Казахский Национальный Университет имени аль-Фараби

УДК 601.4 (043.3)

На правах рукописи

ДОСЖАНОВА БОТАКОЗ НУРМАГАНБЕТОВНА

Полногеномный анализ ассоциаций хозяйственно-ценных признаков сои (*Glycine max* (L.) Merr.) для практической селекции

8D05105 – Биотехнология

Диссертация на соискание степени доктора философии (Ph.D)

Научный консультант:
Туруспеков Е.К., к.б.н., профессор,
зав.лаб. молекулярной генетики ИББР

Зарубежный научный консультант:
Сузуки Т., Ph.D Научной Организации
Хоккайдо, Япония

Республика Казахстан
Алматы, 2025

СОДЕРЖАНИЕ

НОРМАТИВНЫЕ ССЫЛКИ…….....….....4
ОПРЕДЕЛЕНИЯ…….....5
ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ……7
ВВЕДЕНИЕ……..……..8
ОБЗОР ЛИТЕРАТУРЫ.....13
1.1. Соя (*Glycine max* (L.) Merr.) – важная зернобобовая культура…….....13
1.1.1. Происхождение, использование и распространение сои……14
1.1.2. Ботаническое и морфологическое описание сои……...…...15
1.1.3. Строение и химический состав семян сои……16
1.1.4. Производство сои в мире и в Казахстане…….……19
1.2. Хозяйственно-ценные признаки важные для производства сои….....21
1.2.1. Характеристика основных признаков, определяющих продуктивность сои……...22
1.2.2. Основные признаки качества семян сои……....24
1.3. Геном сои…….....25
1.3.1. Структура генома сои……..25
1.3.2. Молекулярная генетика сои……....28
1.3.2.1. Классическое картирование локусов количественных признаков.....30
1.3.2.2. Полногеномный анализ ассоциаций……...32
1.4. Селекция сои в мире и в Казахстане……..……....35
1.4.1. Классические методы селекции……35
1.4.2. Современная молекулярная селекция для создания высокопродуктивных сортов сои.....36
1.4.3. Применение технологии KASP в современной селекции...……38
2. МАТЕРИАЛЫ И МЕТОДЫ……40
2.1. Материалы исследований……...40
2.2. Методы исследований ……41
2.2.1. Полевые работы, погодные условия, оценка хозяйственно-ценных признаков в изучаемой коллекции сои ……41
2.2.1.1. Оценка компонентов урожайности ……42
2.2.1.2. Измерение параметров семян……..……43
2.2.1.3. Оценка растрескиваемости стручков сои……..……43
2.2.1.4. Определение признаков качества семян сои……..….44
2.2.2. Выделение ДНК и полногеномное ресеквенирование образцов.....…...44
2.2.3. Анализ структуры популяции мировой коллекции сои……...…45
2.2.4. Полногеномный анализ ассоциаций хозяйственно-ценных признаков.....46
2.2.5. Анализ гаплотипов значимых ассоциаций, разработка KASP маркеров и KASP-генотипирование ……..……46
2.2.6. Статистическая обработка результатов …….……...…47

3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ……49
3.1. Описание хозяйственно-ценных признаков изучаемой коллекции сои в условиях Алматинской области…….……..49
3.1.1. Компоненты урожайности …….……...…..49
3.1.1.1. Результаты ANOVA анализа для компонентов урожайности.……51
3.1.2. Параметры размеров семян сои – длина, ширина и толщина……..……52
3.1.2.1. Результаты ANOVA анализа для размеров семян 54
3.1.3. Растрескиваемость стручков сои ……...55
3.1.3.1. Результаты ANOVA анализа растрескиваемости стручков сои……56
3.1.4. Описание признаков качества семян сои.....57
3.1.4.1. Результаты ANOVA анализа признаков качества семян сои……59
3.2. Перспективные сорта и линии по изучаемым признакам в мировой коллекции сои…….. ……..……60
3.3. Анализ корреляции хозяйственно-ценных признаков сои……..62
3.4. Ресеквенирование данных мировой коллекции сои ……....64
3.4.1. Анализ плотности покрытия SNP маркеров генома сои……..65
3.4.2. Оценка LD SNP маркеров в геноме изучаемой популяции сои……..…66
3.4.3. Анализ генетической структуры изучаемой популяции сои…..……68
3.5. Полногеномный анализ ассоциаций хозяйственно-ценных признаков…..70
3.5.1. Полногеномный анализ ассоциаций компонентов урожайности сои.....70
3.5.2. Полногеномный анализ ассоциаций размеров семян сои…….....73
3.5.3. Полногеномный анализ ассоциаций признака растрескиваемости сои.....74
3.5.4. Полногеномный анализ ассоциаций признаков качества семян сои……..75
3.6. Аннотация генов-кандидатов хозяйственно-ценных признаков сои……..77
3.6.1. Аннотация генов-кандидатов компонентов урожайности сои……77
3.6.2. Аннотация генов-кандидатов размеров семян сои……...80
3.6.3. Аннотация генов-кандидатов растрескиваемости стручков сои ……81
3.6.4. Аннотация генов-кандидатов признаков качества семян сои……..81
3.7. Анализ гаплотипов значимых ассоциаций хозяйственно-ценных признаков сои ……...84
3.8. Изучение эффективности идентифицированных KASP-маркеров на примере признаков урожайности сои……..88
ЗАКЛЮЧЕНИЕ……...92
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ…...…….....94
ПРИЛОЖЕНИЕ 1…….....112
ПРИЛОЖЕНИЕ 2.....121
ПРИЛОЖЕНИЕ 3……130
ПРИЛОЖНИЕ 4 ……..132

НОРМАТИВНЫЕ ССЫЛКИ

В настоящей диссертации использованы ссылки на следующие стандарты: Закон Республики Казахстан «О науке» от 18.02.2011 г. №407-IVЗРК; ГОСО РК 5.04.034-2011. Государственный общеобразовательный стандарт образования Республики Казахстан. Послевузовское образование. Докторантура. Основные положения (изменения от 23 августа 2012 г. №1080); ГОСТ 7.1-2003. Библиографическая запись. Библиографическое описание. Общие требования и правила составления.

ОПРЕДЕЛЕНИЯ

В настоящей диссертации применяются следующие термины с соответствующими определениями:

Аллель	– различные формы одного и того же гена, расположенные в одинаковых участках (локусах) гомологичных хромосом.
Ген	– участок молекулы ДНК, задающий последовательность мономеров в полипептиде либо в функциональной РНК, детерминирующих определенный признак.
Генетическая карта	– относительное расположение генетических маркеров, генов и локусов на хромосоме.
Генотип	– совокупность аллелей гена или группы генов, контролирующих какой-либо признак.
ДНК маркер	– ген или последовательность ДНК с известным местоположением на хромосоме, используемый для идентификации и/или генетического картирования.
Конкурентная аллель-специфичная ПЦР	– метод аллель-специфической ПЦР, основанной на конкурентной гибридизации праймеров для генотипирования SNP маркеров с использованием флюоресценции продукта полимеразной цепной реакции.
Локус	– местоположение определённого гена на генетической карте хромосомы.
Локусы количественных признаков	– являются участками ДНК, либо содержащими гены, либо сцепленными с генами, которые отвечают за тот или иной количественный признак (quantitative trait loci, QTL).
Маркер-признаковая ассоциация	– связь между маркеров ДНК и количественным признаком, определяемая с помощью статистических моделей ассоциативного картирования.
Неравновесие по сцеплению	– неслучайное сцепление аллелей в разных локусах.
Полногеномный анализ ассоциаций	– метод полногеномного анализ ассоциаций, позволяющий выявить связь между геномными вариантами и фенотипическими проявлениями (GWAS, Genome Wide Association Study).
Поправка Бонферрони	– метод статистической коррекции, который используется при множественном тестировании, чтобы снизить вероятность ложноположительных результатов
Плейотропия	– способность одного гена воздействовать на несколько признаков.
Полиморфизм	– разнообразие популяций по признакам или маркерам генетической природы.

Продуктивность	– это совокупный показатель, отражающий способность растения формировать урожай в определённых условиях окружающей среды, агротехнологий и генетических особенностей.
Сорт	– это совокупность растений, созданная в результате селекционной работы и обладающая ценными биологическими и хозяйственными свойствами, передающимися по наследству.
Среда	– совокупность всех внешних факторов, влияющих на фенотипические признаки.
Структура популяции	– любая закономерность или структуризация генетических вариаций в изучаемой популяции.
Фенотип	– это совокупность всех признаков и свойств организма, которые выявляются в процессе индивидуального развития в конкретных условиях и являются результатом взаимодействия генотипа с комплексом факторов внутренней и внешней среды.

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

ГОСТ	– государственный стандарт
ДНК	– дезоксирибонуклеиновая кислота
КазНИИЗиР	– Казахский Научно-Исследовательский Институт Земледелия и Растениеводства
ПГАА	– Полногеномный анализ ассоциаций (*genome wide association study, GWAS*)
ДС	– длина семян
КПУ	– количество продуктивных узлов
КСР	– количество семян на растение
МСР	– масса семян с растения
МТС	– масса 1000 семян
РС	– растрескиваемость стручков
СП	– содержание сырого протеина
СМ	– содержание сырого масла
ТС	– толщина семян
ШС	– ширина семян
ANOVA	– дисперсионный анализ (*analysis of variance*)
FDR	– доля ложноположительных результатов (*false discovery rate*)
G	– Генотип
E	– Окружающая среда
GxE	– взаимодействие Генотип х Среда
KASP	– конкурентная аллель-специфическая полимеразная цепная реакция (*Kompetitive Allele Specific PCR*)
LD	– сцепление по неравновесию (*linkage disequilibrium*)
MAF	– частота встречаемости менее распространенного аллеля (*minor allele frequency*)
MAS	– маркер-опосредованная селекция (*marker-assisted selection*)
MCMC	– метод статистического моделирования, который позволяет оценивать вероятности и параметры сложных распределений (*Markov Chain Monte Carlo*)
MLMM	– множественная локусная смешанная модель (*multi-locus mixed model*)
MTA	– ассоциация маркер-признак (*marker-trait association*)
PVE	– доля фенотипической изменчивости (*phenotypic variance explained*)
QTL	– локусы количественных признаков (*quantitative trait loci*)
RIL	– рекомбинантные инбредные линии (*recombinant inbred lines*)
SNP	– полиморфизм по единичному нуклеотиду (*Single Nucleotide Polymorphism*)
WGRS	– данные полногеномного ресеквенирования (*whole-genome resequencing*)

ВЕДЕНИЕ

Общая характеристика диссертационного исследования. Работа посвящена идентификации ассоциаций между SNP-маркерами и хозяйственно-ценными признаками сои (*Glycine max* (L.) Merr.) с использованием методологии полногеномного анализа ассоциаций (ПГАА).

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 [1, 2]. Лидерами по производству сои являются Бразилия, США и Аргентина, в общем составляя более 80% мирового производства [3]. В Казахстане, в связи с экономической и практической ценностью данной культуры, ежегодно осваиваются новые посевные площади [4; 5]. Идентификация генетические факторов хозяйственно-ценных признаков, определяющих продуктивность, качество семян и устойчивость к стрессам, открывает новые возможности для ускоренной селекции и создания новых высокопродуктивных сортов сои [6, 7]. Признаки, определяющие продуктивность сои и качество семян, являются сложными признаками и контролируются двумя и более генами, также изменяются в зависимости от условий окружающей среды и их взаимодействия. Одним из эффективных подходов изучения и идентификации таких генетических факторов является методология ПГАА.

ПГАА позволяет идентифицировать генетические локусы, связанные с ценными для сельского хозяйства признаками, обеспечивая знания генетической основы сложных признаков. Идентифицированные ассоциации маркер-признак, способствуют разработке молекулярных маркеров для выведения высококачественных сортов сои, с желаемыми признаками, тем самым повышая продуктивность и обеспечивая устойчивое развитие сельского хозяйства.

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

Предмет исследования. Предметом исследования явился полногеномный анализ ассоциаций (ПГАА) коллекции сои, выращенной в условиях Алматинской области, для идентификации маркер-признаковых ассоциаций (MTA, marker-trait association) 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 продуктивность 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-маркеров (*single nucleotide polymorphism,* полиморфизм по единичному нуклеотиду) 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 (*kompetitive 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 семян (содержание протеина и масла в семенах сои) 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 МТА, связанных с хозяйственно-ценными признаками сои, включая 25 предположительно новых МТА, 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 масла (18), размеры 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 изучение возможности идентификации новых и подтвержденных МТА сои хозяйственно-ценных признаков сои в условиях выращивания мировой коллекции в Алматинской области. Идентифицированные 46 стабильных МТА, включая 25 предположительно новых МТА, имеют потенциал для дальнейшего изучения и возможной идентификации генов-кандидатов, контролирующих хозяйственно-ценные признаки сои. 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 МТА для изученных хозяйственно-ценных признаков. При этом, 21 МТА 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.

Значимые и стабильные ассоциации, связанные с несколькими признаками, определяющих продуктивность сои, были конвертированы в KASP маркеры. Из них три KASP маркера, *ipbb_gm_141, ipbb_sd_16* и *ipbb_sd_17,* ассоциированные с несколькими признаками, показали высокую эффективность для применения их в селекционных 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.).

Апробация 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 Биотехнологии 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 11 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, 4 тезиса в материалах отечественных Международных конференций, 1 постер в зарубежной международной конференции.

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

Объем и структура диссертации. Диссертация изложена на 111 страницах и состоит из обозначений и сокращений, введения, обзора литературы, материалов и методов, результатов и обсуждений, заключения, списка использованных источников из 223 наименований. Содержит 30 рисунков, 25 таблиц и 4 приложения.

1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Соя (*Glycine max* (L.) Merr.) – важная зернобобовая культура

Соя (*Glycine max* (L.) Merr.) – занимает важное место в сельском хозяйстве во всем мире. Среди сельскохозяйственных культур, соя выделяется высоким содержание протеина (до 42%) и масла (до 18%), что делает ее одним из ключевых компонентов продовольственной безопасности и здоровья человека [9]. Мировая практика признала, что соевый протеин является наиболее высококачественным среди растительного протеина и по своему аминокислотному составу близок к животному [10]. Соевое масло является ведущим растительным маслом на мировых рынках, используемым для приготовления пищи, жарки и обработки пищевых продуктов [11]. Кроме того, соевые бобы содержат пищевые волокна, необходимые минералы и антиоксиданты, что еще больше повышает их роль в поддержании здоровья населения [12]. По мере роста потребительского спроса на более сбалансированное питание, продукты из сои, такие как тофу, соевое молоко и текстурированные растительные белки, получили широкое распространение как в развитых, так и в развивающихся странах [13].

Производство кормов для животных является важнейшим направлением рынка сои. Примерно 70% мирового производства сои перерабатывается в соевый шрот – корм с высоким содержанием протеина для птицеводства, свиноводства и аквакультуры. Соевый шрот превосходит другие растительные источники белка благодаря общему содержанию сырого протеина и высокой перевариваемостью аминокислотного состава. Усвоение белка соевого шрота у птицы составляет около 85%, а отдельных аминокислот варьируется от 82% до 94%. Среди растительных источников белка соевый шрот выделяется как экономически эффективный источник лимитирующих аминокислот, дополняя рационы животных на основе злаков и восполняя дефицит ключевых питательных веществ [14-15].

Производство биодизельного топлива из соевого масла является устойчивой альтернативой ископаемому топливу, соответствующей глобальным целям энергетического перехода. Поскольку страны стремятся достичь углеродно-нейтральной экономики, прогнозируется, что спрос на биотопливо на основе сои будет расти [16]. Также, одной из наиболее значимых преимуществ сои является ее способность фиксировать азот. Соя – это бобовое растение, которое образует симбиотические отношения с азотфиксирующими бактериями (ризобиями), снижая потребность в синтетических азотных удобрениях и повышая плодородие почвы [17]. Чередуя соевые бобы с зерновыми, такими как кукуруза и пшеница, фермеры могут диверсифицировать системы земледелия, улучшить структуру почвы и свести к минимуму накопление вредителей и патогенов [18].

Универсальность сои, ее применение в продуктах питания, кормах для животных и промышленных продуктах, повышают экономическую ценность культуры. В 2022 году мировой рынок сои оценивался примерно в 155 миллиардов долларов США [19]. Крупнейшие производители сои в мире – США, Бразилия и Аргентина – получают значительные валютные поступления

от экспорта сои, что способствует укреплению их торгового баланса и экономической стабильности [20].

1.1.1. Происхождение, использование и распространение сои.

Культурная соя происходит от дикорастущей формы *Glycine soja* (Siebold. & Zucc.) в Восточной Азии около 6000-9000 лет назад. Южный Китай, долина Желтой реки, Центральный Китай, северо-восточный Китай и несколько других регионов (Корея и Япония) были определены как регионы-кандидаты, где соевые бобы могли быть культивированы [21]. Дикие растения были выведены в их современные фенотипы сельскохозяйственных культур на протяжении тысячелетий путем многократного культивирования и селекции человеком. Общие черты культивирования сои включают в себя утрату признаков покоя семян и растрескивания бобов, изменение формы роста стебля от вьющегося к прямостоячему, а также изменение размеров семян: от мелких до более крупных. С момента своего первоначального введения в культуру, соя стала неотъемлемой частью в традиционной кухне местного населения и в системе земледелия [22].

В настоящее время глобальное распространение сои значительно расширилось, сою производят в больших объемах и в Северной, и в Южной Америке. Быстрое расширение выращивания сои объясняется ее экономической рентабельностью, адаптируемостью к различным климатическим условиям и растущим спросом на продукты, полученные из сои [19, стр.1; 20, стр. 2]. Международные торговые сети еще больше способствовали глобальному распространению выращивания сои, особенно в Европе и Африке, где соя все чаще используется в кормах для животных и пищевых продуктах [23].

Распространение сои в Центральной Азии произошло сравнительно недавно. Литературные источники утверждают, что первые посевы сои для своих личных нужд проводились в Семиречье дунганами, переселенцами из Западного Китая в конце XIX века. Однако дальнейшее культивирование в Казахстане началось только в конце XX в. Ранее соя была культивирована в регионах современных стран Армении, Азербайджана, Узбекистана, Турции, Таджикистана [24]. Исторически соя не была основной культурой в данных регионах, но питательные качества и экономические преимущества выращивания сои способствовали ее распространению. Казахстан включил сою в программу диверсификации сельскохозяйственных культур, направленной на снижение зависимости от традиционных культур [4, стр.2; 25]. С начала 2000х Казахстан являлся лидером по производству сои в Центральной Азии [24, стр. 7].

В настоящее время выращивание сои охватывает несколько агроэкологических зон, от умеренных до тропических, благодаря достижениям в селекции растений [26]. США, Бразилия и Аргентина остаются ведущими мировыми экспортерами сои, а Китай является крупнейшим импортером, что обусловлено растущим спросом на сою и продукты ее переработки [23, стр.4].

1.1.2. Ботаническое и морфологическое описание сои

Соя (*Glycine max* (L.) Merr.) – однолетнее бобовое растение, принадлежащее семейству Fabaceae [27].

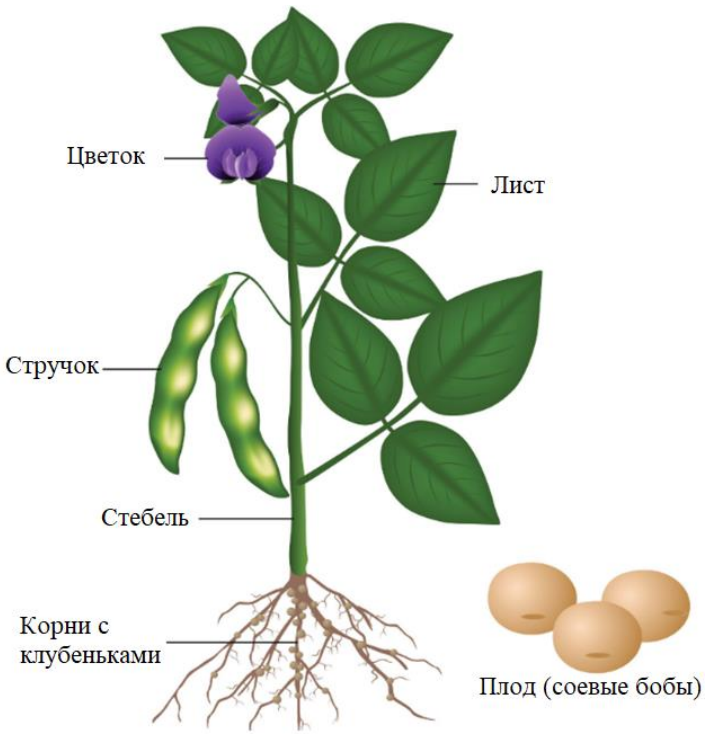


Рисунок 1 – Строение растения сои [28]

Растение сои прямостоячее (реже стелющееся) с опушенным ветвящимся стеблем с узлами, образующим куст высотой от 25 до 150 см [29; Рисунок 1]. Имеет диффузную корневую систему, состоящую из стержневого корня и вторичных корней. На корнях образуются клубеньки с бактериями, фиксирующими атмосферный азот. Листья тройчатые, опушенные овальные, перистосложные с прилистниками. В пазухах листьев расположены соцветия, в которых находятся мелкие цветки, светло-фиолетового, фиолетового или белого цвета. Плод сои представлен в виде опушенного боба, состоящего из двух створок (Рисунок 1). К створкам боба прикреплены от 1 до 4 семян сои [30-31].

Соевые бобы проходят сложный цикл роста и развития, который определяется как экологическими, так и генетическими факторами. Фер и Кэвинесс (1977) разработали систему поэтапного развития сои, получившая широкое признание [32]. Весь цикл роста и развития растения состоит из вегетативной и репродуктивной стадий. В вегетативной стадии роста происходит развитие сои от всходов до развития всех узлов на стебле, также происходит рост и развитие корневой системы. Репродуктивная часть описывает стадии от начала цветения до полного созревания растения сои. Краткое описание репродуктивных стадий представлены в Таблице 1.

Соя культивируется по всему миру, однако это растение короткого дня, чувствительное к фотопериоду [33]. Температура и фотопериод являются двумя важными экологическими факторами, влияющими на вегетативное и репродуктивное развитие [34]. Среди ключевых генетических факторов,

влияющих на развитие и морфологию сои, выделяют гены *Dt* и *E* [35]. Эти гены оказывают воздействие как на вегетативное, так и на репродуктивное развитие, а также на морфологию растения. Гены *Dt* контролируют особенности роста сои и могут влиять на конечное количество узлов главного стебля [36]. Исследования влияния генов *E* на развитие сои после цветения выявили возможные механизмы, которые могут объяснить широкий диапазон географической адаптации растения [35, стр.1031].

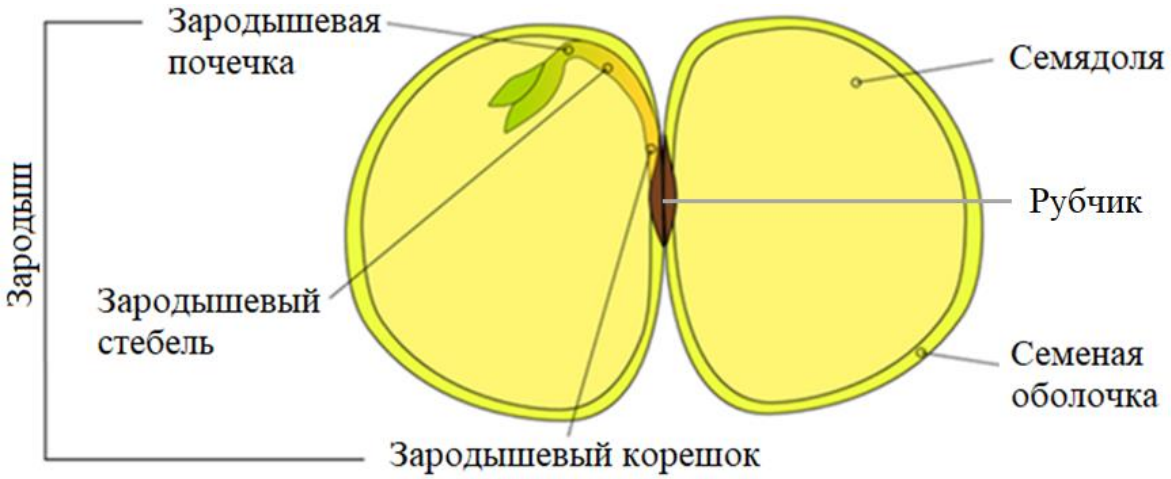
Таблица 1 – Краткое описание репродуктивных стадий развития сои

Стадия	Описание
R1	Начало цветения: один цветок открыт на любом узле главного стебля.
R2	Полное цветение: открыт цветок на одном из двух верхних узлов главного стебля с полностью развитым листом.
R3	Начало формирования бобов: боб длиной около 0,5 см на одном из четырех верхних узлов главного стебля с полностью развитым листом.
R4	Полное формирование бобов: боб длиной около 2 см на одном из четырех верхних узлов главного стебля с полностью развитым листом.
R5	Начало формирования семян: семя длиной около 0,3 см в бобе на одном из четырех верхних узлов главного стебля с полностью развитым листом.
R6	Полное формирование семян: боб содержит зеленое семя, полностью заполняющее полость боба, на одном из четырех верхних узлов главного стебля с полностью развитым листом.
R7	Начало созревания: один нормальный боб на главном стебле достиг зрелого цвета.
R8	Полное созревание: 95% бобов достигли зрелого цвета.

1.1.3. Строение и химический состав семян сои

Структурные и качественные характеристики семян сои имеют решающее значение для их роли в сельскохозяйственном производстве и в продуктах питания. Структура семян напрямую влияет на физиологические процессы, например, прорастание семян сои. Состав же семян сои определяет пищевую, функциональную и экономическую ценность [37].

Рисунок 2 – Строение семени сои [38]



Семя состоит из отдельных слоев, каждый из которых играет определенную роль в развитии семян, прорастании и метаболизме (Рисунок 2). Семенная оболочка образует внешний защитный слой от механических повреждений, патогенов и потери влаги [11, стр.365]. Проницаемость оболочки влияет на способность семян поглощать воду, что имеет решающее значение для прорастания. Оболочка семян, состоящая из фенольных соединений, танинов и флавоноидов, влияет на твердость семян. Твердость семян является ключевой характеристикой, действующую на обработку семян и эффективность измельчения. С другой стороны, более твердая оболочка семени повышает устойчивость к вредителям [40-41].

Семядоли являются наиболее важным компонентом семян сои, являясь органом хранения протеинов, жиров и углеводов. Эти запасы обеспечивают энергию и питательные вещества, необходимые для прорастания и роста семени [11, стр.365]. Семядоли сои в основном хранят два ключевых типа белков: глицинин (11S) и β-конглицинин (7S), которые различаются по своей растворимости и функциональным свойствам [42]. Семядоли составляют бóльшую часть массы семян, и их состав является основным направлением генетической модификации для повышения пищевой ценности семян [43]

Эмбрион (зародыш) – это живая часть семени, которая вырастает в новое растение. Он состоит из корешка (будущего корня), почечки (будущего побега) и гипокотиля. Эта область содержит высокие концентрации белков и жиров, которые поддерживают раннее развитие проростка [10, стр.365]. Когда семя поглощает воду, ферменты в зародыше мобилизуют запасные питательные вещества в семядолях, делая их доступными для роста проростка [44].

Рубчик – это область на семенной оболочке, где семя было прикреплено к стручку. Его положение и цвет могут различаться в зависимости от сорта сои. Рубчик играет значительную роль в поглощении воды во время набухания, и его проницаемость может влиять на скорость прорастания [45].

Состав семян сои является основным фактором, определяющим их ценность для пищевых продуктов, кормов и промышленного использования. Семена

состоят из белков, жиров, углеводов, витаминов, минералов и биоактивных соединений, каждое из которых имеет свою пищевую и функциональную роль [46, Таблица 2].

Таблица 2 – Состав семян сои

Компонент	Приблизительное содержание, %
Протеин	36
Масло	19
Углеводы	9
Клетчатка	19
Влажность	13
Зольные вещества (минералы)	4

Белки составляют примерно 35-41% сухого веса семян сои, что делает ее одной из самых богатых белком бобовых культур. Как уже было сказано, основными белками в сое являются глицинин (глобулин 11S) и β-конглицинин (глобулин 7S) [11, стр.366]. Эти белки ценны в пищевой промышленности своими функциональными свойствами, включая гелеобразование, эмульгирование и пенообразование. Глицинин обеспечивает лучшую термическую стабильность, а β-конглицинин улучшает растворимость белка. Соотношение этих двух белков часто изменяется в продуктах на основе сои, таких как соевое молоко, тофу и изоляты соевого белка, для улучшения текстуры и вкуса [46-47].

Жиры составляют 18-20 % сухого веса семян сои [11, стр.370]. Основными жирными кислотами в соевом масле являются пальмитиновая кислота (16:0), стеариновая кислота (18:0), олеиновая кислота (18:1), линолевая кислота (18:2), и линоленовая кислота (18:3). Многие современные исследования направлены на увеличение содержания олеиновой кислоты для улучшения стабильности масла и срока годности при одновременном снижении содержания линолевой кислоты, которая склонна к окислению. Следовательно семена сои с такими характеристиками будут иметь большую стоимость соевого масла, делая его более подходящим для приготовления пищи, производства биодизеля и косметики [48].

Углеводы составляют около 30-35 % семян сои. К ним относятся полисахариды (компоненты клеточной стенки), олигосахариды (раффиноза, стахиоза) и простые сахара (сахароза, глюкоза). Известно, что олигосахариды, такие как раффиноза и стахиоза, снижают усвояемость сои, и современная селекция направлена на создание сортов сои с более низким содержанием олигосахаридов [11, стр.372].

Кроме того, семена сои являются источником основных витаминов (витамин Е) и минералов (железо, кальций и цинк). Эти микроэлементы играют важную роль в поддержании сбалансированного питания. В семенах сои содержатся биоактивные соединения, такие как изофлавоны, флавоноиды,

фенольные кислоты и сапонины, обладающие антиоксидантными и противовоспалительными свойствами, тем самым способствующие широкому применению в пищевой и фармакологической промышленностях [11, стр.1]. Кроме питательных соединений, соевые бобы также содержат антинутриенты, такие как ингибитор трипсина или фитиновая кислота. Эти соединения могут снижать усвояемость белков и минеральных веществ, а также вызывать аллергические реакции [11, стр.374; 49].

1.1.4. Производство сои в мире и в Казахстане

В мировом производстве сои доминируют несколько стран, среди которых лидируют США, Бразилия и Аргентина. Вместе эти три страны производят более 80% мирового производства сои. В 2023 году США произвели около 120 миллионов тонн сои, большая часть которой экспортируется для удовлетворения международного спроса [19, стр.1]. Бразилия благодаря своим обширным сельскохозяйственным площадям, благоприятному климату и сильным экспортно-ориентированным стратегиям в 2023 году произвела более 152 миллионов тонн сои, что сделало ее крупнейшим производителем в мире. Аргентина занимает третье место, производство сои здесь в основном ориентировано на экспорт, в такие крупные рынки, как Китай и Европа [3] (Таблица 3).

Таблица 3 – Основные страны-производители сои (2023/2024 гг.)

Страна	Доля мирового производства	Производство (2023/2024, тонн)
Бразилия	39%	152 млн
США	29%	113.27 млн
Аргентина	12%	48.21 млн
Китай	5%	20.84
Индия	3%	11.88 млн
Парагвай	3%	11 млн
Канада	2%	6.98 млн
Россия	2%	6.8 млн
Украина	1%	5.2 млн
Боливия	0.96%	3.8 млн

Следующие значимые производители сои – Китай, Индия и Парагвай. Хотя Китай является крупнейшим мировым импортером соевых бобов, он также производит значительные объемы внутри страны. Однако свое производство не покрывает большой спрос страны. Индия производит соевые бобы, уделяя особое внимание внутреннему потреблению в качестве продовольствия и корма для животных. Парагвай, несмотря на свои небольшие территории, является ключевым экспортером, где большая часть его продукции предназначена для рынков Бразилии и Европы. Другие известные производители включают Канаду,

Россию, Украину и Боливию, каждая из которых вносит свой вклад в мировые поставки [19, стр.8-10; 20, стр.15-19].

Производство сои в Казахстане в последние десятилетия набирает обороты, и правительство намерено значительно увеличить площади возделывания [4; 32]. В 2021 году площадь посевов сои в Республике Казахстан составила 113,3 тыс. га, а в 2022 году –128.0 тыс. га с валовым сбором зерна – 250.4 тыс. тонн (Рисунок 3). В целом основная доля посевов сои в республике приходится на Алматинскую (83%), Восточно-Казахстанскую (9,4%) и Костанайскую (3,9%) области [50].



Рисунок 3 – Динамика посевных площадей и урожайности сои по Казахстану, тыс. га, 2009-2022 гг. [50]

В связи с климатическими характеристиками отдельных зон республики, направления селекционных работ по сое различны. На юге востребованы поздние сорта, с вегетационным периодом 135-145 дней (III группа спелости), обладающие засухоустойчивостью и мощной корневой системой. На юго-востоке необходимы среднепоздние сорта, с вегетационным периодом 120-130 дней (II и I группа спелости). На востоке же страны востребованы раннеспелые сорта, с вегетационным периодом около 100-110 дней (0 и 00 группа спелости), с укороченным периодом налив бобов – созревание, и устойчивые к весенним и осенним заморозкам. На возделывания сои на Севере Казахстана необходимы ультраскороспелые сорта с вегетационным периодом 85-95 дней (00 и 000 группа спелости) [50].

В рамках программы диверсификации посевов в Казахстане, исследования по сое зачастую сосредоточены на выведении скороспелых сортов, подходящих для северных и восточных широт. Выведены новые ультраранние и раннеспелые сорта с вегетационным периодом до 85 дней [51-53]. Также имеются полногеномные исследования ассоциаций, выявившие локусы количественных признаков (quantitative trait loci, QTL), связанных с устойчивостью сои к грибковым заболеваниям [54-55]. Эти достижения направлены на расширение производства сои в Казахстане, и расширения его посевных площадей.

1.2. Хозяйственно-ценные признаки важные для производства сои.

Соя является одной из наиболее экономически значимых культур во всем мире благодаря высокому содержанию протеина и масла, а также значимой роли в устойчивых сельскохозяйственных системах за счет биологической фиксации азота. Основные хозяйственно-ценные признаки, на которые нацелена селекция сои, включают компоненты урожайности, качество семян, устойчивость к стрессу и другие ключевые агрономические признаки, такие как время цветения, зрелость, высота растения и устойчивость к растрескиванию стручков и устойчивость к болезням [56-57]. Компоненты урожайности, такие как количество стручков на растении, семян в стручке, размер семян (длина, ширина, толщина) и вес 100 семян, напрямую влияют на общую продуктивность [58].

Признаки качества семян, такие как содержание протеина и масла, являются ключевыми в определении рыночной стоимости сои. Однако генетическое улучшение этих признаков является сложной задачей из-за отрицательной корреляции между ними [59-60]. Содержание незаменимых аминокислот и ненасыщенных жирных кислот также привлекают все большее внимание исследователей из-за их преимуществ для здоровья и применения их фармацевтической промышленности [61].

Кроме урожайности сои и качества семян, под влиянием изменения климата все более важными становятся такие признаки, как устойчивость к стрессу. Признаки цветения, засухоустойчивость, солеустойчивость, а также эффективное использование питательных веществ, стали приоритетными [62-63]. К примеру, изучение признаков адаптации к засухе, включающие более глубокую корневую систему и улучшенную эффективность использования воды, позволяют улучшить морфологию растения и свойства сои для большей продуктивности [64]. Оптимальная высота растения и особенности строения стебля имеют важное значение для предотвращения полегания, которое может значительно снизить урожай культуры [58, стр.1]. Устойчивость к растрескиванию стручков сои имеет решающее значение для минимизации потерь семян во время сбора урожая, особенно в засушливых регионах. Ранее были предприняты значительные усилия для выявления генетических локусов, которые обеспечивают более прочные стенки стручков [65]. Резистентность к основным заболеваниям, таким как соевая цистообразующая нематода (SCN, *Soybean Cyst Nematode*), корневая гниль и ржавчина сои, остаются первоочередной задачей, поскольку вспышки заболеваний могут привести к колосальным потерям урожая [66-68]. Улучшение хозяйственно-ценных признаков сои на генном уровне позволит охватить широкий спектр характеристик, которые в совокупности будут определять урожайность, качество и эффективность культуры.

Далее, более подробно рассмотрены важные хозяйственно-ценные признаки, изученные в условиях Алматинской области. Для облегчения восприятия все признаки сгруппированы в две основные группы: 1) признаки, определяющие продуктивность сои (количество продуктивных узлов, количество семян с растения, масса 1000 семян, масса семян с растения, размеры

семян сои, а также признак растрескиваемости стручков сои); 2) признаки качества семян сои (содержание протеина и масла).

1.2.1. Характеристика основных признаков, определяющих продуктивность сои

Продуктивность семян представляет собой сложный признак, определяемый различными компонентами. Такие характеристики как количество продуктивных узлов, количество семян в бобе, вес 1000 семян, размер семян, являются ключевыми компонентами, напрямую влияющими на урожайность и качество семян сои [58, стр.2].

Количество продуктивных узлов на растение определяется как общее количество продуктивных узлов, на которых развиваются стручки сои. Это важный компонент урожайности, поскольку каждый стручок может содержать несколько семян сои. Количество семян в стручке варьируется в зависимости от сорта сои, но обычно составляет от одного до трех семян [69].

Масса семян – это один из наиболее стабильных компонентов урожайности и часто используется в селекционных программах. Этот показатель регулируется плотностью посева и связан с другими компоненты урожайности, такие как количество продуктивных узлов на растении [58, стр.1].

Признаки компонентов урожайности важны в агрономии для принятия решений о плотности посева, ширине междурядий и внесении удобрений. Увеличение плотности растений может повысить количество растений на единицу площади, но при этом может снизить количество стручков на растении из-за конкуренции за свет и питательные вещества [32, стр.267].

Размер семян сои является важным морфометрическим признаком, который существенно влияет на продуктивность и качество семян сои, что делает его одним из ключевых признаков в селекции сои [Рисунок 4].



Рисунок 4 – Вариации размеров семян в различных сортах [70].

Согласно генетическим исследованиям размеров семян сои несколько генов были выделены и охарактеризованы [70-71]. Такие гены как *GmJAZ3* и *GmPLATZ*

были идентифицированы одними из ключевых для контроля размера и веса семян сои, аллельные вариации гена *GmST05* значительно влияют на толщину и качество семян [72-73]. В основном, локусы или гены размеров семян сои участвуют в важных биохимических процессах, которые имеют решающее значение для развития семян [74].

По данным автора Xu с соавторами (2022), размер семян напрямую влияет на продуктивность сои и коррелирует с содержанием масла и протеина. Более крупные размеры семян, как правило, связаны с повышенной урожайностью и улучшенным качеством семян, хотя эта связь может меняться в зависимости от условий окружающей среды [75]. Продолжение исследований в этой области улучшит стратегии селекции и будет способствовать разработке высокоурожайных и высококачественных сортов сои.

Еще одним из важных морфологических особенностей сои является устойчивость стручков к растрескиванию. Подверженность к растрескиванию стручков сои приводит к существенной потере урожая [76]. Для диких растений растрескиваемость является важным механизмом распространения семян, но для культурных растений эта морфологическая особенность нежелательна, так как зрелые стручки раскрываются перед сбором урожая, тем самым теряя его (Рисунок 5) [77]. Растрескиваемость стручков сои была почти утрачена в процессе культивирования и селекции сои. Тем не менее, в зависимости от генетического фона, факторов окружающей среды, морфологии и анатомии стручка, потери урожая из-за растрескиваемости сегодня могут составлять от 34 до 99 % [78].



Рисунок 5 – Стручки сои, устойчивые (справа) и восприимчивые (слева) к растрескиваемости [65, стр.17798].

Раскрытие стручка является высоко наследуемым и сложным признаком; было показано, что его наследуемость в широком смысле может варьироваться от 90 до 98 % в разных популяциях [79]. Один из основных генов растрескиваемости, *pdh1*, был идентифицирован Фунацуки и соавторами в 2014 году (65, стр.17797). Доминантное состояние гена *Pdh1* кодирует белок, высоко экспрессируемый в створках, увеличивая силу растрескивания. Рецессивный же

pdh1, содержит стоп-кодон, который блокирует правильный синтез белка (65, стр.17797). Несмотря на обнаружение нескольких QTL растрескиваемости стручков сои, некоторые из них могут быть нестабильными в разных средах и могут различаться в разных популяциях [80]. Поэтому, дополнительные исследования по поиску стабильных QTL важны для селекции, особенно при адаптации к новым климатическим регионам.

1.2.2. Основные признаки качества семян сои

Семена сои характеризуются рядом показателей качества, которые определяют их перспективы для использования в пищевой, кормовой, фармацевтической и других промышленностях [19, стр.15]. Среди них ключевыми являются содержание протеина, масла и влажности. Эти параметры регламентируются государственными и международными стандартами качества и зависят от требований конечного использования.

Белок является наиболее значимым компонентом семян сои. Высокое содержание протеина особенно желательно для пищевых соевых продуктов, таких как тофу, соевое молоко, окара, соевые изоляты белка и текстурированный соевый белок [13, стр.1]. Белок из сои применяется как основа для вегетарианских и веганских продуктов. Для пищевых продуктов содержание протеина обычно должно составлять не менее 36% [81]. Высокое содержание протеина является основным критерием для производства соевого шрота, используемого в кормах для птицы, свиней и крупного рогатого скота. Содержание протеина в соевом шроте для кормов – не менее 45% (для обычного шрота) и не менее 48% (для обезжиренного шрота). Допустимые уровни масла в шротах: менее 1% для обезжиренных продуктов [82]. В Казахстане и странах СНГ соевые продукты широко применяют в различных отраслях пищевой промышленности. В настоящее время в Казахстане ведутся работы по импортозамещению и развитию переработки сельскохозяйственной продукции [83].

Масличность семян является ключевым показателем для производства соевого масла, широко используемого в пищевой промышленности для приготовления блюд, производства маргаринов и других продуктов [2, стр.2]. Высокое содержание масла обеспечивает экономическую рентабельность переработки семян, так как позволяет производить больше продукта при минимальных затратах на переработку [3]. В кормовой промышленности остатки масла после переработки семян влияют на энергетическую ценность кормов [15, стр.216]. Международные стандарты рекомендуют содержание масла в семенах для переработки не менее 18%, а для некоторых видов переработки – выше 20%, чтобы удовлетворить потребности как пищевой, так и кормовой промышленности [84].

Влажность семян сои является критическим показателем, влияющим на их качество, хранение и переработку. Оптимальный уровень влажности предотвращает развитие плесени, снижение всхожести и потерю питательных свойств семян, что делает этот показатель ключевым как для производителей, так и для переработчиков. В пищевой промышленности влажность играет

важную роль в переработке соевых бобов: семена с оптимальным уровнем влажности обеспечивают высокую экстракцию масла и минимальные потери при производстве соевых продуктов [85]. Международные стандарты, государственные регламенты требуют, чтобы влажность семян, предназначенных для хранения и переработки, не превышала 12%, обеспечивая их пригодность для всех отраслей использования [86-88].

