АННОТАЦИЯ

диссертации на соискание степени доктора философии (Ph.D.) по специальности «8D05105 – Биотехнология»

Досжановой Ботакоз Нурмаганбетовной

на тему: **«Полногеномный анализ ассоциаций хозяйственно-ценных признаков сои (*Glycine max* (L.) Merr.) для практической селекции»**

Актуальность исследования. Соя (*Glycine max* (L.) Merr.) — одна из наиболее важных зернобобовых культур во всем мире, служащая основным источником белка и масла для продуктов питания, кормов и промышленного применения. В Казахстане, в связи с экономической и практической ценностью данной культуры, ежегодно осваиваются новые посевные площади. Идентификация генетических факторов хозяйственно-ценных признаков сои открывает новые возможности для ускоренной селекции и создания новых высокопродуктивных сортов. Современные подходы биотехнологии с использованием молекулярно-генетических маркеров могут ускорить развитие данного направления селекции. Мировой опыт внедрения результатов таких исследований существенно ускоряет разработку новых сортов сельскохозяйственных культур с желаемыми признаками и экономит финансовые и временные ресурсы. В связи с этим, изучение генетические ресурсов сои казахстанских и зарубежных сортов и линий с использоваием современных генетических методов имеет важное значение как для фундаментальных исследований, так и для практического значения в молекулярной селекции сои.

Объект исследования. Объектом исследования послужила коллекция, состоящая из 252 сортов и линий сои (*Glycine max* (L.) Merr.) различного происхождения, включая 31 сортов и линии казахстанской селекци.

Предмет исследования. Предметом исследования явился полногеномный анализ ассоциаций (ПГАА) коллекции сои, выращенной в условиях Алматинской области, для идентификации маркер-признаковых ассоциаций (MTA, marker-trait associationы) и ДНК-маркеров типа KASP (Kompetitive allele-specific PCR), связанных с признаками, определяющих продуктивность и качество сои.

Цель исследования. Идентифицировать маркер-признаковых ассоциаций признаков, определяющих продуктивности и качество сои (*Glycine max* (L.) Merr.) на основе изучения мировой коллекции различного происхождения и метода ассоциативного картирования генов.

Задачи исследования:

1. Изучить вариабельность хозяйственно-ценных признаков, определяющих продуктивность и качество сои на основе мировой коллекции сои различного географического происхождения в условиях Алматинской области (КазНИЗиР, Алмалыбак, Казахстан);
2. Отобрать перспективные сорта и линии сои на основе изучаемых хозяйственно-ценных признаков для улучшения отечественной селекции сои;

3. Изучить генетическое разнообразие и охарактеризовать генетическую структуру коллекции сои, включая образцы Казахстана, Восточной и Западной Европы, Восточной Азии и Северной Америки с использованием SNP (полиморфизм по единичному нуклеотиду)-маркеров на основе данных ресеквенирования генома сои (WGRS);

4. Идентифицировать маркер-признаковые ассоциаций, определяющих продуктивность и качество семян сои на основе использования технологии ПГАА;

5. Провести анализ гаплотипов значимых ассоциаций и разработать новые ДНК-маркеры для конкурентной аллель-специфической ПЦР (*kompetitve allele-specific PCR*, KASP), связанных с хозяйственно-ценными признаками сои.

