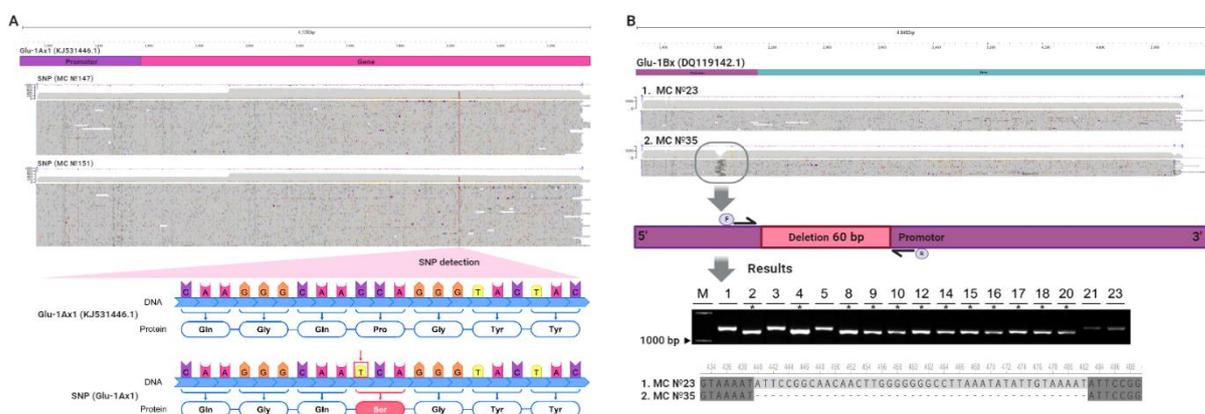


Рисунок 5 – Идентификация структурных вариаций в генах глютенинов. (А) Выявление SNP в кодирующей части гена *Glu-1Ax1* по сравнению с референсной последовательностью. (В) Расположение праймеров, разработанных для анализа делеции промотора *Glu-1Bx7* в коллекции мягкой пшеницы. Представленные сорта: 2. MC № 35, 4. MC № 65, 8. MC № 79, 9. MC № 147, 10. MC № 150, 12. MC № 152, 14. MC № 178, 15. MC № 187, 16. MC № 215, 17. MC № 217, 18. MC № 220, 20. MC № 223 (с делецией); 1. MC № 23, 3. MC № 59, 5. MC № 66, 21. Glenlea, 23. Agatha (без делеции). ДНК-маркер Step50 plus (Biolabmix, Novosibirsk, Russia) (M).

образца). Из пяти образцов, идентифицированных как аллель *Glu-Ax1*, два (МС №147, МС №151) имели SNP (3136С -> Т), приводящий к несинонимичной замене аминокислоты пролин (Pro) на серин (Ser) (рисунок 5А). Новый аллель *Glu-1Ax* был назван *Glu-1Ax1-T*.

Ген *Glu-1Ay* обычно неактивен у культивируемых гексаплоидных сортов пшеницы, что делает невозможным его анализ с помощью SDS-PAGE. Однако ONT Amplicon-Seq позволил проанализировать последовательности этой субъединицы и выявить аллель *Glu-1Ay-d* у 16 образцов, остальные семь образцов несли аллель *Glu-A1-2*, классифицированный как неактивный.

Сравнение последовательностей *Glu-1Bx* с ранее известными выявило разнообразие аллелей: *Bx7* (73%), *Bx17* (9%), *Bx13* (9%) и *Bx14* (9%). Используя данные ONT Amplicon-Seq, была обнаружена делеция ~60 п.н. и SNP (С -> G), расположенные в промоторной области аллеля *Bx7*. Результаты ПЦР и секвенирования по Сэнгеру также подтвердили наличие делеции в аллеле *Glu-1Bx7* (рисунок 5В).

Анализ последовательностей *Glu-1B*, полученных с помощью ONT Amplicon-Seq, показал, что образцы пшеницы МС №151 и МС №67 содержат аллели, сходные с известными аллелями *By16* и *By19\**, соответственно. Эти два аллеля различались на 27 п.н. в кодирующей области гена. Анализ остальных 21 образцов пшеницы показал, что 8 (35%) образцов имели аллель *Glu-B1y8*, 9 образцов (40%) имели аллели *Glu-By9*, а *Glu-By15* (9%) и *Glu-By8* (9%) — по два образца.