1.3. Геном сои

Исследования генома сои начались с разработки первых генетических карт, что позволило изучать наследственные признаки и улучшать сорта с использованием молекулярных маркеров [89]. Геном сои был впервые полностью секвенирован в 2010 году, и его размер составляет около 1,1 гигабазы (Gb) [90]. Геном сои отличается высокой долей повторяющихся элементов, которые составляют около 57% генома, что свойственно культурам с признаками множественной полиплоидии [91]. Обширные дупликации генов, обнаруженные в результате секвенирования, отражают сложные эволюционные события и механизмы адаптации сои до наших дней [90, стр.437]. Такие исследования сельскохозяйственных культур имеют значимый вклад как для фундаментальной науки, так и для аграрных применений.

1.3.1. Структура генома сои

Введение в культуру и диверсификация сои сопровождались значительными генетическими эффектами «бутылочного горлышка». Особенно четко это наблюдалось при культивировании азиатских ландрасов и их последующей интродукции в Северную Америку. Исследования генетического разнообразия дикой сои и ландрасов, с помощью молекулярных маркеров и технологий ресеквенирование геномов сои предположили несколько гипотез эволюции генетического разнообразия сои (Рисунок 6). Однако, в каждом из предлагаемых вариантов, в результате длительных процессов эволюции и селекции было утрачено около половины генетического разнообразия, а также 81% редких аллелей, что отмечалось при переходе от дикого вида *Glycine soja* к культурным ландрасам [39, стр.1](Рисунок 6).

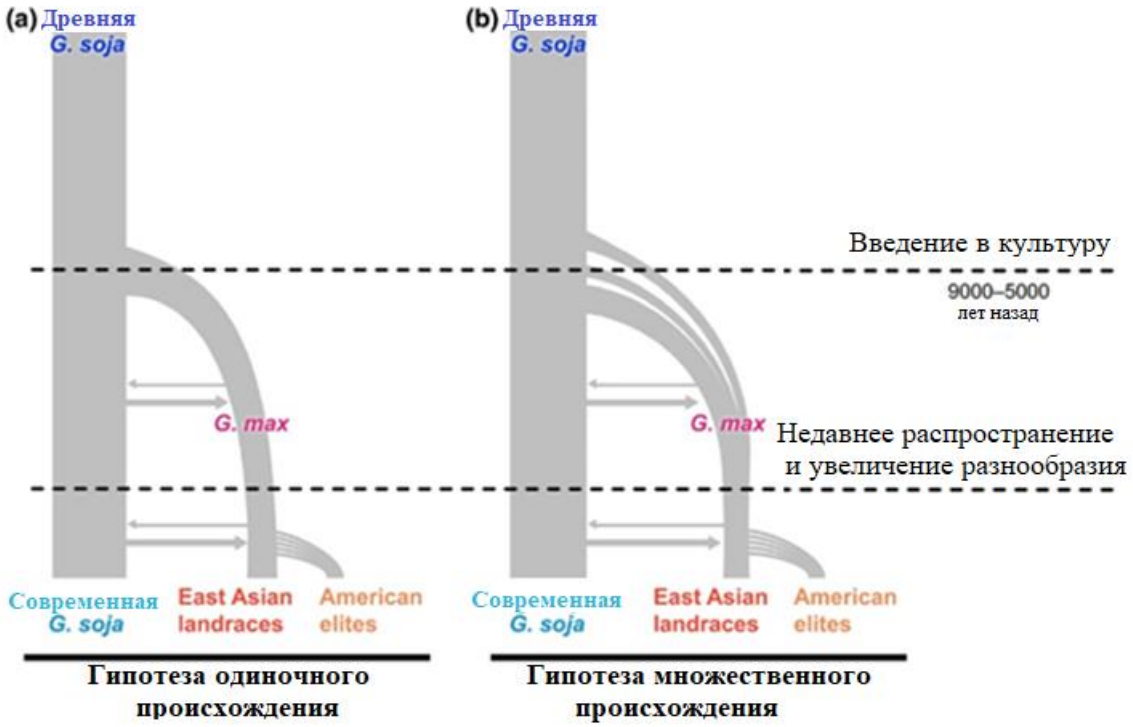


Рисунок 6 – гипотезы эволюции генетического разнообразия сои [21, стр.541]

К примеру, показатель генетического разнообразие у *G. soja* уменьшился с 2.17×10^{-3} до 1.47×10^{-3} у ландрасов и 1.11×10^{-3} у элитных сортов [21, стр. 540]. Эти данные подчеркивают существенные генетические потери в процессе введения в культуру сои и её адаптации в новом регионе.

Геном сои объемом 1,1 Gb был секвенирован методом полного геномного секвенирования и интегрирован с физическими генетическими картами, что позволило создать черновую сборку последовательности на уровне хромосом. Геном включает 46 430 предсказанных генов, кодирующих белки. Около 78% генов сосредоточены в концевых регионах хромосом, составляющих менее половины генома, но ответственных за большую часть генетической рекомбинации [90, стр. 179].

Два события дупликации генома, произошедшие около 59 и 13 миллионов лет назад, привели к высокому уровню плоидности: почти 75% генов представлены в нескольких копиях. Эти события сопровождались диверсификацией, утратой генов и хромосомными перестройками. В результате, геном был собран в 20 групп сцепления, соответствующих хромосомам, 2n = 40 (Рисунок 7). Дупликации генома привели к тому, что многие участки разных хромосом остались очень похожими по своему содержимому и порядку расположения генов [90, стр.178].

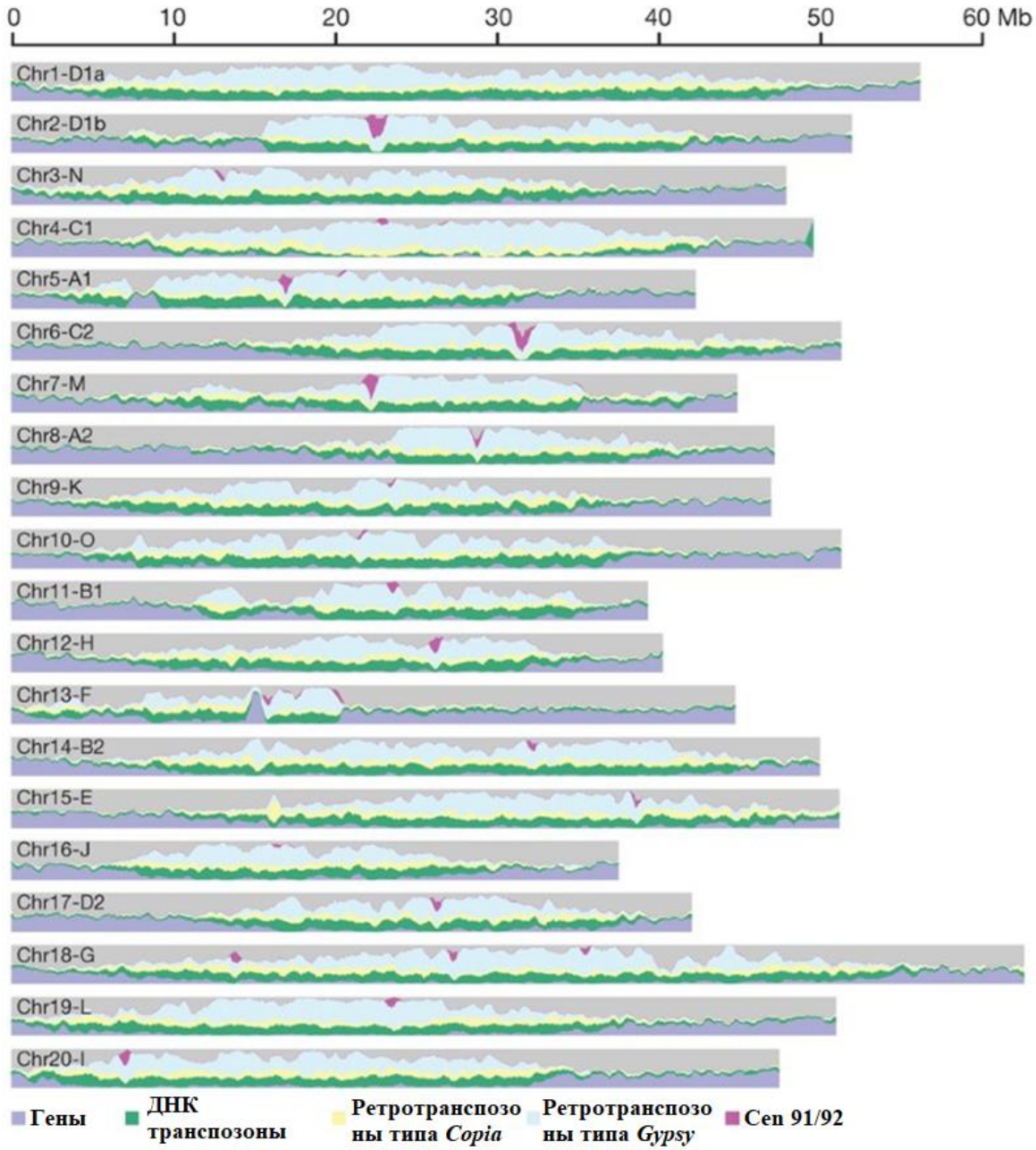


Рисунок 7 – Хромосомы генома сои. Основные компоненты ДНК разделены на гены (синие), ДНК-транспозоны (зеленые), ретротранспозоны (желтые Copia-like, голубые – Gypsy-like), и Cent91/92 (специфичный для сои центромерный повтор, розовые) [90, стр.179]

Центромерные районы хромосом сои богаты ретротранспозонами и тандемными повторами, которые способствуют структурной и функциональной стабильности. Организация этих последовательностей поддерживает правильную сегрегацию хромосом во время деления клеток. Перицентромерные регионы являются горячими точками генетических и эпигенетических модификаций [90, стр.178].

Хромосомы различаются по длине и содержанию генов. Например, некоторые хромосомы имеют более высокую плотность определенных генов, связанных с такими признаками, как устойчивость к болезням, в то время как другие обогащены генами, связанными с метаболическими путями [90, стр.179].

Геном сои является крупнейшим растительным геномом, секвенированным методом дробовика, демонстрируя высокое качество по сравнению с другими растениями. Теломерные повторы (TTTAGGG или CCCTAAA) были обнаружены на 27 из 40 концевых участков хромосом. Кроме того, центромерные повторы длиной 91-92 пар оснований были идентифицированы во внутренних участках большинства хромосом, свидетельствуя о сложности сборки высокоповторяющихся регионов [90, стр.178; 91].

Генетические и физические карты хромосом сои высокого разрешения были построены с использованием различных молекулярных маркеров. Эти карты облегчают идентификацию местоположения генов и их связь с хозяйственно-ценными характеристиками. SoyBase – база данных, объединяющая карты генетических и физических последовательностей с фенотипическими качественными и количественными данными [92]. Аналогичные базы данных, такие как SoyOD, SoybeanGDB и другие, предоставляют ценный ресурс для изучения и понимания структуры и эволюции генома сои, а также для программ селекции или редактирования генома с целью создания улучшенных сортов сои [93-94]

1.3.2. Молекулярная генетика сои

Молекулярная генетика сои (*Glycine max* (L.) Merr.) направлена на изучение механизмов работы генов, их взаимодействие, регуляцию и влияние на ключевые хозяйственно-ценные признаки, включая продуктивность, устойчивость к стрессам и качественные характеристики зерна. Основные инструменты молекулярной генетики, молекулярные маркеры, подразделяются на три основные категории в зависимости от используемых технологий: 1 – маркеры на основе гибридизации, такие как полиморфизм длин рестрикционных фрагментов (RFLP – Restriction Fragment Length Polymorphism); 2 – маркеры на основе полимеразной цепной реакции: RAPD (Random Amplified Polymorphism DNA), AFLP (Amplified Fragment Length Polymorphism) и SSR (Simple Sequence Repeat); 3 – маркеры на основе последовательностей SNP (Single Nucleotide Polymorphism) [95].

Молекулярные маркеры RFLP, отличаются высокой специфичностью благодаря использованию ферментов рестрикции. Эти ферменты, распознающие определенные последовательности ДНК, разрезают геном на фрагменты, длина которых варьируется в зависимости от наличия или отсутствия сайтов рестрикции у различных генотипов. Такой подход позволяет точно идентифицировать генетические различия между близкородственными растениями, что важно для изучения филогенетических связей и картирования генов [96-97]. Кроме того, особенностью метода является его способность детектировать полиморфизм в кодирующих и некодирующих областях генома, включая анализ мутаций и генетическое разнообразие в популяциях растений. Несмотря на точность метода, его высокая стоимость и трудоемкость ограничивают использование в современных программах селекции [95, стр.231].

Одними из первых молекулярных маркеров на основе ПЦР является RAPD – маркеры, разработанные в 1990-х годах. RAPD метод основан на

амплификации случайных сегментов геномной ДНК с помощью одиночных праймеров произвольной последовательности [98]. RAPD были применены для изучения генетической структуры популяции сои и картирования генов устойчивости к стрессам. Например, этот метод использовался в идентификации генов, связанных с засухоустойчивостью, что имело важное значение для регионов с недостатком воды [99]. Однако метод RAPD также имел ограничения, связанные с низкой специфичностью и воспроизводимостью [100].

AFLP маркеры широко применяются в исследованиях растений для оценки генетического разнообразия, анализа структуры популяций и селекции. AFLP представляет собой высокоэффективную и воспроизводимую методику на основе ПЦР, которая сочетает в себе рестрикционное расщепление геномной ДНК с последующей селективной амплификацией рестрикционных фрагментов [101]. Этот метод обеспечивает высокое разрешение генетических различий, что делает его особенно полезным для видов с слабо изученными геномами. Однако, несмотря на свои преимущества, метод AFLP требует высококачественной ДНК и является более трудоемким по сравнению с новыми молекулярно-генетическими технологиями [95, стр. 234].

Маркеры SSR, или микросателлиты, стали следующим значительным этапом в развитии молекулярных технологий. Они представляют собой короткие повторяющиеся последовательности ДНК, которые варьируются между сортами [95, стр.233]. Преимущества SSR включают высокая специфичность, стабильность, воспроизводимость, а также удобство для автоматизированных систем, что сделало их стандартом в селекционных программах [102].

Современные селекционные программы активно используют SNP-маркеры, представляющие собой однонуклеотидные различия в последовательностях ДНК [43]. SNP-маркеры могут обладать высокой плотностью в геноме, что делает их идеальными для геномных исследований, таких как ПГАА. Идентификация SNP начинается секвенированием геномов разнообразных образцов популяции изучаемого растения (Рисунок 8).

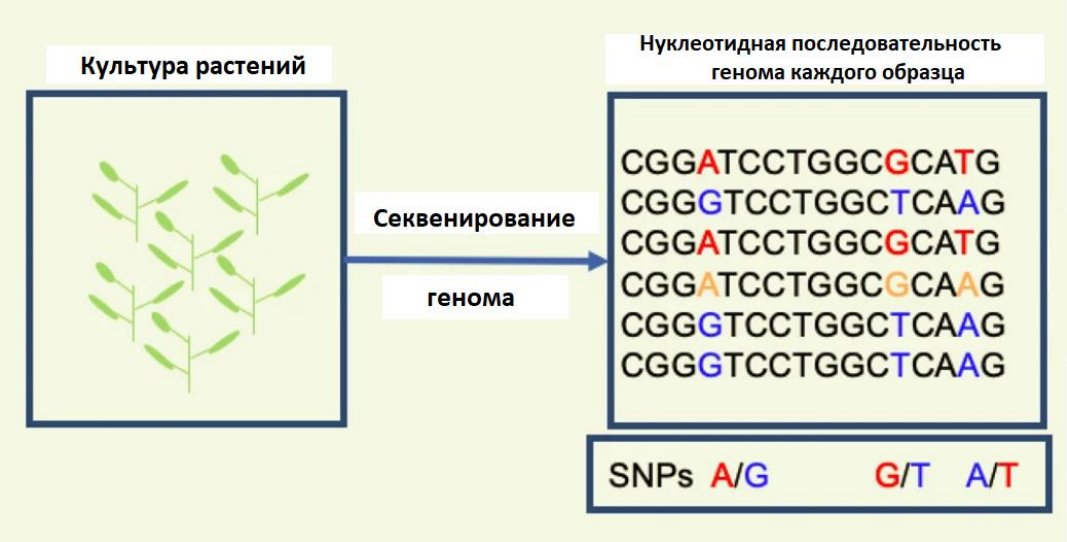


Рисунок 8 – схема идентификации SNP-маркеров [105]

Например, для сои были созданы чипы SNP, состоящие из 50К и 618К маркеров [103-104].

Полученные последовательности выравниваются с референсным геномом для выявления потенциальных SNP. После первичной идентификации полиморфизмы проходят тщательную фильтрацию с целью поиска маркеров с наиболее подходящими критериями, такими как равномерное распределение по геному и высокая частота минорного аллеля. Отобранные SNP используются для разработки и производства микрочипа, при этом производитель проводит дополнительные этапы контроля качества для обеспечения надежности работы тест-системы [103-104]. Развитие технологий высокопроизводительного генотипирования облегчило крупномасштабное открытие и применение SNP. Ожидается, что по мере расширения геномных ресурсов маркеры SNP будут играть все более важную роль в улучшении сельскохозяйственных культур и генетических исследованиях [106].

Эволюция молекулярных маркеров – от RFLP и RAPD до современных SNP – существенно изменила подходы к генетике и селекции сои. Они обеспечили возможность точного картирования генов, идентификации QTL и повышения эффективности селекционных программ, что особенно важно в условиях изменения климата и роста спроса на продуктивные и устойчивые сорта [107-108].

1.3.2.1. Классическое картирование локусов количественных признаков

Картирование количественных признаков – это один из важнейших инструментов генетики растений, который используется для выявления геномных участков, контролирующих количественные признаки, такие как урожайность, устойчивость к стрессам и заболеваниям, состав семян. Метод был впервые предложен в 1980-х годах и применён в сельском хозяйстве для улучшения таких культур, как кукуруза и рис [109]. Основной принцип картирования QTL основан на анализе связи между молекулярными маркерами и фенотипическими данными, что позволяет определить генетические регионы, влияющие на сложные количественные признаки (Рисунок 9). Выявление генов или QTL, контролирующих различные признаки, становится возможным благодаря генетическому анализу неравновесного сцепления (LD, linkage disequilibrium). Анализ LD основан на принципе генетической рекомбинации в процессе мейоза и расположении маркеров на хромосоме [110]. Данный метод позволяет создавать карты сцепления, состоящие из генетических маркеров, специфичных для определенной популяции. Чаще всего для картирования используются картирующие популяции, такие как F2, F3 или обратные скрещивания (Back-cross популяции), рекомбинантные инбредные линии (RIL, recombinant inbred lines), дигаплоидные линии [109, стр.1654].

Этапы картирования QTL включают:

1 – Создание популяции картирования: популяция создаётся путём скрещивания двух генетически различных линий, обладающих контрастными признаками (Рисунок 9 A, B).

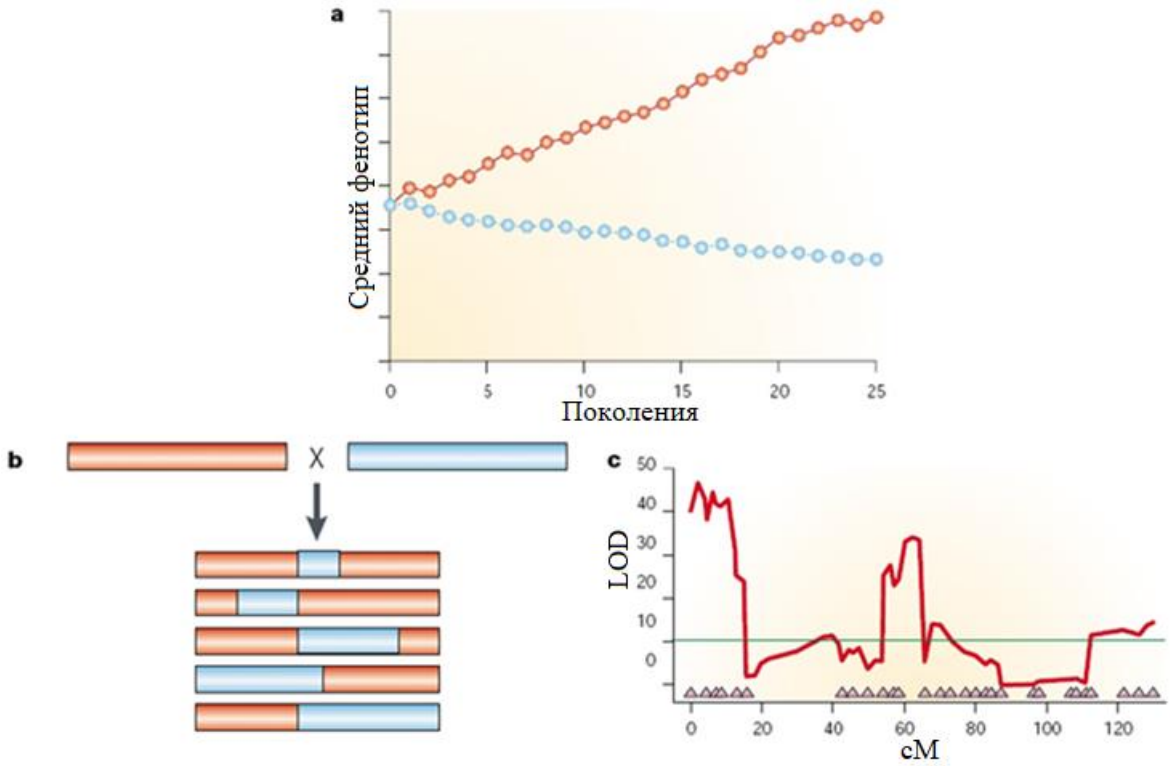


Рисунок 9 – схема QTL анализа [112]. a) две контрастные по фенотипическим признакам родительские линии; b) создание картирующей популяции;c) Результаты QTL анализа представлены в виде графика, где по оси X откладывается положение маркеров на хромосоме (в единицах рекомбинации – cM), а по оси Y – статистическая значимость связи маркера с признаком.Зеленая линия на графике обозначает уровень значимости; участки, превышающие этот порог, считаются достоверными QTL. В приведенном примере на графике выявлено пять потенциальных QTL, с наиболее значимыми участками на расстояниях около 10 и 60 cM.

2 – Генотипирование популяции. Для анализа генотипов используются молекулярные маркеры, такие как SNP и SSR.

3 – Фенотипирование. Изучение проявления фенотипов (например, содержания протеина или массы семян) в разных условиях выращивания.

4 – Анализ ассоциативной связи. Определение достоверной связи между маркерами и фенотипами для локализации QTL на генетической карте с помощью статистических методов (Рисунок 9, C).

5 – Идентификация генов-кандидатов. Поиск генов, экспрессия которых связана с изучаемым признаком [111].

Применение картирования QTL позволило выявить локусы признаков, подверженных изменению в ходе селекции, таких как время цветения, сроки

созревания и растрескивание стручков [22;113]. Большинство из этих признаков контролируются одним или двумя основными QTL и несколькими минорными QTL. Были картированы QTL, ассоциированные с качественными характеристиками семян, такими как содержание протеина, масла и жирных кислот [114]. Идентифицированные QTL и связанные с ними маркеры могут быть использованы в маркер-опосредованной селекции для улучшения сортов сои.

1.3.2.2. Полногеномный анализ ассоциаций

Картирование QTL и ПГАА успешно используются для поиска локусов и генов-кандидатов в сельскохозяйственных растениях. Однако картирование QTL имеет ограничения в виде небольшого аллельного разнообразия и меньшей точности, которые можно компенсировать с помощью ПГАА [115]. Геномные ассоциативные исследования являются мощным инструментом для выявления генетических вариантов, связанных со сложными признаками у сельскохозяйственных культур. Этот метод позволяет установить связь между генетическими полиморфизмами и фенотипическими признаками, что имеет важное значение для программ селекции, включая маркер-опосредованную селекцию [116].

Для успешного проведения ПГАА необходимо выполнение ряда важных условий. Во-первых, наличие генетически разнообразной выборки позволяет выявлять значимые генетические вариации и обеспечивает репрезентативность популяции [115, стр.536]. Во-вторых, важным аспектом является получение качественных генотипических данных с высокой плотностью и точностью SNP-маркеров, что достигается с применением высокопроизводительного секвенирования (new generation sequencing, NGS) или SNP-микрочипов [115, стр.537]. Например, одной из наиболее широко используемых платформ с 2013 года является SoySNP50K iSelect BeadChip, который широко применяется в ПГАА [103]. Позже, в 2021 была разработана платформа SoySNP618K, содержащая около 618 000 SNP, отличающаяся более высокой плотностью покрытия, что повышает разрешение картирования ассоциаций и более точно идентифицирует гены-кандидаты, контролирующие ключевые признаки сои [104]. Третьим важным условием успешного проведения ПГАА являются фенотипические данные исследуемой выборки, которые обеспечиваются путем тщательной оценки признаков в различных экологических условиях для минимизации влияния окружающей среды и повышения наследуемости признаков. Кроме того, необходим тщательный анализ структуры популяции и родственных связей для снижения обнаружения ложных ассоциаций и повышения точности результатов [115, стр.538;116, стр.5744].

Процесс проведения ПГАА для сельскохозяйственных культур включает несколько ключевых этапов, начиная с генотипирования, в ходе которого получают данные о ДНК популяции с использованием методов секвенирования или микрочипов для идентификации SNP. Для обеспечения надежности генетических данных проводится контроль качества, включающий фильтрацию SNP и образцов по таким критериям, как частота минорного аллеля (minor allele

frequency, MAF), доля обнаружения маркеров в генотипах (SNP call rate) [117]. Следующим важным шагом является фенотипирование, предполагающее оценку таких признаков, как урожайность, устойчивость к засухе и болезням, в различных условиях окружающей среды. Затем осуществляется анализ ассоциаций с применением статистических моделей для выявления SNP, связанных с исследуемыми фенотипическими признаками. ПГАА проводится с учетом генетической структуры изучаемой популяции и родственных связей внутри нее. Во избежание ложноположительных результатов проводится коррекция на множественное тестирование, включающая использование поправки Бонферрони и FDR (false discovery rate) [118]. Маркер-признаковые ассоциации (MTA, marker-trait association), обнаруженные в результате ПГАА, показывают статистичекую оценку взаимосвязи между определенным ДНК маркером, в нашем случае SNP, с исследуемым признаком фенотипа. Выявленные статистически значимые MTA формируют пики на графиках Manhattan plot, позволяющие исследователям сосредоточиться на определенных регионах хромосом. Завершающим этапом является интерпретация результатов обнаруженных MTA и последующая идентификация генов - кандидатов, при которой значимые MTA анализируются для поиска генов, ответственных за признак, с последующей их функциональной аннотацией (Рисунок 10) [116, стр.5749].

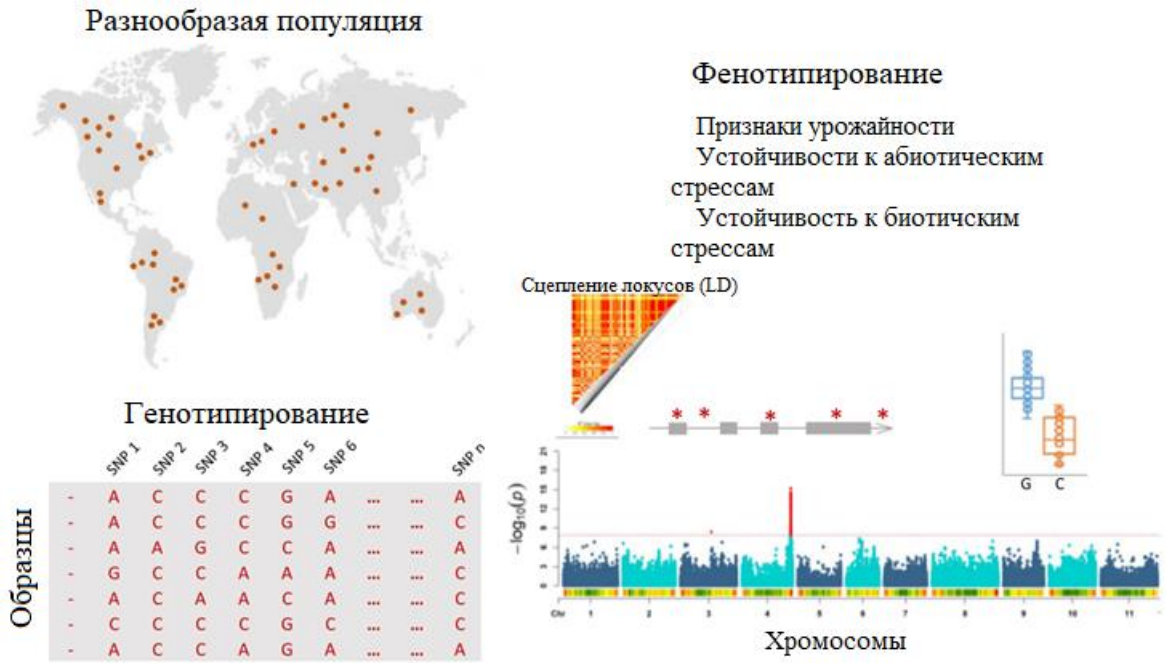


Рисунок 10 – Схема полногеномного анализа [116, стр.5744]

Для выявления значимых ассоциаций в ПГАА применяются различные статистические модели, каждая из которых имеет свои преимущества и ограничения. Модель общей линейной регрессии (general linear model, GLM) представляет собой наиболее простую методику, которая учитывает только фиксированные эффекты. Однако модель GLM не корректирует влияние структуры популяции и родственных связей, что может привести к

ложноположительным результатам. Более точным подходом является модель смешанных линейных эффектов (mixed linear model, MLM), включающая как фиксированные, так и случайные эффекты. Такая модель позволяет учесть структуру популяции и родственные связи, тем самым минимизировать вероятность ложных ассоциаций [119]. Для повышения вычислительной эффективности была разработана сжатая модель смешанных линейных эффектов (compressed mixed linear model, CMLM), которая группирует индивиды на основе их генетического сходства. Так уменьшаются вычислительные затраты без значительного снижения точности анализа. Одним из современных улучшений является множественная модель с множественными локусами (multi loci mixed model, MLMM). Модель MLMM улучшает стандартный подход за счет использования эффективных алгоритмов обработки больших данных и улучшенного учета структуры популяции. Такое преимущество делает данную модель особенно полезной для GWAS с большими выборками и сложными признаками, а также представляет собой перспективный инструмент для геномных исследований в сельском хозяйстве и биотехнологии растений [120].

Для определения значимых MTA между SNP и исследуемым признаком применяются строгие пороговые значения, позволяющие минимизировать вероятность ложноположительных результатов. Одним из наиболее консервативных методов является поправка Бонферрони, которая корректирует уровень значимости. Показатель Бонфррони делит уровень значимости (в большинстве случаев 5%) на общее количество протестированных SNP, что существенно снижает риск случайных ассоциаций. Альтернативный подход – контроль FDR, регулирующий долю ложноположительных результатов, что позволяет балансировать между уровнем значимости и статистической мощностью теста. Для ПГАА в исследованиях сельскохозяйственных культур, из-за различий в геномной архитектуре и высокой степени гетерозиготности, применяются пороги, такие как 1×10^{-5}, что позволяет учитывать специфические особенности геномов растений [121].

Несмотря на высокую информативность, метод ПГАА имеет ряд ограничений. Одной из основных проблем для выявления сложных признаков, обусловленных взаимодействием множества генов с малым эффектом, является необходимость больших выборок и многофакторного фенотипирования для повышения точности результатов. Кроме того, даже при использовании методов деления популяции, остаточные эффекты ее структуры могут приводить к ложноположительным ассоциациям, и затруднять интерпретацию результатов [122]. Существенное влияние факторов окружающей среды на хозяйственно-ценные признаки также усложняет точное выявление генетических эффектов, создавая необходимость в проведении экспериментов в различных условиях. Наконец, анализ больших объемов геномных данных сельскохозяйственых культур требует значительных вычислительных ресурсов и высококвалифицированных специалистов, что может ограничивать широкое применение метода в практических селекционных программах [116, стр.5749].

При правильном подходе вышеперечисленные ограничения преодолимы и полногеномные исследования ассоциаций широко используются для выявления

генетических детерминант важных хозяйственно-ценных признаков сои. Эти исследования выявили многочисленные QTL и гены-кандидаты, связанные с продуктивностью, массой семян, высотой растения, урожайностью [5; 56], и с устойчивостью к болезням, например, к соевой нематоде [6]. Различные модели ПГАА и карты SNP высокой плотности использовались для идентификации значимых маркеров и генов-кандидатов [57;61;66]. Такие исследования расширяют знания о генетики сои и могут быть применены при маркерной селекции для улучшения соевых бобов.

1.4. Селекция сои в мире и в Казахстане

Соя – культура мирового значения, объемы производства которой за последние два десятилетия удвоились [3]. Селекция сои является важным направлением аграрной науки, улучшающее урожайность и качество сои, а также устойчивость к стрессам и изменениям климата. Различные методы селекции, от традиционных до современных биотехнологических, обеспечили значительный прогресс в селекциионных программах сои на мировом уровне [123]. Классические методы селекции основаны на использовании естественного разнообразия видов и генотипов и включают такие подходы, как селекция по фенотипу, использование гибридизации и т.д. [32, стр.328]. Несмотря на фундаментальную значимость классических методов селекции, они всё чаще дополняются современными селекционными подходами, такими как маркер-опосредованная селекция и генетическая инженерия [124]. Эти передовые инструменты ускоряют селекционный процесс и повышают его точность, и вместе с классическими методами продолжают играть важную роль в селекционных программах, направленных на повышение продуктивности и устойчивости сельскохозяйственных культур.

1.4.1. Классические методы селекции

Классические методы селекции основаны на использовании естественного разнообразия видов и генотипов и включают такие подходы, как селекция по фенотипу (отбор), использование гибридизации, мутагенез и обратное скрещивание [32, стр.389].

Метод селекции по фенотипу строится на выборе образца с необходимыми фенотипическими признаками для последующего скрещивания. Это простой, но эффективный подход, позволяющий выбрать наилучшие родительские формы и повышать урожайность или устойчивость последующих поколений к окружающим стрессорам. Недостаток данного метода связан с необходимостью точного изучения всех связей между признаками и условием среды. При гибридизации различные сорта сои скрещиваются для объединения желаемых признаков. Этот метод позволяет селекционерам сочетать дополняющие друг друга ценные признаки путём контролируемого опыления, что приводит к созданию улучшенных сортов [32, стр.146].

При мутагенной селекции генетические вариации индуцируются с использованием химических или радиационных обработок для создания новых признаков, позволяя создать генетическое разнообразие. Обратное скрещивание

особенно полезно для внесения определенных признаков (например, устойчивости к заболеваниям) из родительской формы-донора в элитный высокоурожайный сорт, при этом сохраняется бóльшая часть исходного растения [32, стр.360]. Ряд сортов сои был успешно выведен с использованием вышеперечисленных классических методов. Например, «Williams 82», широко используемый сорт, известный своей высокой урожайностью и устойчивостью к цистообразующей нематоде сои, активно применяется в селекционных программах и молекулярных исследованиях [125].

В Казахстане селекцией сои занимаются несколько ведущих научно-исследовательских институтов и селекционных станций, специализирующихся на разработке сортов, адаптированных к различным климатическим условиям страны. Основными оригинаторами новых сортов являются КазНИИЗиР**,** расположенный в Алматинской области. КазНИИЗиР с 1961 году и по настоящее время занимается селекцией новых сортов сои [126]. На севере, в Петропавловске, Северо-Казахстанская Сельскохозяйственная Опытная Станция (СКСОС) занимается селекцией холодостойких сортов, а в Костанайской области действует Карабалыкская Сельскохозяйственная Опытная Станция**,** специализирующаяся на засухоустойчивых сортах. На востоке страны – Восточно-Казахстанская Сельскохозяйственная Опытная Станция (ВКСОС) в Усть-Каменогорске занимается выведением высокоурожайных сортов сои. Столичный Казахский Агротехнический Университет им. С. Сейфуллина является одним из ведущих аграрных вузов Казахстана, занимающийся научными исследованиями в области сельского хозяйства, в том числе и селекции сои [127-129]. Традиционными методами селекции были созданы многие отечественные сорта разных групп спелости, адаптированные к различным широтам страны, отличающиеся высокой урожайность и устойчивостью к абиотическим и биотическим факторам. Среди них сорта Ласточка, Жансая, Ивушка, Бірлік КВ, Казахстанская 200, Гибридная 670, Казахстанская 2309, Эврика 357 и другие [126]. В Государственный реестр селекционных достижений, рекомендуемых к использованию в Республике Казахстан, включено 56 сортов сои, из них 21 сорт отечественной селекции, 34 сорт зарубежной селекции и 1 сорт совместной селекции [130]. Наиболее используемыми сортами сои по стране являются сорта казахстанской, российской и сербской селекции. В 2021 году наибольшая площадь посева была засеяна сортами: Жансая (Казахстан), Селекта 302 (Россия), Триумф (Сербия и Черногория). Доля сортов отечественной селекции на посевных площадях Казахстана составляет 34,7% [131].

1.4.2. Современная молекулярная селекция для создания высокопродуктивных сортов сои.

Современные методы молекулярной селекции значительно ускорили и упростили процессы создания новых сортов растений. Они основаны на использовании генетической информации для повышения точности и эффективности отбора, что делает их незаменимыми в современном селекционном процессе [95].

QTL-картирование является ключевым инструментом молекулярной селекции, позволяющим выявлять генетические локусы, сцепленные с важными хозяйственно-ценными признаками. Этот метод лежит в основе маркер-опосредованной селекции, которая упрощает отбор генотипов с желаемыми аллелями на основе молекулярных маркеров. Секвенирование генома сои ускорило идентификацию локусов, отвечающих за устойчивость к стрессам, болезням и высокую урожайность, что значительно повысило качество и точность отбора родительских форм для селекционных программ [90].

В отношении стрессоустойчивости на сегодняшний день идентифицировано и зарегистрировано более 500 QTL в базе soybase.org. К примеру, исследование Prince et al. (2015), проведенноев США, было направлено на идентификацию QTL, связанных с устойчивостью сои к засухе. Используя RILs сои, исследователи применили метод QTL-картирования и выявили важный локус *qDT_Gm03*, расположенный на 3-й хромосоме. Этот локус продемонстрировал значительное влияние на физиологические показатели устойчивости к засухе, такие как содержание влаги в листьях и скорость транспирации [132].

Масштабные исследования ведутся в изучении устойчивости сои к соевой цистообразующей нематоде (*Heterodera glycines*), так как этот вредитель является одним из самых опасных для сои. В результате исследований, были выявлены основные локусы устойчивости к патогену *Rhg1* и *Rhg4*, локализованные на 18-й и 8 хромосомах соответственно [133]. Также были идентифицированы дополнительные локусы устойчивости на других хромосомах, например на 11-ой и 15-ой [134]. Данные локусы являются ключевыми мишенями в современных программах селекции для разработки ресистентных к цистообразующей нематоде сортов сои.

Высокая продуктивность является одной из ключевых целей селекционной работы с соей, и на данный момент выявлено более 1200 QTL, связанных с признаками, влияющими на продуктивность сои [92]. Такое количество обнаруженных локусов можно объяснить сложностью признака, который складывается из множества компонентов. Признак продуктивности является количественным признаком и контролируется многочисленными генами, находящимися всложной сети взаимодействий. Расположение генов может варьироваться в зависимости от E и G изучаемой генетической популяции, также G×E играют значительную роль в функционировании генов [114]. Так, исследование Ravemolda et al. (2021), проведенное на базе естественной популяции сои, с использованием методов ПГАА позволило вывить важный регион на 10 хромосоме (11.5Mb), который оказывает значительное влияние на урожайность и массу семян культуры [135].

Селекция сои по качественным признакам семян сои является одной из приоритетных в селекционных программах. На сегодняшний день известно более 250 QTL, связанных с содержанием протеина, и свыше 200 QTL, связанных с содержанием масла в семенах сои. Основные локусы, регулирующие эти признаки, распределены по нескольким хромосомам, при этом ключевыми являются *cqPro-20* и *cqOil-15*, расположенные на 20-й и 15й хромосомах. Данные

локусы оказывают наибольшее влияние на накопление протеина и масла [136]. Основной QTL, ***cqPro-20,*** обладает наиболее значимым аддитивным эффектом среди всех известных QTL, связанных с содержанием протеина, и ген *Glyma.20G85100* считается ответственным за этот ключевой локус [60-61]. Основная проблема производства высококачественной сои является отрицательная корреляция между содержанием протеина и масла [79; 51]. Возможным решением данного противостояния может быть создание сортов с содержанием протеина выше среднего, без ущерба содержанию масла и общей продуктивности [137].

1.4.3. Применение технологии KASP в современной селекции

Маркеры по типу KASP (Kompetitive Allele Specific PCR) впервые были разработаны в начале 2010-х годов как инструмент для точного и экономичного генотипирования SNP и InDel (инсерции/делеции, insertion/deletion) полиморфизмов. Одними из первых культур, для которых начали разрабатывать KASP-маркеры, были пшеница, кукуруза и ячмень, где их использовали для идентификации аллелей, ассоциированных с хозяйственно-ценными признаками [138]. Эти маркеры быстро получили широкое распространение в селекционных программах благодаря их точности, высокой воспроизводимости и возможности масштабирования для разнообразных сельскохозяйственных культур.

Разработка KASP-маркеров включает выбор информативных SNP или InDel, предпочтительно с известной функциональной значимостью. На первом этапе проводится отбор полиморфных позиций из массивов данных, полученных с помощью SNP-чипов или полногеномного секвенирования. Затем для каждой аллели подбираются два аллельно-специфичных праймера, каждый из которых содержит уникальный флуоресцентный хвост. При проведении реакции PCR происходит конкурентное связывание одного из праймеров с матрицей ДНК, в результате чего флуоресцентный сигнал позволяет точно определить аллель [139-140]. На сегодняшний день KASP-маркеры являются важнейшим компонентом маркер-ассоциированной селекции (MAS) и применяются как в предварительном скрининге исходного материала, так и на финальных этапах селекции.

Научные исследования и внедрение современных селекционных практик играют ключевую роль в расширении производства сои в различных климатических зонах Казахстана. Повышение доли посевных площадей страны, засеянных семенами отечественной селекции, адаптированных к местным агроэкологическим условиям является стратегической задачей по продовольственной безопасности страны [141]. Особое внимание уделяется изучению генетической основы качественных признаков семян сои, таких как содержание протеина и масла, которые играют решающую роль в повышении экономической ценности и эффективности культуры. Поиск и идентификация новых QTL, связанных с этими признаками, является приоритетным направлением исследований в Казахстане. Исследования, проведённые в странах с различными климатическими условиями, продемонстрировали, что проявление генетических факторов, определяющих качественные признаки,

может изменяться в зависимости от окружающей среды [142]. Это открывает возможности для создания новых высокопродуктивных и качественных сортов сои, адаптированных к региональным особенностям Казахстана.

Целью настоящего исследования является поиск и идентификация ассоциации между SNP маркерами и хозяйственно-ценными признаками сои методом ПГАА. Для достижения поставленной цели были использованы данные ресеквенирования коллекции сои, выращенной в условиях Алматинской области. Алматинская область является основным регионом производства сои в стране, и изучение влияние местной среды на проявление признаков увеличит прикладное значение полученных результатов для будущих селекционных программ на основе маркер-опосредованной селекции. Кроме того, учитывая различную реакцию изученных коллекций сои на факторы окружающей среды при идентификации МТА, результаты могут служить ценным ориентиром для поиска генов и QTL хозяйственно-ценных признаков культуры в других регионах мира.