Методы исследования. В настоящем исследовании проводились полевые эксперименты, биохимические, молекулярно-генетические и статистические методы изучения. Полевые эксперименты проводились в Казахском Научно-Исследовательском Институте Земледелия и Растениеводства (КазНИИЗиР). Определение качества зерна проводилось согласно ГОСТ 32749-2014 в лаборатории биотехнологии, биохимии и качества зерна КазНИИЗиР с использованием анализатора зерна NIRS DS2500 (FOSS, Хиллерод, Дания). Молекулярно-генетические методы исследования, в частности процедура выделения ДНК из проростков сои, проверка ее качества и визуализация проводилась с использованием соответствующей методологии и материально-технической базы лаборатории молекулярной генетики Института Биологии и Биотехнологии Растений (ИББР) КН МНВО РК. Работы по ресеквенированию были проведены с использованием системы Illumina HiSeq X Ten на кафедре Школы естественных наук Университета Гуанчжоу, Гунчжоу, Китай. Методы биометрической статистики включали дисперсионный анализ, корреляционный анализ по Пирсону, t-тест Стьюдента и визуализацию результатов с использованием пакетов для RStudio. Также, статистические методы включали модели регрессионного анализа (смешанная модель множественных локусов (*multi locus mixed model*, *MLMM*), с использованием пакета GAPIT для платформы R, кластеризации методом объединением соседей (*neighbor joining, NJ*), метод Байесовской цепи Маркова Монте-Карло (*Markov chain Monte Carlo, MCMC*) в программе STRUCTURE. В ходе выполнения диссертационной работы проводилось активное сотрудничество с коллегами из Японии и Китая. В 2019 году докторантом была осуществлена научная стажировка по теме исследований в Исследовательский Центр Хоккайдо, под руководством зарубежного научного консультанта.

Научная новизна исследования. На основе использования SNP-маркеров проведено сравнение генетического профиля сортов и линий сои Казахстана с зарубежными образцами, изучена структура популяции мировой коллекции сои. Научная новизна исследования заключается в идентификации 46 стабильных MTA, связанных с хозяйственно-ценными признаками сои, включая 25 предположительно новых MTA, на основе использования

современных молекулярно-генетических методов. Среди стабильных МТА были обнаружены 6 плейотропных МТА, ассоциированных с несколькими изученными признаками. Наибольшее количество МТА было обнаружено для компонентов урожайности (20), далее следовали признаки содержания протеина и жира (18), размеров семян (5), и признака растрескиваемости стручков сои (3). Полученные результаты свидетельствуют о важности генетико-селекционных исследований в определенных условиях выращивания, где должны быть учтены особенности взаимодействия генотипа (*G, genotype*) и окружающей среды (*E, environment*). Идентифицированные ассоциации имеют потенциал для дальнейшего изучения и детальной идентификации генов-кандидатов, контролирующих хозяйственно-ценные признаки сои. Полученные результаты обеспечивают теоретическую основу для будущих стратегий селекции, направленных на создание улучшенных сортов сои.

Практическая значимость исследования. На основе идентифицированных значимых и стабильных ассоциаций были выявлены эффективные гаплотипы, связанные с важными хозяйственно-ценными признаками. Идентифицированные ассоциации являются основой для создания новых информативных KASP-маркеров, ценных инструментов для селекции сои, направленных на улучшение качества семян и общей продуктивности сельскохозяйственных культур. Был разработан и зарегистрирован патент на полезную модель «Способ идентификации селекционно-ценных линий сои (*Glycine max* (L.) Merr.) с использованием набора KASP-маркеров» (№9289), позволяющий эффективно отобрать ценные линии для дальнейшей селекции. Полученные результаты могут быть внедрены в селекционные программы с целью улучшения показателей желаемых признаков сои в Казахстане.

Теоретическая значимость. Одной из главных гипотез настоящей работы явилось изучения возможности идентификации новых и подтвержденных МТА сои хозяйственно-ценных признаков сои в условиях выращивания мировой коллекции в Алматинской области. Идентифицированные МТА имеют потенциал для дальнейшего изучения и возможной идентификации генов-кандидатов. Полученные результаты могут быть эффективно использованы в генетико-селекционных исследованиях сои, с учетом специфики взаимодействия генотипа и окружающей среды, и могут служить теоретической основой для будущих стратегий селекции, направленных на создание улучшенных сортов сои.

Основные положения диссертации, выносимые на защиту:

1. Исследуемая коллекция, состоящая из 252 сортов и линий сои Казахстана, Восточной и Западной Европы, Восточной Азии и Северной Америки, обладает высоким уровнем фенотипического разнообразия в условиях Алматинской области. По совокупности десяти изученных хозяйственно-ценных признаков сои коллекция представляет собой эффективный генетический ресурс для проведения ПГАА с целью выявления

новых генетических факторов, влияющих на вариабельность хозяйственно-ценных признаков сои.