Для генов *Glu-1D* были выявлены следующие распределения по аллелям: 8 образцов (35%) имели аллели *Glu-Dx2* и *Glu-Dy12* и 15 образцов (65%) имели аллели *Glu-Dx5* и *Glu-Dy10*.

### **3. Поиск ранее неаннотированных генов зерновки тритикале на разных стадиях развития с помощью нанопорового секвенирования.**

Целью данной части диссертационной работы была идентификация ранее неаннотированных генов, экспрессирующихся в процессе развития зерновки в разных ее частях. Для этого было выполнено ONT-секвенирование кДНК, полученной из РНК, выделенной из двух частей развивающейся зерновки: зародыш + эндосперм (ЕЕ) и эндосперм (Ен). Были проанализированы данные секвенирования РНК, выделенной на ранней стадии (10 dpa), а также секвенирования кДНК зерновок на 15 dpa (ЕЕ\_15 и Ен\_15) и 20 dpa (ЕЕ\_20 и Ен\_20).

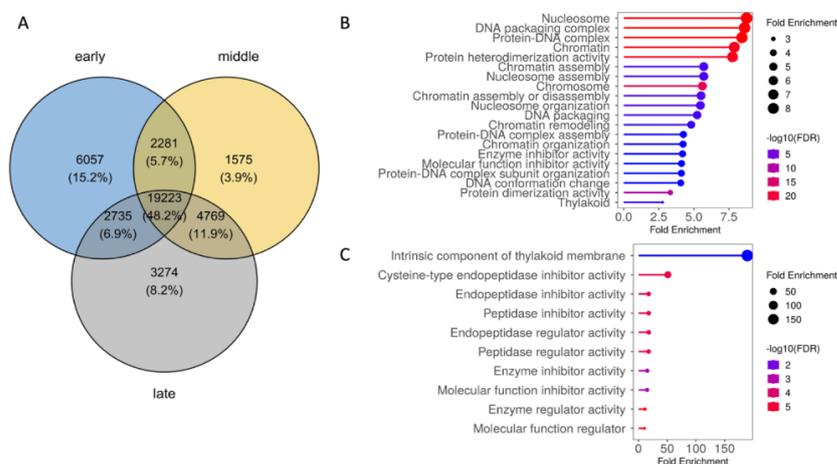
В соответствии с последней аннотацией геномов А, В (IWGS54) и R (Lo7) были определены экспрессирующиеся гены, расположенные в межгеномном пространстве. В общей сложности, используя кДНК и РНК риды

нанопорового секвенирования, было идентифицировано 7128 (17.6%) генов как ранее неаннотированные ('новые'), из которых 2340, 2947 и 1841 генов были локализованы в геномах А, В, и R, соответственно.

Среди ранее неаннотированных генов более 40% (3 174) представлены днРНК, в то время как среди аннотированных их доля не превышает 5%. Доля днРНК среди транскриптов «новых» генов была в 10 раз выше (критерий  $\chi^2$  Пирсона с поправкой Йетса,  $p\text{-value} < 2,2 \times 10^{-16}$ ), чем в наборе аннотированных генов. Это свидетельствует о том, что днРНК плохо аннотированы из-за высокого содержания повторяющихся элементов. А следовательно, данный эксперимент подчеркивает эффективность нанопорового секвенирования для поиска любых экспрессирующихся генов, как белок кодирующих, так и не кодирующих генов.

Был проведен анализ экспрессии генов на ранней, средней и поздней стадиях развития зерновки. Учитывая значение экспрессии  $\text{TPM} > 2$ , был получен набор, состоящий из 6057, 1575 и 3274 генов, специфически экспрессирующихся на ранней, средней и поздней стадиях, соответственно (рисунок 6А). Из всех 39914 выявленных генов 19223 генов экспрессировались на всех стадиях развития зерновки. Доля днРНК среди аннотированных транскриптов составляла 4,6% (879) на ранней, 42,7% (1756) на средней и 31,6% (1739) на поздней стадиях развития зерновки.

Рисунок 6 – Характеристики транскриптов, экспрессируемых на ранней, средней и поздней стадиях развития зерновок. (А) Диаграмма Венна генов, экспрессируемых ( $\text{TPM} > 2$ ) на определенных стадиях развития зерновок тритикале: ранней (10 день после цветения), средней (15 день после цветения) и поздней (20 день после цветения). (В) и (С) Анализ обогащения аннотированных генов, экспрессируемых на ранней и средней стадиях развития семян, соответственно.