2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. Материалы исследований

Материалом исследования послужила мировая коллекция сои, состоящая из 252 селекционных линий и сортов, что обеспечивает высокий уровень генетического разнообразия и репрезентативность выборки для комплексного анализа хозяйственно-ценных признаков сои. Происхождение и страны-оригинаторы изучаемой коллекции сои представлены в Таблице 4, более подробное описание содержит Приложение 1. Данная коллекция была использована для фенотипического анализа с 2018 по 2022 год, анализов структуры популяции и генотипирования, а также для проведения ПГАА.

Таблица 4 – Мировая коллекция сои

Происхождение	Страна-оригинатор	Количество образцов, шт
Центральная Азия	Казахстан	31
Восточная Азия	Китай	32
	Южная Корея	16
	Филиппины	1
	Вьетнам	1
Восточная Европа	Россия	61
	Украина	30
	Польша	5
	Беларусь	4
	Чехия	4
	Молдова	2
	Бельгия	1
	Болгария	1
Западная Европа	Франция	12
	Сербия	6
	Швеция	2
	Швейцария	1
Северная Америка	Канада	24
	США	16
Всего		252

Из казахстанских сортов сои в коллекции представлены 31 отечественный сорт и линия сои, таких как Жалпаксай, Алматы, Мисула, Роза, Искра, Перизат, Вита, Даная, Сабира, Ласточка, Жансая и др. Районированным сортом-стандартом сои по Алматинской области является сорт «Жансая» [130].

2.2. Методы исследований

2.2.1. Полевые работы, погодные условия, оценка хозяйственно-ценных признаков в изучаемой коллекции сои.

Мировая коллекция сои выращивалась на экспериментальных полях КазНИИЗиР (Алматинская область, 43°15' с.ш., 76°54' в.д., 740 м над уровнем моря) с 2018 по 2022 годы, в двух повторностях. Тип почвы – светло-каштановая, характерная для степных и полузасушливых регионов, с умеренным содержанием гумуса (2,0-2,5%). Изучение проявления хозяйственно-ценных признаков сои в условиях Алматинской области является важным этапом исследования, так как Алматинская область является основным регионом производства сои в стране.

Алматинская область характеризуется континентальным климатом, с холодной зимой и жарким летом. Зимние температуры могут опускаться ниже -10°C, а летние максимумы часто превышают 30°C. Осадки умеренные, пик приходится на весну и начало лета, тогда как конец лета и осень более сухие. Ветры, как правило, слабые или умеренные, но могут быть и сильные порывы, особенно в предгорьях и горных районах из-за местного рельефа. С 2018 по 2022 экспериментальные года погодные условия в Алматинской области демонстрировали типичную сезонную закономерность: температура повышалась с мая, достигала пика в июле и постепенно снижалась к сентябрю (Таблица 5). В большинстве лет самым теплым месяцем был июль, когда температура достигала около 25-30 °C. Однако 2020 год выдался более прохладным, особенно в мае и июне, с более низкими температурами по сравнению с другими годами.

Таблица 5 – Погодные условия в Алматинской области с 2018 по 2022 годы

Год испытаний		Май	Июнь	Июль	Август	Сентябрь
2018	Т, °C	16.34	22.32	25.15	24.43	17.34
	Осадки, мм	124.90	28.70	32.30	43.50	18.90
2019	Т, °C	17.00	22.26	26.96	24.88	18.55
	Осадки, мм	39.3	72.7	22.6	67.7	67.2
2020	Т, °C	19.52	15.50	24.00	16.70	9.70
	Осадки, мм	73.50	42.60	38.10	43.70	21.20
2021	Т, °C	19.48	22.44	26.00	24.10	20.53
	Осадки, мм	81.6	20.9	22.8	27.2	1.6
2022	Т, °C	18.96	24.34	26.51	22.69	10.94
	Осадки, мм	145.40	35.90	16.20	8.20	42.20

Уровень осадков значительно различался в разные годы. В 2018 и 2022 годах количество осадков было относительно высоким в мае, а затем резко снизилось в последующие месяцы. Напротив, в 2019 году распределение было более сбалансированным, с умеренными осадками в течение всего лета. В 2020 году в целом было относительно мало осадков, тогда как в 2021 году наблюдался устойчивый, но более низкий характер осадков. В августе и сентябре в большинстве лет в этот период наблюдалось меньше осадков. Зафиксированные

среднесуточная температура и количество осадков в КазНИИЗиР представлены в Таблице 5.

Полевые опыты и сбор урожая проводились с мая по сентябрь по методу Доспехова [143]. Данный метод широко используется в полевых исследованиях для стандартизации сбора данных по урожайности. Образцы высаживались в четыре ряда на делянке, расстояние между растениями – 25 см, между рядами – 50 см, длина ряда – 1 м. Повторности высевались ежегодно по принципу блоковой рандомизации, в неконтролируемых погодных условиях, с использованием орошения и без каких-либо вмешательств (удобрений, гербицидов, фунгицидов и т.д.). Сорт-стандарт по региону высевался в числе общей коллекции.

Фенологические наблюдения проводились по методу Фера и Кэвинесса [32]. Сбор урожая коллекции сои на экспериментальных полях проводился на стадии R8, стадии полной спелости 95% бобов растения (Рисунок 11).



Рисунок 11 – Коллекция сои в период формирования урожая на экспериментальных полях КазНИИЗиР (слева), растение сои на стадии полного созревания (справа), 2018 год

2.2.1.1. Оценка компонентов урожайности

Коллекция сои была проанализирована по компонентам урожайности, таким как количество продуктивных узлов (КПУ, шт.), количество семян на растение (КСР, шт.), масса тысячи семян (МТС, г), масса семян с растения (МСР, г) по методу Корсакова [144]. Среднее значение трех случайных растений с одной делянки использовали для фенотипической оценки признака каждого отдельного генотипа.

2.2.1.2. Измерение параметров семян

Оценка размеров семян сои изучаемой коллекции проводилась в 2018 и 2019 годах. Такие параметры как длина, ширина и толщина семян сои измерялись ручным методом 10 рандомных семян с образца использованием штангенциркуля (Рисунок 12).

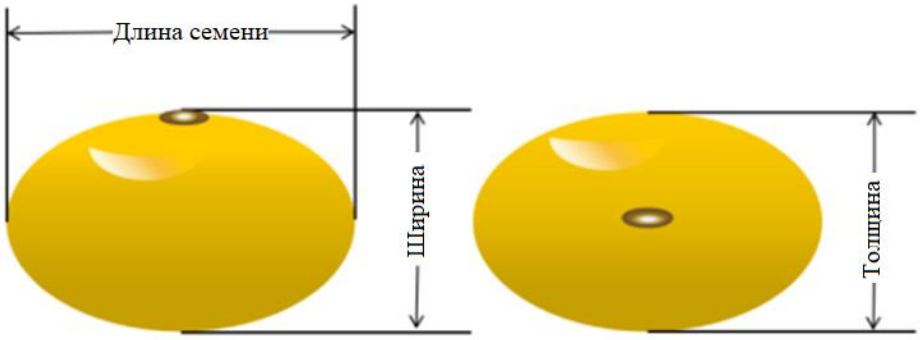


Рисунок 12 – Параметры размеров семян сои [145].

Среднее арифметическое значение 10 показаний высчитывалось для каждого образца. Данные средних значений за два года также использовались в полногеномном анализе и представлены в Приложении 2. На Рисунке 13 представлена часть разнообразия изучаемой коллекции сои.



Рисунок 13 – Разнообразие семян коллекции сои

2.2.1.3. Оценка растрескиваемости стручков сои

Оценка растрескивания стручков сои проводилась в 2019 и в 2021 годах, где изучаемая коллекция была проанализирована также по компонентам урожайности (КПУ, КСР, МТС и МСР) для выявления связи с растрескиваемостью стручков сои. Данные растрескиваемости стручков сои собирались путем визуальной оценки стручков на стадии полной зрелости (R8),

по шкале от 1 до 5, где 1 ≤ 1-20 %, 2 ≤ 21-40 %, 3 ≤ 41-60 %, 4 ≤ 61-80 % и 5 ≤ 81-100 % растреснутых стручков и, по итогу, рассыпанных семян сои (Приложение 2).

2.2.1.4. Определение признаков качества семян сои

Семена коллекции сои были собраны для анализа показателей качества – содержание протеина и масла. Анализ проводился согласно ГОСТу 18057-64 «Семена масличные. Методы определения масличности» и ГОСТу 53600-2009 «Семена масличные, жмыхи и шроты. Определение влаги, масла, протеина и клетчатки методом спектроскопии в ближней инфракрасной области» [146-147] в лаборатории биотехнологии, физиологии, биохимии и оценки качества продукции ТОО «Казахский научно-исследовательский земледелия и растениеводства» с использованием анализатора зерна NIRS DS2500 (FOSS, Хиллерод, Дания) с калибровкой, предоставленной производителем. Анализатор измеряет пробы в широком диапазоне длин волн от 700 до 2500 нм. Для обеспечения надежности и точности результатов анализа, были соблюдены все рекомендации производителя. Многочисленными международными исследованиям были подтверждены высокоточные результаты по содержанию протеина и масла в семенах, полученные с помощью данного анализатора [148].

2.2.2. Выделение ДНК и полногеномное ресеквенирование образцов

Образцы ДНК выделяли из проростков сои. Протокол выделения ДНК сои состоял из этапов измельчения проростков, лизис клеток буфером на основе CTAB, депротеинизация хлороформом и осаждение ДНК изопропанолом. Последующий контроль качества и чистоты ДНК от примесей белковых и сторонних органических соединений проводился измерением соотношения оптической плотности при A260/A280 и A260/A230 на спектрофотометре (Nanodrop 2000, THERMO SCIENTIFIC), визуализация продуктов выделения проводилась с помощью гель-электрофореза.

Полногеномное ресеквенирование (WGRS, whole-genime resequencing) проводилось в несколько этапов. Для начала – создание библиотеки данных 252 линий коллекции сои с помощью набора Illumina TruSeq в соответствии с инструкциями производителя. Далее, было проведено полногеномное секвенирование с использованием системы Illumina HiSeq X Ten с глубиной покрытия более чем 10x [56;152]. Результаты чтения геномов всей коллекции были ресеквенированы с помощью биоинформатической программы BWA (Burrows-Wheeler Aligner) сопоставлением с референсным геномом китайского сорта сои Zhonghuang 13, что позволило получить 4 923 660 SNP маркеров [153]. Последовательности, близкие к вставкам или делециям, были выровнены с помощью функции IndelRealigner с помощью GATK [154]. SNP маркеры, вставки и делеции идентифицировали с помощью программного обеспечения GATK и SAMtools. Работы по полногеномному секвенированию/ресеквенированию мировой коллекции сои были проведены в коллаборации с кафедрой Школы естественных наук Университета Гуанчжоу, Гунчжоу, Китай.

Для обеспечения надежности последующего анализа был проведен контроль качества генетических данных. Были исключены SNP маркеры с долей пропущенных генотипов более 10%, а также маркеры с частотой минорного аллеля (MAF, minor allele frequece) менее 10 %, и в общей сложности было отобрано 44 385 маркеров SNP, которые далее использовались для определения структуры популяции и связи маркера с признаком. Подобные строгие фильтры QC необходимы для минимизации возможных ошибок и повышения воспроизводимости результатов в исследованиях ассоциаций по всему геному [155-156]. Высококачественные данные генотипирования являются основой для корректного анализа структуры популяции и выявления ассоциаций между маркерами и признаками.

2.2.3. Анализ структуры популяции мировой коллекции сои.

Структура популяции изучаемой коллекции сои изучалась с помощью метода Neighbor-joining tree [157], который широко используется для построения дендрограмм на основе генетических расстояний. Для расчета расстояния между генотипами использовался алгоритм Identity-by-state (IBS) [158] на основе данных 44385 SNP маркеров 252 генотипов. Для визуализации древа использовалась программа iTOL [159], представившая коллекцию сои в трех отдельных кластерах (красная, зеленая и синяя ветви) на основе их генетических связей. Также для анализа структуры популяции использовалась программа STRUCTURE, которая помогает минимизировать влияние стратификации популяции на результаты ПГАА [160]. В STRUCTURE использовался метод Байесовской кластеризации с применением оценки Марковской цепи Монте-Карло (MCMC). Была применена модель ADMIXTURE, допускающая обмен генетической информации между популяциями. Количество гипотетических групп в диапазоне от k = 1 до 10 оценивалось с использованием 100 000 начальных итераций (burn-in iterations), за которыми следовали 100 000 репликации MCMC с количеством итераций, равным 3. Дельта K(ΔK), которая определяет наиболее подходящее количество кластеров (K) в наборах данных были определены согласно Evanno et al.[161] с использованием веб-программы STRUCTURE HARVESTER (версия 0.6.94) [162]. Файлы, содержащие результаты STRUCTURE, использовались в качестве входных файлов для STRUCTURE HARVESTER, где использовались настройки по умолчанию для определения наиболее вероятного количества кластеров. Полученные значения далее были преобразованы в матрицу структуры популяции (Q). Матрица kingship (K), описывающая наиболее вероятное сходство каждого аллеля между образцами, была создана с помощью TASSEL (версия 5.0) [163].

Определение параметров LD был выполнен для оценки уровня сцепления между молекулярными маркерами, связанных с изучаемыми признаками. Степень LD можно оценить с помощью статистического параметра r^2 [164], представляющий собой квадрат коэффициента корреляции аллельных состояний двух полиморфных локусов. Параметр r^2 дает прямую меру доли дисперсии в локусе признака, которую можно предсказать с помощью маркера. Значение r^2, равное 1, указывает на полное неравновесное сцепление, при которой аллели в

двух локусах идеально связаны, а значение r^2, равное 0, означает равновесие сцепления, то есть аллели несцеплены. В генетических исследованиях принято оценивать LD, определяя расстояние, на котором r^2 падает ниже определенного порога, на уровне 0,1 когда сцепленность двух маркеров значительно ослабевает [165].

2.2.4. Полногеномный анализ ассоциаций хозяйственно-ценных признаков.

Ассоциации между генотипическими и фенотипическими данными по изучаемому признаку были идентифицированы моделью MLMM с пакетом GAPIT (версия 3) в RStudio (версия 2024.04.2) [119, 166]. Фенотипические данные включали показатели, определяющие продуктивность сои и качество семян, а генотипические данные, как уже было описано выше, были представлены данными ресеквенирования мировой коллекции сои – 44385 SNP-маркеров. Для оценки корректности модели и эффективности поправки, основанной на матрицах K и Q, использовались Q-Q графики, сравнивающие наблюдаемые и ожидаемые P-значения в логарифмическом масштабе.

База данных SoyBase [92] использовалась для поиска генов-кандидатов для выявленных ассоциаций маркер-признак с референсным геномом сои Cv. Zhonghuang 13, версия 1 (Glyma.Zh13. gnm1). Величина LD применялась в качестве границы для идентификации генов-кандидатов рядом с обнаруженными значимыми ассоциациями маркер-признак.

Всем идентифицированным ассоциациям маркер-признак были присвоены названия согласно их признаку, названию лаборатории и номера хромосомы в геноме сои, например *q.NFN.ipbb.7* (NFN, number of fertile nodes – КПУ, количество продуктивных узлов). Если на одной хромосоме были обнаружены более одной ассоциации, нумерация продолжалась порядковым номером с увеличением позиции маркера.

2.2.5. Анализ гаплотипов значимых ассоциаций, разработка KASP маркеров и KASP генотипирование.

Анализ гаплотипов был проведен для оценки связи между значимыми ассоциациями и фенотипической изменчивостью хозяйственно-ценных признаков. Анализ гаплотипов позволяет выявить более сложные и информативные генетические структуры, чем анализ отдельных SNP-маркеров, поскольку учитывает взаимодействие между локусами и их совместное влияние на фенотипические признаки [167]. Во многих исследованиях анализ гаплотипов выполняется с использованием специализированного программного обеспечения [168], в настоящем исследовании был применен упрощенный подход с использованием Microsoft Excel, что обусловлено ограниченным числом значимых ассоциаций изучаемых признаков и удобством визуального контроля группировок. Процедура проводилась путем группировки образцов на основе аллельных комбинаций SNP и сравнения средних фенотипических значений. Для каждого признака были выбраны SNP, идентифицированные как значимо ассоциированные в ПГАА, с наивысшими значениями $-\log10(P)$,

объясняющие наибольшую долю фенотипической дисперсии по каждому признаку.

Затем фенотипические данные сравнивались по группам гаплотипов для оценки влияния различных аллельных вариантов на проявление признака. Средние значения показателей признаков для каждого гаплотипа были рассчитаны. Выявленные гаплотипы, ассоциированные с благоприятными показателями признаков сои, могут быть использованы в селекционных программах для создания новых конкурентоспособных сортов сои с улучшенными показателями качества и продуктивности.

Для проверки эффективности маркер-признаковых ассоциаций, выявленных в результате полногеномного анализа и связанных с хозяйственно-ценными признаками, была проведена разработка KASP-маркеров [138]. На основании нуклеотидных последовательностей SNP-маркеров был осуществлён дизайн двух аллель-специфических прямых праймеров и одного общего обратного праймера для каждого маркера [108]. Дизайн праймеров осуществлялся с использованием онлайн-инструмента Primer3 и внутренних рекомендаций компании LGC Genomics (Beverly, MA, США), где в последующем заказывались готовые праймеры.

Генотипирование по технологии KASP выполнялось согласно стандартному протоколу, разработанному LGC Genomics. Реакции ПЦР проводились в 96-луночных планшетах с применением KASPMix, который содержит флуоресцентные метки (FAM и HEX) для аллельного разделения.

Флуоресцентная детекция проводилась с использованием прибора Fluoroskan Ascent (Thermo Fisher Scientific, США). Результаты генотипирования анализировались с помощью специализированного программного обеспечения KlusterCaller™ (LGC Genomics), позволяющего автоматически формировать кластеры и визуализировать разделение образцов по генотипам (гомозиготные по аллелю 1, гомозиготные по аллелю 2 и гетерозиготные).

Полученные KASP-маркеры, обладающие различием аллелей и высокой воспроизводимостью, могут быть использованы для дальнейшей селекционной работы, в том числе в программах MAS.

2.2.6. Статистическая обработка результатов

Для каждого года экспериментов рассчитывались показания всех признаков - минимальные, максимальные, средние, средние значения, стандартная ошибка и диапазон. Для оценки вклада генотипа (G, genotype), среды (E, environment) и их взаимодействия, G×E, в фенотипическую изменчивость признаков был применен анализ ANOVA (analysis of variance), позволяющий определить основную структуру влияющих факторов [32;149]. Анализ для каждого года с влиянием G, E и взаимодействия G x E рассчитывался в программном обеспечении SPSS 22 [150]. Расчет коэффициента наследуемости (h^2) дает возможность оценить степень генетического контроля изучаемых признаков. Коэффициент h2 оценивался по следующей формуле:

$$h^2 = VG/ (V_G + V_E + V_{GxE} + V_{error}),$$

где V_G – генотипическая дисперсия, V_E – дисперсия среды, V_{GxE} – дисперсия, обусловленная взаимодействием генотипа и среды, а V_{error} – дисперсия ошибки. Значение h^2 интерпретируется как доля фенотипической изменчивости, обусловленной генетическими факторами, и может варьироваться от 0 (полное влияние среды) до 1 (полное влияние генетики) [151]. Корреляционный анализ позволяющий установить взаимосвязь между хозяйственно-ценными и качественными показателями семян сои изучаемых признаков также проводился в SPSS 22.

Вычисление r^2 для LD проводилось с помощью программы TASSEL (версия 5.0), визуализация – в RStudio (версия 2024.04.2) [166]. При проведении ПГАА значение P $<1x10^{-4}$ использовалось в качестве порога значимости для идентифицированной ассоциации маркер-признак.

3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ

3.1. Описание хозяйственно-ценных признаков изучаемой коллекции сои в условиях Алматинской области

3.1.1. Компоненты урожайности

Вся коллекция сои была оценена по компонентам урожайности, таких как количество продуктивных узлов, количество семян на растение, масса тысячи семян и урожайности с растения. За пять лет эксперимента компоненты урожайности показали значительное разнообразие значений (Таблица 6).

Таблица 6 – Фенотипическая изменчивость компонентов урожайности коллекции сои за пять лет исследования

Признак	Показатели	2018	2019	2020	2021	2022	Среднее значение за 2018-2022
КПУ, шт	Мин	6.9	4.9	4.6	4.7	4	8.32
	Макс	41.2	31.3	34.9	52	60	30.19
	Сред ±станд. ошибка	19.59 ±0.41	14.45 ±0.34	17.45 ±0.35	15.87 ±0.47	24.77 ±0.66	18.42 ±0.29
КСР, шт.	Мин	15	8.4	6.7	8	5.8	17.04
	Макс	105.1	87.7	84.2	126	182.5	81.07
	Сред ±станд. ошибка	48.075 ±1.12	36.51 ±0.98	39.06 ±0.86	40.27 ±1.33	66.15 ±1.84	46.03 ±0.83
МТС, г	Мин	124	18.9	11.85	118	114	106.12
	Макс	287	310.9	361.1	276	283.2	227.87
	Сред ±станд. ошибка	175.97 ±1.51	125.97 ±3.5	145.17 ±4.99	173.52 ±1.61	187.59 ±1.83	161.56 ±1.49
МСР, г	Мин	3	1.5	0.4	0.3	1.12	3.82
	Макс	44.15	39.3	40.65	55.6	68.33	30.22
	Сред ±станд. ошибка	16.54 ±0.46	10.50 ±0.36	10.68 ±0.34	9.11 ±0.45	20.45 ±0.73	13.45 ±0.31
КПУ – количество продуктивных узлов, КСР – количество семян с растения, МТС – масса тысячи семян, МСР – масса семян с растения							

Среднее значение компонентов урожайности составляли 18,42 шт. для КПУ, 46,03шт для КСР, 161,56гр для МТС и 13,45 гр. для МСР (Таблица 5), однако

разброс данных свидетельствует о сильном влиянии погодных условий, почвенных характеристик и генетических особенностей сортов. Стоит отметить, что 2022 год был наиболее урожайным, когда коллекция показала максимальные значения компонентов урожайности, что может быть связано с благоприятными погодными условиями. Наименьшие показатели для КПУ, КСР и МТС (14,45 шт., 36,51шт и 125,97гр) были зафиксированы в 2019 году, а для МСР – в 2021 году (9,11гр) [169-171]. На Рисунке 14 изображены гистограммы распределения показателей компонентов урожайности в условиях Алматинской области с сравнением показателей сорта-стандарта «Жансая».

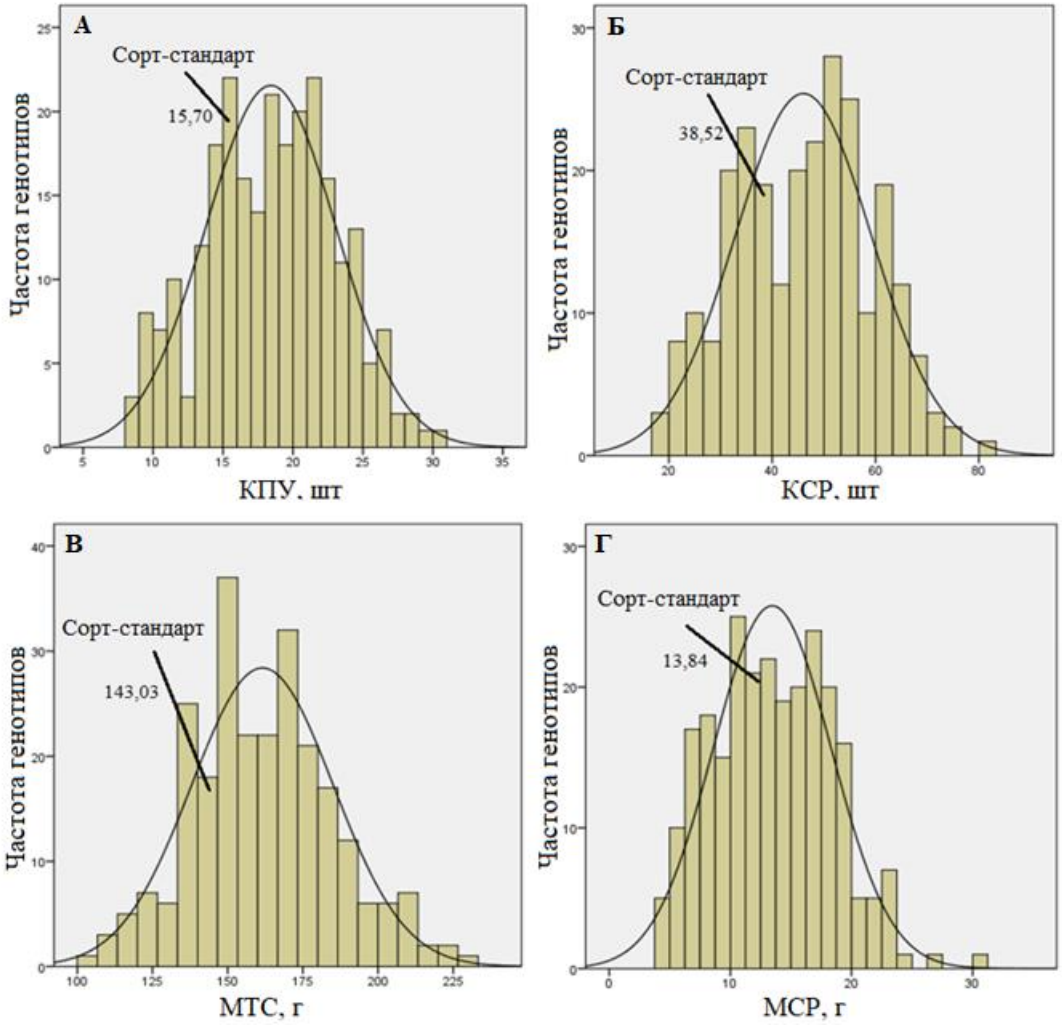


Рисунок 14 – Гистограмма нормального распределения компонентов урожайности сои в условиях Алматинской области, черная линия – показатель сорта-стандарта

Анализ компонентов урожайности выявил наиболее продуктивные сорта среди южно-азиатских образцов (6877, Hei feng 50), американских (Agassis) и европейских (Черемош, Веста). Казахстанские сорта также отличились высокими показателями по урожайности, особенно по массе 1000 семян. Такие сорта как Алматы, Мисула и Роза вошли в число наиболее продуктивных сортов. Изучаемая коллекция сои в условиях Алматинской области превзошла

показатели компонентов урожайности сорта стандарта «Жансая», что говорит о большом генетическом потенциале в поиске новых факторов высокого урожая.

3.1.1.1. Результаты ANOVA анализа для компонентов урожайности

Результаты ANOVA, включая значения вероятности P-value и коэффициента наследования (h2) компонентов урожайности, представлены в Таблице 7.

Признаки, связанные с продуктивностью растения, в большей степени зависели от факторов среды. Например, количество продуктивных узлов имело низкую наследуемость ($h^2 = 0,12$), а основной вклад в изменчивость вносила E ($P < 2x10^{-16}$). Значимость взаимодействия G × E ($P = 0.000432$) говорит о разной реакции сортов на условия выращивания. Количество семян с растения имело аналогичный характер изменчивости ($h^2 = 0,12$), при этом влияние E было доминирующим.

Таблица 7 – Анализ ANOVA и коэффициент наследования (h2) компонентов урожайности изучаемой коллекции сои

Признак	Фактор	Степень свободы	Сумма квадратов	Среднее квадратов	Вероятность P	h^2
КПУ	G	4	2260	565	$4.99x10^{-09}$ ***	0.12
	E	4	16120	4030	$<2x10^{-16}$ ***	
	G x E	16	2105	132	0.000432***	
	Ошибка	1219	60794	50		
КСР	G	4	20589	5147	$7.91x10^{-11}$***	0.12
	E	4	143650	35913	$<2x10^{-16}$ ***	
	G x E	16	17503	1094	0.000114***	
	Ошибка	1219	60794	50		
МТС	G	4	73688	18422	$4.14x10^{-07}$ ***	0.10
	E	4	635799	158950	$<2x10^{-16}$ ***	
	G x E	16	149338	9334	$7.42x10^{-09}$ ***	
	Ошибка	1221	2499565	2062		
МСР	G	4	4168	1042	$3.02x10^{-15}$***	0.15
	E	4	23148	5787	$<2x10^{-16}$ ***	
	G x E	16	2658	166	$4.7x10^{-05}$***	
	Ошибка	1220	66570	55		

КПУ – количество продуктивных узлов, КСР – количество семян с растения, МТС – масса тысячи семян, МСР – масса семян с растения
h^2 – коэффициент наследования признака, G – генотип, E – среда/год, GxE – взаимодействие генотипа и среды/года, *** – высокая значимость вероятности, *– значимость вероятности

Анализ литературы свидетельствует о том, что продуктивные характеристики сои, такие как число узлов и семян, имеют сильную зависимость от окружающих условий, что подтверждается полученными результатами [69]. Масса 1000 семян оказалась одним из наиболее изменчивых признаков, где G был наименее значимым ($h^2 = 0,10$), а E имела максимальное влияние ($P < 2x10^{-16}$). Существенная значимость взаимодействия G × E ($P = 7.42x10^{-09}$) подтверждает высокую степень адаптивной изменчивости сортов. Признак MCP также зависел в большей степени от окружающей среды ($P < 2x10^{-16}$), но имел умеренную наследуемость ($h^2 = 0,15$), что указывает на возможность селекционного улучшения данного показателя.

3.1.2. Параметры размеров семян – длина, ширина и толщина

Для каждого экспериментального года и средние данные за два года показатели размеров семян сои – длина, ширина и толщина – всей коллекции и для сорта-стандарта представлены в Таблице 8.

Таблица 8 – Фенотипическая изменчивость размеров семенах за два исследования и показатели сорта-стандарта Жансая

Показатели всей коллекции	Признаки размеров семян сои, мм								
	ДС 2018	ШС 2018	ТС 2018	ДС 2019	ШС 2019	ТС 2019	ДС сред.	ШС сред.	ТС сред.
Сорт-стандарт Жансая	6,91	6,53	5,35	7,02	6,65	5,49	6,97	6,59	5,51
Макс.	9,67	7,92	6,37	9,26	8,25	7,06	9,31	8,09	6,71
Мин.	6,45	5,41	4,49	6,43	5,31	4,48	6,58	5,57	4,64
Среднее± станд. Ошибка	7,65± 0,03	6,48± 0,02	5,40± 0,02	7,64± 0,03	6,42± 0,02	5,39± 0,02	7,65± 0,03	6,45± 0,02	5,40± 0,02
Длина семян – ДС, ширина – ШС, толщина – ТС									

Образцы из Восточной Европы показали наивысшие значения по длине семян сои, например, российский сорт Фора отличился не только по признаку длины, но и по ширине семян, 8,68мм и 7,64 мм, соответственно (Приложение 2). По признаку ширины и толщины семян образцы из Восточной Азии показали наивысшие показатели, например китайский сорт Dong Doe показал наибольшую ширину и толщину семени, 8,09 и 6,71 мм, соответственно. По признаку толщины семян максимальные показатели также принадлежат Восточноазиатским сортам, следуя за Dong Doe, еще несколько китайских образцов сои показали наивысшие результаты (Рисунок 15, Приложение 2).

Наименьшие показатели по трем параметрам были зафиксированы в основном у восточноевропейских сортов сои. Например, российский сорт

Быстрица, показал одни из минимальных размеров по длине и толщине, 6,63 и 4,80 мм, соответственно. Линия СН 147020-1 также была зафиксирована с наименьшими показателями по ширине и толщине, 5,87 и 4,62 мм (Приложение 2).

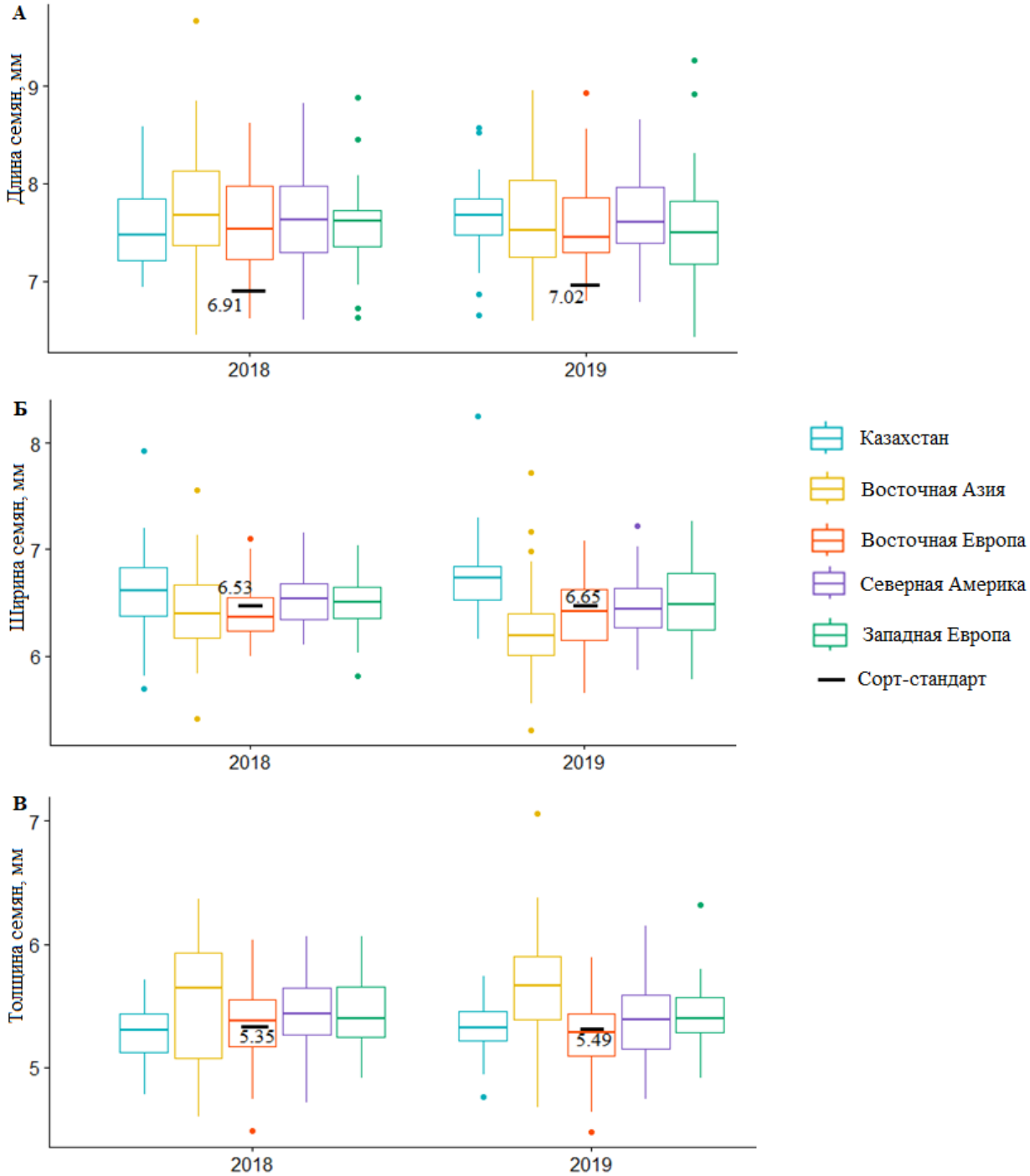


Рисунок 15 – График распределения признаков размера семян в зависимости от происхождения образцов коллекции (Лист 1)

При сравнении всей коллекции с сортом-стандартом, результаты измерений показали, что по длине семян, к примеру в 2018 году, вся коллекция оказалась намного длиннее, со средним значением 7,65 мм, тогда как у сорта Жансая – 6,91мм. По признаку ширины и толщины семян показатель сорта стандарта в оба

года примерно совпадает со средними данными размеров семян коллекции (Рисунок 16, Приложение 2).

В целом, вся изучаемая коллекция сои показала различные данные по параметрам семян по сравнению с сортом стандартом, что подчеркивает огромный потенциал коллекции для поиска новых генетических факторов, влияющих на размеры семян сои (Рисунок 16).

3.1.2.1. Результаты ANOVA анализа для размеров семян

Результаты ANOVA анализа размеров семян сои показало, что различные признаки размера семян контролируются как генетическими, так и факторами окружающей среды и их взаимодействием (Таблица 9).

Таблица 9 – Анализ ANOVA размеров семян сои и растрескиваемости стручков сои

Признак	Фактор	Df	Sum Sq	Mean Sq	F value	Pr (> F)
ДС	G	4	0.70	0,18	0.708	0.587
	E	1	0.08	0,08	0.316	0.574
	G x E	4	0.86	0,22	0.872	0.480
	Res.	472	116.72	0,25		
ТС	G	4	5.83	1,458	13.930	$9.47x10^{-11}$ ***
	E	1	0.05	0,0517	0.494	0.4827
	G x E	4	1.03	0,2582	2.467	0.0442 *
	Res.	472	49.40	0,1047		
ШС	G	4	8.01	2,0037	18.520	$3.77x10^{-14}$ ***
	E	1	0.45	0,4465	4.127	0.04277 *
	G x E	4	1.58	0,3955	3.655	0.00604 **
	Res.	472	51.07	0,1082		
РС	G	4	28.3	7.067	3.794	0.00476 **
	E	1	0.1	0.050	0.027	0.86968
	G x E	4	3.2	0.799	0.429	0.78773
	Res.	488	909.1	1.863		

ДС – длина семян, ШС – ширина семян, ТС – толщина семян сои, РС – растрескиваемость стручков;
Уровни значимостей: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Результаты ANOVA показали, что на размеры семян существенно влияет и G, и E, а также их взаимодействие, в особенности для параметров ширины и толщины семян. Так, на ширину семени все три фактора оказали статистически значимое влияние. Это предполагает многофакторную природу признака, чувствительного как к G, так и к E и их взаимодействию. Признак толщины семян также имел значимое влияние G ($P < 9.47x10^{-11}$), вместе с тем, G×E также оказалось значимым ($P = 0.044$), что свидетельствует о различной реакции генотипов на условия выращивания.

3.1.3. Растрескиваемость стручков сои

В 2019 и 2021 экспериментальных годах изучаемая коллекция сои была проанализирована по признаку растрескиваемости стручков. Результаты наблюдение в полевых условиях Алматинской области показали, что подавляющее большинство коллекции устойчиво к растрескиванию стручков (Рисунок 16). Однако 23 образца в 2019 году и 21 образец в 2021 году оказались восприимчивы к растрескиванию, среди них 12 образцов повторно полностью или почти полностью подверглись растрескиванию стручков (Приложение 2, Рисунок 16) [172-173].

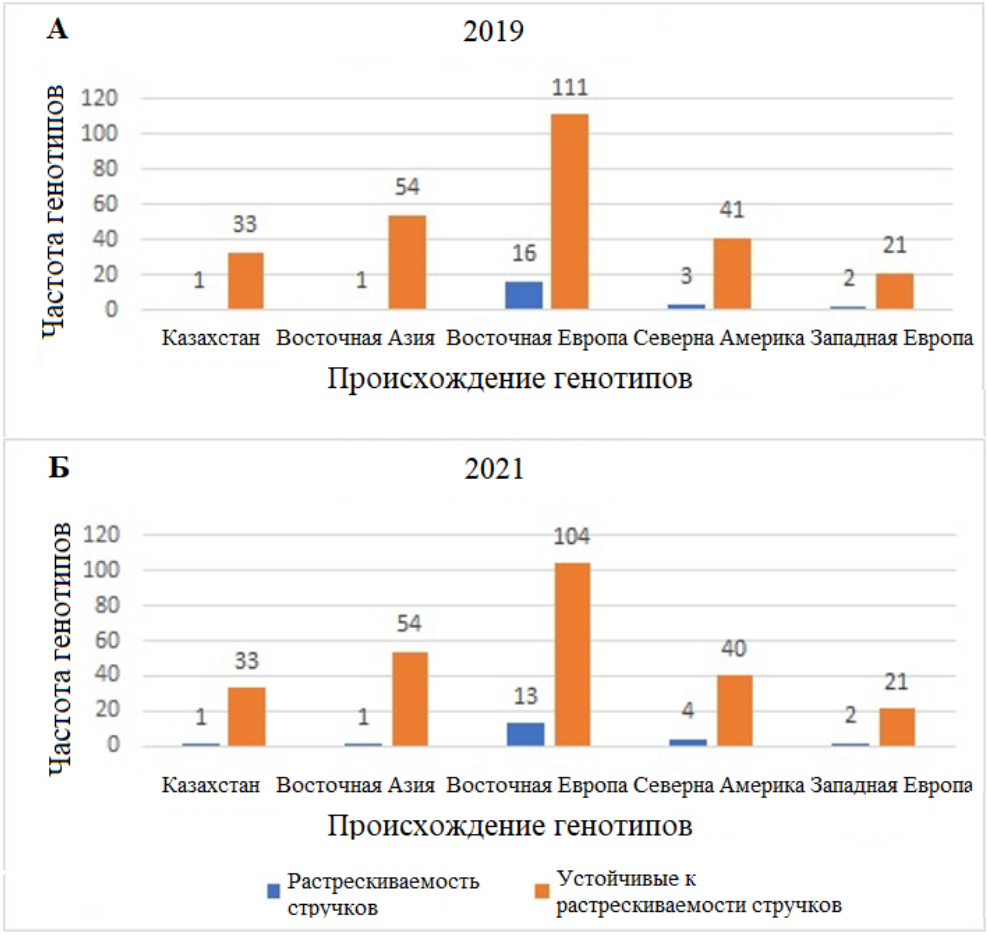


Рисунок 16 – Результаты полевых наблюдений признака растрескиваемости стручков (PC, pod dehiscence) сои (a – 2019, b – 2021 года).

Сорта, восприимчивые к растрескиванию в условиях Алматинской области в двух годах эксперимента, оказались преимущественно восточноевропейского происхождения, такие как Ясельда (Беларусь), Arctic, Nawiko (Польша), Toury (Чехия), ПЭП27, Алтом, Белор, Фора, Октябрь 70 (Россия). Также с признаком растрескивания стручков были зафиксированы два североамериканских сорта Accord (Канада) и Carola (США). Из казахстанских сортов только два образца, Жалпаксай (2021г.) и линия 308/1(2019г.), были зафиксированы с растрескаваемостью стручков в одном из годов эксперимента. Сорт стандарт Жансая показал устойчивость к растрескиванию во все года эксперимента. В Таблице 10 представлен список образцов сои с растрескиваемостью стручков во

все года эксперимента и показателями компонентов урожайности в условиях Алматинской области в сравнении с сортом-стандартом.

Таблица 10 – Список образцов сои с растрескиваемостью стручков во все года эксперимента и показателями компонентов урожайности в условиях Алматинской области в сравнении с сортом-стандартом Жансая.

№	Сорт/линия	Страна	Происхождение	МСР, г	шкала РС, балл
1	Ясельда	Беларусь	Восточная Европа	6,7	5
2	Аккорд	Канада	Северная Америка	8,9	5
3	Арктик	Польша	Восточная Европа	5,5	5
4	Навико	Польша	Восточная Европа	5,4	5
5	ПЭП27	Россия	Восточная Европа	4,7	5
6	Алтом	Россия	Восточная Европа	9,1	5
7	Белор	Россия	Восточная Европа	10,3	5
8	Карола	США	Северная Америка	8,2	5
9	Таури	Чехи	Восточная Европа	12,0	5
10	Фора	Россия	Восточная Европа	15,4	5
11	Октябрь	Россия	Восточная Европа	13,4	5
12	**Жансая**	**Казахстан**	**Казахстан**	**13,8**	**1**
Примечание: МСР – масса семян с растения, г; РС – растрескиваемость стручков, баллы 1 ≤ 0-20 %, 2 ≤ 21-40 %, 3 ≤ 41-60 %, 4 ≤ 61-80 % и 5 ≤ 81-100 % растреснутых стручков сои					

3.1.3.1. Результаты ANOVA анализа растрескиваемости стручков сои

Для признака растрескиваемости влияние G было весьма значимым (P = 0,00476), что указывает на наличие существенной генетической изменчивости среди изучаемой коллекции (Таблица 11).