2. В результате фенотипической оценки коллекции сои по 10 хозяйственно-ценным признакам, были выделены 9 отечественных и 42 зарубежных генотипов с высокими показателями продуктивности и качества сои, превосходящие показатели сорта-стандарта по Алматинской области. Данные генотипы являются ценным ресурсом для дальнейшего использования в селекционных программах по созданию улучшенных сортов культуры.

3. На основе данных полногеномного ресеквенирования генетическая структура мировой коллекции сои подразделяется на три основные субпопуляции, разделенных по географическому происхождению и характеризуется высоким уровнем генетического разнообразия. Выявленная генетическая структура и степень разнообразия популяции соответствуют ключевым требованиям, предъявляемым к объектам для проведения ПГАА.

4. Идентифицированные в результате ПГАА 106 значимых ассоциаций признаков, определяющих продуктивность и качество семян сои, включая 46 стабильных ассоциаций, представляют собой важную информацию для поиска генетических факторов, контролирующих сложные количественные признаки сельскохозяйственной культуры. Из 46 стабильных ассоциаций, ПГАА позволил выявить 25 предположительно новых MTA для изученных хозяйственно-ценных признаков. При этом, 21 MTA были ранее идентифицированы в других исследованиях, что подтверждает надежность полученных результатов в данной работе.

5. На основе значимых ассоциаций были выявлены эффективные гаплотипы для признаков, определяющих продуктивность и качество сои. Данные гаплотипы могут быть применимы на ранних стадиях селекционных программ по созданию отечественных сортов сои высокого урожая и качества. Также, наиболее значимые ассоциации, связаные с несколькими признаками, определяющих урожайность сои, были конвертированы в KASP маркеры *ipbb_gm_141, ipbb_sd_16* и *ipbb_sd_17* показавшие высокую эффективность для применения их в слекционных программах.

Связь с планом основных научных работ. Диссертационная работа была выполнена в рамках проектов грантового финансирования МНВО РК AP05131592 «Полногеномное исследование ассоциаций устойчивости к грибковым болезням сои в Казахстане» (2018-2020 гг.), AP13068118 «Поиск генов, связанных с хозяйственно-ценными признаками сои (*Glycine max* (L.) Merr.), на основе использования ресеквенирования полного генома мировой коллекции» (2022-2024 гг.) и научно-технической программы BR24992903 «Практическое внедрение современных молекулярно-генетических, физиолого-биохимических, биотехнологических методов и цифрового фенотипирования в селекцию экономически значимых сельскохозяйственных культур» (2024-2026 гг.).

Личный вклад докторанта в подготовку каждой публикации включает сбор данных по теме исследования, выполнении теоритических и экспериментальных исследований, а также подготовку статей к публикации.

Аппробация научно-исследовательской работы. Результаты диссертационной работы доложены и опубликованы на международных научно-практических конференциях: Международная научная конференция студентов и молодых ученых «Фараби әлемі» (Алматы, Казахстан, 2019,2020), Международная научная конференция молодых ученых, приуроченная в честь 40-летия Института Молекулярной биологии и биохимии (Алматы, Казахстан, 2023), Международная конференция биологии и биотехнологии растений (Алматы, Казахстан, 2024), 4-ая Школа молодых ученых «Генетика, геномика, биоинформатика и биотехнология растений» PlantGen School 2024 (Новосибирск, Россия, 2024). Основные результаты диссертационной работы ежегодно заслушивались на научно-техническом совете факультета «Биологии и биотехнологии», на заседаниях кафедры «Биотехнология» КазНУ имени аль-Фараби, на научных семинарах лаборатории молекулярной генетики, на международных конференциях молодых ученых и на Ученом Совете РГП на ПХВ «Института биологии и биотехнологии растений». Результаты диссертационной работы частично включены в отчеты о научно-исследовательской работе по проектам AP05131592, AP13068118 и BR24992903.