Среди генов, специфически экспрессирующихся на ранней стадии, в значительной мере ( $p\text{-value} < 0.0001$ ) представлены гены, участвующие в организации, упаковке и ремоделировании хроматина (рисунок 6В). Это отражает интенсивную транскрипцию в процессе дифференцировки клеток на ранней стадии развития зерновок, когда происходит дифференцировка зародыша. Анализ обогащения генов, экспрессирующихся на средних и поздних стадиях развития зерновок, показал сходные закономерности: наиболее значительное обогащение было обнаружено для генов, связанных с

ингибиторами пептидаз и катаболизмом глюкозы. Кроме того, было обнаружено, что гены, кодирующие белки, расположенные в апопласте (поздняя стадия), и белки, являющиеся компонентами тилакоидной мембраны (средняя стадия), чрезмерно представлены на средней и поздней стадиях развития (рисунок 6С).

На следующем этапе исследования была задача идентифицировать различия между наборами генов, экспрессирующихся в зародыше и эндосперме зерновки. Для этого было выполнено ONT-секвенирование кДНК, полученной из РНК, выделенной из двух частей зерновки на разных этапах развития (15 dpa и 20 dpa): зародыш + эндосперм (EE\_15 и EE\_20) и эндосперм (En\_15 и En\_20).

Используя ограничение значения уровня экспрессии  $TPM > 2$ , было определено, что 11250 генов экспрессировались во частях зерновки. Кроме того, 1604, 2238 и 903 гена экспрессировались в EE\_15, EE\_20 и в обоих образцах, соответственно. Эти гены не имели достаточной экспрессии в образцах эндосперма (En), а значит они могут составлять транскриптом зародыша на 15 и 20 день после цветения. Анализ обогащения EE\_15 выявил наборы генов, кодирующие белки комплекса оксидоредуктазы и комплекса мембранных белков, которые вовлечены в процессы фосфорорганического метаболизма. Наиболее значимо в наборе генов EE\_20 были представлены гены, кодирующие белки рибонуклеопротеинового комплекса, комплекса убиквитинлигазы, и вовлеченные в сайленсинг генов, биогенез рибосом, процессинг РНК и трансляцию (коэффициент обогащения  $FDR < 0,003$ ).

Хорошо известно, что эндосперм представляет собой гипометилированную ткань, а метилирование ДНК важно для накопления запасных белков и биогенеза эндосперма (Zemach et al., 2010). Однако пониженное метилирование ДНК способствует активации ТЕ, создавая опасную среду для зародыша, поскольку активация ТЕs может привести к летальным мутациям и снижению жизнеспособности зерновок. Ранее был идентифицирован LTR-ретротранспозон (TE) MIG Ty1/Copia, экспрессирующийся на ранней стадии развития зерновок, хотя, в какой именно ткани экспрессируется MIG, было неизвестно (Kirov et al., 2020). Используя профилирование транскриптома для зародыша и эндосперма, было показано, что MIG экспрессируется, с высокой долей вероятности, в эндосперме или одновременно в тканях эндосперма и зародыша. В зародышевой части на поздней стадии развития зерновок выявлена сниженная представленность ТЕ-РНК по сравнению со средней стадией. В то же время в наборе генов, экспрессирующихся на этой стадии, представлены гены

транскрипционного сайленсинга. Эти результаты могут свидетельствовать о том, что экспрессия ТЕ на средней стадии развития зерновок запускает активацию посттранскрипционного сайленсинга генов на более поздних стадиях развития зерновок. Последнее может приводить к метилированию ТЕ и защите зародыша от инсерций ТЕ. Таким образом, наши данные согласуются с гипотезой о том, что гипометилирование ДНК эндосперма играет защитную роль для зародыша (Zemach et al., 2010, Hsieh et al., 2009).

Анализ транскриптома развивающихся зерновок яровой тритикале на ранней и средней/поздней стадиях сильно отличаются с точки зрения экспрессируемых генов и их функций. Выявленные днРНК, экспрессирующихся на этих стадиях, представляют собой интересную мишень для селекции с помощью маркеров. Анализ полиморфизма показал, что наиболее дивергентным вариантом в коллекции 23 сортообразцов яровой тритикале была линия С95 (рисунок 7) ввиду того, что при амплификации семи пар праймеров отклонение от референсной ДНК-матрицы (Chinese Spring) наблюдалось в пяти случаях (1А59, 6А12, 3В75, 3В82, 1А2902).