Таблица 11 – Анализ ANOVA растрескиваемости стручков сои

Признак	Фактор	Df	Sum Sq	Mean Sq	F value	Pr(>F)
РС	G	4	28.3	7.067	3.794	0.00476 **
	E	1	0.1	0.050	0.027	0.86968
	G x E	4	3.2	0.799	0.429	0.78773
	Res.	488	909.1	1.863		
ДС – длина семян, ШС – ширина семян, ТС – толщина семян сои, РС – растрескиваемость стручков; Уровни значимости: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1						

Фактор G объяснил значительную часть общей дисперсии (F = 3,794), что говорит о том, что растрескивание стручка находится под сильным генетическим контролем, что согласуется с предыдущими выводами [65]. Однако, влияние E и G×E не было значимым (P = 0,86968 и P = 0,78773, соответственно), что говорит о том, что признак устойчивости к растрескиванию стручка относительно стабильный в различных условиях окружающей среды. Эта стабильность в различных условиях подчеркивает потенциал эффективного отбора и использования эффективных аллелей в селекционных программах, что также наблюдалось в других исследованиях [80].

3.1.4. Описание признаков качества семян сои.

Признаки содержания протеина и масла в семенах, определяющие пищевую и кормовую ценность сои, показали широкий разброс в изучаемой коллекции. В 2018 году было зарегистрирован наибольший средний показатель содержания масла среди всех испытанных лет (21,01%), возможно погодные условия благоприятно воздействовали накоплению масла в семенах сои (Таблица 12). Содержание масла варьировало от 17.3% до 24.4%, наибольшее накопление масла (>24%) было отмечено в канадских сортах Buster и Emerson.

Таблица 12 - Фенотипическая изменчивость признаков качества семян коллекции сои за пять лет исследования

Признак	Показатели	2018	2019	2020	2021	2022	Среднее значение за 2018-2022
СП, %	Мин	35.85	35.09	37.85	35.1	36.04	37.07
	Макс	48.59	48.22	48.6	49.3	47.7	48.36
	Сред ±станд. ошибка	41.67 ±0,19	40.95 ±0.19	43.29 ±0.18	41.96 ±0.19	42.33 ±0.15	42.12 ±0.16
СМ, %	Мин	16.45	15.87	16.6	16.2	16.89	17.05
	Макс	24.38	23.96	23.65	24.3	24.08	23.78
	Сред ±станд. ошибка	21.01 ±0.10	20.87 ±0.11	20.70 ±0.11	20.97 ±0.09	20.31 ±0.09	20.70 ±0.08
СП – содержание протеина в семенах, СМ – содержание масла в семенах							

Рекордным же по содержанию протеина стал 2020 год – 43,29%. В 2020 году наблюдались более высокие температуры летом (22,3-26,9°C) и относительно равномерное распределение осадков, что, вероятно, благоприятно способствовало накоплению белка.

На Рисунке 17 представлен график, отражающий изменения содержания протеина и масла в соевых образцах различных географических регионов происхождения за период 2018-2022 годов.

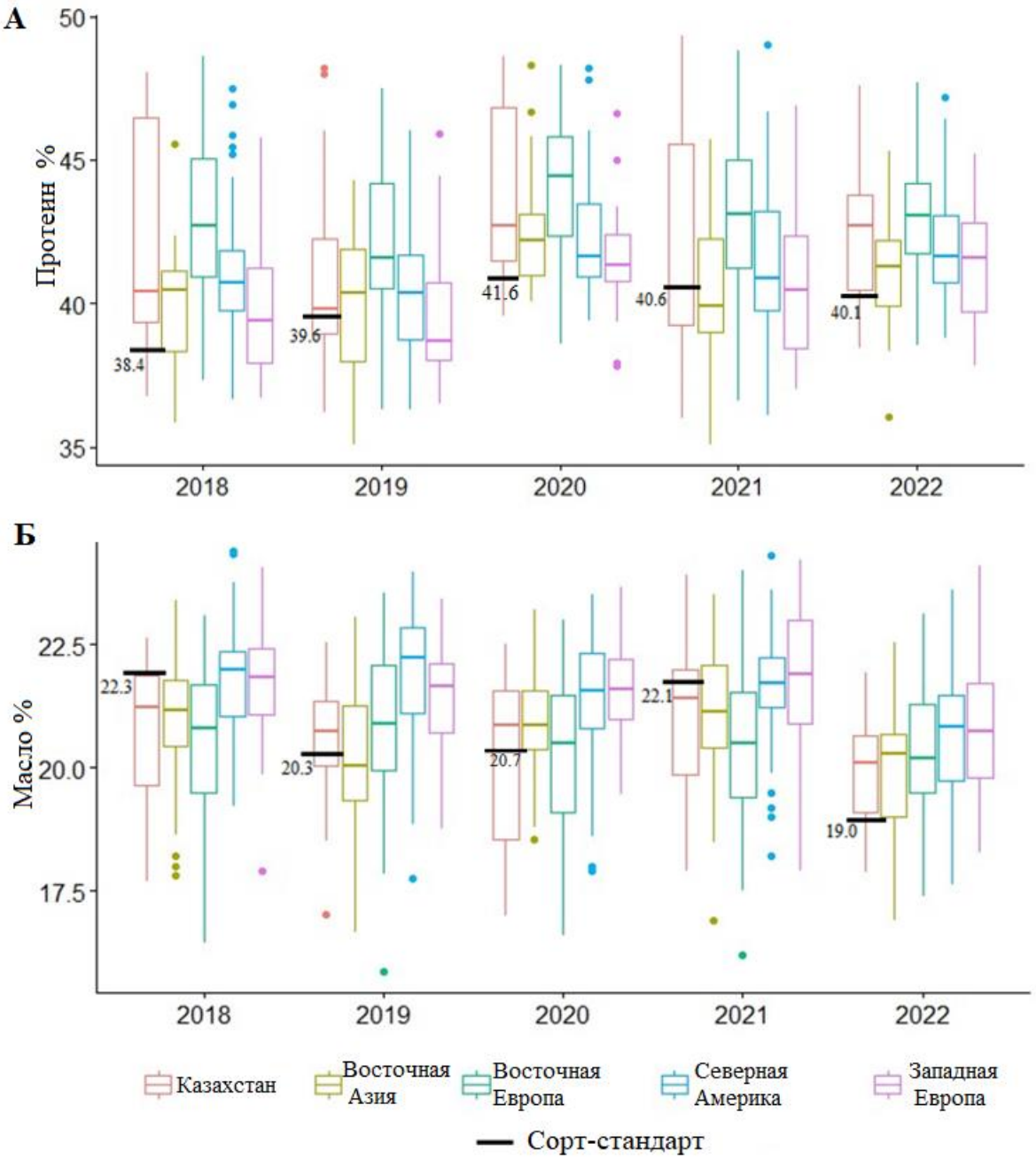


Рисунок 17 – Бокс плот график распределения признаков качества семян в зависимости от происхождения образцов коллекции

Данные на графике сгруппированы по годам, различными цветами обозначены регионы: Центральная Азия, Восточная Азия, Восточная Европа, Северная Америка и Западная Европа. Чёрной линией выделен сорт-стандарт Жансая. Во всей коллекции высокие показатели по содержанию протеина в основном преобладали в сортах восточноевропейского происхождения и Казахстана. Эти регионы показали наибольшую медиану и широкий диапазон значений протеина и масла, что свидетельствует о высокой генетической ценности образцов. Разброс значений среди исследованных образцов указывает на влияние как генетических, так и внешних факторов окружающей среды на формирование признаков.

Вся изучаемая коллекция сои по сравнению со стандартом Жансая показала превышающие данные по компонентам урожайности, что еще раз подчеркивает огромный потенциал коллекции для поиска новых генетических факторов урожайности (Рисунок 17, 18).

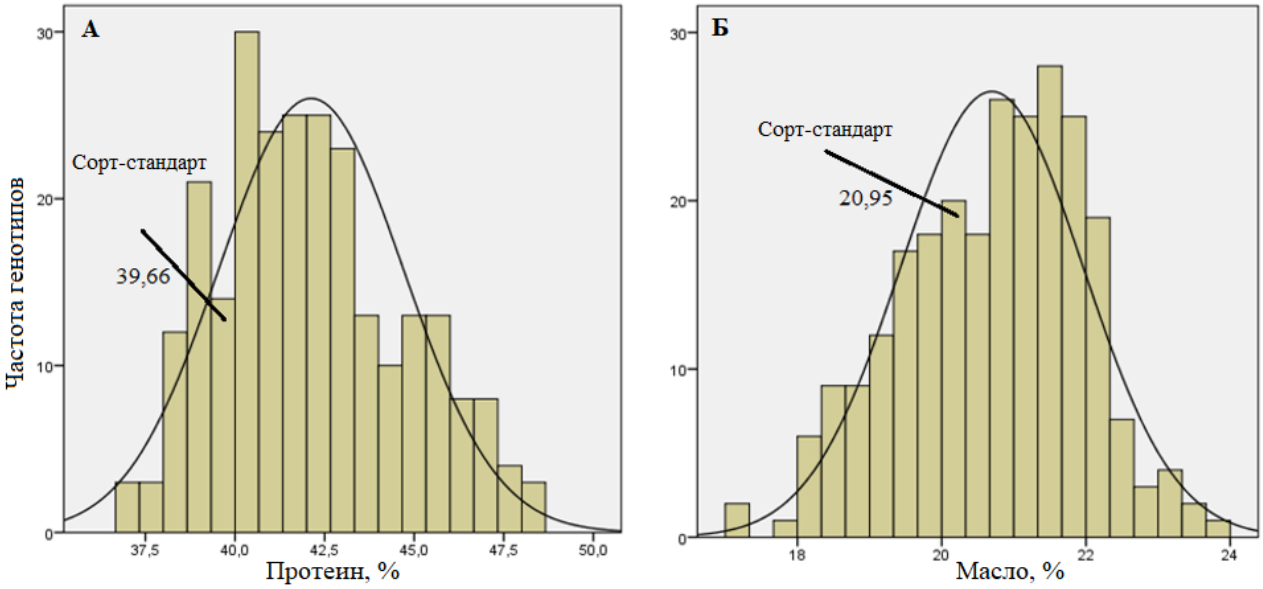


Рисунок 18 – Гистограмма распределения изучаемых признаков качества в коллекции с указанием значения сорта-стандарта (черные линии)

Средний показатель содержания протеина всей коллекции за пять лет оказался выше (42,12%) по сравнению Жансая (39,66%). По содержанию же масла, средний показатель всей коллекции сравним с показателем Жансая – 20,70% и 20,95% соответственно (Рисунок 19, Таблица 9).

3.1.4.1. Результаты ANOVA анализа признаков качества семян сои

Результаты исследования с использованием ANOVA, свидетельствуют о том, что содержание протеина в семенах являются генетически обусловленным (h^2 = 0,64), что подтверждает высокую значимость G ($P < 2x10^{-16}$). Также, E оказывала сильное влияние ($P < 2x10^{-16}$) на признак СП. Однако, взаимодействие G × E было незначимым (P = 0.858), что указывает на стабильность генотипов по данному признаку в различных условиях. Аналогичная ситуация наблюдалась для признака СМ (h^2 = 0,65), при этом значимость взаимодействия G × E была умеренной (P = 0.0106), что указывает на различную адаптивность сортов в зависимости от условий выращивания. Эти данные подчеркивают важность генетического контроля на влияние показателей вариабельности признаков качества семян (Таблица 13).

Таблица 13 – Анализ ANOVA и коэффициент наследования (h2) признаков качества семян сои

Признак	Фактор	Df	Сумма квадратов	Среднее квадратов	Вероятность P	h^2
СП	G	4	1091	272.87	$<2x10^{-16}$***	0.64
	E	4	585	146.13	$<2x10^{-16}$***	
	G x E	16	63	3.95	0.858	
	Ошибка	1051	6542	6.22		
СМ	G	4	182.0	45.49	$<2x10^{-16}$ ***	0.65
	E	4	74.2	18.55	$1.61x10^{-07}$***	
	G x E	16	62.9	3.93	0.0106*	
	Ошибка	1050	2057.9	1.96		

СП – содержание протеина, СМ – содержание масла; Df (degree of freedom) – степень свободы, Sum Sq – сумма квадратов, Mean Sq – средний квадрат отклонений, F value – критерий Фишера, Pr (>F) – вероятность
h^2 – коэффициент наследования признака, G – генотип, E – среда/год, GxE – взаимодействие генотипа и среды/года, *** – высокая значимость вероятности, *- значимость вероятности

Таким образом, G значительно влияет на вариабельность содержания протеина и масла, что позволяет вести целенаправленную селекцию на улучшение этих признаков на основе использования ценных сортов и линий с различным происхождением. В то же время, ключевые признаки продуктивности, такие как урожайность и масса 1000 семян, в большей степени зависят от E, что требует более тщательного подбора сортов для конкретных условий выращивания.

3.2. Перспективные сорта и линии по изучаемым признакам в мировой коллекции сои

Анализ коллекции сои по хозяйственно-ценным признакам в условиях Алматинской области, позволил выявить, что многие сорта и линии сои различного происхождения превышают показатели местного сорта-стандарта Жансая. Это свидетельствует об высоком потенциале генетического фонда изучаемой коллекции. В Таблице 14 представлены сорта с наивысшими показателями по каждому признаку, а также показатели стандарта.

Таблица 14 – Список перспективных сортов и линии сои по признакам качества и урожайности по сравнению с сортом-стандартом по Алматинской области

№	Признак	Сорт/линия	Происхождение	Значение
1	2	3	4	5
1	СП	186/1	Казахстан	48,36%
2	СП	126/1	Казахстан	48,13%
3	СП	Рана	Восточная Европа	48,10%
4	СП	Злата	Восточная Европа	47,81%
5	СП	Picket K6298	Северная Америка	47,78%
6	СП	173/1	Казахстан	47,68%
7	СП	261/1	Казахстан	47,61%
8	СП	Смена	Восточная Европа	47,24%
9	СП	391	Казахстан	47,05%
10	СП	209/1	Казахстан	46,88%
11	**СП**	**Жансая**	**Казахстан**	**39,66%**
1	СМ	Emerson	Северная Америка	23,78%
2	СМ	Sepia	Западная Европа	23,66%
3	СМ	1044	Восточная Азия	23,40%
4	СМ	Buster	Северная Америка	23,28%
5	СМ	Sava	Западная Европа	23,26%
6	СМ	Астра	Восточная Европа	23,07%
7	СМ	Venera	Западная Европа	23,03%
8	СМ	ОАО Wallace	Северная Америка	22,91%
9	СМ	Аннушка	Восточная Европа	22,71%
10	СМ	Надежда	Восточная Европа	22,69%
11	**СМ**	**Жансая**	**Казахстан**	**20,95%**
1	КПУ	1674	Восточная Азия	30,19
2	КПУ	Linkoln	Северная Америка	29,51
3	КПУ	1022	Восточная Азия	28,78
4	КПУ	Nhat 11	Восточная Азия	28,52
5	КПУ	Веста	Восточная Европа	27,50
6	**КПУ**	**Жансая**	**Казахстан**	**15,70**
1	КСР	Даная	Казахстан	81,07
2	КСР	Isidor	Западная Европа	74,84
3	КСР	Веста	Восточная Европа	74,06
4	КСР	Быстрица 2	Восточная Европа	72,93
5	КСР	Linkoln	Северная Америка	72,62
6	**КСР**	**Жансая**	**Казахстан**	**38,52**
1	МТС	Xinjiang D10-51	Восточная Азия	227,87 г
2	МТС	Nikko	Западная Европа	223,07 г
3	МТС	Память ЮГК	Казахстан	220,26 г
4	МТС	Dong doe 027	Восточная Азия	219,95 г
5	МТС	1054	Восточная Азия	216,52 г
6	**МТС**	**Жансая**	**Казахстан**	**143,03 г**
1	МСР	Xinjiang D10-51	Восточная Азия	30,22 г
2	МСР	Веста	Восточная Европа	26,71 г
3	МСР	Хабаровская 4429	Восточная Европа	23,95 г
4	МСР	Воеводжанка	Западная Европа	23,45 г

Продолжение Таблицы 14

1	**2**	**3**	**4**	**5**
5	МСР	Isidor	Западная Европа	23,40 г
6	**МСР**	**Жансая**	**Казахстан**	**13,84 г**
1	ДС	Алтом	Восточная Европа	9,31 мм
2	ДС	Fiskeby5	Западная Европа	9,07 мм
3	ДС	551	Восточная Азия	8,82 мм
4	ДС	350/1	Казахстан	8,77 мм
5	ДС	Фора	Восточная Европа	8,68 мм
6	**ДС**	**Жансая**	**Казахстан**	**6.97 мм**
1	ШС	Dong doe 027	Восточная Азия	8,09 мм
2	ШС	Фора	Восточная Европа	7,64 мм
3	ШС	551	Восточная Азия	7,32 мм
4	ШС	Wilstar 194	Северная Америка	7,19 мм
5	ШС	Safrfna	Западная Европа	7,15 мм
6	**ШС**	**Жансая**	**Казахстан**	**6.59 мм**
1	ТС	Dong doe 027	Восточная Азия	6,71 мм
2	ТС	551	Восточная Азия	6,64 мм
3	ТС	Bey jiang 91	Восточная Азия	6,35 мм
4	ТС	Hei lun 48	Восточная Азия	6,21 мм
5	ТС	Safrfna	Западная Европа	6,19 мм
6	**ТС**	**Жансая**	**Казахстан**	**5.42 мм**
СП – содержание протеина, СМ– содержание масла, КПУ – число продуктивных узлов, КСР – число семян с растения, МТС – масса тысячи семян, МСР – масса семян с растения, ДС – длина семян, ШС – ширина семян, ТС – толщина семян сои; Жансая – сорт-стандарт сои Алматинской области				

Стоить отметить, что среди 10 перспективных сортов сои, отличившихся высокими значениями по протеину, шесть сортов являются сортами казахстанской селекции. Также, местные сорта, такие как Даная и Память ЮГК отличались высоким количеством семян на растение и массе 1000 семян, соответственно.

По признаку растрескивания большая часть коллекции оказалась устойчивой по данному признаку и сохраняли свой урожай в целых стручках (Рисунок 16, Приложение 2). Среди казахстанских сортов только линия 308/1 и Жалпаксай были зафиксированы с растрескаваемостью стручков в одном из годов эксперимента (Приложение 2).

3.3. Анализ корреляции хозяйственно-ценных признаков

Для всех изученных признаков качества и урожайности сои был проведен анализ корреляции по Пирсону. На Рисунке 20 представлена корреляционная матрица, которая показывает степень связи между 10 хозяйственно-ценными признаками сои (КПУ, КСР, МТС, МСР, ДС, ШС, ТС, РС, СП, СМ.). Корреляционные коэффициенты варьируются от -1 (полная отрицательная связь) до +1 (полная положительная связь). Красные цвета обозначают положительную корреляцию, а синие – отрицательную корреляцию (Рисунок 19).

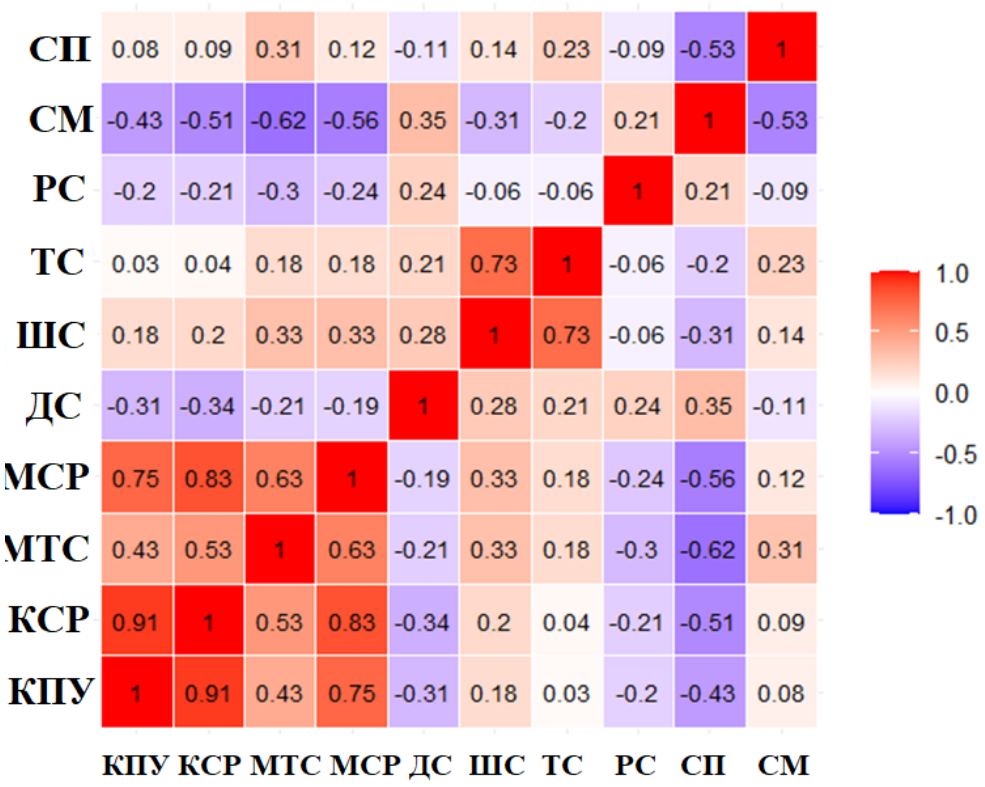


Рисунок 19 – Корреляция 10 хозяйственно-ценных признаков по Пирсону

Корреляционный анализ показал существенную отрицательную связь между содержанием протеина (СП) и масла (СМ) (r = -0.53), что подтверждает известный антагонизм между этими компонентами в семенах сои [97], Рисунок 19]. Также выявлена высокая положительная корреляция урожайности (МСР) с количеством семян (КСР, r = 0.83) и числом продуктивных узлов (КПУ, r = 0.75), что подтверждает, что ключевыми факторами продуктивности являются генеративные органы растения [174]. Кроме того, умеренная положительная связь между урожайностью и массой 1000 семян (МТС, r = 0.63) говорит о том, что увеличение размеров семян также способствует повышению урожайности [175].

Признак МТС показал слабую положительную корреляцию с содержанием масла (r = 0.31) и значимую отрицательную связь с СП (r = -0.62), что указывает на тенденцию снижения массы семян при высоком белковом содержании [122]. Также выявлена очень высокая корреляция между КСР и КПУ (r = 0.91), что ожидаемо, так как большее количество продуктивных узлов напрямую увеличивает количество семян на растении [176].

Размеры семян, частности ШС и ТС, показали слабовыраженную или умеренную корреляцию с компонентами урожайности. Длина же семян, отрицательно коррелировала с компонентами урожайности (r = -0,19 до r = -0.31).

Признак растрескиваемости показал отрицательную корреляцию со всеми компонентами урожайности, что также наблюдалось в предыдущих исследованиях [76]. В работах Kang 2020 показано, что раскрытие стручков сои может привести к значительной потере урожая, поскольку приводит к

преждевременному рассеиванию семян. Это отрицательно влияет как на количество, так и на качество урожая [177]. Стоит отметить также что с длиной семян РС показал положительную корреляцию (r = 0.24)

Признак СП показывает отрицательную корреляцию с шириной и толщиной семян (r = -0,31 и r = -0,20, соответственно). Содержание масла же, наоборот, имеет положительную связь с указанными параметрами (r = 0,14 и r = 0,23 соответственно). Современные исследования показывают, что увеличение размера семян сои тесно связано с повышением содержания масла [178]. Это объясняется тем, что в процессе селекции происходил отбор генотипов с крупными семенами, которые, как правило, содержат больше масла. Одним из примеров является ген *GmSWEET10a*, мутации в котором приводят к увеличению размера семян и содержания масла, но сопровождаются снижением содержания белка. Это свидетельствует о существовании общих метаболических путях, где увеличение одного компонента может происходить за счет уменьшения другого [179].

Анализ корреляции также показал, что ширина семян сои вносит больший вклад в урожайность (r = 0,18 для КПУ, 0,20 для КСР, 0,33 для МТС и 0,33 для МСР). Shillin Luo в своей работе показал, что ширина семян имела наибольшее коэффициент корреляции с массой семян в большинстве экспериментов, что говорит о том, что выбор более широких семян может повысить урожайность [145].

Интересно, что в нашем исследовании устойчивость к растрескиванию семян имела положительную значимую корреляцию с содержанием протеина в семенах сои. Подобная взаимосвязь наблюдалось и в других исследованиях, например Bingjie Tu 2025, в работе со 140 сортами сои обнаружил положительную корреляцию содержания протеина в семенах и устойчивостью к растрескиванию стручков [180]. Возможно, анатомические особенности стручков, такие как содержание лигнина, имеют решающее значение для раскрытия. Более высокое содержание лигнина, часто связанное с повышенным содержанием протеина в семенах, может повысить прочность стручка и устойчивость к раскрытию [181].

3.4. Ресеквенирование данных мировой коллекции сои

В растениеводстве, включая генетику и селекцию сои (*Glycine max* (L.) Merr.), WGRS широко используется для изучения структуры популяций, картирования признаков и поиска ассоциаций важных сельскохозяйственных признаков. Особое значение метод имеет при построении всего генома, ПГАА-анализах и выявлении генов-кандидатов желаемых признаков для молекулярной селекции [56].

В рамках настоящего исследования был выполнен WGRS 252 образцов мировой коллекции сои, включающей отечественные сорта, а также сорта и линии из других 4 географических происхождений. ДНК из проростков семян сои выделяли СТАБ методом, и не менее 5 мкг использовалось для создания библиотеки геномных данных всей коллекции на платформе Illumina TruSeq.

Секвенирование геномов проводилось на платформе Illumina HiSeq X Ten с глубиной покрытия более 10× для каждого образца, что обеспечило высокую достоверность данных. Сопоставление с референсным геномом китайского сорта сои Zhonghuang 13 (Glyma.Zh13. gnm1) позволило получить массив данных, включающий 4 923 660 однонуклеотидных полиморфизмов (SNP), который в дальнейшем прошел строгий контроль качества и 44385 высококачественных SNP-маркеров использовались для анализа популяционной структуры, оценки разнообразия и ПГАА. Также данные ресеквенирования позволили выявить различия между группами образцов, и стать основой для ПГАА и разработки функциональных молекулярных маркеров.

3.4.1. Анализ плотности покрытия SNP маркеров генома сои

В исследованиях по сое использование высокой плотности SNP-маркеров позволило выявить значимые ассоциации между генотипом и фенотипом. Обнаруженные значимые ассоциации между маркерами и признаками были использованы для создания молекулярных маркеров и их применения в селекционных программах [182].

В настоящем исследовании использовались данные 44385 SNP маркеров, что позволило получить более плотное покрытие генома, повышающее разрешающую способность генетического анализа и облегчая идентификацию геномных регионов, связанных с ключевыми хозяйственно-ценными признаками. Данных распределения SNP маркеров по геному сои представлены на Рисунке 20.

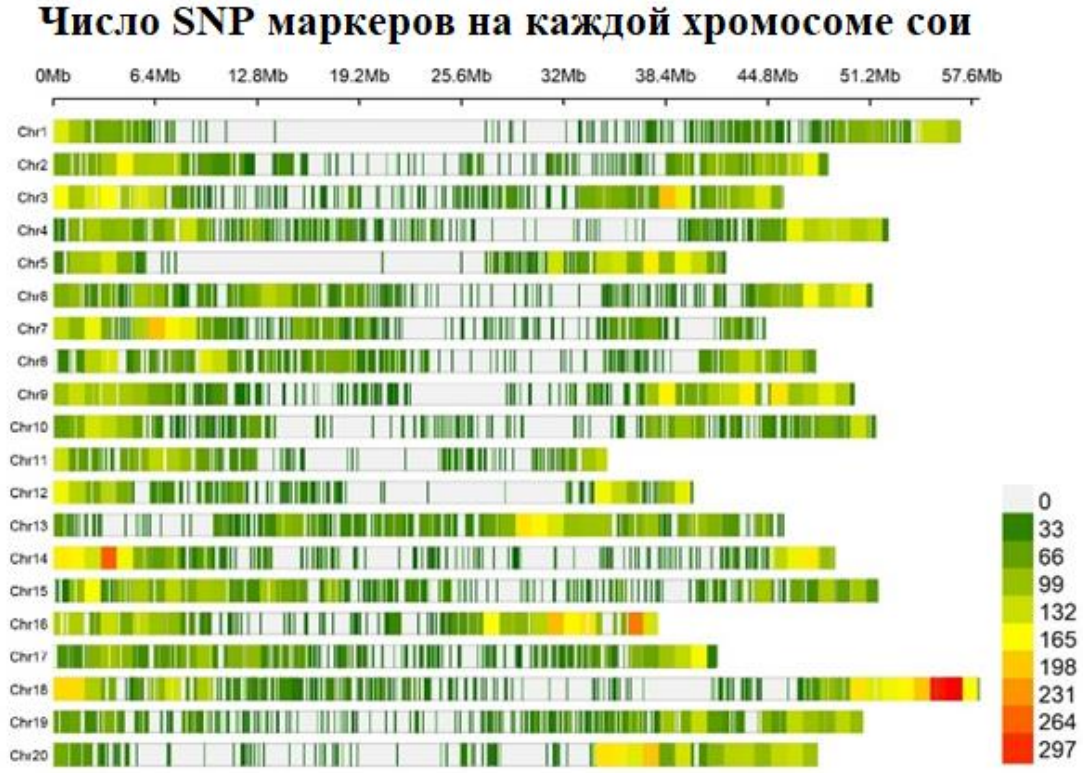


Рисунок 20 – Плотность 44385 SNP маркеров по геному сои

Представленный график иллюстрирует распределение SNP в пределах окна размером 1 Мб по хромосомам генома сои. Каждая горизонтальная строка представляет хромосому с цветовой кодировкой плотности SNP. Ось X обозначает размер хромосом в мегабазах (Мб). Вариабельность плотности SNP очевидна по хромосомам: в некоторых областях наблюдаются высокие концентрации SNP (желто-красные зоны), а в других – бледные (области с низкой плотностью, белые области) (Рисунок 20).

Заметной закономерностью является более низкая плотность SNP в центральных областях хромосом. Это явление объясняется наличием центромерных и прицентромерных областей, которые характеризуются низкой частотой рекомбинаций, высокой плотностью мобильных элементов (TE) и более низким содержанием генов. В этих консервативных последовательностях ограничен полиморфизм и накопление SNP, как и отражено на графике бледной окраской в центре хромосом, также такой эффект наблюдался в предыдущих исследованиях полногеномного секвенирования сои и других видов растений [183].

Наиболее высокая плотность расположения отобранных SNP-маркеров наблюдается ближе к дистальным концам хромосом, где скорость рекомбинации обычно повышена. Эта закономерность характерна для самоопыляющихся видов растений, таких как соя, где рекомбинация происходит чаще в богатых генами, эухроматических регионах. Повышенное количество SNP в этих дистальных областях хромосом имеет важное значение для генетического картирования и отбора с помощью MAS, поскольку эти области часто содержат QTL, связанные с важными хозяйственно-ценными признаками [183].

Неравномерное распределение SNP в геноме сои имеет важное значение для генетических исследований и программ селекции. Отчетливая кластеризация SNP в некоторых хромосомах, таких как Chr 3, Chr 14, Chr16 и Chr18, предполагает наличие “hotspots” или горячих точек генетического разнообразия, возможно это может быть связано с эволюционной селекцией или внедрением генов других сортов сои с течением времени (Рисунок 20). Наблюдаемое распределение SNP отражает особенности геномной структуры и частоты рекомбинации. Эти результаты согласуются с предыдущими геномными исследованиями сои и дают ценную информацию для исследований в области молекулярной селекции и функциональной геномики [184].

3.4.2. Оценка LD SNP маркеров в геноме изучаемой популяции сои

Оценка LD играет решающую роль в полногеномных исследованиях ассоциаций сельскохозяйственных культур, помогая идентифицировать генетические маркеры, связанные со сложными признаками [185-186]. Всего для анализа LD было использовано 44 385 SNP (Рисунок 21).

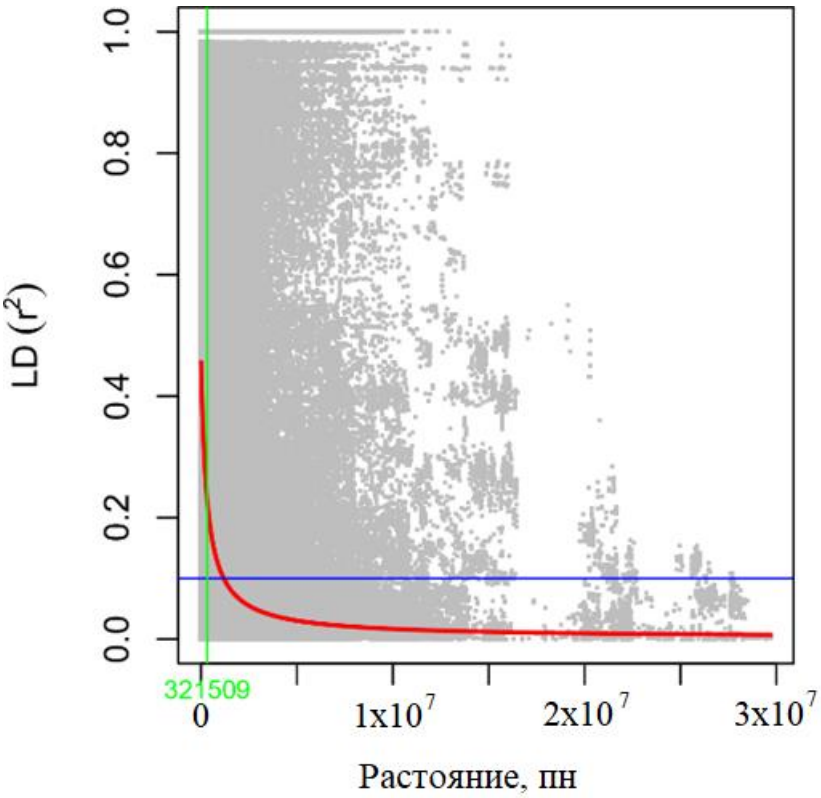


Рисунок 21 – График снижения LD по мере увеличения расстояния между маркерами (в п.н., красная линия – снижение r^2). Порог LD (r^2 = 0,1) обозначен синей линией; расстояние, при котором r^2 снижается до 0,1 показано зеленой линией. Серые точки на графике – пары SNP маркеров

Анализ LD, проведенный в рамках исследования, показывает, что расстояние, на котором коэффициент корреляции (r^2) снижается до 0,1, составляет примерно 321 509 пар оснований (п.н.) (Рисунок 21). Данное значение LD в нашем исследовании предполагает, что сцепленность значимых ассоциации между аллелями сохраняется на расстояние до 321509 пн. Наблюдаемое расстояние снижения LD в нашем исследовании находится в пределах диапазона, указанного в исследовании Hyten et al. (2007), где анализировались несколько популяций сои и значение LD варьировало от 100 тыс. п.н. (у дикой сои) до 574 тыс. п.н. (в культивируемых сортах) [164]. В исследовании подчеркивается, что история культивирования и селекции существенно влияет на структуру LD сои, при этом культивируемые сорта демонстрируют гораздо более крупные блоки LD [187]. Эта популяционно-зависимая LD также наблюдается у других видов растений, таких как фасоль обыкновенная и маш [188-189]. Графики снижения LD для каждой хромосомы генома сои проиллюстрированы в Приложении 3.

3.4.3. Анализ генетической структуры изучаемой популяции сои

Для понимания генетической структуры внутри изучаемой коллекции сои была построена дендрограмма взаимосвязей образцов коллекции на основе данных SNP маркеров (Рисунок 22).

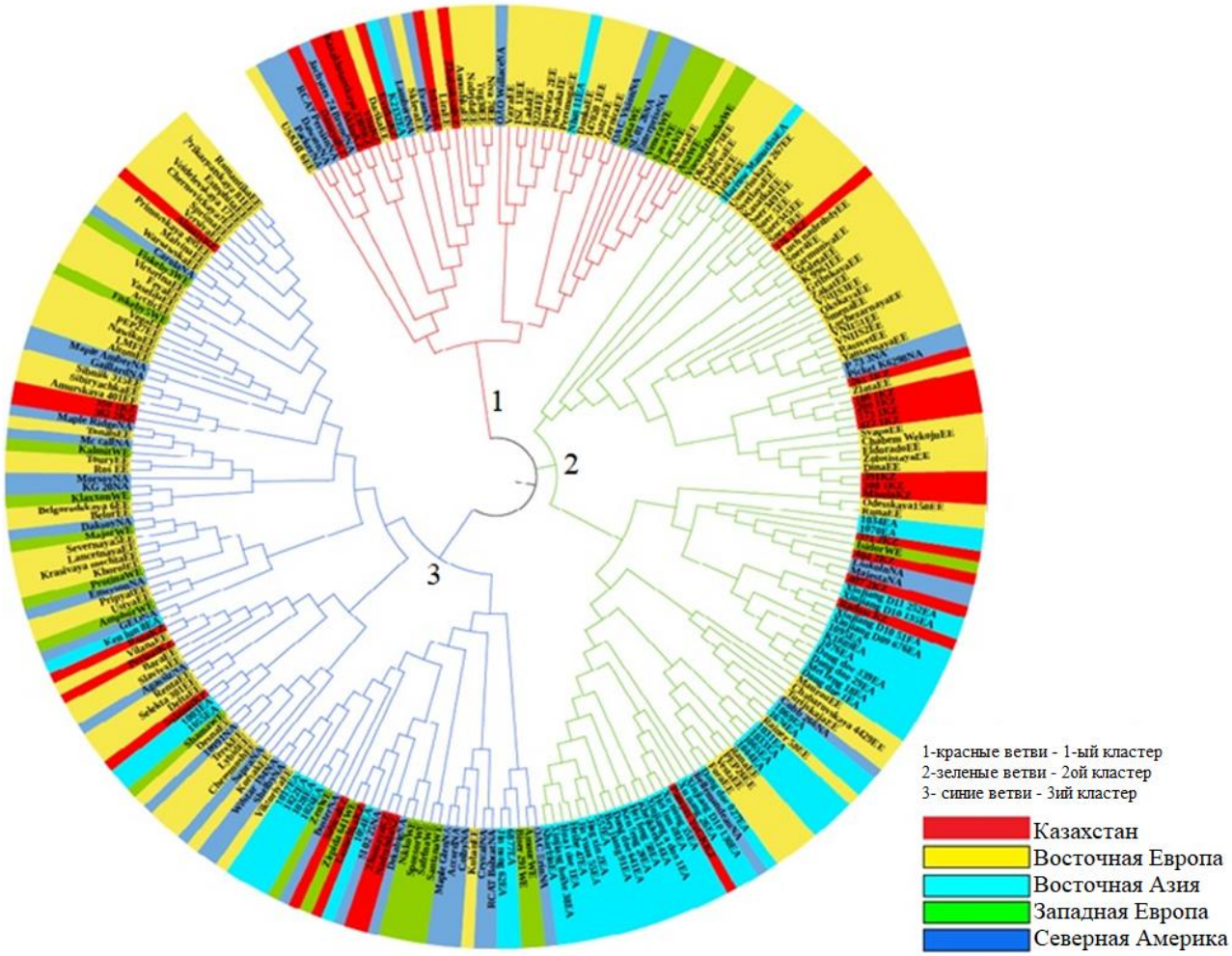


Рисунок 22 – Дендрограмма изучаемой коллекции сои на основе данных 44385 SNP маркеров 252 генотипов коллекции сои (метод Neighbor-joining tree)

Внешнее кольцо с цветовой кодировкой показывает географическое происхождение каждого образца (Казахстан, Восточная Европа, Восточная Азия, Западная Европа и Северная Америка) (Рисунок 22). Первый кластер (красный цвет) состоит в основном из восточноевропейских и казахстанских образцов, что указывает на общий генетический фон. Второй кластер (зеленый) выглядит более разнообразным, так как в нем представлены образцы из Восточной Европы и Азии и Казахстана, что позволяет предположить общий поток генов между этими регионами. Третий кластер (синий) является наиболее генетически разнообразным, с большим вкладом из Северной Америки, Восточной Европы и Восточной Азии, что позволяет предположить влияние международных программ селекции (Рисунок 22). Эта филогенетическая кластеризация предполагает, что историческая селекция и отбор сформировали генетическое разнообразие генотипов сои в разных регионах.

Результаты, полученные с помощью пакетных приложений STRUCTURE и STRUCTURE HARVESTER, показали, что K = 3 было оптимальным значением для изучаемой коллекции, поскольку иллюстрировала максимальную дельту K при K = 3 [161] (Рисунок 23 А). Выявление оптимального количества кластеров равное трем указывает на наличие трех генетически различных субпопуляций [189-190]. Данные результаты подчеркивают генетическую дифференциацию образцов сои из Восточной Европы, Северной Америки и Азии, демонстрируя их историю селекции и адаптацию к местной среде.

На Рисунке 23 (график Б, В) первая субпопуляция (Q1 или Cluster 1) в основном состоит из восточноевропейских образцов (81%), что указывает на общий генетический фон, вероятно, сформированный местными программами селекции и адаптацией к умеренному климату. Вторая субпопуляция также содержит более высокую долю образцов из Восточной Европы, однако увеличивается доля казахстанских образцов (19%) и образцов Восточной Азии (13%), что позволяет предположить исторический генетический обмен между этими регионами.