Публикации. Основное содержание диссертации опубликовано в 12 печатных работах, в том числе 5 статей в международных рецензируемых журналах с импакт-фактором, индексируемых в SCOPUS и/или Web of Science, 1 патента на полезную модель от РГП на ПХВ «Национальный Институт Интеллектуальной Собственности» МЮ РК, 5 тезисов в материалах отечественных Международных конференций, 1 постер в зарубежной международной конференции.

Выводы. В настоящем исследовании была выполнена комплексная работа по идентификации ассоциаций маркер-признак, связанных с важными хозяйственно-ценными признаками сои, включающие компоненты продуктивности, признаки качества и морфологические особенности сои (*Glycine max* (L.) Merr.). На основании полученных результатов, были сделаны следующие выводы:

1) Исследуемая коллекция, состоящая из 252 сортов и линий сои Казахстана, Восточной и Западной Европы, Восточной Азии и Северной Америки, обладает высоким уровнем фенотипического разнообразия в условиях Алматинской области. По совокупности десяти изученных хозяйственно-ценных признаков сои коллекция представляет собой эффективный генетический ресурс для проведения полногеномного анализа ассоциаций с целью выявления новых генетических факторов, влияющих на вариабельность хозяйственно-ценных признаков сои. Анализ ANOVA показал, что среди всех изученных признаков, наибольшее влияние G оказывалось на признаки качества семян ($P<2x10^{-16}$). Среда в большей степени обуславливала компоненты урожайности ($P<2x10^{-16}$). Взаимодействие G × E значимо для большинства признаков, что подтверждает необходимость тестирования сортов в различных агроэкологических зонах перед их внедрением в производство ($P<0.01$). Анализ корреляции выявил обратную

значимую корреляцию между признаком содержания протеина и жира, а также с компонентами урожайности, что имеет важное значение для селекционных программ ($P<0.05$). Также размеры семян, как длина и толщина, имели положительную взаимосвязь с содержанием жира и урожайностью ($P<0.05$).

2) Фенотипическая оценка коллекции сои по 10 хозяйственно-ценным признакам выделила 9 отечественных (сорта Даная и Память ЮГК, линии 186/1, 126/1, 173/1, 261/1, 391, 209/1, 350/1) и 42 зарубежных генотипов с высокими показателями продуктивности и качества сои, превосходящие показатели сорта-стандарта по Алматинской области. Данные генотипы являются ценным ресурсом для дальнейшего использования в селекционных программах по созданию улучшенных сортов культуры.

3) Генотипирование коллекции 252 образцов сои с использованием данных ресеквенирования 4 923 660 SNP маркеров позволило выявить 44 385 SNP-маркеров для проведения дальнейшего изучения генетического разнообразия и проведения ПГАА в высоком разрешении. Мировая коллекци коллекция состоит из основных трех субпопуляций, разделенных по географическому происхождению. Образцы из Казахстана (31 образец) были представлены во всех трех субпопуляциях, что свидетельствует о широком использовании гермоплазмы из других регионов мира в селекции сои в Казахстане.

4) ПГАА выявил 106 значимых ассоциаций, среди которых 46 стабильных и превосходящих статистический порог. Из 46 стабильных ассоциаций были обнаружены 25 потенциальны новых МТА, которые ранее не были описаны в литературе. При этом, 21 МТА были ранее идентифицированы в других исследованиях, что подтверждает надежность полученных результатов в данной работе. Среди 25 потенциально новых МТА, идентифицированы 15 ассоциаций для компонентов урожайности, 5 ассоциаций для качества, 3 для размеров семян, 2 ассоциации для признака растрескиваемости стручков.

5) На основе значимых ассоциаций выявлены эффективные гаплотипы для признаков, определяющих продуктивность и качество сои ($P<1x10^{-4}$). Данные гаплотипы могут быть применимы на ранних стадиях селекционных программ по созданию отечественных сортов сои высокого урожая и качества. Наиболее значимые ассоциации, связаные с несколькими признаками, компонентами урожайности сои, конвертированы в KASP маркеры - *ipbb_gm_141, ipbb_sd_16* и *ipbb_sd_17,* показавшие высокую эффективность для применения их в слекционных программах ($P<0,05$).