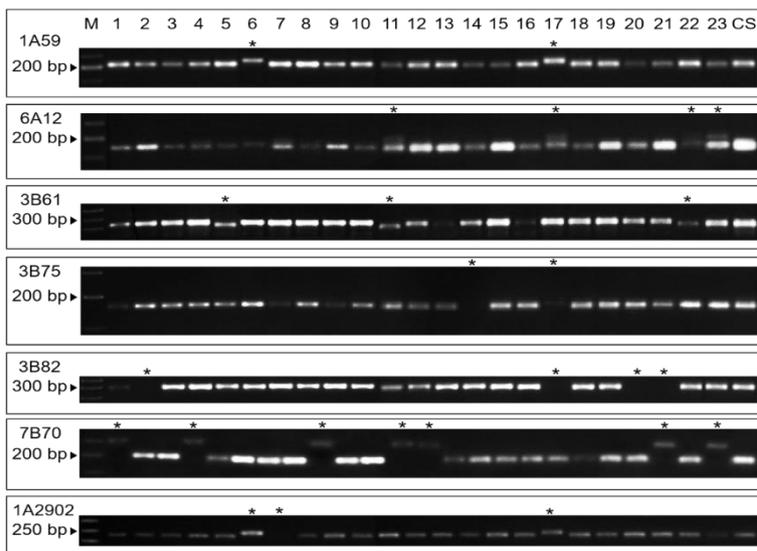


Рисунок 7 – Результаты анализа полиморфизма генов lncRNA в коллекции яровой тритикале: Dublet (1); Legalo (2); Sandro (3); Гребешок (4); Лана (5); Памяти Мережко (6); Укро (7); Ульяна (8); Хлебодар харьковский (9); Ярило (10); 131/1656 (11); 131/7 (12); 6-35-5 (13); C238 (14); C245 (15); C259 (16); C95 (17); V17/50 (18); Л8665 (19); P2-16-20 (20); P 13-5-2 (21); P13-5-13 (22); PRAG 551 (23); Chinese Spring (CS). ДНК-маркер Step 50 plus (Biolabmix, Новосибирск, Россия) (M). Нереперентные аллели обозначены звездочкой (\*).

## ЗАКЛЮЧЕНИЕ

1. Использование метода nCATS позволяет получить значительное обогащение (до 200×) ридами для целевых генов (*Glu-1Ax*, *Glu-1Bx* и *Glu-1By*) и их промоторов у гексаплоидной тритикале, предоставляя уникальную информацию для детекции структурных вариаций и профиля метилирования, но данный метод не подходит для высокопоточного секвенирования генов у видов с большими геномами.

2. Профилирование метилирования ДНК показало, что гены глютенинов имеют значительно более высокую плотность метилирования цитозинов в кодирующей области по сравнению с промоторным регионом.

3. С помощью ONT Amplicon-Seq охарактеризован полиморфизм и определены аллели шести полноразмерных генов глютенинов (*Glu-1Ax*, *Glu-1Ay*, *Glu-1Bx*, *Glu-1By*, *Glu-1Dx*, *Glu-1Dy*) у 23 сортов пшеницы, идентифицирован новый аллель (*Glu-A1x1-T*) гена *Glu-1Ax* и показано, что ONT Amplicon-Seq является точным и быстрым методом генотипирования полноразмерных генов у растений с большими и сложными геномами.

4. С помощью нанопорового секвенирования кДНК было выявлено 39914 генов, которые экспрессируются на ранней (6057), средней (1575), поздней (3274) стадиях развития зерновки, и 19223 гена, которые экспрессируются на всех стадиях развития зерновки, а также 1604 экспрессировались в образце с зародышевой частью зерновки на 15 дpa, 2238 – в образце с зародышевой частью зерновки на 20 дpa.

5. Нанопоровое секвенирование кДНК позволило идентифицировать 7128 генов как ранее неаннотированных, локализованных в А- (2340), В-(2947) и R- (1847) субгеномах, из которых 3174 (более 40%) кодируют днРНК, что свидетельствует о их значимой роли в процессе развития зерновки.

6. Гены, кодирующие длинные некодирующие РНК, полиморфны в коллекции яровой тритикале, и линия С95 является наиболее дивергентной.

## СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ, ВХОДЯЩИХ В ПЕРЕЧЕНЬ ВАК МОН РФ

1. Polkhovskaya E. et al. Nanopore Amplicon Sequencing Allows Rapid Identification of Glutenin Allelic Variants in a Wheat Collection //Agronomy.– 2023.–Т. 14. – №. 1. – С. 13. doi.org/10.3390/agronomy14010013 (Q1, WoS, Scopus)

2. Polkhovskaya E. et al. Long-Read cDNA Sequencing Revealed Novel Expressed Genes and Dynamic Transcriptome Landscape of Triticale (x Triticosecale Wittmack) Seed at Different Developing Stages //Agronomy. – 2023. – Т. 13. – №. 2. – С. 292. doi.org/10.3390/agronomy13020292 (Q1, WoS, Scopus)

3. Kirov I., Polkhovskaya E. et al. Searching for a needle in a haystack: Cas9-targeted Nanopore sequencing and DNA methylation profiling of full-length glutenin genes in a big cereal genome //Plants. – 2021. – Т. 11. – №. 1. – С. 5. doi.org/10.3390/plants11010005 (Q1, WoS, Scopus, PubMed)

**Авторское свидетельство:**

1. Патент №2785924 С1 от 15.12.2022 г. Способы одновременного нанопорового секвенирования полных последовательностей значимых генов твердой пшеницы // Патентообладатель ФГБНУ "Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии" / Авторы: Киров И.В., Меркулов П.Ю., Полховская Е.С., Власова А.В., Дудников М.В., Соловьев А.А., Карлов Г.И., Дивашук М.Г.

**Тезисы докладов и другие публикации в зарубежных изданиях, входящих в международные реферативные базы данных и системы цитирования Scopus/WoS, и в прочих изданиях:**

1. Пырников А., Сыксин С., Милюкова Н., Груздев И., Полховская Е., Гарибян Ц., Соловьев А. (2022) Применение нового молекулярного маркера и метода SDS-PAGE для установления взаимосвязи аллельного состава генов глютенинов с качественными показателями зерна тритикале. Федеральный научный центр кормопроизводства и агроэкологии имени В.Р. Вильямса. Кормопроизводство, (9), 27-33.
2. Щуклина О., Соловьев А., Полховская Е., Квитко В., Клименкова И., Завгородний С. (2021). Тимирязевская 42-новый сорт яровой тритикале (x Triticosecale Wittm. Ex. Camus). Кормопроизводство, (8), 43.
3. Энзекрей Е. (Полховская Е.), Щуклина О., Завгородний С. Влияние метеорологических условий и азотных удобрений на биологическую урожайность яровой тритикале сорта Тимирязевская 42 //Зерновое хозяйство России. – 2021. – №. 2. – С. 88-93.
4. Polkhovskaya E. et al. Identification of non-annotated genes involved in the grain development //Plant Genetics, Genomics, Bioinformatics, and Biotechnology. – 2023. – p. 313 – тезисы.
5. Полховская Е. С. и др. Транскриптомный анализ развивающейся зерновки яровой тритикале //Генетические и радиационные технологии в сельском хозяйстве. – 2023. – С. 72-74 – тезисы.
6. Полховская Е. и др. Идентификация неаннотированных генов развития зерновки с помощью нанопорового секвенирования //Генетические и радиационные технологии в сельском хозяйстве. – 2022. – С. 35-37– тезисы.
7. Болотина А., Полховская Е. и др. Поиск новых генов, экспрессирующихся в процессе развития зерновки, с помощью нанопорового секвенирования //Генетические и радиационные технологии в сельском хозяйстве. – 2022. – С. 16-17 – тезисы.

8. Болотина А., Полховская Е. и др. Идентификация и анализ неаннотированных генов в развивающейся зерновке тритикале //Биотехнология в растениеводстве, животноводстве и сельскохозяйственной микробиологии. – 2022. – Т. 7. – С. 20 – тезисы.
9. Полховская Е. и др. Анализ генов глютеинов яровой тритикале с помощью нанопорового секвенирования //Биотехнология в растениеводстве, животноводстве и сельскохозяйственной микробиологии. – 2021. – С. 64-65 – тезисы.
10. Pirsikov A., Enzekrey, E. (Polkhovskaya E.), Siksina S., Milyukova N. (2021). Development of molecular markers for the identification of prolamins genes and their correlation with baking qualities of grain. In Plant Genetics, Genomics, Bioinformatics, and Biotechnology (p. 173) – тезисы.
11. Болотина А., Полховская Е., Киров И. Эпигенетическая регуляция генов запасных белков яровой тритикале //Биотехнология в растениеводстве, животноводстве и сельскохозяйственной микробиологии. – 2021. – С. 72-72 – тезисы.