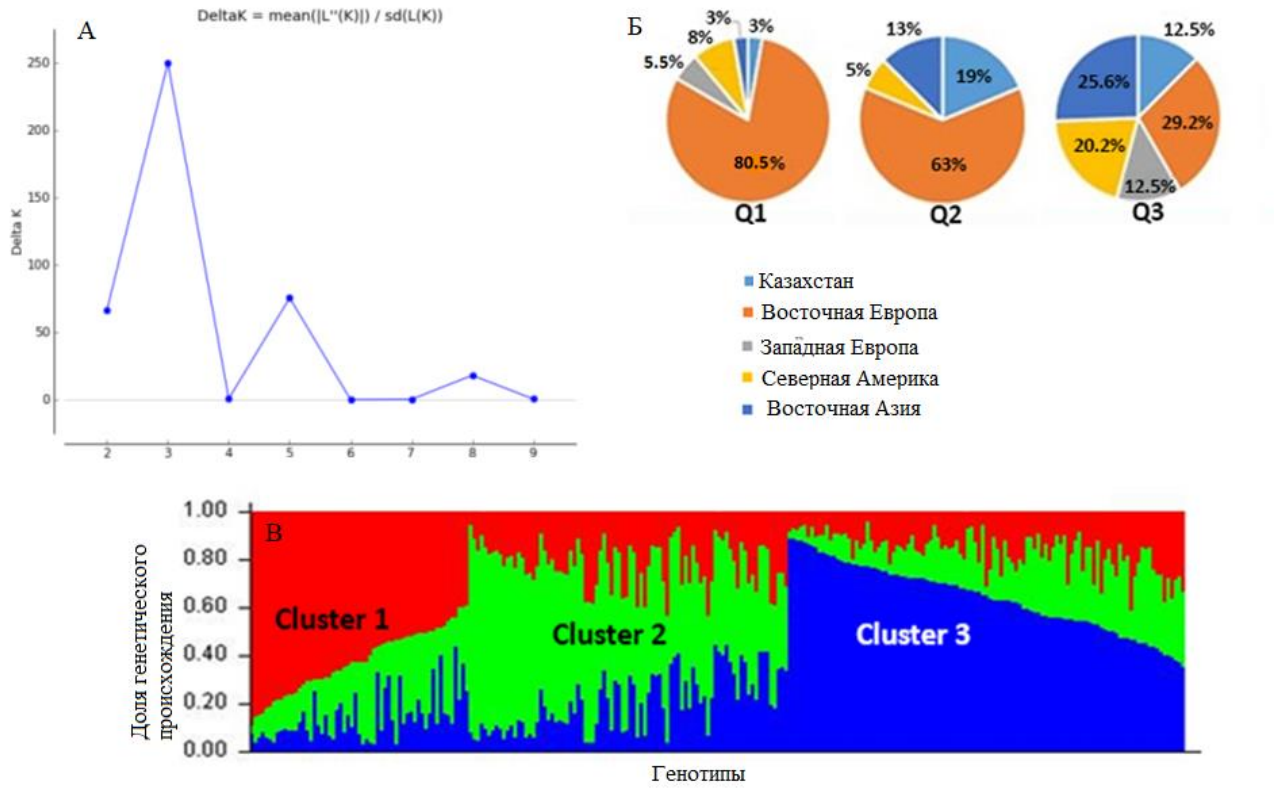


Рисунок 23 – Результаты анализа структуры популяции. А – определение оптимального числа генетических кластеров с использованием среднего L(K) ±SD и ΔK (Evano); Б – диаграмма на основе Q-матрицы; В – график распределения изучаемой популяции сои по трем кластерам

Третья субпопуляция имеет наиболее разнообразный состав: почти равные доли представителей Восточной Европы (29,2%), Восточной Азии (25,6%) и Северной Америки (20,2%), и малые доли казахстанских и западноевропейских

представителей (по 12,5%). Эта группа, вероятно, представляет собой смешанные линии, отражающие прошлую гибридизацию и международные селекционные программы.

Примечательно, что анализ STRUCTURE и NJ-дерева взаимно подтверждают друг друга: генотипы, сгруппированные в одну кластерную группу на основе SNP-данных, также демонстрируют высокую степень принадлежности к определенной субпопуляции в ковариационной матрице (Рисунок 22 и 23). Данное деление на кластеры отражает как эволюционные связи между линиями, так и исторические пути интродукции и селекции.

Распределение казахстанских образцов по нескольким кластерам еще раз подтверждает гипотезу о генетической разнородности и использовании в отечественной селекции сои генетического материала различного происхождения. Образцы казахстанской сои были распределены по всем трем субпопуляциям: 68% в третьей, 29% во второй и 3% в первой субпопуляции, что отражает сложный генетический фон, на который, возможно, повлияли как интродукции культивируемой сои, так и местная адаптация.

3.5. Полногеномный анализ ассоциаций хозяйственно-ценных признаков сои.

3.5.1. Полногеномный анализ ассоциаций компонентов урожайности сои

В результате ПГАА по средним данным за 5 лет по признакам, определяющим продуктивность сои, было идентифицировано 78 ассоциаций, среди них для КПУ, КСР, МТС и МСР было идентифицировано 56 МТА. Из них 20 ассоциаций повторно выявлялись как минимум в двух годах испытания, и считались стабильными, среди них семь – по количеству продуктивных узлов, шесть – по количеству семян с растения, четыре – по массе тысячи семян и три – по урожайности с растения. В Таблице 15 представлена полная информация о каждом маркере, его позиции на хромосоме, P-value, аллеле, эффекта и доля объяснения фенотипической вариации, PVE (phenotypic variance explained).

По компонентам урожайности ассоциации S20_34832847, S10_45323571, S13_21819620 и S08_14486855 показали максимальные доли PVE 9,02%, 8,27%, 6,54% и 11,98% для количества продуктивных узлов, количества семян на растение, массы 1000 семян и урожайности на растение соответственно. Эффект идентифицированных маркеров по направленности проявления признака варьировал от -39,87(отрицательная направленность при маркере S15_50483491) до 16.22 (положительная направленность при S13_21819620) [171]. Ассоциации S07_17041516, S04_48109794 и S13_21819620 показали наибольшие эффекты (Таблица 15).

Таблица 15 – Перечень стабильных значимых ассоциаций маркер-признак, связанных с компонентами урожайности средних данных за 2018-2022 гг.

Признак	SNP маркер	Хромосома	Позиция (пн)	P	Позитивный аллель	Эффект	PVE (%)
КПУ	**S07_17041516**	**7**	**16978432**	**$6.77x10^{-06}$**	**T**	**4.24**	**6.35**
КПУ	S11_8211474	11	8211474	$6.01x10^{-5}$	G	3.48	6.81
КПУ	S13_28224204	13	28224204	$2.28x10^{-06}$	C	-3.68	7.50
КПУ	S18_55651846	18	55651846	4.34x10-5	G	1.44	7.01
КПУ	S19_38590123	19	38590123	$5.59x10^{-06}$	T	4.37	8.12
КПУ	S20_11932337	20	11932337	$6.92x10^{-5}$	C	-2.05	6.41
КПУ	**S20_34832847**	**20**	**34832847**	**$1.37x10^{-5}$**	**A**	**-1.36**	**9.02**
КСР	S04_48109794	4	48109794	$2.27x10^{-5}$	G	13.54	6.32
КСР	**S07_17041516**	**7**	**17041516**	**$2.32x10^{-5}$**	**T**	**10.72**	**5.08**
КСР	**S10_45323571**	**10**	**45323571**	**$4.08x10^{-06}$**	**T**	**5.59**	**8.27**
КСР	S16_8084401	16	8084401	$5.45x10^{-5}$	T	-9.19	4.97
КСР	**S16_36572386**	**16**	**36572386**	**$5.91x10^{-5}$**	**C**	**6.94**	**6.78**
КСР	**S20_34832847**	**20**	**34832847**	**$2.70x10^{-5}$**	**A**	**-3.58**	**8.14**
МТС	S06_3119005	6	3119005	$8.94x10^{-5}$	G	10.85	5.48
МТС	S10_5908937	10	5908937	$8.83x10^{-5}$	A	-9.66	4.61
МТС	S13_21819620	13	21819620	$2.67x10^{-5}$	T	16.22	6.54
МТС	S15_50483491	15	50483491	$9.11x10^{-5}$	C	-39.87	5.75
МСР	S03_580652	3	580652	$4.16x10^{-5}$	T	1.87	6.78
МСР	S06_3459276	6	3459276	$2.60x10^{-5}$	C	-2.81	7.48
МСР	S08_14486855	8	14486855	$4.08x10^{-09}$	G	3.96	11.98

СП – содержание протеина, СМ – содержание масла, КПУ – число продуктивных узлов, КСР – число семян с растения, МТС – масса тысячи семян, МСР – масса семян с растения;
Маркеры, выделенные жирным шрифтом – ассоциированы с несколькими признаками;
PVE – phenotypic variance explained, доля объяснения фенотипической вариации

Примечательно, что несколько SNP в LD были связаны с двумя и более признаками одновременно. Имеются исследования, изучающие плейотропные эффекты маркеров SNP у растений, и их влияние на множество признаков [192]. Например, плейотропные эффекты наблюдались у нута в отношении признаков, связанных с засухоустойчивостью [193], и у вигны в отношении биомассы и урожайности зерна в условиях стресса засухи [194]. У кукурузы плейотропные эффекты были также обнаружены в отношении времени цветения, высоты растения и количества листьев [195]. В Таблице 15 жирным шрифтом выделены маркеры, обладающих плейотропным эффектом. SNP маркеры S10_45323571 и S10_45325143 на 10 хромосоме ассоциированы с количеством семян на растение и содержанием протеина, но наблюдается обратные эффекты, 5,59 для компонента урожайности и -0,92 для содержания протеина (Таблица 15). Маркеры S16_36471151 и S16_36572386 на 16 хромосоме связаны как с содержанием масла, так и с количеством семян на растение, однако также наблюдаются обратные эффекты, 6,94 для компонента урожайности и -0,55 для масла. Также, были идентифицированы плейотропные ассоциации, к примеру,

S07_17041516 на 7 хромосоме ассоциирован с признаком количества продуктивных узлов и количеством семян на растение, для обоих признаков маркер проявил положительный эффект. Маркер S20_34832847 на 20 хромосоме связан с двумя компонентами урожайности, количество продуктивных узлов и количество семян на растение, однако с отрицательным эффектом для каждого признака (Таблица 15) [171].

Результаты полногеномного анализа представлены на Рисунке 24 с помощью графиков Manhattan и QQ.

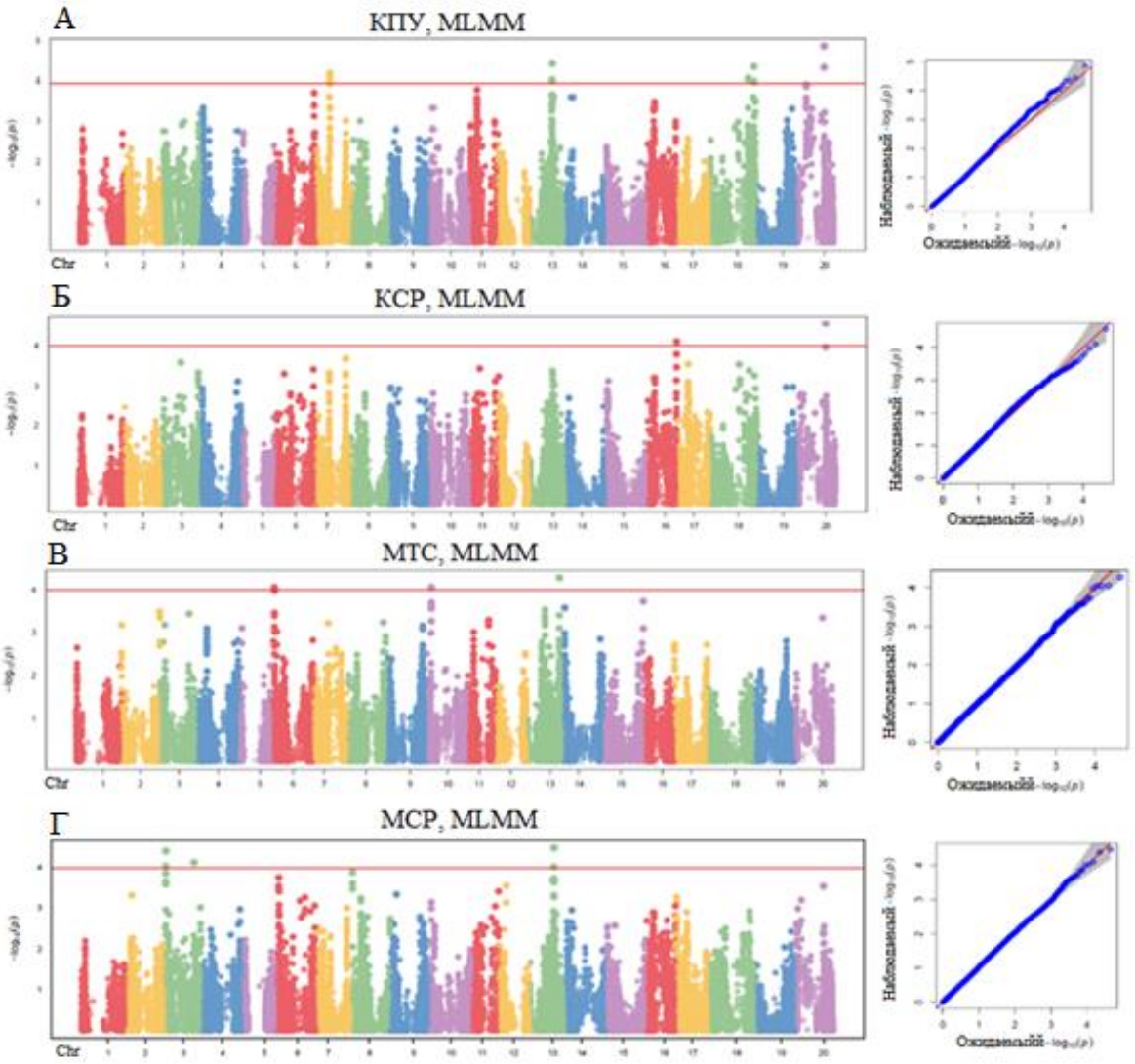


Рисунок 24 – графики Manhattan и QQ полногеномного анализа компонентов урожайности за средние данные 5 лет (2018-2022гг.)

Результаты полногеномного анализа средних данных показывают, что хромосома 20 содержит ключевой локус S20_34832847, влияющий на все компоненты урожайности, что делает его сильным кандидатом для дальнейшего изучения. Другие хромосомы, включая 6, 7 и 13, также показывают дополнительные локусы, обладающие потенциальным плейотропным эффектом.

3.5.2. Полногеномный анализ ассоциаций размеров семян сои

ПГАА также был проведен для размеров семян сои – длина, ширина и толщина. В результате анализа было идентифицировано 16 значимых ассоциаций ($P<1,0x10^{-4}$), среди них для длины семян – шесть, для ширины – четыре и для толщины семян – пять ассоциаций. Стабильно проявились 5 MTA (marker-trait association), по одной для длины и толщины, S02_2485010 и S08_47353557 соответственно и три ассоциации – для ширины семян сои, S06_593504, S09_34467480 и S14_2765409 (Таблица 16) [195].

Таблица 16 – Идентифицированные MTA с признаками размеров семян сои

Признак	SNP маркер	Хромосома	Позиция	P value	Эффект
ДС	S02_2485010	2	2485010	$4,03x10^{-05}$	-0,15
ШС	S06_593504	6	593504	$2,00x10^{-05}$	-0,17
	S09_34467480	9	34467480	$1,30x10^{-05}$	-0,11
	S14_2765409	14	2765409	$9,12x10^{-06}$	-0,17
ТС	S08_47353557	8	47353557	$2,43x10^{-06}$	-0,20
Длина семени – ДС, ширина – ШС, толщина – ТС					

Стоит отметить, что маркер S14_2765409 проявил себя во всех средах, что говорит о высокой стабильности и важности генетического маркера для признака ширины семян сои. На Рисунке 25 представлены графики Manhattan и QQ значимых ассоциаций маркер-признак размеров семян сои, статистический порог значимости изображен красной линией ($P<1,0x10^{-4}$).

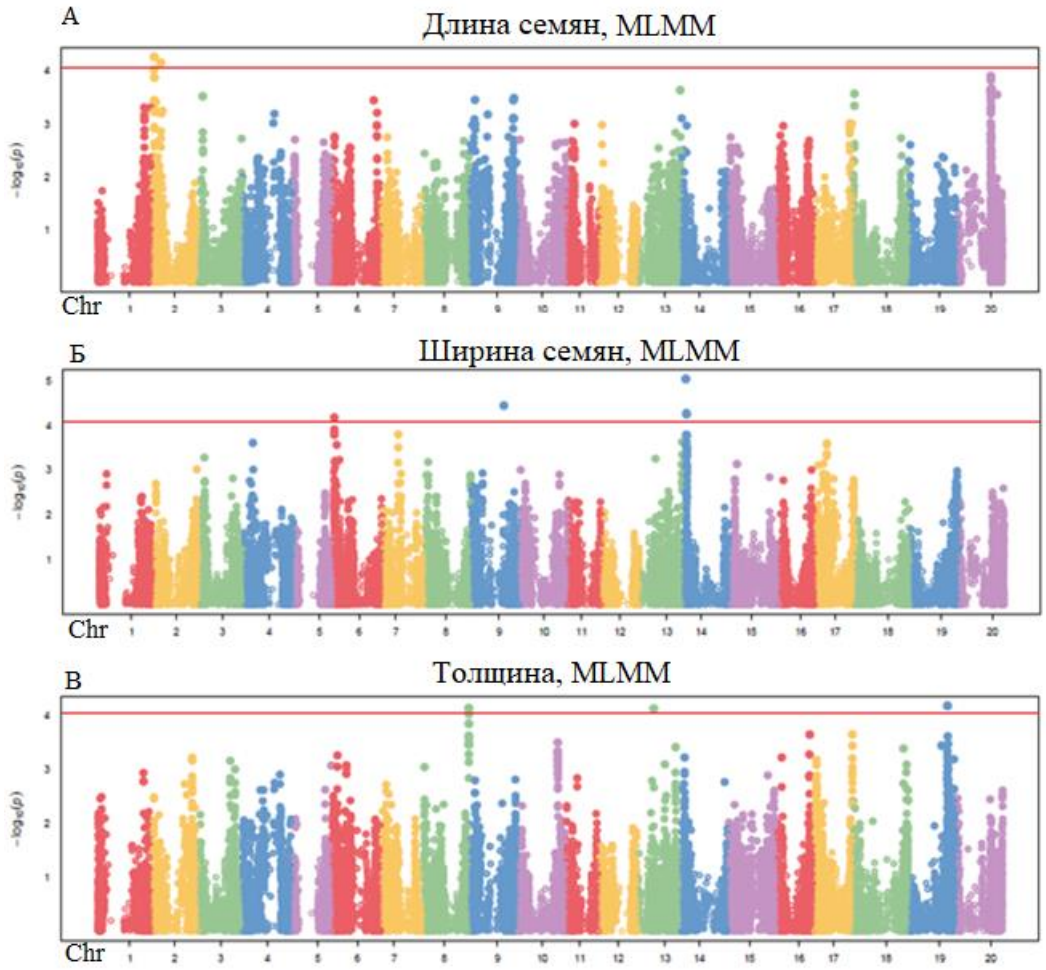


Рисунок 25 – Графики Manhattan и QQ полногеномного анализа признаков размеров семян сои

Рисунок 25 иллюстрирует, что хромосомы 2, 6, 8,9 и 14 имеют локусы, влияющие на размеры семян сои. Все идентифицированные маркеры могут быть полезны в селекционных программах по улучшению качественных характеристик семян сои в Алматинском регионе (Рисунок 25).

3.5.3. Полногеномный анализ ассоциаций признака растрескиваемости сои

ПГАА позволил идентифицировать три MTA Gm10_47774781, Gm13_6207590 и Gm16_29681065 для РС на хромосомах 10, 13 и 16 соответственно [173]. Информация о положениях маркеров на хромосомах, p-значениях, эффектах и фенотипических вариациях для аллелей представлена в Таблице 17.

Таблица 17 – Список идентифицированных значимых SNP маркеров, ассоциированных с признаком растрескиваемости стручков сои

Признак	SNP маркер	Хромосома	Позиция	P value	Эффект	PVE
PC	Gm10_47774781	10	47774781	$3.81x10^{-5}$	0.20783	0.06
	Gm13_6207590	13	6207590	$1.01x10^{-4}$	0.1743	0.06
	Gm16_29681065	16	29681065	$4.71x10^{-10}$	0.45428	0.15
PC – растрескиваемость стручков, PVE-проявление фенотипической вариации, %						

Gm16_29681065 был расположен в непосредственной близости от известного гена растрескивания стручков сои *Pdh1* на хромосоме 16 (Funatsuki et al., 2014). Два других минорных QTL были идентифицированы на хромосомах 10 и 13. Идентифицированные SNP с наиболее значимыми p-значениями Gm16_29681065, Gm10_47774781 и Gm13_6207590 были обозначены как *qPD16-1, qPD10-1* и *qPD13-1* соответственно. На Рисунке 26 представлена визуализация ассоциаций в виде графика Manhattan и QQ.

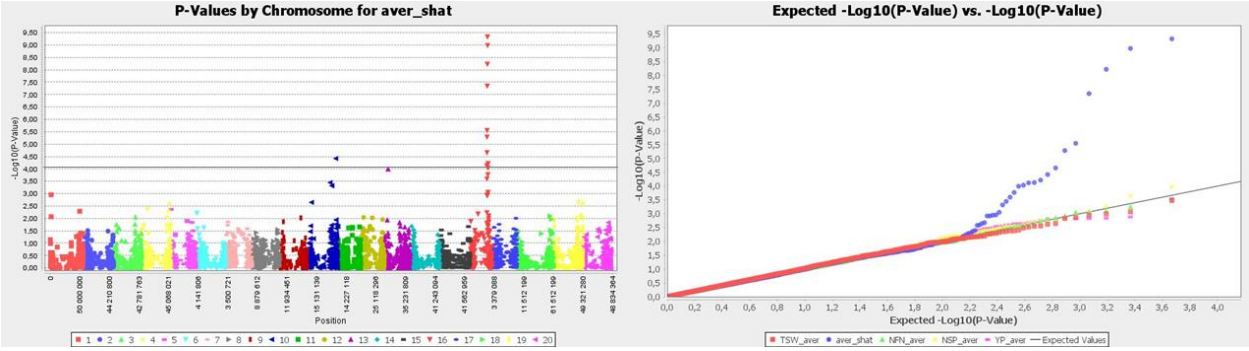


Рисунок 26 – графики Manhattan и QQ обнаруженных маркер-признак ассоциаций с признаком растрескиваемости стручков сои средних данных 2х лет [173]

3.5.4. Полногеномный анализ ассоциаций признаков качества сои

По результатам ПГАА, на основе MLMM модели в GAPIT и 44385 SNP маркеров, выявлено 28 значимых MTA, связанных с признаками содержания протеина и масла в семенах сои [171]. Среди всех 28 значимых MTA, обнаруженных в различных средах, 18 были идентифицированы как минимум в двух годах эксперимента со значимыми значениями вероятности P ($P < 1x10^{-4}$) и считались стабильными. Маркеры находящиеся на LD (321509 п.н.) обобщались в один маркер. В результате исследования выявилось девять стабильных ассоциаций по содержанию протеина, девять – по содержанию масла (Таблица 18).

В Таблице 18 представлена полная информация о каждом маркере, его позиции на хромосоме, P-value, эффекта и доля объяснения фенотипической вариации, PVE (phenotypic variance explained).

Таблица 18 – Перечень стабильных значимых ассоциаций маркер-признак, связанных с содержанием протеина и масла средних данных за 2018-2022 гг.

Признак	SNP маркер	Хромосома	Позиция (пн)	P	Эффект	PVE (%)
СП	S02_40916569	2	40916569	$2.80x10^{-5}$	-0.73	6.98
СП	S05_3377432	5	3377432	$5.06x10^{-5}$	0.75	4.81
СП	**S06_5653082**	**6**	**5700406**	**$1.92x10^{-5}$**	**0.88**	**7.36**
СП	S06_48327505	6	48327505	$2.66x10^{-5}$	0.69	8.23
СП	**S10_45325143**	**10**	**45325143**	**$3.67x10^{-5}$**	**-0.92**	**7.24**
СП	S14_2073248	14	2073248	$2.21x10^{-5}$	-0.87	8.45
СП	**S14_45537479**	**14**	**45537479**	**$5.53x10^{-5}$**	**-0.85**	**4.52**
СП	S17_40642517	17	40642517	$1.26x10^{-5}$	-0.7	8.04
СП	S19_48095581	19	48015856	$1.04x10^{-06}$	0.92	11.15
СМ	S03_3327118	3	3349008	$4.03x10^{-5}$	0.35	6.38
СМ	**S06_5537821**	**6**	**5653082**	**$2.04x10^{-5}$**	**-1.29**	**5.56**
СМ	S07_42119515	7	42119515	$9.80x10^{-06}$	-0.5	7.81
СМ	S09_41221274	9	41221274	$7.71x10^{-05}$	-0.49	4.97
СМ	**S14_45542107**	**14**	**45542107**	**$6.91x10^{-05}$**	**0.59**	**6.94**
СМ	S16_4427570	16	4427570	$6.67x10^{-05}$	0.55	6.76
СМ	S16_6256537	16	6256537	$4.17x10^{-05}$	0.59	9.48
СМ	**S16_36471151**	**16**	**36461023**	**$4.63x10^{-05}$**	**-0.55**	**5.81**
СМ	S17_5368514	17	5368514	$8.13x10^{-05}$	0.48	5.48
СП – содержание протеина, СМ – содержание масла						

Ассоциации маркер-признак по всем изучаемым признакам объясняют от 4,52 (S14_45537479, СП) до 11,15 % (S19_48095581, также СП) фенотипических вариаций признаков. Маркеры S19_48095581 и S16_6256537 показали самый высокий PVE – 11,15% и 9,48% для протеина и масла соответственно (Таблица 18).

Эффект идентифицированных маркеров на направленность проявления признака варьировала от -1,29 (отрицательная направленность ассоциации СМ S06_5537821) до 0.92 (положительная направленность ассоциации СП S19_48095581) [171].

Графики Манхэттен и Q-Q полногеномного анализа для средних значений за пять лет признаков содержания протеина и масла показаны на Рисунке 27.

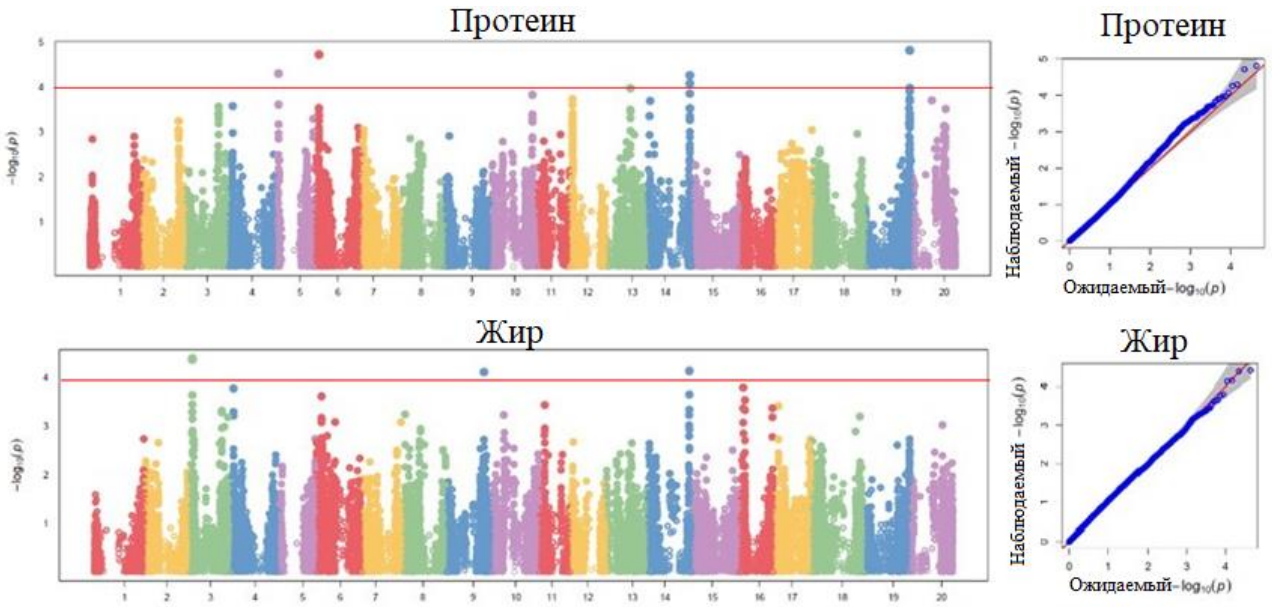


Рисунок 27 – графики Manhattan и QQ полногеномного анализа признаков качества за средние данные 5 лет (2018-2022гг.)

Результаты средних данных за 5 лет показывают, что хромосомы 14 и 19 является основным генетическим регионом, влияющим как на содержание протеина, так и на содержание масла, в то время как хромосомы 3 и 6 демонстрируют сильные эффекты, специфичные для признаков. Эти результаты предоставляют ценные данные для селекции сортов сои с высоким содержанием протеина или масла и для дальнейшей функциональной проверки генов-кандидатов. Кроме того, маркеры S06_5653082 и S14_45537479 на 6 и 14 хромосомах – плейотропные, так как ассоциированы как с содержанием протеина, так и с содержанием масла, однако с противоположными эффектами. Графики Manhattan и QQ плоты каждого признака за 5 лет (2018-2022 гг.) предоставлены в Приложении 4.

3.6. Аннотация генов-кандидатов идентифицированных ассоциаций хозяйственно-ценных признаков сои

3.6.1. Аннотация генов-кандидатов компонентов урожайности сои

Идентификация генов-кандидатов, связанных с хозяйственно-ценными признаками, имеет важное значение для повышения урожайности и пищевой ценности сельскохозяйственных культур. В Таблице 18 представлены гены-кандидаты идентифицированных MTA признаков КПУ, КСР, МТС и МСР, расположенные в определенных геномных позициях, каждый из которых аннотирован соответствующим белком или функциональным доменом.

Таблица 19 – Гены-кандидаты идентифицированных МТА по компонентам урожайности сои

SNP маркер	ID	Условие, где была обнаружена ассоциация	Публикации о ближлежайших локусах или генах	Расположение в геноме	Гены/локусы
S07_17041516	***q.NFN.ipbb.7***	**2022, среднее за 5 лет**		**Межгенное пространство**	**SoyZh13_07G132900**
S11_8211474	***q.NFN.ipbb.11***	**2022, среднее за 5 лет**		**Межгенное пространство**	**SoyZH13_11G105000**
S13_28224204	*q.NFN.ipbb.13*	2022, среднее за 5 лет	[197]	Межгенное пространство	*Rab5a2*
S18_55651846	*q.NFN.ipbb.18*	2022, среднее за 5 лет	[198]	Межгенное пространство	*Dt2*
S19_38590123	*q.NFN.ipbb.19*	2022, среднее за 5 лет	[199]	Кодирующий регион	SoyZH13_19G101500
S20_11932337	***q.NFN.ipbb.20.1***	**2019,2022, среднее за 5 лет**		**Межгенное пространство**	**SoyZH13_20G046400**
S20_34832847	*q.NFN.ipbb.20.2*	2020, среднее за 5 лет	[200]	Межгенное пространство	*POWR1*
S04_48109794	*q.NSP.ipbb.4*	2022, среднее за 5 лет	[201]	Кодирующий регион	SoyZH13_04G177900
S07_17041516	***q.NSP.ipbb.7***	**2022, среднее за 5 лет**		**Межгенное пространство**	**SoyZh13_07G132900**
S10_45323571	*q.NSP.ipbb.10*	2019,2021	[202]	Межгенное пространство	*GA20OX*
S16_8084401	***q.NSP.ipbb.16.1***	**2018, среднее за 5 лет**		**Кодирующий регион**	**SoyZH13_16G072400**
S16_36572386	***q.NSP.ipbb.16.2***	**2019,2021, среднее за 5 лет**		**Межгенное пространство**	**SoyZH13_16G178000**
S20_34832847	*q.NSP.ipbb.20*	2020, среднее за 5 лет	[200]	Межгенное пространство	*POWR1*
S06_3119005	***q.TSW.ipbb.6***	**2020, среднее за 5 лет**		**Кодирующий регион**	**SoyZH13_06G039000**
S10_5908937	***q.TSW.ipbb.10***	**2020, среднее за 5 лет**		**Межгенное пространство**	**SoyZH13_10G059400**
S13_21819620	*q.TSW.ipbb.13*	2019, среднее за 5 лет	[178]	Межгенное пространство	*GmPDAT*
S15_50483491	***q.TSW.ipbb.15***	**2020, среднее за 5 лет**		**Кодирующий регион**	**SoyZH13_15G233400**
S03_580652	***q.YP.ipbb.3***	**2019,2021, среднее за 5 лет**		**Межгенное пространство**	**SoyZH13_03G005900**
S06_3459276	*q.YP.ipbb.6*	2018,2020	[203]	Межгенное пространство	SoyZH13_06G043000
S08_14486855	***q.YP.ipbb.8***	**2019**		**Межгенное пространство**	**SoyZH13_08G170800**
*Жирным шрифтом выделены предполагаемые новые МТА влияющие на изучаемые признаки NFN (number of fertile nodes) – число продуктивных узлов, NSP (number of seed per plant) – число семян с растения, TSW (thousand seed weight) – масса тысячи семян, YP (yield per plant) – масса семян с растения					

ПГАА позволил идентифицировать 20 стабильных ассоциаций по признакам КПУ, КСР, МТС и МСР, на 13 различных хромосомах сои. Среди них 9 идентифицированных ассоциаций были ранее опубликованы [197-203]. К примеру, маркер *q.NFN.ipbb.18* расположен в межгенном пространстве вблизи с геном *Dt2,* функция которого заключается в кодировании фактора транскрипции MADS-box, контролирующего терминацию стебля [198]. Присутствие *q.NFN.ipbb.18* рядом с *Dt2* предполагает возможную связь между количеством узлов и регуляцией роста растений. Вариации *Dt2* могут контролировать развитие узлов, в конечном итоге, влияя на количество продуктивных узлов и, следовательно, на урожайность.

Следующая ассоциация признака КПУ, *q.NFN.ipbb.13,* близко расположена к гену *Rab5a2* на 13 хромосоме. Ген *Rab5a2* кодирует ГТФазу, участвующую во внутриклеточном транспорте и мембранном транспорте. *Rab*-белки регулируют транспорт, опосредованный везикулами, который играет важную роль в передаче сигналов гормонов и транспортировке питательных веществ во время развития семян [197]. Расположение *q.NFN.ipbb.13* рядом с геном *Rab5a2* предполагает, что этот ген может влиять на количество продуктивных узлов, контролируя распределение питательных веществ и дифференцировку клеток во время формирования семян**.**

Стоит отметить обнаружение плейотропного маркера S20_34832847, ассоциированного с двумя признаками урожайности КСР и КПУ и расположенного в межгенном пространстве, в близи известного гена *POWR1* (Protein, Oil, Weight) на 20 хромосоме [200]. Ген *POWR1* – важный ген соевых бобов, влияющий на качество семян и урожайность. Исследования показали, что вставка мобильного элемента в ген *POWR1* нарушает его функцию, что приводит к увеличению содержания масла в семенах, их веса и общего урожая при одновременном снижении содержания протеина [200]. Хотя прямых доказательств влияния *POWR1* на количество продуктивных узлов и семян на растение в литературе недостаточно, его известная роль в регуляции транспорта питательных веществ и метаболизма липидов может опосредованно влиять на развитие растений и формирование урожая. Например, изменения в экспрессии *POWR1* могут приводить к перераспределению ресурсов, что, в свою очередь, может отражаться на количестве узлов и семян. Учитывая, что *POWR1* также контролирует массу семян, можно предположить, что он может влияет на другие компоненты урожайности, которые положительно коррелируют друг с другом и отрицательно – с содержанием протеина. Следовательно, *POWR1* может играть комплексную роль в определении урожайных и качественных характеристик сои, что и может объяснить его возможную связь с обнаруженными маркерами *q.NSP.ipbb.20* и *q.NFN.ipbb.20.2*.

Еще одна ассоциация признака МТС, *q.TSW.ipbb.13,* расположена в ближайшем соседстве с известным геном *GmPDAT* на 13 хромосоме [178]. Его функция – кодировать фосфолипид диацилглицерин-ацилтрансферазу, фермент, ответственный за биосинтез триацилглицерина. Учитывая его решающую роль в производстве масла, *GmPDAT* определенно может влиять на массу семян

посредством накопления липидов, что важно для увеличения содержания соевого масла и биомассы семян.

Также стоит отметить маркер *q.NFN.ipbb.19*, который перекрывается с QTL *qSW-19-3*, ассоциированный с массой семян [199], что позволяет предположить, что данный регион генома может иметь генетический фактор с плейотропным эффектом.

Множественное совпадение результатов полногеномного анализа в наших исследованиях, в сравнении с ранее опубликованными данными, позволяет предположить о высоком уровне достоверности обнаруженных ассоциаций. Литературный обзор, а также поиски в базе данных Soybase.org, не дали информации об оставшихся значимых ассоциациях компонентов урожайности на 3,6,7,8,10,11,15,16 и 20 хромосомах, предполагая, что данные ассоциации могут рассматриваться как новые генетические факторы для компонентов урожайности, а также качества семян сои (Таблица 18).

3.6.2. Аннотация генов-кандидатов размеров семян сои

Идентификация генов-кандидатов основывалось на физическом положении каждого SNP с использованием референсного генома сои. Важно отметить, что два из этих SNP – S08_47353557 и S14_2765409 – были расположены в геномных регионах, ранее описанных в независимых исследованиях. SNP S08_47353557, связанный с размерами семян, перекрывался с регионом гена *Glyma.08g363300,* идентифицированным Ли и др. (2021) в исследовании, посвященном весу и форме семян сои [204]. Аналогичным образом, маркер S14_2765409, связанный с шириной семян, был расположен в регионе, ассоциированный с размерами семян, *qSLT-14-1(seed length-to-thickness ratio)* [205]. Анализ литературы не дал дополнительной информации о трех оставшихся ассоциациях S02_2485010 – для длины семян, S06_593504 и S09_34467480 – для ширины семян сои, предполагая, что данные ассоциации могут быть потенциально новыми генетическими факторами размеров семян сои в условиях Алматинской области. Так, MTA длины семени, S02_2485010 перекрывается с геном SoyZH13_02G027200, участвующий в связывании ионов цина, что может быть связано с регуляцией экспрессии генов роста через цинк-зависимые транскрипционные факторы [206-207]. MTA ширины семян находятся вблизи с генами SoyZH13_09G115300 и SoyZH13_06G006800, отвечающие за перенос дыхательных электронов и каталитической активности соответственно (soybase.org). Поскольку активный митохондриальный метаболизм и энергетический обмен играют ключевую роль в клеточном делении и росте тканей, эти гены могут быть функционально связаны и с развитием семян [208].

3.6.3. Аннотация генов-кандидатов растрескиваемости стручков сои

Три идентифицированных MTA для признака растрескиваемости стручков сои, *qPD10-1, qPD13-1 и qPD16-1*были расположены на хромосомах 10, 13 и 16 соответственно (Таблица 17). Поскольку MTA *qPD16-1* был весьма значимым как в 2019, так и в 2021 году, его можно считать основным генетическим фактором, оказывающим заметное влияние на PD. Расположение MTA *qPD16-1* перекрывалось с генетическим положением гена *Pdh1* (Таблица 17). Ген *Pdh1* кодирует белок, участвующий в лигнификации, тем самым способствуя скручиванию сухих створок стручка. Потери функции гена, *pdh1,* широко использовались в селекции сои в качестве гена устойчивости к растрескиванию стручка [65].

Другой значимый MTA *qPD10-1* для PD, идентифицированный на 10 хромосоме, был расположен в гене *Glyma10g40330* [90], который отвечает за экспрессию растительного белка, связывающего кальмодулин [92]. Ранее, еще один QTL для PD, *Satt243*, был идентифицирован на хромосоме 10, который был расположен в пределах LD [79], что предполагает общий локус двух ассоциаций. Интересно, что значимый QTL, *qPD13-1,* идентифицированный на хромосоме 13, был расположен вблизи *Glyma13g05890*, который также экспрессирует растительный белок, связывающий кальмодулин [90; 92].

Результаты влияния всех трех выявленных генетических факторов на показатели PD предполагают, что, хотя роль *qPD16-1* значительна, аллельные статусы Gm10_47774781 и Gm13_6207590 также имеют важное значение (Таблица 17). Следовательно, можно предположить, что белок, связывающий кальмодулин, является частью генной сети, контролирующей растрескиваемость стручков сои.

3.6.4. Аннотация генов-кандидатов признаков качества семян сои

Среди выявленных 18 стабильных ассоциаций маркер-признак по качеству семян, 13 ассоциаций соответствовали уже ранее опубликованным QTL в предыдущих публикаций. Более подробная информация по обнаруженным маркерам, условиям, где они были обнаружены, локализация в референсном геноме, а также уже опубликованные локусы и референсы к статьям представлены в Таблице 20.

Таблица 20 – Список идентифицированных МТА по признакам качества семян сои в результате полногеномного анализа и информация об генах кандидатов

SNP маркер	ID	Условие, где была обнаружена ассоциация	Публикации о близлежащих локусов или генах	Расположение в геноме	Гены/локусы
S02_40916569	*q.Prt.ipbb.2*	2018, среднее за 5 лет	[209]	Кодирующий регион	SoyZH13_02G180200
S05_3377432	*q.Prt.ipbb.5*	2020, среднее за 5 лет	[210]	Межгенное пространство	SoyZH13_05G037300
S06_5653082	*q.Prt.ipbb.6.1*	2018, 2021,2022, среднее за 5 лет	[211]	Кодирующий регион	SoyZH13_06G068400
S06_48327505	*q.Prt.ipbb.6.2*	2020,2022, среднее за 5 лет	[212-214]	Межгенное пространство	*GmZF351, GmMYB73*
S10_45325143	*q.Prt.ipbb.10*	2018,2019, 2021, среднее за 5 лет	[202]	Межгенное пространство	*GA20OX*
S14_2073248	*q.Prt.ipbb.14.1*	2019, 2020, среднее за 5 лет	[215]	Межгенное пространство	SoyZH13_14G025700
S14_45537479*	***q.Prt.ipbb.14.2***	**2018,2019,2021, среднее за 5 лет**		**Межгенное пространство**	**SoyZH13_14G158400**
S17_40642517	*q.Prt.ipbb.17*	2022, среднее за 5 лет	[216]	Межгенное пространство	SoyZH13_17G229400
S19_48095581	*q.Prt.ipbb.19*	2018, 2021, среднее за 5 лет	[217]	Межгенное пространство	*GmCYP78A72*
S03_3327118	***q.Oil.ipbb.3***	**2019,2020,2022, среднее за 5 лет**		**Межгенное пространство**	**SoyZH13_03G028300**
S06_5537821	*q.Oil.ipbb.6*	2018, среднее за 5 лет	[218]	Кодирующий регион	SoyZH13_06G067300
S07_42119515	***q.Oil.ipbb.7***	**2018,2022, среднее за 5 лет**		**Кодирующий регион**	**SoyZH13_07G202800**
S09_41221274	*q.Oil.ipbb.9*	2021, среднее за 5 лет	[216]	Межгенное пространство	SoyZH13_09G146600
S14_45542107	*q.Oil.ipbb.14*	2019,2021, среднее за 5 лет	[219]	Межгенное пространство	SoyZH13_14G158400
S16_4427570	***q.Oil.ipbb.16.1***	**2018, среднее за 5 лет**		**Межгенное пространство**	**SoyZH13_16G043200**
S16_6256537	*q.Oil.ipbb.16.2*	2021, среднее за 5 лет	[202]	Межгенное пространство	SoyZH13_16G059500
S16_36471151	*q.Oil.ipbb.16.3*	2018,2021, среднее за 5 лет	[202]	Кодирующий регион	SoyZH13_16G177500
S17_5368514	*q.Oil.ipbb.17*	2021, среднее за 5 лет	[220]	Межгенное пространство	*GmDof4*

*Жирным шрифтом выделены предполагаемые новые МТА влияющие на изучаемые признаки
Prt (protein content) – содержание протеина, *Oil* (oil content) – содержание масла

В Таблице 20 представлены гены-кандидаты идентифицированных МТА, расположенные в определенных геномных позициях, каждый из которых аннотирован соответствующим белком или функциональным доменом.

Содержание протеина и масла являются важнейшими компонентами качества семян и важными характеристиками соевых бобов. При достижении высоких урожаев, качество семян, в частности содержание протеина, часто снижено, из-за отрицательной генетической корреляции между содержанием протеина и масла с урожайностью [96-97], что также подтвердилось и в нашем исследовании (Рисунок 15). В результате проведения полногеномного анализа было выявлено 9 значимых МТА по признаку содержания протеина, расположенных на семи хромосомах сои (Таблица 20). Из этих девяти МТА, о восьми уже сообщались ранее в научных публикациях [202; 209-220]. Результаты анализа показывают связь между большинством идентифицированных МТА и важными генами, которые участвуют в гормональной регуляции, контроле транскрипции, липидном обмене и транспорте питательных веществ, что позволяет предположить их потенциальное влияние на состав семян сои и урожайность (Таблица 20, Soybase.org). Так, к примеру, *q.Prt.ipbb.6.2* был расположен в межгенном пространстве рядом с двумя генами кандидатами. Первый, ген *GmZF351,* кодирует белок цинковых пальцев и участвует в биосинтезе липидов и накоплении масла семян. Исследования показали, что *GmZF351* положительно регулирует пути биосинтеза масла, а его сверхэкспрессия приводит к увеличению содержания масла в семенах [213]. Присутствие *q.Prt.ipbb.6.2* рядом с этим геном позволяет предположить его потенциальное участие в регуляции состава семян, в частности влияя на накопление масла в семенах, косвенно отражается на содержании протеина. Второй, ген *GmMYB73,* фактор транскрипции MYB, идентифицированный в соевых бобах (*Glycine max* (L.) Merr.), играет значительную роль в накоплении жиров. Исследования показали, что взаимодействие *GmMYB73* с другими белками приводит к усилению биосинтеза жиров [214]. Вполне возможно, что сверхэкспрессия *GmMYB73*, которая увеличивает накопление жиров, может привести к снижению содержания протеина в семенах из-за компенсаторного механизма.

Ассоциация S10_45325143 на 10 хромосоме, связанная с СП и КСР, расположена вблизи гена *GA20OX*, регулятора развития семян посредством биосинтеза гиббереллиновой кислоты [202]. Исследование арабидопсиса и сои показало, что мутанты по биосинтезу гиббереллина демонстрируют измененный состав семян как протеина, так и масла [221]. Кроме того, известно, что гиббереллины влияют на обмен углерода и азота. Исследования показали, что рост, вызванный гиббереллиновой кислотой, может увеличить поглощение и ассимиляцию азота, что приводит к более высокому накоплению протеина в семенах [222].

В пределах LD, охватывающий маркер S19_48095581 (*q.Prt.ipbb.19*) на хромосоме 19, находится ген-кандидат *GmCYP78A72,* кодирующий фермент цитохрома P450. Предыдущие исследования показали, что сверхэкспрессия генов *CYP78A72* приводит к увеличению размера семян и увеличению веса [217].

Поскольку размер и вес семян часто связаны с развитием семян, вероятно, они также могут прямыми и непрямыми путями регулировать состав семян, в данном случае содержание протеина.

Среди всех 9 идентифицированных МТА для признака содержания протеина, не найдено информации в литературе для ассоциации *q.Prt.ipbb.14.2* (S14_45537479*)*, расположенной на хромосоме 14~~,~~. Данный маркер согласно базе данных Soybase.org расположен в межгенном пространстве, вблизи генной модели SoyZH13_14G158400. Функция данной модели – регуляция транспорта через липидный бислой [92] и, следовательно, возможное участие в контроле содержания протеина посредством влияния транспорта питательных веществ.

Для признака содержания масла в семенах сои ПГАА выявил девять МТА на хромосомах 3, 6,7,9,14,16,17 (Таблица 20). Расположение МТА на хромосомах 6, 9, 14, 16 и 17 перекрывалось с ранее опубликованными QTL и генами сои [202; 216; 218-219]. Два МТА S16_6256537 и S16_36471151 на 16 хромосоме совпали с QTL содержания масла, обнаруженными Jin и соавторами (2023) [202]. Более того, S14_45542107 на хромосоме 14 в этом исследовании располагался рядом с QTL, ассоциированным с содержанием масла [219]. МТА S17_5368514, *q.Oil. ipbb.17,* на 17 хромосоме был близко расположен к гену *GmDof4*, влияющего на содержание липидов и массы семян, как это было показано в трансгенных семенах арабидопсиса [220]. Это позволяет предположить, что *q.Oil. ipbb.17* может регулировать метаболический путь, влияющие на содержание масла и массу семян. Обзор литературы показывает, что три МТА для СМ на хромосомах 3, 7 и 16 предположительно являются новыми факторами для признака содержания масла в семенах сои. Среди них МТА *q.Oil. ipbb.7* был обнаружен в кодирующей последовательности генной модели SoyZH13_07G202800 [92], функция которого связана с белками, подобные дифтинсинтазе. Эти белки участвуют в биосинтезе дифтамида, уникальной посттрансляционной модификации фактора элонгации 2 (eEF2) [223]. Правильное функционирование eEF2 имеет решающее значение для синтеза запасных белков во время развития семян. Нарушения в этом процессе вероятнее всего могут привести к изменениям в накоплении протеина в семенах и, следовательно, масла.

МТА *q.Oil.ipbb.3* и *q.Oil.ipbb.16.1* согласно базе данных Soybase.org были близки к генам *SoyZH13_03G028300* и *SoyZH13_16G043200,* соответственно [92]. Эти гены участвуют в важных биологических функциях, таких как обогащение лейцином, повторный синтез белка и активность цистеиновой протеазы [92]. Перекрытие этих МТА с вышеуказанными генами предполагает возможное участие этих генов также в регулировании содержания масла в сое.

3.7. Анализ гаплотипов значимых ассоциаций хозяйственно-ценных признаков сои

В результате ПГАА были идентифицированы 46 значимых стабильных ассоциации хозяйственно-ценных признаков, где для каждого SNP-маркера были выявлены позитивные аллели, ассоциированные с благоприятным эффектом изучаемых признаков. В Таблице 21 представлена информация для каждой МТА, включая позитивные аллели и их эффекты.

Таблица 21 – Позитивные аллели, ассоциированные с благоприятным эффектом изучаемых признаков

Признак	Маркер	Аллели	Эффект	ПА	Признак	Маркер	Аллели	Эффект	ПА
КПУ	S07_17041516	C/T	4.24	T	МСР	S08_14486855	A/G	3.96	G
	S11_8211474	A/G	3.48	G	СП	S02_40916569	C/G	-0.73	C
	S13_28224204	T/C	-3.68	C		S05_3377432	C/T	0.75	T
	S18_55651846	G/C	1.44	G		S06_5653082	C/T	0.88	T
	S19_38590123	C/T	4.37	T		S06_48327505	T/G	0.69	T
	S20_11932337	T/C	-2.05	C		S10_45325143	A/G	-0.92	A
	S20_34832847	A/C	-1.36	A		S14_2073248	G/A	-0.87	A
КСР	S04_48109794	A/G	13.54	G		S14_45537479	G/A	-0.85	A
	S07_17041516	C/T	10.72	T		S17_40642517	G/A	-0.7	A
	S10_45323571	C/T	5.59	T		S19_48095581	G/A	0.92	G
	S16_36572386	C/T	-9.19	C	СМ	S03_3327118	G/A	0.35	G
	S16_8084401	T/C	6.94	T		S06_5537821	G/A	-1.29	A
	S20_34832847	A/C	-3.58	A		S07_42119515	G/T	-0.5	G
МТС	S06_3119005	A/G	10.85	G		S09_41221274	G/A	-0.49	A
	S10_5908937	G/A	-9.66	A		S14_45542107	C/T	0.59	T
	S13_21819620	T/C	16.22	T		S16_4427570	G/A	0.55	G
	S15_50483491	T/C	-39.87	C		S16_6256537	T/C	0.59	T
YP	S03_580652	G/T	1.87	T		S16_36471151	A/G	-0.55	A
	S06_3459276	T/C	-2.81	C		S17_5368514	T/C	0.48	T

СП – содержание протеина, СМ – содержание масла, КПУ – число продуктивных узлов, КСР – число семян с растения, МТС – масса тысячи семян, МСР - масса семян с растения; ПА-позитивный аллель

Идентифицированные эффективные гаплотипы, состоящие из комбинаций позитивных аллелей значимых ассоциаций хозяйственно-ценных признаков приведены в Таблице 22. Также, для сравнения, приведены показатели данных гаплотипов в сравнении со средним данными признака всей коллекции в условиях Алматинской области.

Таблица 22 – Эффективные гаплотипы и показатели изучаемых признаков в условиях Алматинской области.

Признак	Гаплотипы	Количество позитивных аллелей	Показатель признака для гаплотипа	Средний показатель признака в популяции
КПУ	TGCGTCA	7	26,7 шт.	18.42±0.29
КСР	GCTTCA	5	81,7 шт.	46.03±0.83
МТС	GATC	4	191,64 г	161.56±1.49
МСР	TCA	3	13,86 г	13.45±0.31
СП	CTTTAAAAG	7	46.75 %	42.12±0.16
СМ	GAGATGTAT	7	21.50%	20.70±0.08
СП – содержание протеина, СМ – содержание масла, КПУ – число продуктивных узлов, КСР – число семян с растения, МТС – масса тысячи семян, МСР – масса семян с растения				

Наиболее эффективными для признаков компонентов урожайности оказались следующие сочетания аллелей: КПУ – TGCGTCA, КСР – GTTCTA, МТС – GATC, МСР – TCA (Таблица 22). Для КПУ показатели признака варьировали от 13,73 шт. у образцов без единого эффективного аллеля до более 26,7 шт. у образцов с 6 и 7 позитивными аллелями (1033(Корея), Rainer 58(Молдова), 1674(Китай), Linkoln и Cobb 266 (США). Высокие показатели КСР с 5 из 6 позитивных аллелей достигали в среднем до 81,7 шт, для признака МТС наличие всех 4-х позитивных аллелей значимых ассоциаций в среднем также показали высокое значение признака по сравнению с остальной коллекцией, в среднем 191,64 грамм у образцов как Приморская 495(Россия) и Hei feng 50 (Китай). Для признака урожайности с растения также наблюдается улучшение фенотипа с увеличением числа позитивных аллелей до 13,86 грамм.

Наиболее эффективными гаплотипами, связанными с повышенным содержанием протеина и масла, оказались комбинации аллелей CTTTAAAAG и GAGATGTAT соответственно. Самый минимальный показатель был отмечен у канадского сорта Supra, 38,45% (гаплотип GCCGGGGGA), наибольшие же показатели, более 45%, были отмечены у трех казахстанских линий 308/1, 407/2, 404/2, а также Warsawska (Польша), Подяка (Украина) и Majesta (Канада). По признаку содержания масла минимальный показатель лишь с двумя эффектами позитивных маркеров был отмечен в основном у восточноевропейских сортов со средним значением 18,84%, выше же показатели были отмечены для количества 5 и 6 позитивных аллелей (Рисунок 29). В группах 5-6 (количество позитивных аллелей) наблюдается преобладание маркера S07_42119515 с большей частотой, который отличился наименьшим P значением и одним из больших PVE, 7,81%. Данный результат предполагает, что MTA S07_42119515 – *q.Oil.ipbb.7*– является преобладающим генетическим фактором содержания масла в семенах сои в изучаемой коллекции сои. На Рисунке 28 изображен график эффектов гаплотипов с позитивными аллелями и фенотипическим проявлением признаков содержания протеина и масла.

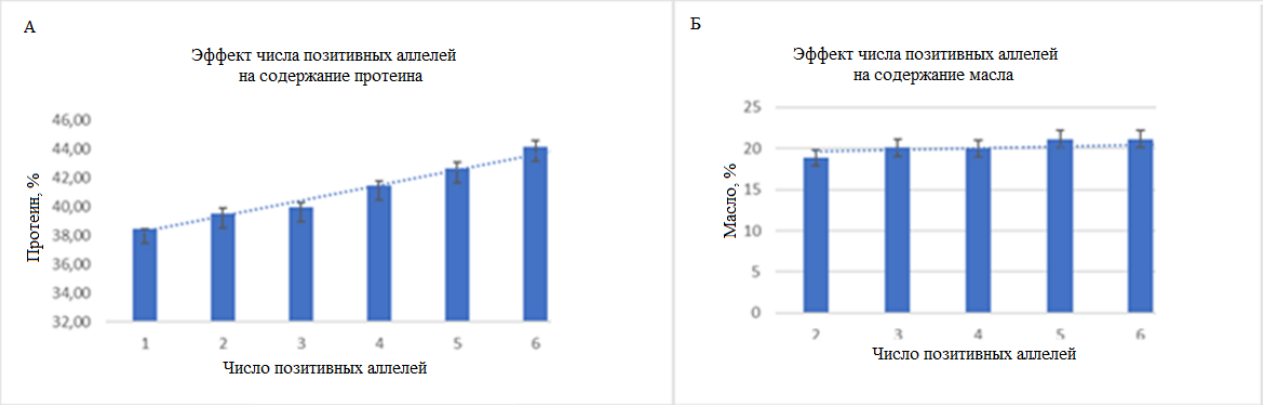


Рисунок 28 – Эффекты гаплотипов значимых ассоциаций признаков качества с разным количеством позитивных аллелей

Что касается признака растрескиваемости стручков семян, влияние аллельного статуса наиболее значимых ассоциаций на признак показывают, что комбинация аллелей TTG в гаплотипах привела к устойчивости к РС со средним баллом растрескиваемости 0,1 (Таблица 22).

Таблица 23 – Средние значения показателей растрескивания стручка (РС) в полевых условиях и аллельные комбинаций трех значимых ассоциаций признаков РС

Gm16_29681065	Gm10_47774781	Gm13_6207590	Количество растений	Балл РС
T	T	G	192	0,1
T	T	A	29	0,3
T	C	G	19	0,4
T	C	A	2	4,5
G	T	G	9	1,5
G	C	G	4	4,5
G	C	A	4	3,9

Напротив же, комбинация альтернативных аллелей (GCA) показала восприимчивость к РС со значением 3,9. Интересно, что два растения с комбинацией TCA (устойчивый аллель для Gm16_29681065 и два восприимчивых аллеля для Gm10_47774781 и Gm13_6207590) показали фенотип РС со значением 4,5 (Таблица 23), что позволяет предположить, что эффективный аллель Gm16_29681065 сам по себе недостаточен для устойчивости к РС.

По результатам проведения анализа, обнаруженные гаплотипы, связанные с желаемыми хозяйственно-ценными признаками могут быть применимы на ранних стадиях селекционных программ по созданию отечественных сортов сои высокого урожая и качества. Также были выделены перспективные сорта и линии, показавшие наивысшие значения хозяйственно-ценных признаков сои в

условиях Алматинской области, и являющиеся потенциальным ресурсом для селекции новых сортов.

3.8. Изучение эффективности идентифицированных KASP-маркеров на примере признаков урожайности сои

В целом, ПГАА позволил выявить 106 значимых ассоциаций, среди них 78 – для признаков, определяющих продуктивность (56 ассоциаций для КПУ, КСР, МТС и МСР, 16 – для размеров семян, 6 – для признака растрескиваемости стручков сои) и 28 – для признаков качества семян сои. Из них 46 значимых и стабильных ассоциаций, проявившиеся как минимум в двух годах испытаний (Таблица 24).

Таблица 24 – Общее количество идентифицированных значимых маркер-признак ассоциаций хозяйственно-ценных признаков сои

Признак		Количество значимых MTA	Количество значимых стабильных MTA	Количество ранее опубликованных MTA	Количество предполагаемых новых MTA
Признаки, определяющие продуктивность сои	КПУ	14	7	1	6
	КСР	16	6	2	4
	МТС	14	4	1	3
	МСР	12	3	1	2
	ДС	6	1	-	1
	ШС	5	3	1	2
	ТС	4	1	1	-
	РС	6	3	1	2
	Всего	78	28	8	20
Признаки качества	СП	16	9	8	1
	СМ	12	9	5	4
	Всего	28	18	13	5
Всего		106	46	21	25

На Рисунке 29 представлена карта хромосом сои с позициями идентифицированных ассоциаций в результате проведенного ПГАА (выделены красным и зеленым цветами), а также уже известные гены/QTL (синий цвет).

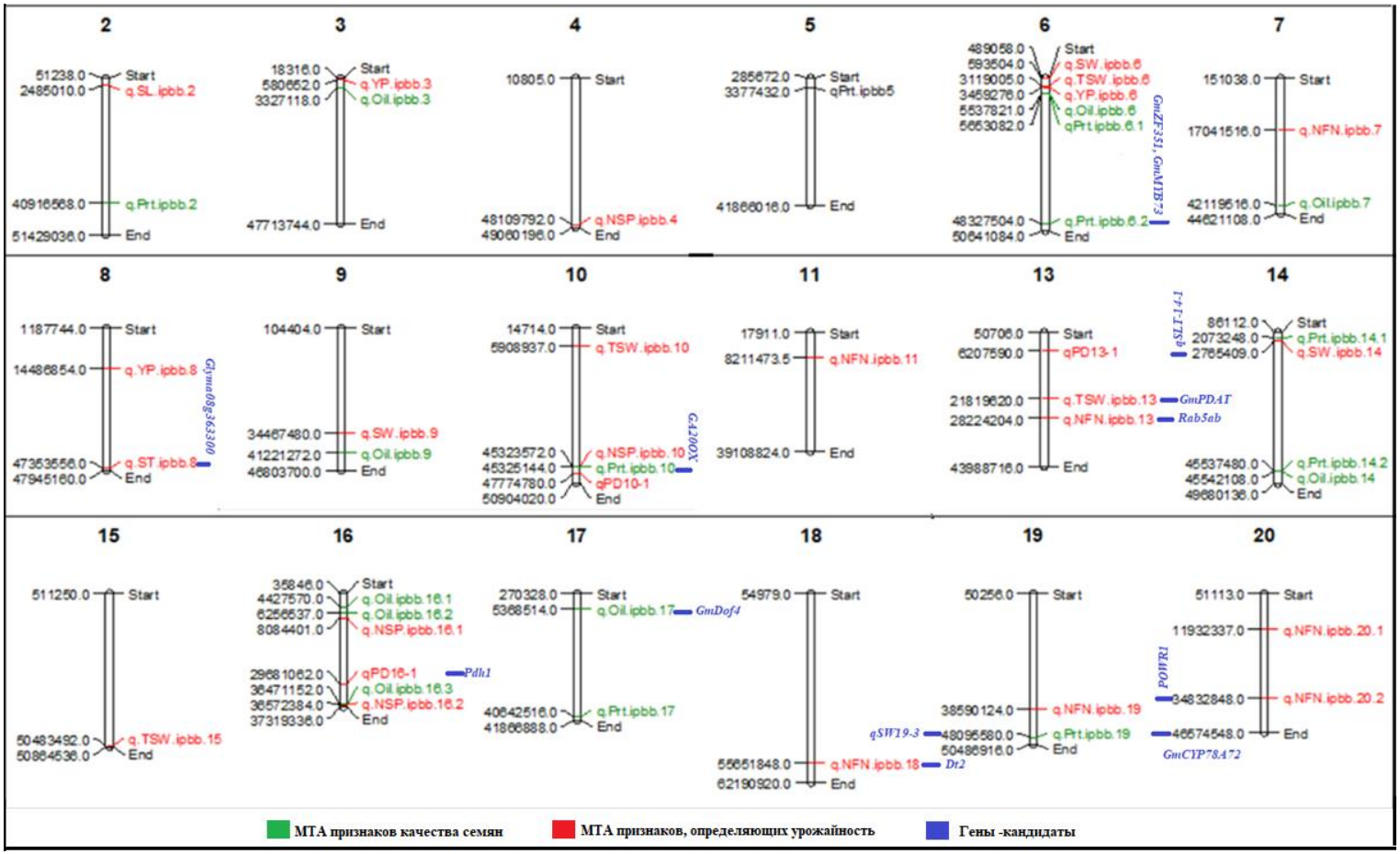


Рисунок 29 – Идентифицированные МТА хозяйственно-ценных признаков на хромосомах сои с генами – кандидатами / QTL. Примечание к рисунку – числами изображены номера хромосом в геноме сои, на хромосоме нанесены позиции идетифицированных МТА относительно друг – друга (п.н.). Сокращения наименования признаков SPC (seed protein content) – содержание протеина, SOC (seed oil content) – содержание масла, NFN (number of fertile nodes) – число продуктивных узлов, NSP (number of seed per plant) – число семян с растения, TSW (thousand seed weight) – масса тысячи семян, YP (yield per plant) – масса семян с растения, SL (seed length) – длина семян, SW (seed width) – ширина семян, ST (seed thickness) – толщина семян сои, PD (pod dehiscence) – растрескиваемость стручков.

Идентифицированные в настоящем исследовании MTA имеют значительные перспективы для применения в MAS сои, способствуя повышению ее урожайности и качества, в особенности юго-восточного региона страны, являющегося основным производителем сои. Предполагаемые новые ассоциации открывают возможности для более глубокого изучения генетической структуры признаков, определяющих продуктивность и качество сои.

Для подтверждения результатов ПГАА были отобраны высокозначимые MTA урожайности и разработаны KASP маркеры. Данные маркеры были проанализированы для дальнейшей валидации на независимой выборке, включающей в себя 94 линии гибридной популяции сортов Зара х Малета. Данные, полученные в ходе полевых экспериментов и структурного анализа гибридной популяции, выращенной в юго-восточном регионе Казахстана в 2023 году, были использованы для статистического анализа с целью подтверждения эффективности маркеров в выявлении генотипов с высоким потенциалом хозяйственно-ценных признаков сои. Анализ *t-mecma* выявил статистически значимую разницу по нескольким признакам урожайности для мировой коллекции сои в зависимости от двух аллелей, соответствующих KASP-маркеров. В Таблице 25 представлена информация о KASP маркерах и связь с хозяйственно-ценными признаками, на Рисунке 30 изображено результаты генотипирования независимой популяции Зара х Малета, с четким разделением генотипов по аллельным состояниям каждого KASP маркера.

Таблица 25 – Ассоциации набора KASP маркеров с хозяйственно-ценными признаками

№	Название KASP маркера	SNP-маркер	Признаки
1	*ipbb_gm_141*	S04_48109794	количество семян с растения, масса семян
2	*ipbb_sd_16*	S07_17041516	количество семян с растения, количество продуктивных узлов, масса семян
3	*ipbb_sd_17*	S20_3483847	количество семян с растения, количество продуктивных узлов, масса семян

Обнаруженные ассоциации S04_48109794, S07_17041516 и S20_3483847 были связаны с несколькими признаками, определяющих продуктивность сои, и были конвертированы в KASP маркеры *ipbb_gm_141, ipbb_sd_16* и *ipbb_sd_17.* Стоит отметить маркер *ipbb_sd_17,* который перекрывается с хорошо известным геном *POWR1*, влияющий на содержания протеина, масла семян сои, а также веса семян [200].

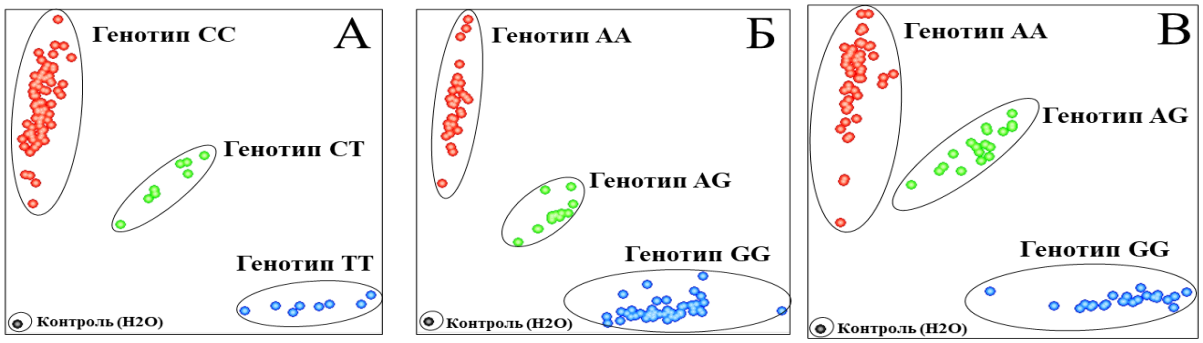


Рисунок 30 – Результаты амплификации 94 линий сои по 4 KASP-маркерам. А – *ipbb_sd_16*, Б – *ipbb_sd_17*, В – *ipbb_gm_141*. Отдельные кластеры точек на диаграмме отражают гомозиготные и гетерозиготные состояния образцов

Применение набора KASP-маркеров позволило выявить линии сои, которые обладали благоприятными аллелями SNP-маркеров, и показали повышенные показатели урожайности. Полученные результаты свидетельствуют о высокой эффективности применения набора KASP-маркеров для улучшения процесса селекции сои и создания новых сортов с повышенной урожайностью. В результате проведенных исследований был зарегистрирован патент на полезную модель №9289 [8]. Использование полезной модели позволит ускорить идентификацию высокопродуктивных и скороспелых образцов сои, уменьшить затраты на идентификацию селекционно-ценных линий сои и увеличить эффективность селекции.

ЗАКЛЮЧНИЕ

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

1) В ходе исследования было выявлено, что 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 Пирсона выявил обратную значимую корреляцию между признаком содержания протеина и масла, а также с компонентами урожайности, что имеет важное значение для селекционных программ ($P<0.05$). Также размеры семян имели положительную взаимосвязь с содержанием масла и урожайностью с растения ($P<0.05$).

2) В результате фенотипической оценки коллекции сои по 10 хозяйственно-ценным признакам, были выделены 9 отечественных и 42 зарубежных генотипов с высокими показателями урожайности 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, что позволило выявить 44 385 SNP-маркеров, для проведения дальнейшего изучения генетического разнообразия и проведения полногеномного анализа ассоциаций в высоком разрешении. Было установлено, что в основном 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) В результате ПГАА выявлены 106 значимых ассоциаций, среди которых 46 стабильных и превосходящих статистический порог. Из 46 стабильных ассоциаций были выявлены 25 потенциальны новых MTA, которые ранее не были описаны в литературе. При этом, 21 из 46 стабильных MTA 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 были 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. Также, наиболее значимые ассоциации, связанные с несколькими признаками, определяющих продуктивность сои, были конвертированы в KASP маркеры - *ipbb_gm_141, ipbb_sd_16* и *ipbb_sd_17,* показавшие высокую эффективность для применения их в селекционных программах ($P<0,05$).

Таким образом, полученные результаты имеют высокую теоретическую и практическую значимость, и могут быть эффективно использованы в отечественных селекционных программах для ускоренного отбора перспективных линий и повышения урожайности и качества сои.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Vollmann, J., Fritz, C. N., Wagentristl, H., & Ruckenbauer, P. Environmental and genetic variation of soybean seed protein content under Central European growing conditions // Journal of the Science of Food and Agriculture. - 2000**.** - Vol**.** 80, №9.- P.1300-1306.
2. Zhou, Y., Zhao, W., Lai, Y., Zhang, B., & Zhang, D. (2020). Edible plant oil: Global status, health issues, and perspectives. Frontiers in Plant Science. -2020. - Vol. 11**.**-P.1315. https://doi.org/10.3389/fpls.2020.01315.
3. Production, Supply and Distribution (PSD) Online. https://ipad.fas.usda.gov. 22 Апрель, 2025.
4. Abugalieva, S., Didorenko S., Anuarbek S., Volkova L., Gerasimova Y., Sidorik I., Turuspekov Y. Assessment of Soybean Flowering and Seed Maturation Time in Different Latitude Regions of Kazakhstan // PLOS ONE**.** -2016. DOI: 10.1371/journal.pone.0166894.
5. Zatybekov A., Abugalieva S., Didorenko S., Rsaliyev A., Maulenbay A., Fang C., Turuspekov Y. Genome-wide association study for charcoal rot resistance in soybean harvested in Kazakhstan // Vavilovskii Zhurnal Genet Selektsii. - 2023. - Vol.27, №6.-P.565-571. DOI 10.18699/VJGB-23-68.
6. Shi A., Gepts P., Song Q., Xiong H., Michaels T.E. Genome-wide association study and genomic prediction for soybean cyst nematode resistance in USDA common bean (*Phaseolus vulgaris*) core collection // Frontiers in Plant Science. – 2021. – Vol.12. – P.624156. DOI: 10.3389/fpls.2021.624156.
7. Wang J., Hu B., Jing Y., Hu X., Guo Y., Chen J., Liu Y., Hao J., Li W.-X., Ning H. Detecting QTL and candidate genes for plant height in soybean via linkage analysis and GWAS // Frontiers in Plant Science. – 2022. – Vol.12. – P.803820. DOI: 10.3389/fpls.2021.803820.
8. Пат. №9289 РК. Способ идентификации селекционно-ценных линий сои (*Glycine max* (L.) Merr.) с использованием набора KASP-маркеров. Zatybekov A., Doszhanova B., Turuspekov Y., Abugaliyeva S.
9. Carpenter, A. C., & Board, J. E. Branch yield components controlling soybean yield stability across plant populations // Crop Science. - 1997. - Vol. 37, №3**.** - P.885-891. https://doi.org/10.2135/cropsci1997.0011183X003700030031x.
10. Kudełka, W., Kowalska, M., & Popis, M. Quality of soybean products in terms of essential amino acids composition // Molecules. - 2021.- Vol.26**,** №16.- P.5071. https://doi.org/10.3390/molecules26165071.
11. Jelena Medic, Christine Atkinson, Charles R. Hurburgh Jr. Current Knowledge in Soybean Composition. JAOCS 100th Anniversary Virtual Issue. – 2014.-Vol.91, №3.- P.363-384.
12. Kim, I.-S.; Kim, C.-H.; Yang, W.-S. Physiologically Active Molecules and Functional Properties of Soybeans in Human Health. A Current Perspective // Int. J. Mol. Sci. -2021.- Vol. 22.-P. 4054. https://doi.org/ 10.3390/ijms22084054.
13. Qin P., Wang T., Luo Y. A review on plant-based proteins from soybean: Health benefits and soy product development // Journal of Agriculture and Food Research. - 2022. -Vol.7. – P.100265, https://doi.org/10.1016/j.jafr.2021.100265.

14. Guo B, Sun L, Jiang S, Ren H, Sun R, Wei Z, Hong H, Luan X, Wang J, Wang X, Xu D, Li W, Guo C, Qiu LJ. Soybean genetic resources contributing to sustainable protein production // Theor Appl Genet. – 2022. – Vol.135№11. - P.4095-4121. doi: 10.1007/s00122-022-04222-9.

15. Herbert Dei. Soybean as a Feed Ingredient for Livestock and Poultry. Recent Trends for Enhancing the Diversity and Quality of Soybean Products // Recent Trends for Enhancing the Diversity and Quality of Soybean Products. - 2011. DOI:10.5772/17601.

16. Chen, R., Qin, Z., Han, J., Wang, M., Taheripour, F., Tyner, W., O'Connor, D., & Duffield, J. Life cycle energy and greenhouse gas emission effects of biodiesel in the United States with induced land use change impacts // Bioresource Technology. - 2018. - Vol. 251. - P. 249–258. https://doi.org/10.1016/j.biortech.2017.12.031.

17. Rodríguez-Navarro, D. N., Oliver, I. M., Albareda, M., & Ruiz-Sainz, J. E. Soybean interactions with soil microbes: agronomical and molecular aspects // Agronomy for Sustainable Development. – 2011. – Vol. 31, №1. – P.173–190.

18. Kabita Kumari Shah, Bindu Modi, Hari Prasad Pandey, Arjun Subedi, Geeta Aryal, Meena Pandey, Jiban Shrestha. Diversified Crop Rotation: An Approach for Sustainable Agriculture Production // Advances in agriculture. - 2021. DOI:10.1155/2021/8924087.

19. Global market report: Soybean 2024**.** https://www.iisd.org/publications/report/2024-global-market-report-soybean. 20.04.2025.

20. Soybean prices, economic growth and poverty in Argentina and Brazil. Food and Agriculture Organization of the United Nations Rome. - 2017. http://openknowledge.fao.org. 20.04.2025.

21. Eric J. Sedivy, Faqiang Wu and Yoshie Hanzawa. Soybean domestication: the origin, genetic architecture and molecular bases // New Phytologist. – 2017. – 214. 539–553 doi: 10.1111/nph.14418.-2016.

22. Baohui Liu, Toshiro Fujita, Ze-Hong Yan, Shinichi Sakamoto, Donghe Xu, Jun Abe. QTL Mapping of Domestication-related Traits in Soybean (*Glycine max*) // Annals of Botany. – 2007. – Vol.100, №5. – P. 1027–1038. https://doi.org/10.1093/aob/mcm149.

23. De Maria, M., Robinson, E. J. Z., Kangile, J. R., Kadigi, R. M. J., & others. Global soybean trade – The geopolitics of a bean (Technical report). – 2020. https://doi.org/10.34892/7yn1-k494.

24. Shurtleff, W., & Aoyagi, A. History of soybeans and soyfoods in Central Asia (1876–2008) // Extensively annotated bibliography and sourcebook. Soyinfo Center. – 2008.

25. Didorenko, S., Araliev, Z., Alenkhanova, I., Sidorkin, I., et al. Diversification of crop production by means of spreading soybeans to the northern regions of the Republic of Kazakhstan // Biosciences Biotechnology Research Asia. – 2016.- Vol.13, №1.-P. 23–30. https://doi.org/10.13005/bbra/1998.

26. Mishra R., Tripathi M., Sikarwar R., Singh Y. and Tripathi N. Soybean (*Glycine max* L. Merrill): A Multipurpose Legume Shaping Our World // Plant Cell

Biotechnology and Molecular Biology. - 2024. - Vol. 25. – P.17-37. DOI: 10.56557/PCBMB/2024/v25i3-48643.

27. Корсаков Н.И. Определение видов и разновидностей сои: Метод. указания. Л.: ВНИИР. - 1972. - 189 с.

28. Soybeans. https://aitcla.org/soybeans. 25.04.2025.

29. Енкен, В. Б. Соя / В. Б. Енкен. - М.: Сельхоз.- 1959. - 622 с.

30. Федотов, В.А. Соя в России: (монография). - Москва: Агролига России. - 2013. - 432 с.

31. Черноголовин В. П., Бакаева Е. В. Соя: в Казахстане. - Зерновое хозяйство. - 1975.- с. 41-42.

32. Fehr W. R., Caviness C. E. Stages of soybean development. - Iowa Coop. Ext. Serv. Iowa Agric. Home Econ. Exp. Stn. Spec. Rep. – 1977.

33. Destro D., Carpentieri-Pípolo V., Afonso R., Kiihl S., Alves De Almeida L. Photoperiodism and Genetic Control of the Long Juvenile Period in Soybean: A Review // Crop Breeding and Applied Biotechnology. - 2001.- Vol. 1, №1. DOI:10.13082/1984-7033.v01n01a10.

34. Lin, X., Liu, B., Weller James, L., Abe, J., and Kong, F. Molecular mechanisms for the photoperiodic regulation of flowering in soybean // *J. Integr. Plant Biol*. – 2021.- Vol.63.-P.981–994.

35. Zhai H., Wan Z., Jiao S., Zhou J., Xu K., Nan H., Liu Y., Xiong S., Fan R., Zhu J., Jiang W., Pang T., Luo X., Wu H., Yang G., Bai X., Kong F., Xia Z. GmMDE genes bridge the maturity gene E1 and florigens in photoperiodic regulation of flowering in soybean // Plant Physiology. – 2022. – Vol.189, №2. – P.1021–1036. DOI: 10.1093/plphys/kiac092.

36. Kou, K., Su, T., Wang, Y. *et al*. Natural variation of the *Dt2* promoter controls plant height and node number in semi-determinant soybean // Mol Breeding. – 2021. – Vol.41, №40. https://doi.org/10.1007/s11032-021-01235-y.

37. Xiaoyu Saldivar, Ya-Jane Wang, Pengying Chen, Anfu Hou. Changes in chemical composition during soybean seed development // Food Chemistry. – 2011.- Vol.124, № 4. - P. 1369-1375. https://doi.org/10.1016/j.foodchem.2010.07.091.

38. Kawamura, S. Quantitative paper chromatography of sugars of the cotyledon, hull, and hypocotyl of soybeans of selected varieties. - Tech. Bull. Fac. Agr.: Kagawa Univ. – 1967. -Vol.18.- P.117–131.

39. Srii, V.S., Nagarajappa, N., & Ranjitha, H.P. Seed coat composition in black and white soybean seeds with differential water permeability// Plant biology. - 2023. DOI:10.1111/plb.13551.

40. Sun Y., Gong Y. Research advances on the hard seededness trait of soybean and the underlying regulatory mechanisms // Front. Plant Sci.-2024.-Vol.15. https://doi.org/10.3389/fpls.2024.1419962.

41. William M. Harris. 1987. Comparative Ultrastructure of Developing Seed Coats of "Hard-Seeded" and "Soft-Seeded" Varieties of Soybean, *Glycine max* (L.) Merr. // International Journal of Plant Sciences. – 1987. – Vol.148**,** № 3. https://doi.org/10.1086/337660.

42. Seed Proteins Biochemistry, Genetics. - Advances in Agricultural Biotechnology / editors Werner Gottschalk, Hermann P. Müller. - Nutritive Value Book© 1983.

43. Patil, Gunvant, Chaudhary, Juhi, Vuong, TriD., Jenkins, Brian, Qiu, Dan , Kadam, Suhas, Shannon, Grover J., Nguyen, Henry T. Development of SNP Genotyping Assays for Seed Composition Traits in Soybean // International Journal of Plant Genomics. - 2017.- 12 pages. https://doi.org/10.1155/2017/6572969.

44. Welison Andrade Pereira et al. Dynamics of reserves of soybean seeds during the development of seedlings of different commercial cultivars // J. Seed Sci. – 2015. -Vol.37, №01. https://doi.org/10.1590/2317-1545v37n1142202.

45. Pietrzak, L.N., Frégeau-Reid, J.A., Chatson, B., & Blackwell, B.D. Observations on water distribution in soybean seed during hydration processes using nuclear magnetic resonance imaging // Canadian Journal of Plant Science. – 2002.- Vol. 82.-P. 513-519.

46. Bruce German, Srinivasan Damodaran, John E. Kinsella. Thermal dissociation and association behavior of soy proteins // J. Agric. Food Chem. -1982. - Vol. 30, №5.- P.807–811. https://doi.org/10.1021/jf00113a002.

47. Yang Yuan, Zhi-Li Wan, Xiao-Quan Yang, Shou-Wei Yin. Associative interactions between chitosan and soy protein fractions: Effects of pH, mixing ratio, heat treatment and ionic strength // Food Research International. -2014.- Vol.55. -P. 207-214. https://doi.org/10.1016/j.foodres.2013.11.016.

48. Clemente, T. E. and E. B. Cahoon. Soybean Oil: Genetic approaches for modification of functionality and total content // Plant Physiol.-2009.-Vol.151. - P.1030-1040.

49. Doszhanova B.N., Satou E., Suzuki T., Turuspekov Y. Phytic acid analysis of soybean seeds grown in the south-east of Kazakhstan // thesis in the International Scientific Conference of Young Scientists. – 2023.

50. Дидоренко С.В., Кисетова Э.М., Касенов Р.Ж., Байжанов Ж.Р., Кушанова Р.Ж., Сагит И. Продуктивность и качество сортов сои созданных на разных этапах селекционных работ в казахском научно – исследовательском институте земледелия и растениеводства // Исследования, результаты. - 2024.- №2. https://doi.org/10.37884/2-1-2024/542.

51. Didorenko S., Zakiyeva A., Sidorik I., Abuglieva A, Kudaibergenov M, Iskakov A. Diversification of crop production by means of spreading soybeans to the northern regions of the Republic of Kazakhstan // Biosciences Biotechnology Research Asia**.** – 2016.- Vol. 13, №1. http://dx.doi.org/10.13005/bbra/1998.

52. Didorenko S.V., Sidorik I.V., Zinchenko A.V., Zakieva A.A. Creation of ultra-ripening and early maturing soybean varieties for conditions of Northern Kazakhstan. - 2024. DOI:10.32523/2616-7034-2024-147-2-7-25.

53. Yerzhebayeva, R., Didorenko, S., Amangeldiyeva, A., Daniyarova, A., Mazkirat, S., Zinchenko, A., & Shavrukov, Y. Marker-Assisted Selection for Early Maturing *E* Loci in Soybean Yielded Prospective Breeding Lines for High Latitudes of Northern Kazakhstan // Biomolecules. – 2023. -Vol. 13*, №*7. – P.1146. https://doi.org/10.3390/biom13071146.

54. Пат. № 2020/0108.2. Способ идентификации высокопродуктивных и устойчивых к грибным болезням линий сои (Glycine max) с использованием набора ДНК-маркеров по технологии KASP. Затыбеков А. К., Абугалиева С.И, Досжанова Б. Н., Туруспеков Е.К. – 2020.

55. Zatybekov, A., Abugalieva, S., Didorenko, S., Rsaliyev, A, Turuspekov, Y. GWAS of a soybean breeding collection from South East and South Kazakhstan for resistance to fungal diseases // Vavilovskii Zhurnal Genetiki i Selektsii. 2018. – Vol. 22, №5. - P.: 536-543. DOI: 10.18699/VJ18.392.

56. Fang, C., Ma, Y., Wu, S., Liu, Z., Wang, Z., Yang, R., … & Tian, Z. Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean // Genome Biology. -2017. – Vol. 18, **№**1.-P. 161. https://doi.org/10.1186/s13059-017-1289-9rex.

57. Nakano, Y., & Kobayashi, Y. Genome-wide association studies of agronomic traits consisting of field - and molecular-based phenotypes // Reviews in Agricultural Science. – 2020. - Vol. 8**.** - P.28–45. https://doi.org/10.7831/ras.8.0_28

58. Assefa, T., Otyama, P. I., Brown, A. V., Kalberer, S. R., Kulkarni, R. S., & Cannon, S. B. Genome-wide associations and epistatic interactions for internode number, plant height, seed weight and seed yield in soybean // BMC Genomics. – 2019. -Vol. 20, **№** 527. https://doi.org/10.1186/s12864-019-5907-7

59. Patil, G., Mian, R., Vuong, T., Pantalone, V., Song, Q., Chen, P., Shannon, G. J., Carter, T. C., & Nguyen, H. T. Molecular mapping and genomics of soybean seed protein: A review and perspective for the future // Theoretical and Applied Genetics. – 2017. Vol. 130, **№**10. – P.1975–1991. https://doi.org/10.1007/s00122-017-2955-8

60. Diers, B. W., Specht, J. E., Graef, G. L., Song, Q., Rainey, K. M., Ramasubramanian, V., Liu, X., Myers, C. L., Stupar, R. M., An, Y. Q. C., & Beavis, W. D. Genetic architecture of protein and oil content in soybean seed and meal // The Plant Genome. – 2023. - Vol.16, №1. https://doi.org/10.1002/tpg2.20308

61. Leamy, L. J., Zhang, H., Li, C., Chen, C. Y., & Song, B. H. (2017). A genome-wide association study of seed composition traits in wild soybean (Glycine soja) // BMC Genomics. – 2017. – Vol. 18. https://doi.org/10.1186/s12864-016-3397-4

62. Yang C, Yan J, Jiang S, Li X, Min H, Wang X, Hao D. Resequencing 250 Soybean Accessions: New Insights into Genes Associated with Agronomic Traits and Genetic Networks // Genomics Proteomics Bioinformatics. – 2022. Vol.20, №1. – P.29-41. doi: 10.1016/j.gpb.2021.02.009.

63. Beche, E., Gillman, J. D., Song, Q., Nelson, R., Beissinger, T., Decker, J., Shannon, G., & Scaboo, A. M. Nested association mapping of important agronomic traits in three interspecific soybean populations // Theoretical and Applied Genetics. - 2020.- Vol. 133, **№**4. - P.1227–1241. https://doi.org/10.1007/s00122-019-03529-4.

64. Hina Arya, Mohan B. Singh, Prem L. Bhalla. Towards Developing Drought-smart Soybeans // Front. Plant Sci. – 2021. -Vol.12, № 6. https://doi.org/10.3389/fpls.2021.750664

65. Funatsuki H, Suzuki M, Hirose A, Inaba H, Yamada T, Hajika M, Komatsu K, Katayama T, Sayama T, Ishimoto M, Fujino K. Molecular basis of a

shattering resistance boosting global dissemination of soybean // Proc Natl Acad Sci U S A. - 2014. – Vol.111, №50. -P.17797-802. doi: 10.1073/pnas.1417282111.

66. Zatybekov, A., Abugalieva, S., Didorenko, S., Rsaliyev, A., Maulenbay, A., Fang, C., & Turuspekov, Y. Genome-wide association study for charcoal rot resistance in soybean harvested in Kazakhstan // Vavilov Journal of Genetics and Breeding. 2023.-Vol. 27*, №* 6. -P. 565–571. https://doi.org/10.18699/VJGB-23-68

67. Liu, S., Kandoth, P. K., Warren, S. D., Yeckel, G., Heinz, R., Alden, J., Yang, C., Jamai, A., El-Mellouki, T., Juvale, P. S., Hill, J., Baum, T. J., Cianzio, S., Whitham, S. A., Korkin, D., Mitchum, M. G., & Meksem, K. A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens // Nature. – 2012. – Vol.492*, №*7428. – P.256–260. https://doi.org/10.1038/nature11651

68. Doszhanova B.N., Suzuki T., Zatybekov A.K., Turuspekov Y.K. Genotyping soybean collection for resistance to *Heterodera Glycines ichinohe*. – thesis in the International Conference on plant biology and biotechnology. - 2024

69. Qiong Wang, Wei Zhang, Wenjing Xu, Hongmei Zhang, Xiaoqing Liu, Xin Chen, Huatao Chen. Genome-Wide Association Study and Identification of Candidate Genes Associated with Seed Number per Pod in Soybean // International Journal of Molecular Sciences. – 2024. https://doi.org/10.3390/ijms25052536

70. Zhang, Y., Bhat, J. A., Zhang, Y., & Yang, S. Understanding the Molecular Regulatory Networks of Seed Size in Soybean // International Journal of Molecular Sciences. – 2024. – Vol. 25*, №*3. – P.1441. https://doi.org/10.3390/ijms25031441

71. Xu, Y., Li, H., Li, G., Wang, X., Cheng, L., & Zhang, Y. Mapping quantitative trait loci for seed size traits in soybean (*Glycine max* L. Merr.) // Theoretical and Applied Genetics**. –** 2011. – Vol.122, №3. – P.581–594. https://doi.org/10.1007/s00122-010-1471-xP

72. Hu, Y., Liu, Y., Tao, J.-J., Lu, L., Jiang, Z.-H., Wei, J.-J., Wu, C.-M., Yin, C.-C., Li, W., Bi, Y.-D., Lai, Y.-C., Wei, W., Zhang, W.-K., Chen, S.-Y., & Zhang, J.-S. GmJAZ3 interacts with GmRR18a and GmMYC2a to regulate seed traits in soybean // Journal of Integrative Plant Biology**. –** 2023. -Vol.65, №8. – P. 1983–2000. https://doi.org/10.1111/jipb.13494

73. Duan Z, Zhang M, Zhang Z, Liang S, Fan L, Yang X, Yuan Y, Pan Y, Zhou G, Liu S, Tian Z. Natural allelic variation of GmST05 controlling seed size and quality in soybean // Plant Biotechnol J. – 2022. – Vol. 20. P.1807-1818. doi: 10.1111/pbi.13865

74. Chen Y, Xiong Y, Hong H, et al. Genetic dissection of and genomic selection for seed weight, pod length, and pod width in soybean // The Crop Journal. – 2023. – Vol. 11, №3. – P. 832-841. https://doi.org/10.1016/j.cj.2022.11.006

75. Xu, C., Wu, T., Yuan, S., Sun, S., Han, T., Song, W., & Wu, C. Can Soybean Cultivars with Larger Seed Size Produce More Protein, Lipids, and Seed Yield? A Meta-Analysis // *Foods.* – 2022. – Vol. 11, № 24. - P.4059. https://doi.org/10.3390/foods11244059

76. Zhang Q., Tu B., Liu C., Liu X. Pod anatomy, morphology and dehiscing forces in pod dehiscence of soybean (*Glycine max* (L.) Merrill) // Flora. - 2018.-Vol.248. – P. 48-53. DOI 10.1016/j.flora.2018.08.014

77. Funatsuki H., Hajika M., Hagihara S., Yamada T., Tanaka Y., Tsuji H., Ishimoto M., Fujino K. Confirmation of the location and the effects of a major QTL controlling pod dehiscence, *qPDH1*, in soybean // Breed. Sci. 2008.- Vol.58.-P.63-69. DOI 10.1270/jsbbs.58.63

78. Parker T.A., Lo S., Gepts P. Pod shattering in grain legumes: emerging genetic and environment-related patterns // Plant Cell. – 2021. – Vol.33, №2. – P.179-199. DOI 10.1093/plcell/koaa025

79. Kang S.T., Kwak M., Kim H.K., Choung M.G., Han W.Y., Baek I.Y., Kim M.Y., Van K., Lee S.H. Population-specific QTLs and their different epistatic interactions for pod dehiscence in soybean [*Glycine max* (L.) Merr.] // Euphytica. – 2009.-Vol.166, №1. -P.15-24. DOI 10.1007/s10681-008-9810-6

80. Seo J.H., Kang B.K., Dhungana S.K., Oh J.H., Choi M.S., Park J.H., Shin S.O., Kim H.S., Baek I.Y., Sung J.S., Jung C.S., Kim K.S., Jun T.H. QTL mapping and candidate gene analysis for pod shattering tolerance in soybean (*Glycine max*) // Plants. – 2020. – Vol.9, №9. – P.1163. DOI 10.3390/plants9091163

81. Proposed draft regional standard for non-fermented soybean products (N06-2007) // Joint FAO/WHO food standards programme. FAO/WHO coordinating committee for Asia. - 2012

82. ГОСТ 12220–96. Шрот соевый. Технические условия. – М.: Изд-во стандартов, 1996. – 6 с.

83. https://www.apk-inform.com. 1.05.2025

84. FAO/WHO. Codex Alimentarius. Food and Agriculture Organization of the United Nations and World Health Organization. https://www.fao.org/fao-who-codexalimentarius. 1.05.2025.

85. Toomer, O.T.; Oviedo, E.O.; Ali, M.; Patino, D.; Joseph, M.; Frinsko, M.; Vu, T.; Maharjan, P.; Fallen, B.; Mian, R. Current Agronomic Practices, Harvest & Post-Harvest Processing of Soybeans (Glycine max)–A Review // Agronomy. -2023. – Vol.13, № 427. https://doi.org/ 10.3390/agronomy13020427

86. ГОСТ 17109–88. Соя. Требования при заготовках и поставках. – М.: Госстандарт СССР. - 1988.

87. ГОСТ 12044–93. Зерно. Методы определения влажности. – М.: Изд-во стандартов. - 1993.

88. China Notified Draft National Standard for Soybeans // Foreign Agricultural Service (FAS), USDA. – 2023. – https://www.fas.usda.gov/data/china-china-notified-draft-national-standard-soybeans. 11.04.2025.

89. P.B. Cregan, T. Jarvik, A. L. Bush, R. C. Shoemaker, K. G. Lark, A. L. Kahler, N. Kaya, T. T. VanToai, D. G. Lohnes, J. Chung, J. E. Specht. An Integrated Genetic Linkage Map of the Soybean Genome // Crop science. – 1999. – Vol.39, № 5. P.1464-1490. https://doi.org/10.2135/cropsci1999.3951464x

90. Schmutz, J., Cannon, S., Schlueter, J. *et al*. Genome sequence of the palaeopolyploid soybean // Nature. – 2010. – Vol. 463. – P. 178–183. https://doi.org/10.1038/nature08670

91. Shoemaker RC, Schlueter J, Doyle JJ. Paleopolyploidy and gene duplication in soybean and other legumes // Curr Opin Plant Biol. – 2006. – Vol.9, №2. doi: 10.1016/j.pbi.2006.01.007.

92. Grant D, Nelson RT, Cannon SB, Shoemaker RC. SoyBase, the USDA-ARS soybean genetics and genomics database // Nucleic Acids Res. – 2010. – Vol.38. doi: 10.1093/nar/gkp798.

93. Jie Li, Qingyang Ni, Guangqi He, Jiale Huang, Haoyu Chao, Sida Li, Ming Chen, Guoyu Hu, James Whelan, Huixia Shou, SoyOD: An Integrated Soybean Multi-omics Database for Mining Genes and Biological Research // Genomics, Proteomics & Bioinformatics. – 2024. https://doi.org/10.1093/gpbjnl/qzae080.

94. Li H, Chen T, Jia L, Wang Z, Li J, Wang Y, Fu M, Chen M, Wang Y, Huang F, Jiang Y, Li T, Zhou Z, Li Y, Yao W, Wang Y. SoybeanGDB: A comprehensive genomic and bioinformatic platform for soybean genetics and genomics // Comput Struct Biotechnol J. – 2023. – Vol.21. - P.3327-3338. doi: 10.1016/j.csbj.2023.06.012.

95. Reyazul RoufMir, Pavana J. Hiremath, Oscar Riera-Lizarazu, and Rajeev K. Varshney. Diagnostics in Plant Breeding. Evolving Molecular Marker Technologies in Plants: From RFLPs to GBS. – 2013. – P. 229-246. DOI:10.1007/978-94-007-5687-8

96. B.W. Diers t, p. Keim 2, W.R. Fehr 1, and R. C. Shoemaker. RFLP analysis of soybean seed protein and oil content // Theor Appl Genet. – 1992. – Vol. 83. – P. 608 – 612. DOI:10.1007/BF00226905

97. Fasoula, Vasilia A. Harris, Donna K. Boerma, H. Roger. 2004. Validation and Designation of Quantitative Trait Loci for Seed Protein, Seed Oil, and Seed Weight from Two Soybean Populations // Crop Science. – 2004. DOI:10.2135/cropsci2004.1218

98. Tingey SV, del Tufo JP. Genetic analysis with random amplified polymorphic DNA markers // Plant Physiol. – 1993. – Vol.101, №2. doi: 10.1104/pp.101.2.349.

99. Estri Laras Arumingtyas, Wahyu Widoretno, Serafinah Indriyani. Somaclonal variations of Soybeans (*Glycine Max*. L. Merr) stimulated by drought stress based on random amplified polymorphic DNAs (RAPDs) //American Journal of Molecular Biology. – 2012. - Vol.2 No.1. doi:10.4236/AJMB.2012.21009

100. Jafar Mammadov,1 Rajat Aggarwal,1 Ramesh Buyyarapu,1 and Siva Kumpatla2. SNP Markers and Their Impact on Plant Breeding // International Journal of Plant Genomics. – 2012. doi:10.1155/2012/728398

101. Mba C and Tohme J. Use of AFLP markers in surveys of plant diversity // Methods in Enzymology. – 2005. – Vol. 395. – P.177–201

102. Singh RK, Bhatia VS, Bhat KV, Mohapatra T, Singh NK, Bansal KC, Koundal KR. SSR and AFLP based genetic diversity of soybean germplasm differing in photoperiod sensitivity // Genet Mol Biol. - 2010. – Vol.33, №2. doi: 10.1590/S1415-47572010005000024.

103. Song, Q., Hyten, D. L., Jia, G., Quigley, C. V., Fickus, E. W., Nelson, R. L., & Cregan, P. B. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean // *PLOS ONE. – 2013. – Vol. 8, №*1. https://doi.org/10.1371/journal.pone.0054985

104. Zhang, H., Li, Y., Wang, Y., Wang, L., Zhang, Y., Wang, X., ... & Li, Y. SoySNP618K array: A high-resolution single nucleotide polymorphism platform as a

valuable genomic resource for soybean genetics and breeding // Journal of Integrative Plant Biology. – 2021. – Vol.63, №12. – P.2089–2101. https://doi.org/10.1111/jipb.13202

105. Bhat, J.A., Yu, D., Bohra, A. *et al.* Features and applications of haplotypes in crop breeding // Commun Biol. – 2021. – Vol.4. https://doi.org/10.1038/s42003-021-02782-y

106. Martin W Ganal, Thomas Altmann, Marion S Röder. SNP identification in crop plants // Current Opinion in Plant Biology. – 2009. – Vol. 12, №2. – P. 211-217.10.1016/j.pbi.2008.12.009

107. Uzma Majeed1, Essam Darwish2, Shoaib Ur Rehman1 & Xueyong Zhang1. Kompetitive Allele Specific PCR (KASP): A Singleplex Genotyping Platform and Its Application // Journal of Agricultural Science. – 2019. -Vol. 11, No. 1. doi:10.5539/jas.v11n1p11

108. Rosso, M. L., Shang, C., Song, Q., Escamilla, D., Gillenwater, J., & Zhang, B. Development of Breeder-Friendly KASP Markers for Low Concentration of Kunitz Trypsin Inhibitor in Soybean Seeds // International Journal of Molecular Sciences. – 2021. - Vol.22, №5. https://doi.org/10.3390/ijms22052675

109. Rex Bernardo. Molecular Markers and Selection for Complex Traits in Plants: Learning from the Last 20 Years // Crop science. – 2008. – Vol.48, Issue 5. – P.1649-1664. https://doi.org/10.2135/cropsci2008.03.0131

110. Mackay I, Powell W. Methods for linkage disequilibrium mapping in crops // Trends Plant Sci. – 2007. – Vol.12, №2. – P.57-63. doi: 10.1016/j.tplants.2006.12.001.

111. Alonso-Blanco C, Koornneef M, van Ooijen JW. QTL analysis // Methods Mol Biol. – 2006. – Vol.323. -P. 79-99. doi: 10.1385/1-59745-003-0:79. PMID: 16739569

112. Mackay TFC. Quantitative trait loci in Drosophila // Nat Rev Genet. – 2001. – Vol. 2.- P. 11-20. DOI:10.1038/35047544

113. Funatsuki H, Hajika M, Yamada T, Suzuki M, Hagihara S, Tanaka Y, Fujita S, Ishimoto M, Fujino K. Mapping and use of QTLs controlling pod dehiscence in soybean // Breed Sci. – 2012. – Vol.61, №5. doi: 10.1270/jsbbs.61.554.

114. Li, X., Shao, Z., Tian, R. *et al.* Mining QTLs and candidate genes for seed protein and oil contents across multiple environments and backgrounds in soybean // Mol Breeding. -2019. – Vol. 39, №139. https://doi.org/10.1007/s11032-019-1055-7

115. Huang X, Han B. Natural variations and genome-wide association studies in crop plants // Annu Rev Plant Biol. – 2014. – Vol. 65. doi: 10.1146/annurev-arplant-050213-035715.

116. Alseekh, S., Kostova, D., Bulut, M. *et al.* Genome-wide association studies: assessing trait characteristics in model and crop plants // Cell. Mol. Life Sci. – 2021. – Vol.78. - P.5743–5754. https://doi.org/10.1007/s00018-021-03868-w

117. Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, Derks EM. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis // Int J Methods Psychiatr Res. – 2018. – Vol. 27, №2. doi: 10.1002/mpr.1608.

118. Kaler, A.S., Purcell, L.C. Estimation of a significance threshold for genome-wide association studies // BMC Genomics. – 2019. – Vol.20, № 618. https://doi.org/10.1186/s12864-019-5992-7

119. Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., & Zhang, Z. GAPIT: Genome association and prediction integrated tool // Bioinformatics. – 2012. – Vol. 28, №18. – P.2397–2399. https://doi.org/10.1093/bioinformatics/bts444

120. Segura, V., Vilhjálmsson, B., Platt, A. *et al*. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat Genet. – 2012. – Vol. 44. P.825–830. https://doi.org/10.1038/ng.2314

121. Susmitha P, Kumar P, Yadav P, Sahoo S, Kaur G, Pandey MK, Singh V, Tseng TM, Gangurde SS. Genome-wide association study as a powerful tool for dissecting competitive traits in legumes // Front Plant Sci. – 2023. – Vol. 14, №14. doi: 10.3389/fpls.2023.1123631.

122. Pavan S, Delvento C, Ricciardi L, Lotti C, Ciani E, D'Agostino N. Recommendations for Choosing the Genotyping Method and Best Practices for Quality Control in Crop Genome-Wide Association Studies // Front Genet. – 2020. – Vol. 5, №11. doi: 10.3389/fgene.2020.00447.

123. Liu, X., Song, Y., Jiang, W., et al. Advances in soybean breeding strategies over the past decade // Theoretical and Applied Genetics. -2020.- Vol.133, №5. – P. 1249–1262.

124. Singh, R. J. (Ed.). Genetic resources, chromosome engineering, and crop improvement: Oilseed crops. – CRC Press. -2007. -Vol. 4.

125. Haun WJ, Hyten DL, Xu WW, Gerhardt DJ, Albert TJ, Richmond T, Jeddeloh JA, Jia G, Springer NM, Vance CP, Stupar RM. The composition and origins of genomic variation among individuals of the soybean reference cultivar Williams 82 // Plant Physiol. – 2011. – Vol.155, №2. doi: 10.1104/pp.110.166736.

126. www.kzniizr.kz. 1.03.2025

127. https://www.gov.kz/memleket/entities/moa. 10.04.2025

128. https://agroinfo.kz. 10.04.2025

129. https://kazatu.edu.kz. 10.04.2025

130. Государственный реестр селекционных достижений, рекомендуемых к использованию в Республике Казахстан, и Перечень перспективных сортов сельскохозяйственных растений. https://adilet.zan.kz/rus/docs/V090005759. 10.04.2025

131. Кипшакбаева, Г., & Тлеулина, З. Влияние условий возделывания сортов сои различного происхождения на формирование показателей продуктивности и качества: влияние условий возделывания сортов сои различного происхождения на формирование показателей продуктивности и качества // Ġylym žáne bìlìm. – 2022. – Vol.2, №67. – P. 212–223.

132. Prince SJ, Joshi T, Mutava RN, Syed N, Joao Vitor Mdos S, Patil G, Song L, Wang J, Lin L, Chen W, Shannon JG, Valliyodan B, Xu D, Nguyen HT. Comparative analysis of the drought-responsive transcriptome in soybean lines contrasting for canopy wilting // Plant Sci. – 2015. – Vol. 240. P.65-78. doi: 10.1016/j.plantsci.2015.08.017.

133. Liu, S., Kandoth, P. K., Warren, S. D., et al. A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens // Nature. – 2012. – Vol. 492.-P. 256–260.

134. Abdulwahab S. Shaibu, Bin Li, Shengrui Zhang, Junming Sun. Soybean cyst nematode-resistance: Gene identification and breeding strategies // The Crop Journal. – 2020. -Vol 8, Issue 6.- P. 892-904. https://doi.org/10.1016/j.cj.2020.03.001.

135. Ravelombola, W., Qin, J., Shi, A., Song, Q., Yuan, J., Wang, F., Chen, P., Yan, L., Feng, Y., Zhao, T., Meng, Y., Guan, K., Yang, C., & Zhang, M. Genome-wide association study and genomic selection for yield and related traits in soybean // PLOS ONE. – 2021. – Vol.16, №8. https://doi.org/10.1371/journal.pone.0255761

136. Fang, C., Ma, Y., Wu, S., et al. Genome-wide association studies for protein and oil content in soybean // Theoretical and Applied Genetics. - *2017.- Vol. 130, №4. P. 841-856*

137. Eskandari, M.; Cober, E.R.; Rajcan, I. Genetic Control of Soybean Seed Oil: II. QTL and Genes That Increase Oil Concentration without Decreasing Protein or with Increased Seed Yield // Theor. Appl. Genet. – 2013. – Vol.126. P. 1677–1687. https://doi.org/10.1007/s00122-013-2083-z

138. He, C., Holme, J., Anthony, J. SNP Genotyping: The KASP Assay / In: Fleury, D., Whitford, R. (eds) Crop Breeding. - Methods in Molecular Biology. 2014. – Vol. 1145. - Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-0446-4_7

139. Semagn, K., Babu, R., Hearne, S., & Olsen, M. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement // *Molecular Breeding. – 2014. – Vol. 33, №*1. – P. 1–14. https://doi.org/10.1007/s11032-013-9917-x

140. Rasheed, A., et al. Crop Breeding Chips and Genotyping Platforms: Progress, Challenges, and Perspectives // Molecular Plant. – 2017. – Vol.10*, №*8. – P. 1047–1064. https://doi.org/10.1016/j.molp.2017.06.008

141. Об утверждении Комплексного плана по развитию селекции и семеноводства сельскохозяйственных культур Республики Казахстан на 2024 – 2028 годы. https://adilet.zan.kz/rus/docs/P2400000427. 10.04.2025

142. Bellaloui, Nacer, Reddy, Krishna N.,Bruns, H. Arnold,Gillen, Anne M.,Mengistu, Alemu, Zobiole, Luiz H.S., Fisher, Daniel K.,Abbas, Hamed K,Zablotowicz, Robert M.,Kremer, Robert J. Soybean seed composition and quality: Interactions of environment, genotype, and management practices. - Soybeans: Cultivation, Uses and Nutrition. – 2011.

143. Доспехов Б.А., 1979. Методика полевого опыта

144. Корсаков, Н.И. Соя (систематика и основы селекции). - автореф. докт. дисс.-Л. -1973

145. Luo S, Jia J, Liu R, Wei R, Guo Z, Cai Z, Chen B, Liang F, Xia Q, Nian H, Cheng Y. Identification of major QTLs for soybean seed size and seed weight traits using a RIL population in different environments // Front Plant Sci. – 2023. – Vol.11, №13. doi: 10.3389/fpls.2022.

146. ГОСТ 18057-64 Семена масличные. Методы определения масличности. – М.: Изд-во стандартов, 1964. – 6 с.

147. ГОСТ 53600-2009 Семена масличные, жмыхи и шроты. Определение влаги, жира, протеина и клетчатки методом спектроскопии в ближней инфракрасной области. Стандартинформ, 2009. – 12 с.

148. Williams, P. and Norris, K. Near-Infrared Technology in the Agricultural and Food Industries // American Association of Cereal Chemists. – 2001.

149. Montgomery, D.C. Design and Analysis of Experiments. - John Wiley & Sons, Inc.- 2013.

150. IBM Corp. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp. – 2021.

151. Holland, J. B., Nyquist, W. E., & Cervantes-Martínez, C. T. Estimating and interpreting heritability for plant breeding: An update // Plant Breeding Reviews. – 2003.- Vol. 22. DOI: 10.1002/9780470650202.ch2

152. Lu, S., Dong, L., Fang, C. et al. Stepwise selection on homeologous PRR genes controlling flowering and maturity during soybean domestication // Nat Genet. – 2020.- Vol. 52.-P. 428–436. https://doi.org/10.1038/s41588-020-0604-7

153. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform // Bioinformatics. – 2009. – Vol. 25.

154. https://gatk.broadinstitute.org/hc/en-us. 10.10.2022

155. Stephen Turner. Quality Control Procedures for Genome Wide Association Studies. – 2011.

156. David W. Fardo. On Quality Control Measures in Genome-Wide Association Studies: A Test to Assess the Genotyping Quality of Individual Probands in Family-Based Association Studies and an Application to the HapMap. – 2009.

157. Saitou, N., & Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees // Molecular Biology and Evolution. -1987.- Vol. 4.- P. 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

158. Weir et al. Genetic similarity and relationships in association studies // Nature Reviews Genetics. – 2006.- Vol. 7, №11. -P. 812–821. DOI: 10.1038/nrg1961

159. https://itol.embl.de/. 10.02.2023

160. Pritchard, J. K., Stephens, M., & Donnelly, P. Inference of population structure using multilocus genotype data // Genetics. – 2000.- Vol. 155, №2. P. 945–959. https://doi.org/10.1093/genetics/155.2.945

161. Evanno, G., Regnaut, S., & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study // Molecular Ecology. - 2005. – Vol. 14, №8. – P. 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

162. Earl, D. A., & vonHoldt, B. M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method // Conservation Genetics Resources. – 2012. – Vol. 4, №2. - P. 359–361. https://doi.org/10.1007/s12686-011-9548-7

163. Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. TASSEL: Software for association mapping of complex traits in diverse samples // Bioinformatics. – 2007. – Vol.23, №19. P.2633–2635. https://doi.org/10.1093/bioinformatics/btm308

164. Hyten, D. L., et al. Examination of Genetic Diversity and Linkage Disequilibrium in Soybean // Crop Science. – 2007.- Vol. 47.

165. Vos PG, Paulo MJ, Voorrips RE, Visser RGF, van Eck HJ, van Eeuwijk FA. Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato // Theor Appl Genet. – 2017.

166. RStudio Team. RStudio: Integrated development environment for R (Version 2022.07.1). - 2022. https://posit.co/

167. Patil, G., Do, T., Vuong, T. et al. Genomic-assisted haplotype analysis and the development of high-throughput SNP markers for salinity tolerance in soybean // Sci Rep. - 2016. – Vol. 6. https://doi.org/10.1038/srep19199

168. Bhat JA, Adeboye KA, Ganie SA, Barmukh R, Hu D, Varshney RK, Yu D. Genome-wide association study, haplotype analysis, and genomic prediction reveal the genetic basis of yield-related traits in soybean (Glycine max L.) // Front Genet. – 2022. -Vol. 17, №13. doi: 10.3389/fgene.2022.953833.

169. B.N. Doszhanova, S.V. Didorenko, A.K. Zatybekov, Y.K. Turuspekov, S.I. Abugalieva. Analysis of soybean world collection in conditions of south-eastern Kazakhstan // International Journal of Biology and Chemistry. - 2019.-Vol. 12.

170. Досжанова Б.Н., Затыбеков А.К., Дидоренко С.В., Туруспеков Е.К. Изучение коллекции сои казахстанской и зарубежной селекции в условиях юго-востока Казахстана // Тезис в конференции «Аль фараби әлемі». – 2019.

171. Doszhanova, B.; Zatybekov, A.; Didorenko, S.; Fang, C.; Abugalieva, S.; Turuspekov, Y. Genome-Wide Association Study of Seed Quality and Yield Traits in a Soybean Collection from Southeast Kazakhstan // Agronomy. – 2024. – Vol. 14, № 2746. https://doi.org/10.3390/agronomy14112746

172. Досжанова Б.Н., Дидоренко С.В., Сузуки Т., Туруспеков Е.К. Изучение мировой коллекции сои в условиях юго-востока Казахстана по признаку растрескиваемости бобов // Тезис в конференции «Материалы международной научной конференции студентов и молодых ученых «Фараби әлемі» Алматы, Казахстан. - 2020

173. Doszhanova B.N., Zatybekov A.K., Didorenko S.V., Suzuki T., Yamashita Y., Turuspekov Y. Identification of quantitative trait loci of pod dehiscence in a collection of soybean grown in the southeast of Kazakhstan // Vavilovskii Zhurnal Genetiki i Selektsii. – 2024. – Vol. 28, №5. -P.515-522. DOI 10.18699/vjgb-24-58

174. Kezar, S., Ballagh, A., Kankarla, V., Sharma, S., Sharry, R., & Lofton, J. Response of Soybean Yield and Certain Growth Parameters to Simulated Reproductive Structure Removal // Agronomy. – 2023. – Vol. 13, №3. https://doi.org/10.3390/agronomy13030927

175. Rupesh Tayade et al. Molecular, genetic, and genomic basis of seed size and yield characteristics in soybean // Front. Plant Sci. - 2023. – Vol. 14. https://doi.org/10.3389/fpls.2023.1195210

176. Zhao, X., Han, Y., Li, Y., Liu, D., Sun, M., Zhao, Y., Teng, W., & Li, W. Genome-wide association studies for dynamic plant height and number of nodes on the main stem in summer sowing soybeans // Frontiers in Plant Science. - 2018. – Vol. 9, №1184. https://doi.org/10.3389/fpls.2018.01184

177. Kang, X., Cai, J., Chen, Y., Yan, Y., Songtao, Y., He, R., Wang, D., & Zhu, Y. Pod-shattering characteristics differences between two groups of soybeans are associated with specific changes in gene expression. Functional & Integrative Genomics. – 2020. – Vol. 20, №2. – P. 201–210. https://doi.org/10.1007/S10142-019-00702-2

178. Liu, J.-Y., Liu, J.-Y., Zhang, Y.-W., Han, X., Zuo, J.-F., Zhang, Z., Shang, H., Song, Q., & Zhang, Y.-M. An evolutionary population structure model reveals pleiotropic effects of *GmPDAT* for traits related to seed size and oil content in soybean. Journal of Experimental Botany. – 2020. – Vol.71, №22. – P. 6988–7002. https://doi.org/10.1093/JXB/ERAA426

179. Wang, S., Liu, S.-L., Wang, J., Yokosho, K., Zhou, B., Yu, Y. C., Liu, Z., Frommer, W. B., Ma, J. F., Chen, L. Q., Guan, Y., Shou, H., & Tian, Z. Simultaneous changes in seed size, oil content and protein content driven by selection of *SWEET* homologues during soybean domestication // National Science Review. - 2020. – Vol. 7, №11. – P. 1776–1786. https://doi.org/10.1093/NSR/NWAA110

180. Bingjie Tu, Qiuying Zhang, Xiaobing Liu, Shaopeng Yu, Nan Xu, Jia Liu, Changkai Liu. Agronomic and pod traits in relation to pod shattering in cultivated soybeans // Czech Journal of Genetics and Plant Breeding. – 2025. – Vol. 61, №2. – P. 67–76. DOI:10.17221/119/2024-CJGPB

181. Krzyzanowski, F. C., França-Neto, J. de B, Henning, F. A. Importance of the lignin content in the pod wall and seed coat on soybean seed physiological and health performances // Journal of Seed Science. - 2023. https://doi.org/10.1590/2317-1545v45268562

182. Государственный реестр селекционных достижений, рекомендуемых к использованию в Республике Казахстан. https://sortcom.kz. 10.04.2025

183. Jianchang Du, Zhixi Tian, Yi Sui, Meixia Zhao, Qijian Song, Steven B Cannon, Perry Cregan, Jianxin Ma. Pericentromeric Effects Shape the Patterns of Divergence, Retention, and Expression of Duplicated Genes in the Paleopolyploid Soybean. – 2012.

184. Kim, D.-G., Lyu, J. I., Kim, J. M., Seo, J. S., Choi, H.-I., Jo, Y. D., Kim, S. H., Eom, S. H., Ahn, J.-W., Bae, C.-H., & Kwon, S.-J. Identification of Loci Governing Agronomic Traits and Mutation Hotspots via a GBS-Based Genome-Wide Association Study in a Soybean Mutant Diversity Pool // International Journal of Molecular Sciences. – 2022. – Vol. *23, №*18. https://doi.org/10.3390/ijms231810441

185. Mather KA, Caicedo AL, Polato NR, Olsen KM, McCouch S, Purugganan MD. The extent of linkage disequilibrium in rice (Oryza sativa L.) // Genetics. – 2007. – Vol.177, №4. doi: 10.1534/genetics.107.079616.

186. Flint-Garcia SA, Thornsberry JM, Buckler ES. Structure of linkage disequilibrium in plants // Annu Rev Plant Biol. – 2003. – Vol.54. doi: 10.1146/annurev.arplant.54.031902.134907.

187. Hyten DL, Choi IY, Song Q, Shoemaker RC, Nelson RL, Costa JM, Specht JE, Cregan PB. Highly variable patterns of linkage disequilibrium in multiple soybean populations // Genetics. – 2007. -Vol. 175, №4. doi: 10.1534/genetics.106.069740.

188. Noble TJ, Tao Y, Mace ES, Williams B, Jordan DR, Douglas CA, Mundree SG. Characterization of Linkage Disequilibrium and Population Structure in a Mungbean Diversity Panel // Front Plant Sci.- 2018. – Vol.12, №8. doi: 10.3389/fpls.2017.02102.

189. Caldwell et al. Extreme Population-Dependent Linkage Disequilibrium Detected in an Inbreeding Plant Species, *Hordeum vulgare*. – 2006.

190. Zatybekov, A., Turuspekov, Y., Doszhanova, B., Didorenko, S. & Abugalieva, S. Effect of Population Size on Genome-Wide Association Study of Agronomic Traits in Soybean // Proceedings of the Latvian Academy of Sciences. – *2020*. – Vol.74, *№*4. – P. 244-251. https://doi.org/10.2478/prolas-2020-0039

191. Zatybekov A.K., Turuspekov Y.T., Doszhanova B.N., Abugalieva S.I. A study of the genetic diversity in the world soybean collection using microsatellite markers associated with fungal disease resistance // Proceedings on applied botany, genetics and breeding. – 2020. – Vol.181, №3. P.81-90. https://doi.org/10.30901/2227-8834-2020-3-81-90

192. Ataul Haleem, Selina Klees, A. Schmitt, Mehmet Gültas. Deciphering Pleiotropic Signatures of Regulatory SNPs in Zea mays L. Using Multi-Omics Data and Machine Learning Algorithms // International Journal of Molecular Sciences. - 2022.

193. Yongle Li et al. Investigating Drought Tolerance in Chickpea Using Genome-Wide Association Mapping and Genomic Selection Based on Whole-Genome Resequencing Data. - 2018

194. Muchero, W., Ehlers, J. D., Close, T. J., & Roberts, P. A. Genetic architecture of delayed senescence, biomass, and grain yield under drought stress in cowpea// PLoS ONE. – 2013. – Vol. 8, №7. https://doi.org/10.1371/journal.pone.0070041

195. Durand et al. Flowering Time in Maize: Linkage and Epistasis at a Major Effect Locus. - 2012

196. Досжанова Б.Н., Затыбеков А.К., Туруспеков Е.К. Identification of genetic variations associated with soybean seed size traits // Тезис в конференции «Аль фараби әлемі». – 2025

197. Zhongyan Wei, Tian Pan, Yuyang Zhao, Bohong Su, Yulong Ren, Lijuan Qiu. The small GTPase Rab5a and its guanine nucleotide exchange factors are involved in post-Golgi trafficking of storage proteins in developing soybean cotyledon // Journal of Experimental Botany. – 2020. – Vol. 71, № 3. – P. 808–822. https://doi.org/10.1093/jxb/erz454

198. Liang, Q., Chen, L., Yang, X. *et al.* Natural variation of *Dt2* determines branching in soybean // Nat Commun. – *2022*. – Vol. 13, №6429. https://doi.org/10.1038/s41467-022-34153-4

199. Karikari, B., Chen, S., Xiao, Y., Chang, F., Zhou, Y., Kong, J., & Wang, Y. Utilization of interspecific high-density genetic map of RIL population for the QTL detection and candidate gene mining for 100-seed weight in soybean // Frontiers in Plant Science. 2019. – Vol.10**,** №1001. https://doi.org/10.3389/fpls.2019.01001

200. Goettel, W., Zhang, H., Li, Y. *et al*. *POWR1* is a domestication gene pleiotropically regulating seed quality and yield in soybean // Nat Commun. – 2022. – Vol.13, №3051. https://doi.org/10.1038/s41467-022-30314-7

201. Alencar Xavier, Diego Jarquin, Reka Howard, Vishnu Ramasubramanian, James E Specht, George L Graef, William D Beavis, Brian W Diers, Qijian Song, Perry B Cregan, Randall Nelson, Rouf Mian, J Grover Shannon, Leah McHale, Dechun Wang, William Schapaugh, Aaron J Lorenz, Shizhong Xu, William M Muir, Katy M Rainey. Genome-Wide Analysis of Grain Yield Stability and Environmental Interactions in a Multiparental Soybean Population // G3 Genes Genomes Genetics. -2018. – Vol.8, №2. – P. 519–529. https://doi.org/10.1534/g3.117.300300

202. Jin, H., Yang, X., Zhao, H., Song, X., Tsvetkov, Y. D., Wu, Y., Gao, Q., Zhang, R., & Zhang, J. Genetic analysis of protein content and oil content in soybean by genome-wide association study // Frontiers in Plant Science**.** – 2023. – Vol. ***14***. https://doi.org/10.3389/fpls.2023.1182771

203. Zhang, H., Hao, D., Sitoe, H. M., Xu, C., Hu, Z.-B., Zhang, G., & Yu, D. Genetic dissection of the relationship between plant architecture and yield component traits in soybean (*Glycine max*) by association analysis across multiple environments // Plant Breeding. – 2015. – Vol. 134, №5. – P. 564–572. https://doi.org/10.1111/pbr.12305

204. Li, Y., Zhang, X., Zhao, J., Zhang, Y., Wang, Y., Wang, Y., ... & Liu, B. Genome-wide association study and genomic selection for yield and related traits in soybean // PLOS ONE. – 2021. – Vol.16, №8. https://doi.org/10.1371/journal.pone.0255761

205. Yoosefzadeh-Najafabadi, M., Torabi, S., Tulpan, D., Rajcan, I., & Eskandari, M. Genome-wide association studies of soybean yield-related hyperspectral reflectance bands using machine learning-mediated data integration methods // Frontiers in Plant Science. – 2021. – Vol. 12**.** https://doi.org/10.3389/fpls.2021.777028

206. Palmgren MG, Clemens S, Williams LE, Krämer U, Borg S, Schjørring JK, Sanders D. Zinc biofortification of cereals: problems and solutions// Trends Plant Sci. -2008. – Vol.13, №9. doi: 10.1016/j.tplants.2008.06.005.

207. Nadia Bouain, Zaigham Shahzad, Aida Rouached, Ghazanfar Abbas Khan, Pierre Berthomieu, Chedly Abdelly, Yves Poirier, Hatem Rouached, Phosphate and zinc transport and signalling in plants: toward a better understanding of their homeostasis interaction// Journal of Experimental Botany. – 2014. - Vol. 65, № 20.- P. 5725–5741. https://doi.org/10.1093/jxb/eru314

208. Farooq MA, Zhang X, Zafar MM, Ma W, Zhao J. Roles of Reactive Oxygen Species and Mitochondria in Seed Germination // Front Plant Sci. – 2021. – Vol.9, №12. doi: 10.3389/fpls.2021.781734.

209. Whiting, R.M., Torabi, S., Lukens, L. *et al*. Genomic regions associated with important seed quality traits in food-grade soybeans //BMC Plant Biol. -*2020*.- *Vol*. 20, № 485. https://doi.org/10.1186/s12870-020-02681-0

210. Priyanatha, C., Torkamaneh, D., & Rajcan, I. Genome-wide association study of soybean germplasm derived from Canadian × Chinese crosses to mine for novel alleles to improve seed yield and seed quality traits // Frontiers in Plant Science. – 2022. – Vol. 13. https://doi.org/10.3389/fpls.2022.866300

211. Bandillo, N., Jarquin, D., Song, Q., Cregan, P., Nelson, R. L., Specht, J., & Lorenz, A. A population structure and genome-wide association analysis on the USDA soybean germplasm collection // The Plant Genome. – 2015. – Vol. 8**,** ***№***3. https://doi.org/10.3835/plantgenome2015.04.0024

212. Zhang, K., Liu, S., Li, W., Liu, S., Li, X., Fang, Y., Zhang, J., Wang, Y., Xu, S., Zhang, J., Song, J., Qi, Z., Tian, X., Tian, Z., Li, W.-X., & Ning, H. Identification of QTNs controlling seed protein content in soybean using multi-locus genome-wide association studies // Frontiers in Plant Science. – 2018. – Vol. 9, № 1690. https://doi.org/10.3389/fpls.2018.01690

213. Qing-Tian Li, Xiang Lu, Qing-Xin Song, Hao-Wei Chen, Wei Wei, Jian-Jun Tao, Xiao-Hua Bian, Ming Shen, Biao Ma, Wan-Ke Zhang, Ying-Dong Bi, Wei Li, Yong-Cai Lai, Sin-Man Lam, Guang-Hou Shui, Shou-Yi Chen, Jin-Song Zhang, Selection for a Zinc-Finger Protein Contributes to Seed Oil Increase during Soybean Domestication // Plant Physiology**.** – 2017. - Volume 173, Issue 4. – P. 2208–2224. https://doi.org/10.1104/pp.16.01610

214. Liu, YF., Li, QT., Lu, X. *et al.* Soybean *GmMYB73* promotes lipid accumulation in transgenic plants // *BMC Plant Biol*. - 2014. – Vol.14, №73. https://doi.org/10.1186/1471-2229-14-73.

215. Qin, J., Wang, F., Zhao, Q., Shi, A., Zhao, T., Song, Q., Ravelombola, W., An, H., Yan, L., Yang, C., & Zhang, M. Identification of candidate genes and genomic selection for seed protein in soybean breeding pipeline // Frontiers in Plant Science. – 2022. – Vol. 13. https://doi.org/10.3389/fpls.2022.882732

216. Zhu, X., Leiser, W. L., Hahn, V., & Würschum, T. Identification of seed protein and oil related QTL in 944 RILs from a diallel of early-maturing European soybean // The Crop Journal. – 2021. – Vol. 9, №2. – P. 238–247. https://doi.org/10.1016/j.cj.2020.06.006

217. Zhao B, Dai A, Wei H, Yang S, Wang B, Jiang N, Feng X. Arabidopsis KLU homologue GmCYP78A72 regulates seed size in soybean // Plant Mol Biol. – 2016.- Vol.90. – P.33-47. doi: 10.1007/s11103-015-0392-0.

218. Pathan, S. M., Vuong, T., Clark, K., Lee, J.-D., Shannon, J. G., Roberts, C. A., Ellersieck, M. R., Burton, J. W., Cregan, P. B., Hyten, D. L., Nguyen, H. T., & Sleper, D. A. Genetic mapping and confirmation of quantitative trait loci for seed protein and oil contents and seed weight in soybean // *Crop Science. – 2013. – Vol.53, №*3. – P. 765–774. https://doi.org/10.2135/cropsci2012.03.0153

219. Zhang J, Wang X, Lu Y, Bhusal SJ, Song Q, Cregan PB, Yen Y, Brown M, Jiang GL. Genome-wide Scan for Seed Composition Provides Insights into Soybean Quality Improvement and the Impacts of Domestication and Breeding // Mol Plant. – 2018. – Vol.5, №11. P.460-472. doi: 10.1016/j.molp.2017.12.016.

220. Wang, H.-W., Zhang, B., Hao, Y.-J., Huang, J., Tian, A.-G., Liao, Y., Zhang, J.-S., & Chen, S.-Y. The soybean Dof-type transcription factor genes, *GmDof4* and *GmDof11*, enhance lipid content in the seeds of transgenic *Arabidopsis* plants // The Plant Journal. – *2007*.- *Vol. 52, №*4. -P. 716–729. https://doi.org/10.1111/j.1365-313X.2007.03268.x

221. Shaopei Gao, Chengcai Chu, Gibberellin Metabolism and Signaling: Targets for Improving Agronomic Performance of Crops // Plant and Cell Physiology. – 2020**. –** Vol. 61, №11. - P.1902–1911. https://doi.org/10.1093/pcp/pcaa104

222. Huang Y, Ji Z, Zhang S, Li S. Function of hormone signaling in regulating nitrogen-use efficiency in plants // J Plant Physiol. – 2024. – Vol.294. doi: 10.1016/j.jplph.2024.154191.

223. Zhang H, Quintana J, Ütkür K, Adrian L, Hawer H, Mayer K, Gong X, Castanedo L, Schulten A, Janina N, Peters M, Wirtz M, Brinkmann U, Schaffrath R, Krämer U. Translational fidelity and growth of Arabidopsis require stress-sensitive diphthamide biosynthesis // Nat Commun. – 2022. -Vol.11, №13. doi: 10.1038/s41467-022-31712-7.

Приложение 1

ID	Name	Country	origin	Protein, %	Oil, %	КПУ, count	КСР, count	МТС, g	МСР, g
SK001	Yaselda	Belarus	Восточная Европа	43,13	21,55	13,58	38,18	123,84	6,68
SK002	Pripyat	Belarus	Восточная Европа	43,93	20,73	13,02	32,14	174,13	7,77
SK003	OAC Vision	Canada	Северная Америка	45,49	19,19	14,03	28,51	146,83	7,47
SK004	Emerson	Canada	Северная Америка	39,28	23,78	16,72	44,24	176,96	9,86
SK005	P-73-3	Canada	Северная Америка	42,51	21,09	14,90	30,11	150,45	9,08
SK006	Supra	Canada	Северная Америка	38,45	22,62	20,69	46,29	186,24	15,14
SK007	Maple Ridge	Canada	Северная Америка	44,31	20,26	11,00	26,28	142,62	7,33
SK008	Maple Glen	Canada	Северная Америка	41,04	22,55	17,06	41,86	182,59	12,80
SK009	Maple Amber	Canada	Северная Америка	44,00	21,10	12,14	30,22	142,98	8,30
SK010	Gaillard	Canada	Северная Америка	44,11	21,60	15,23	32,71	167,81	7,98
SK011	KG 20	Canada	Северная Америка	42,04	21,48	15,34	31,26	159,96	9,47
SK012	Accord	Canada	Северная Америка	43,60	22,50	16,44	35,14	161,23	8,85
SK013	Kharbin	China	Восточная Азия	41,02	20,98	14,79	34,81	150,94	11,78
SK014	Kheikhek14	China	Восточная Азия	45,86	19,39	9,04	23,12	136,99	6,42
SK015	LMF	Poland	Восточная Европа	44,33	20,40	11,42	24,92	128,90	6,49
SK016	Chabem Wekoju	Poland	Восточная Европа	44,94	20,16	11,84	26,46	149,46	7,30
SK017	Arctic	Poland	Восточная Европа	43,65	19,49	9,12	21,02	122,64	5,46
SK018	Nawiko	Poland	Восточная Европа	45,64	19,81	10,85	27,82	130,71	5,37
SK019	Warsawska	Poland	Восточная Европа	45,55	17,05	15,22	31,22	140,72	8,33
SK020	Kasatka	Russia	Восточная Европа	45,67	18,01	8,76	17,04	112,59	4,65
SK021	Severnaya5	Russia	Восточная Европа	43,72	20,88	9,64	26,04	116,25	5,01
SK022	Smena	Russia	Восточная Европа	47,24	18,43	11,33	25,16	135,85	5,71
SK023	Rassvet	Russia	Восточная Европа	43,98	20,60	15,41	42,33	124,78	8,03
SK024	Amurskaya 401	Russia	Восточная Европа	46,29	19,34	10,64	20,60	149,09	5,64
SK025	Soer3491	Russia	Восточная Европа	43,09	19,70	13,74	34,58	136,83	7,17
SK026	Soer-3	Russia	Восточная Европа	44,82	20,05	13,85	32,64	149,52	6,87

SK027	Nadejda	Russia	Восточная Европа	39,80	22,69	11,12	30,46	165,47	10,08
SK028	Lidiya	Russia	Восточная Европа	43,64	21,64	19,64	40,94	134,98	10,82
SK029	VNIIS 1	Russia	Восточная Европа	44,46	19,94	17,72	38,72	148,86	10,96
SK030	Luchezarnaya	Russia	Восточная Европа	42,97	20,91	16,22	31,87	142,09	10,33
SK031	PEP27	Russia	Восточная Европа	44,76	19,16	9,06	20,30	137,82	4,74
SK032	Sibniik 315	Russia	Восточная Европа	44,63	19,76	10,21	25,51	137,03	6,36
SK033	VNIIS 2	Russia	Восточная Европа	42,21	21,56	22,33	48,69	174,51	12,21
SK034	Soer4	Russia	Восточная Европа	46,52	19,30	13,58	32,58	167,99	13,01
SK035	Bara	Russia	Восточная Европа	43,95	21,35	14,94	34,40	155,97	8,86
SK036	Zolotistaya	Russia	Восточная Европа	45,90	19,30	10,95	27,02	134,77	8,07
SK037	Zlata	Russia	Восточная Европа	47,81	19,70	13,58	31,24	134,75	5,92
SK038	Soer 5	Russia	Восточная Европа	44,07	19,67	10,98	28,70	136,68	7,35
SK039	Okskaya	Russia	Восточная Европа	46,79	18,14	12,26	33,10	115,51	5,27
SK040	Svetlaya	Russia	Восточная Европа	46,85	18,79	9,84	25,48	110,08	5,04
SK041	Maleta	Russia	Восточная Европа	41,75	19,69	9,26	20,78	122,63	6,28
SK042	Vega	Russia	Восточная Европа	42,02	20,64	17,00	35,50	168,10	11,47
SK043	Luch nadezhdy	Russia	Восточная Европа	46,00	19,30	8,32	17,60	132,93	4,49
SK044	Zakat	Russia	Восточная Европа	46,80	19,06	11,94	25,96	151,89	6,26
SK045	Svapa	Russia	Восточная Европа	45,10	21,50	9,38	21,32	106,98	5,75
SK046	Lancetnaya	Russia	Восточная Европа	43,28	21,99	12,01	28,86	143,80	6,70
SK047	Zernica	Russia	Восточная Европа	46,10	19,81	15,66	35,92	178,53	11,67
SK048	Niva 70	Russia	Восточная Европа	41,18	21,79	17,02	38,49	153,58	11,58
SK049	Sibiryachka	Russia	Восточная Европа	44,50	19,30	10,13	21,28	138,04	4,59
SK050	Eldorado	Russia	Восточная Европа	46,22	18,78	9,67	23,31	133,71	7,02
SK051	Garmoniya	Russia	Восточная Европа	42,85	20,68	14,54	32,98	155,78	9,68
SK052	Romantika	Russia	Восточная Европа	41,84	20,29	18,08	43,78	159,96	11,51
SK053	Belgorodskaya 6	Russia	Восточная Европа	43,87	21,39	21,06	47,53	150,17	13,37
SK054	Veidelevskaya 17	Russia	Восточная Европа	41,75	19,54	14,50	35,07	161,84	7,41
SK055	Yantarnaya	Russia	Восточная Европа	43,70	20,50	19,28	38,14	158,33	10,16
SK056	Altom	Russia	Восточная Европа	42,62	20,86	14,48	36,14	150,15	9,06

SK057	PEP26	Russia	Восточная Европа	41,40	18,90	8,32	19,86	106,12	3,82
SK058	Belor	Russia	Восточная Европа	43,26	22,00	18,06	37,56	166,54	10,31
SK059	Soer345	Russia	Восточная Европа	45,22	19,38	13,96	38,04	159,24	8,74
SK060	Gribskaya	Russia	Восточная Европа	45,83	18,78	15,62	38,80	148,50	7,77
SK061	Krasivaya mechta	Russia	Восточная Европа	45,73	20,27	11,04	29,02	135,42	7,02
SK062	Carola	USA	Северная Америка	43,29	18,88	16,36	36,16	137,15	8,22
SK063	Daksoy	USA	Северная Америка	42,00	21,65	15,13	34,49	173,79	10,42
SK064	Dawson	USA	Северная Америка	39,72	22,42	21,66	56,56	208,55	17,80
SK065	USKHI 6	Ukraine	Восточная Европа	46,12	18,97	13,65	30,57	168,93	8,00
SK066	Prikarpatskaya 81	Ukraine	Восточная Европа	43,24	20,50	13,12	32,77	149,21	9,72
SK067	Chernovickaya7	Ukraine	Восточная Европа	43,30	20,12	17,00	47,00	165,96	9,29
SK068	Terek	Ukraine	Восточная Европа	41,56	22,22	22,26	49,88	172,08	19,21
SK069	Ustya	Ukraine	Восточная Европа	45,34	20,91	14,19	38,26	168,58	9,80
SK070	Khorol	Ukraine	Восточная Европа	45,04	20,10	16,22	36,40	173,48	10,41
SK071	Yug30	Ukraine	Восточная Европа	40,54	22,36	15,43	35,12	152,25	10,78
SK072	Estophita	Ukraine	Восточная Европа	41,19	20,60	14,95	42,02	165,75	11,67
SK073	Podyaka	Ukraine	Восточная Европа	45,15	18,39	24,34	61,88	162,18	9,85
SK074	Victorina	Ukraine	Восточная Европа	43,91	19,37	11,54	33,58	165,45	8,56
SK075	Annushka	Ukraine	Восточная Европа	40,42	22,71	15,14	30,94	130,87	6,92
SK076	Amour	France	Западная Европа	39,71	22,11	23,60	52,47	170,38	17,56
SK077	Kalmit	France	Западная Европа	42,81	21,34	22,56	47,14	157,75	11,01
SK078	Sepia	France	Западная Европа	40,72	23,66	23,00	61,12	191,56	15,38
SK079	Amphor	France	Западная Европа	41,73	21,33	13,54	32,31	171,07	10,31
SK080	Toury	Czech Republic	Восточная Европа	42,45	21,69	18,86	40,74	147,69	12,05
SK081	Rana	Czech Republic	Восточная Европа	48,10	18,38	14,26	34,84	150,10	8,26
SK082	Turijskaja	Czech Republic	Восточная Европа	46,75	17,97	20,73	43,24	152,30	11,53
SK083	Fiskeby5	Sweden	Западная Европа	42,71	19,23	10,84	26,47	118,46	7,94
SK084	Fiskeby3	Sweden	Западная Европа	44,44	18,49	15,74	34,84	150,12	10,36

SK085	308/1	Kazakhstan	Казахстан	45,09	19,88	16,06	33,86	116,57	7,61
SK086	422/1	Kazakhstan	Казахстан	46,37	19,61	19,77	44,59	162,20	10,35
SK087	186/1	Kazakhstan	Казахстан	48,36	18,08	14,78	33,72	139,21	8,89
SK088	173/1	Kazakhstan	Казахстан	47,68	18,46	16,86	40,48	131,73	9,44
SK089	126/1	Kazakhstan	Казахстан	48,13	18,37	14,98	32,24	137,38	8,25
SK090	209/1	Kazakhstan	Казахстан	46,88	19,69	16,58	38,13	155,31	10,85
SK091	261/1	Kazakhstan	Казахстан	47,61	19,42	13,97	32,18	144,05	8,32
SK092	350/1	Kazakhstan	Казахстан	46,01	18,83	17,50	46,22	168,23	19,37
SK093	362/2	Kazakhstan	Казахстан	42,94	20,69	26,62	65,86	166,52	17,10
SK094	371/2	Kazakhstan	Казахстан	41,02	21,09	20,78	60,32	173,99	21,23
SK095	407/2	Kazakhstan	Казахстан	45,20	20,70	14,41	35,21	144,90	12,66
SK096	404/2	Kazakhstan	Казахстан	43,30	21,14	19,62	46,77	168,15	14,41
SK097	Roza	Kazakhstan	Казахстан	40,01	21,17	18,86	47,78	159,46	16,62
SK098	Misula	Kazakhstan	Казахстан	42,01	21,27	23,26	59,92	183,97	16,58
SK099	Almaty	Kazakhstan	Казахстан	40,63	21,10	18,21	52,77	199,94	15,02
SK100	Zhalpaksai	Kazakhstan	Казахстан	41,14	22,07	19,28	53,86	167,93	14,82
SK101	Agassiz	USA	Северная Америка	41,22	22,18	23,38	67,34	190,85	16,55
SK102	Evans	USA	Северная Америка	41,75	21,79	18,92	51,96	177,97	13,45
SK103	Lambert	USA	Северная Америка	40,74	22,26	20,62	47,10	186,89	13,06
SK104	Mc call	USA	Северная Америка	42,37	21,08	19,20	45,86	178,01	11,47
SK105	Parker	USA	Северная Америка	38,97	21,74	26,86	60,52	147,31	17,10
SK106	Dekabig	USA	Северная Америка	40,05	21,39	20,56	60,50	199,51	17,21
SK107	Jachynes 74 Brond	USA	Северная Америка	39,56	21,70	22,06	59,73	151,84	16,05
SK108	Wilstar 194	USA	Северная Америка	41,62	21,13	18,33	51,08	208,13	16,79
SK109	Lara	Serbia	Западная Европа	39,13	22,10	21,26	49,38	188,24	15,17
SK110	Cobb 266	USA	Северная Америка	46,78	19,48	26,70	63,78	168,56	12,54
SK111	OAO Wallace	Canada	Северная Америка	39,44	22,91	15,93	43,60	155,89	10,23
SK112	GEO	Canada	Северная Америка	40,13	22,28	20,24	51,24	197,83	16,38
SK113	Buster	Canada	Северная Америка	38,08	23,28	20,74	52,20	188,06	17,53
SK114	SL 01 26	Canada	Северная Америка	40,65	22,03	14,68	38,30	179,91	11,48

SK115	Sl 02 25	Canada	Северная Америка	41,39	21,56	23,68	64,34	173,74	19,52
SK116	RCAT Persian	Canada	Северная Америка	40,47	20,19	19,79	46,60	153,40	11,72
SK117	Enterprise	Canada	Северная Америка	41,31	21,71	21,50	55,78	204,68	17,62
SK118	Voevodzhanka	Serbia	Западная Европа	38,35	21,34	21,45	60,89	188,49	23,45
SK119	Ana	Serbia	Западная Европа	39,32	20,82	18,74	50,10	205,51	20,42
SK120	Nikko	Serbia	Западная Европа	41,26	21,71	23,64	60,40	223,07	19,69
SK121	Sava	Serbia	Западная Европа	37,67	23,26	22,12	58,57	203,50	16,91
SK122	Venera	Serbia	Западная Европа	37,73	23,03	24,54	67,70	166,92	15,96
SK123	Zen	Switzerland	Западная Европа	39,09	22,20	22,27	63,05	167,73	17,40
SK124	Protina	France	Западная Европа	45,94	19,71	18,54	53,12	155,89	13,81
SK125	Sponsor	France	Западная Европа	38,07	21,83	21,08	62,00	200,24	18,81
SK126	Isidor	France	Западная Европа	40,47	20,64	22,24	74,84	178,03	23,40
SK127	Shama	France	Западная Европа	42,30	20,32	18,66	54,20	148,06	16,32
SK128	Safrfna	France	Западная Европа	40,01	22,19	20,32	60,40	208,84	22,45
SK129	Santana	France	Западная Европа	38,72	21,53	19,96	50,74	171,28	14,49
SK130	Lada	Russia	Восточная Европа	42,32	21,68	20,98	51,06	150,43	15,22
SK131	Lira	Russia	Восточная Европа	38,91	22,40	21,71	55,69	165,08	18,54
SK132	Bystrica 2	Russia	Восточная Европа	42,19	21,72	26,39	72,93	148,82	17,76
SK133	Vesta	Russia	Восточная Европа	40,36	20,01	27,50	74,06	147,36	26,71
SK134	Renta	Russia	Восточная Европа	40,47	20,79	20,67	52,08	181,80	21,56
SK135	Vilana	Russia	Восточная Европа	41,06	21,82	22,42	58,16	172,06	17,34
SK136	Del’ta	Russia	Восточная Европа	40,73	21,87	19,37	49,13	165,76	17,22
SK137	Runa	Russia	Восточная Европа	38,70	21,06	21,19	48,49	144,42	13,94
SK138	Astra	Russia	Восточная Европа	40,99	23,07	24,60	60,80	179,90	23,01
SK139	Slaviya	Russia	Восточная Европа	38,87	22,18	16,52	46,90	134,68	12,48
SK140	Biser 291	Bulgaria	Западная Европа	40,28	21,13	20,13	46,71	168,73	12,67
SK141	Zispida 641	Belgium	Западная Европа	40,09	19,93	21,04	52,88	176,33	13,29
SK142	Iskra	Kazakhstan	Казахстан	40,60	21,77	24,46	68,42	162,45	22,87
SK143	Perizat	Kazakhstan	Казахстан	39,40	21,45	21,08	57,96	174,19	17,09
SK144	Zhansaya	Kazakhstan	Казахстан	39,66	20,95	15,70	38,52	143,03	13,84

SK145	Vita	Kazakhstan	Казахстан	40,41	21,61	18,36	51,32	169,96	17,79
SK146	Danaya	Kazakhstan	Казахстан	40,33	21,26	26,03	81,07	161,27	18,82
SK147	Sabira	Kazakhstan	Казахстан	37,98	21,66	16,54	44,64	174,31	15,31
SK148	Pamyat'YuGK	Kazakhstan	Казахстан	39,09	22,04	21,88	54,46	220,26	17,36
SK149	Evrika	Kazakhstan	Казахстан	39,36	21,38	22,92	56,66	135,30	17,80
SK150	Lastochka	Kazakhstan	Казахстан	38,93	20,14	19,98	55,90	151,34	14,00
SK151	Sulamit	Kazakhstan	Казахстан	41,64	19,77	20,40	48,96	156,65	16,94
SK152	Kazakhstanskaya 2309	Kazakhstan	Казахстан	39,52	20,88	20,92	43,36	137,64	10,68
SK153	Akku	Kazakhstan	Казахстан	38,91	21,58	24,78	67,94	178,94	18,89
SK154	Radost'	Kazakhstan	Казахстан	41,92	18,81	23,74	57,16	141,45	18,04
SK155	Nadezhda	Kazakhstan	Казахстан	41,41	19,59	20,97	57,25	146,52	14,24
SK156	Xinjiang a don 1	China	Восточная Азия	42,66	21,43	17,53	41,88	181,38	10,90
SK157	Xinjiang heihe 38	China	Восточная Азия	42,12	20,22	15,44	38,02	146,92	8,64
SK158	Xinjiang D09-676	China	Восточная Азия	39,45	20,36	17,99	50,51	164,77	19,28
SK159	Xinjiang D10-51	China	Восточная Азия	40,34	21,22	21,01	69,41	227,87	30,22
SK160	Xinjiang D10-130	China	Восточная Азия	39,54	21,24	18,94	53,16	135,32	15,50
SK161	Xinjiang D10-135	China	Восточная Азия	39,91	20,39	19,84	58,60	183,32	19,80
SK162	Xinjiang D11-252	China	Восточная Азия	38,51	22,24	19,34	57,24	157,70	18,37
SK163	Лыбидь	Ukraine	Восточная Европа	44,64	21,19	17,20	43,64	187,29	17,22
SK164	Cheremosh	Ukraine	Восточная Европа	45,07	20,77	20,02	44,04	182,77	14,28
SK165	Korsak	Ukraine	Восточная Европа	38,92	21,89	21,40	52,18	168,00	12,04
SK166	Tanais	Ukraine	Восточная Европа	42,38	19,71	19,14	43,68	182,38	12,83
SK167	Desna	Ukraine	Восточная Европа	40,62	22,16	23,62	53,76	182,09	19,72
SK168	1054	Korea	Восточная Азия	38,46	20,57	21,54	53,68	216,52	17,11
SK169	1031	Korea	Восточная Азия	42,04	18,37	24,85	63,25	168,98	20,09
SK170	1070	Korea	Восточная Азия	38,53	20,81	20,71	55,51	147,48	17,44
SK171	1034	Korea	Восточная Азия	40,48	20,80	20,43	43,93	126,00	13,03
SK172	1076	Korea	Восточная Азия	38,70	19,73	18,68	48,83	162,63	13,32
SK173	1065	Korea	Восточная Азия	42,73	18,41	25,21	60,87	145,62	16,79

SK174	1017	Korea	Восточная Азия	40,29	20,20	18,96	50,06	187,64	17,07
SK175	1044	Korea	Восточная Азия	39,00	23,40	23,25	48,57	135,75	11,80
SK176	1069	Korea	Восточная Азия	40,95	19,59	21,33	51,63	148,50	14,20
SK177	1028	Korea	Восточная Азия	39,76	20,79	22,90	53,04	200,57	21,77
SK178	1055	Korea	Восточная Азия	37,32	20,48	25,18	51,64	149,48	16,07
SK179	1095	Korea	Восточная Азия	41,64	18,28	22,52	57,94	186,26	18,83
SK180	1026	Korea	Восточная Азия	40,34	21,23	23,72	54,62	122,86	18,97
SK181	1003	Korea	Восточная Азия	40,64	20,23	21,96	50,66	138,32	13,26
SK182	1022	Korea	Восточная Азия	38,83	20,12	28,78	67,28	180,66	16,64
SK183	1033	Korea	Восточная Азия	41,93	18,03	24,33	62,00	158,43	17,50
SK184	Hei lun 48	China	Восточная Азия	43,07	20,58	15,36	39,74	167,32	16,65
SK185	Hua i doe 1	China	Восточная Азия	42,56	21,60	15,01	38,06	153,10	13,56
SK186	Xu nong 35	China	Восточная Азия	42,50	21,00	16,83	39,97	156,36	9,79
SK187	Xu nong 26	China	Восточная Азия	40,62	21,23	17,00	43,28	188,31	14,30
SK188	Bey jiang 91	China	Восточная Азия	42,91	19,57	14,23	34,18	165,06	11,26
SK189	Dong doe 641	China	Восточная Азия	41,66	20,94	16,52	35,92	166,08	11,28
SK190	Jin nong 62	China	Восточная Азия	38,92	21,32	19,75	54,80	197,61	22,09
SK191	Dong doe 1	China	Восточная Азия	37,07	21,13	15,73	54,36	147,05	11,24
SK192	Ken lun 8	China	Восточная Азия	40,28	21,39	15,69	43,92	167,35	14,98
SK193	Jin yuan 55	China	Восточная Азия	42,64	21,70	16,66	38,00	154,59	14,13
SK194	Ken feng 20	China	Восточная Азия	41,47	20,79	14,97	41,21	174,64	14,04
SK195	551	China	Восточная Азия	42,53	21,01	15,77	37,48	171,41	13,97
SK196	Dong doe 29	China	Восточная Азия	39,98	19,44	24,85	68,84	168,63	23,26
SK197	Jing xin 2	China	Восточная Азия	41,80	21,56	11,79	27,51	143,31	11,73
SK198	Dong doe 339	China	Восточная Азия	39,04	19,86	17,25	55,44	187,03	14,16
SK199	Heihe 47	China	Восточная Азия	42,37	19,95	14,66	33,49	148,90	12,02
SK200	Hei feng 50	China	Восточная Азия	38,18	22,56	18,28	54,74	212,70	23,25
SK201	Dong doe 027	China	Восточная Азия	40,84	20,10	16,85	44,43	219,95	18,06
SK202	Mei feng 18	China	Восточная Азия	37,23	20,93	18,21	50,46	169,93	15,38
SK203	Klaxxon	France	Западная Европа	43,39	21,39	17,65	38,11	167,80	10,84

SK204	Major	France	Западная Европа	42,03	21,23	15,26	35,32	167,19	10,98
SK205	Versiya	Ukraine	Восточная Европа	41,61	19,51	19,56	48,86	142,88	13,06
SK206	Dina	Russia	Восточная Европа	45,35	18,76	14,73	31,92	152,40	9,40
SK207	Picket K6298	USA	Северная Америка	47,78	19,10	11,90	26,12	145,60	7,87
SK208	Mriya	Ukraine	Восточная Европа	42,12	20,98	20,32	55,98	148,92	14,78
SK209	K 9961	Russia	Восточная Европа	43,20	20,70	19,09	47,11	150,96	12,58
SK210	Ros’	Belarus	Восточная Европа	43,71	19,63	15,59	36,97	153,13	10,71
SK211	Spritna	Ukraine	Восточная Европа	42,79	19,29	16,41	44,49	158,92	15,39
SK212	Fora	Russia	Восточная Европа	45,32	17,28	18,52	45,22	181,40	15,41
SK213	Mal’vina	Ukraine	Восточная Европа	46,02	18,64	17,94	51,62	196,06	12,18
SK214	OAC Erin	Canada	Северная Америка	41,09	22,15	22,94	50,94	153,90	14,74
SK215	10991	Canada	Северная Америка	40,82	21,73	19,20	43,60	197,67	13,11
SK216	Vatra	Ukraine	Восточная Европа	42,62	21,40	21,29	55,67	209,14	15,23
SK217	Feya	Ukraine	Восточная Европа	42,89	20,60	25,41	60,68	174,45	15,21
SK218	Odesskaya150	Ukraine	Восточная Европа	44,95	19,72	16,36	49,04	187,76	13,48
SK219	9224	Hungary	Восточная Европа	42,32	21,75	21,48	55,73	154,07	15,85
SK220	Darika	Moldova	Восточная Европа	41,10	22,18	20,14	54,50	174,81	15,01
SK221	Osobliva	Ukraine	Восточная Европа	43,10	20,21	18,14	44,23	136,22	12,83
SK222	VNIIS 1	Russia	Восточная Европа	41,25	21,63	22,20	52,58	155,98	12,26
SK223	ISZ 13	Hungary	Восточная Европа	42,82	21,87	18,50	55,92	136,00	11,20
SK224	Rainer 58	Mo ova	Восточная Европа	44,35	19,51	26,42	65,93	165,30	18,37
SK225	Oktyabr’ 70	Russia	Восточная Европа	42,15	20,22	19,99	51,62	146,14	13,39
SK226	Selekta 301	Russia	Восточная Европа	39,00	21,90	24,38	60,64	146,82	18,76
SK227	Majesta	Canada	Северная Америка	44,97	20,56	15,32	42,30	142,38	10,09
SK228	Poltava	Ukraine	Восточная Европа	40,61	21,40	22,45	64,12	170,94	18,01
SK229	Don’ka	Ukraine	Восточная Европа	41,35	21,75	17,00	61,58	200,57	21,31
SK230	Skleya	Ukraine	Восточная Европа	40,80	21,96	22,72	53,70	180,45	18,92
SK231	K2132	China	Восточная Азия	40,80	21,67	21,78	54,90	166,08	10,36
SK232	Kuban’	Ukraine	Восточная Европа	41,45	22,13	15,50	33,96	178,14	11,21
SK233	1674	China	Восточная Азия	45,50	19,49	30,19	62,35	168,35	18,31

SK234	Shelby	USA	Северная Америка	38,98	20,71	20,94	51,91	141,64	12,87
SK235	Maurau	Czech Republic	Восточная Европа	41,50	20,91	25,09	53,73	167,54	15,50
SK236	Bellemondeau	Canada	Северная Америка	40,31	20,90	27,37	65,93	182,06	18,07
SK237	Ussurisskaya 267	Russia	Восточная Европа	41,10	19,27	24,60	56,60	182,76	18,29
SK238	Primorskaya 495	Russia	Восточная Европа	41,37	20,44	18,50	61,89	170,58	19,05
SK239	Crystal	Canada	Северная Америка	40,27	20,78	21,05	47,79	183,73	17,25
SK240	RCAT Bobcat	Canada	Северная Америка	38,10	22,32	21,78	64,38	209,98	20,05
SK241	6877	Philippines	Восточная Азия	40,75	21,15	25,36	63,87	161,19	22,53
SK242	Morsoy	USA	Северная Америка	42,91	21,74	22,02	47,33	150,03	18,11
SK243	391	Kazakhstan	Казахстан	47,05	18,27	13,18	29,30	113,44	5,38
SK244	Colby	Canada	Северная Америка	40,56	21,78	21,01	46,98	207,45	15,50
SK245	Linkoln	USA	Северная Америка	41,93	19,16	29,51	72,62	126,08	14,47
SK246	Viktoriya	Ukraine	Восточная Европа	38,99	20,31	23,02	63,62	137,20	13,16
SK247	Chabarovskaya 4429	Russia	Восточная Европа	40,09	21,04	24,97	64,97	187,78	23,95
SK248	K1889	China	Восточная Азия	38,58	20,58	18,06	42,99	164,20	17,63
SK249	Harrow Manuchu	China	Восточная Азия	41,36	20,65	26,42	64,94	132,74	20,04
SK250	Nhat 11	Vietnam	Восточная Азия	38,34	21,56	28,52	65,24	167,13	19,67
SK251	Peremoga	Ukraine	Восточная Европа	38,34	20,61	24,32	52,29	186,26	16,08
SK252	CH 147020-1	Belarus	Восточная Европа	42,85	20,16	24,60	70,78	136,45	18,65

ПРИЛОЖИЕ 2

ID	ДС2018	ШС2018	ТС2018	ДС2019	ШС2019	ТС2019	ДСmean	ШСmean	ТСmean	РС2019	РС2021	РСmean
SK001	7,43	6,33	5,27	7,28	5,99	5,06	7,35	6,16	5,17	5,00	5,00	5,00
SK002	7,61	6,52	5,32	7,32	6,45	5,54	7,47	6,48	5,43	5,00	0,00	2,50
SK003	8,82	6,71	5,83	8,39	6,27	5,39	8,60	6,49	5,61	0,00	0,00	0,00
SK004	7,51	6,17	5,21	7,57	6,26	5,34	7,54	6,21	5,27	0,00	0,00	0,00
SK005	7,16	6,57	5,56	7,24	6,44	5,44	7,20	6,51	5,50	0,00	0,00	0,00
SK006	7,94	6,73	5,66	7,65	6,39	5,47	7,79	6,56	5,57	0,00	0,00	0,00
SK007	7,53	6,53	5,59	7,44	6,30	5,31	7,48	6,41	5,45	0,00	0,00	0,00

SK008	8,01	6,61	5,42	7,45	6,29	5,18	7,73	6,45	5,30	0,00	0,00	0,00
SK009	8,57	6,45	5,28	7,79	6,32	5,00	8,18	6,38	5,14	5,00	0,00	2,50
SK010	8,25	6,66	5,66	7,99	6,55	5,40	8,12	6,61	5,53	0,00	0,00	0,00
SK011	7,94	6,31	5,04	7,63	6,36	5,11	7,79	6,34	5,07	0,00	5,00	2,50
SK012	8,22	6,86	5,62	8,17	6,27	5,08	8,20	6,57	5,35	5,00	5,00	5,00
SK013	6,99	6,53	5,65	7,62	6,98	5,96	7,30	6,76	5,81	0,00	0,00	0,00
SK014	NA	NA	NA	7,76	6,59	5,97	7,76	6,59	5,97	0,00	0,00	0,00
SK015	8,22	6,00	5,39	8,29	5,81	5,29	8,26	5,91	5,34	0,00	0,00	0,00
SK016	7,94	6,18	5,03	8,23	6,33	5,37	8,09	6,26	5,20	0,00	0,00	0,00
SK017	8,26	5,98	5,55	8,22	6,12	5,84	8,24	6,05	5,69	5,00	5,00	5,00
SK018	7,41	6,31	5,05	7,34	6,04	4,64	7,37	6,18	4,84	5,00	5,00	5,00
SK019	8,46	6,29	5,27	7,46	6,32	5,39	7,96	6,30	5,33	5,00	0,00	2,50
SK020	7,57	5,90	5,06	6,88	5,75	4,74	7,23	5,83	4,90	0,00	0,00	0,00
SK021	NA	NA	NA	7,57	6,59	5,50	7,57	6,59	5,50	0,00	0,00	0,00
SK022	7,21	6,04	5,11	7,07	6,16	5,36	7,14	6,10	5,24	0,00	0,00	0,00
SK023	6,81	5,93	5,13	6,97	6,07	5,22	6,89	6,00	5,18	0,00	0,00	0,00
SK024	8,66	6,40	5,82	8,30	6,26	5,58	8,48	6,33	5,70	0,00	0,00	0,00
SK025	7,51	6,16	5,04	7,15	5,92	4,93	7,33	6,04	4,98	0,00	0,00	0,00
SK026	7,70	6,43	5,64	7,16	6,17	5,21	7,43	6,30	5,42	0,00	0,00	0,00
SK027	8,25	6,67	5,84	7,99	6,14	5,43	8,12	6,40	5,64	0,00	0,00	0,00
SK028	7,39	6,06	5,13	7,33	6,12	5,45	7,36	6,09	5,29	5,00	0,00	2,50
SK029	7,52	6,52	5,89	7,29	6,12	5,22	7,41	6,32	5,55	0,00	0,00	0,00
SK030	7,41	6,08	5,22	7,42	5,99	5,31	7,41	6,04	5,26	0,00	0,00	0,00
SK031	8,78	7,05	5,50	7,87	6,63	5,16	8,33	6,84	5,33	5,00	5,00	5,00
SK032	7,41	6,06	5,15	8,28	5,88	5,16	7,84	5,97	5,15	0,00	0,00	0,00
SK033	7,24	6,08	5,41	6,70	5,77	4,92	6,97	5,92	5,16	0,00	0,00	0,00
SK034	8,30	6,82	5,88	8,44	6,74	5,55	8,37	6,78	5,71	0,00	0,00	0,00
SK035	7,95	6,72	5,66	7,85	6,53	5,55	7,90	6,62	5,61	0,00	0,00	0,00
SK036	8,06	6,03	5,11	8,18	6,09	4,82	8,12	6,06	4,96	0,00	0,00	0,00
SK037	8,02	6,10	5,10	8,03	5,94	5,13	8,03	6,02	5,12	0,00	0,00	0,00

SK038	NA	NA	NA	7,34	6,03	5,19	7,34	6,03	5,19	0,00	0,00	0,00
SK039	7,94	6,06	5,31	7,82	5,72	4,87	7,88	5,89	5,09	5,00	0,00	2,50
SK040	6,82	5,41	4,49	7,34	5,99	5,28	7,08	5,70	4,88	0,00	0,00	0,00
SK041	8,00	6,36	5,53	7,59	6,17	5,33	7,79	6,26	5,43	0,00	0,00	0,00
SK042	7,67	6,83	6,01	8,06	6,51	5,60	7,86	6,67	5,80	0,00	0,00	0,00
SK043	7,30	6,57	5,66	7,19	6,42	5,56	7,25	6,49	5,61	0,00	5,00	2,50
SK044	NA	NA	NA	8,30	5,89	5,05	8,30	5,89	5,05	0,00	0,00	0,00
SK045	NA	NA	NA	8,60	6,46	4,87	8,60	6,46	4,87	0,00	0,00	0,00
SK046	6,99	6,18	5,48	6,59	5,95	5,02	6,79	6,06	5,25	0,00	0,00	0,00
SK047	8,07	6,20	5,36	8,03	6,36	5,45	8,05	6,28	5,41	0,00	0,00	0,00
SK048	8,24	6,41	5,48	8,24	6,33	5,33	8,24	6,37	5,40	0,00	0,00	0,00
SK049	NA	NA	NA	8,28	6,24	5,57	8,28	6,24	5,57	0,00	0,00	0,00
SK050	7,92	6,29	5,38	7,16	5,55	4,70	7,54	5,92	5,04	0,00	0,00	0,00
SK051	7,37	6,39	5,62	7,32	5,80	5,24	7,35	6,09	5,43	0,00	0,00	0,00
SK052	7,58	6,44	5,38	7,95	6,10	5,20	7,76	6,27	5,29	0,00	0,00	0,00
SK053	8,17	6,80	6,03	7,15	6,13	5,29	7,66	6,47	5,66	0,00	0,00	0,00
SK054	7,87	6,63	5,59	7,59	6,22	5,45	7,73	6,42	5,52	0,00	0,00	0,00
SK055	7,73	6,82	5,80	6,98	6,17	5,06	7,35	6,50	5,43	0,00	0,00	0,00
SK056	9,67	6,40	5,15	8,95	5,97	5,08	9,31	6,18	5,12	5,00	5,00	5,00
SK057	8,31	6,10	4,79	7,68	5,91	4,76	8,00	6,00	4,77	5,00	0,00	2,50
SK058	8,44	6,43	5,58	8,02	6,08	5,36	8,23	6,25	5,47	5,00	5,00	5,00
SK059	7,61	6,39	5,54	7,33	6,07	5,13	7,47	6,23	5,33	0,00	0,00	0,00
SK060	7,36	6,29	5,51	7,47	5,88	5,38	7,42	6,08	5,45	0,00	0,00	0,00
SK061	7,12	6,21	5,47	6,94	6,12	5,49	7,03	6,17	5,48	0,00	0,00	0,00
SK062	7,66	6,33	5,10	7,62	6,08	5,19	7,64	6,21	5,15	5,00	5,00	5,00
SK063	7,52	6,48	5,26	8,13	6,70	5,92	7,83	6,59	5,59	0,00	0,00	0,00
SK064	7,34	6,22	5,09	7,66	6,23	5,11	7,50	6,22	5,10	0,00	0,00	0,00
SK065	8,46	6,66	5,56	8,21	6,39	5,38	8,34	6,52	5,47	0,00	5,00	2,50
SK066	8,85	6,72	5,23	8,28	6,08	5,36	8,56	6,40	5,30	0,00	0,00	0,00
SK067	8,22	5,84	5,04	7,52	5,31	4,79	7,87	5,57	4,92	0,00	0,00	0,00

SK068	7,29	6,10	4,95	7,36	6,37	5,18	7,33	6,24	5,06	0,00	0,00	0,00
SK069	7,35	6,52	5,37	7,06	6,19	5,24	7,21	6,36	5,31	0,00	0,00	0,00
SK070	7,91	6,26	4,94	7,96	6,30	5,05	7,94	6,28	4,99	0,00	0,00	0,00
SK071	8,15	6,54	5,53	8,25	6,31	5,42	8,20	6,43	5,48	0,00	0,00	0,00
SK072	7,67	6,34	5,51	7,50	6,19	5,27	7,59	6,26	5,39	0,00	0,00	0,00
SK073	6,87	6,29	5,38	6,63	6,03	4,96	6,75	6,16	5,17	0,00	0,00	0,00
SK074	7,79	6,61	5,46	7,44	6,02	4,87	7,61	6,31	5,16	0,00	0,00	0,00
SK075	7,15	6,25	5,45	7,40	6,40	5,60	7,28	6,32	5,53	0,00	0,00	0,00
SK076	7,67	6,88	5,65	7,32	6,66	5,40	7,49	6,77	5,53	0,00	0,00	0,00
SK077	7,62	6,61	5,82	7,44	5,99	5,29	7,53	6,30	5,55	0,00	5,00	2,50
SK078	7,74	6,44	5,66	6,64	5,78	5,04	7,19	6,11	5,35	5,00	0,00	2,50
SK079	7,47	6,56	5,78	7,51	6,21	5,67	7,49	6,38	5,72	0,00	0,00	0,00
SK080	7,91	6,58	5,67	8,29	6,33	5,36	8,10	6,46	5,51	5,00	5,00	5,00
SK081	7,41	6,04	5,29	7,53	6,04	5,42	7,47	6,04	5,35	5,00	0,00	2,50
SK082	7,71	6,71	5,01	8,05	6,60	5,25	7,88	6,66	5,13	0,00	0,00	0,00
SK083	8,88	6,36	5,77	9,26	6,21	5,56	9,07	6,28	5,66	5,00	0,00	2,50
SK084	8,08	7,04	5,11	8,92	6,91	5,17	8,50	6,98	5,14	0,00	0,00	0,00
SK085	7,24	6,06	4,82	7,28	6,11	5,10	7,26	6,09	4,96	0,00	5,00	2,50
SK086	7,95	6,61	5,28	7,89	6,48	5,28	7,92	6,55	5,28	0,00	0,00	0,00
SK087	8,28	6,40	5,17	7,84	6,15	5,25	8,06	6,28	5,21	0,00	0,00	0,00
SK088	7,83	6,01	5,00	7,83	6,07	5,12	7,83	6,04	5,06	0,00	0,00	0,00
SK089	7,63	6,26	5,07	7,69	6,13	5,03	7,66	6,20	5,05	0,00	0,00	0,00
SK090	8,42	6,44	5,38	8,42	6,12	5,36	8,42	6,28	5,37	0,00	0,00	0,00
SK091	8,22	6,22	4,93	8,31	6,26	5,26	8,27	6,24	5,09	0,00	0,00	0,00
SK092	8,62	6,36	5,59	8,93	6,25	5,29	8,77	6,30	5,44	0,00	0,00	0,00
SK093	7,70	6,35	5,26	7,54	6,56	5,40	7,62	6,46	5,33	0,00	0,00	0,00
SK094	7,98	7,10	5,71	7,67	6,99	5,66	7,83	7,04	5,69	0,00	0,00	0,00
SK095	8,02	6,46	5,30	7,77	6,15	5,12	7,90	6,31	5,21	0,00	0,00	0,00
SK096	7,36	6,32	5,54	7,25	6,15	5,37	7,31	6,24	5,46	0,00	0,00	0,00
SK097	7,50	6,58	5,45	7,29	6,46	5,39	7,39	6,52	5,42	0,00	0,00	0,00

SK098	7,53	6,00	4,78	7,87	6,59	5,32	7,70	6,30	5,05	0,00	0,00	0,00
SK099	8,19	6,35	5,43	8,56	6,34	5,09	8,38	6,35	5,26	0,00	0,00	0,00
SK100	8,16	6,99	5,71	7,93	6,81	5,55	8,04	6,90	5,63	0,00	5,00	2,50
SK101	7,06	6,50	5,84	7,08	6,48	5,60	7,07	6,49	5,72	0,00	0,00	0,00
SK102	7,64	6,40	5,36	7,59	6,69	5,62	7,62	6,55	5,49	0,00	0,00	0,00
SK103	7,91	6,61	5,44	8,00	6,95	5,65	7,96	6,78	5,55	0,00	0,00	0,00
SK104	7,12	6,36	5,14	7,55	6,27	5,27	7,34	6,32	5,20	0,00	0,00	0,00
SK105	7,41	6,28	5,06	8,05	6,55	5,75	7,73	6,41	5,40	0,00	0,00	0,00
SK106	7,15	6,67	5,76	6,88	6,41	5,53	7,01	6,54	5,65	0,00	0,00	0,00
SK107	7,53	6,43	5,35	7,26	6,16	5,12	7,40	6,30	5,23	0,00	0,00	0,00
SK108	7,72	7,16	6,06	8,02	7,22	6,15	7,87	7,19	6,11	0,00	0,00	0,00
SK109	7,32	6,51	4,98	7,74	6,80	5,52	7,53	6,66	5,25	0,00	0,00	0,00
SK110	7,46	6,11	5,37	7,17	5,87	4,74	7,32	5,99	5,05	0,00	0,00	0,00
SK111	7,63	6,61	5,45	7,48	6,56	5,30	7,56	6,58	5,37	0,00	0,00	0,00
SK112	NA	NA	NA	7,66	6,10	5,32	7,66	6,10	5,32	0,00	0,00	0,00
SK113	8,02	6,79	5,41	7,94	6,83	5,40	7,98	6,81	5,40	0,00	0,00	0,00
SK114	8,06	6,70	5,27	7,62	6,61	5,25	7,84	6,65	5,26	0,00	0,00	0,00
SK115	7,13	6,31	5,59	7,35	6,53	5,57	7,24	6,42	5,58	0,00	0,00	0,00
SK116	7,72	6,60	5,57	7,51	6,38	5,38	7,61	6,49	5,48	0,00	0,00	0,00
SK117	7,24	6,69	5,70	7,61	7,00	5,66	7,42	6,85	5,68	0,00	0,00	0,00
SK118	7,39	6,35	5,27	7,18	6,28	5,35	7,28	6,32	5,31	0,00	0,00	0,00
SK119	7,71	6,43	5,40	7,50	6,86	5,64	7,60	6,64	5,52	0,00	0,00	0,00
SK120	7,21	6,28	5,41	7,71	6,96	5,80	7,46	6,62	5,61	0,00	0,00	0,00
SK121	7,50	6,62	5,25	7,35	6,61	5,29	7,43	6,61	5,27	0,00	0,00	0,00
SK122	7,71	6,58	5,39	7,17	6,48	5,24	7,44	6,53	5,32	0,00	0,00	0,00
SK123	6,63	5,81	4,91	6,87	6,10	5,02	6,75	5,96	4,97	0,00	0,00	0,00
SK124	7,63	6,29	4,99	7,48	6,40	5,28	7,55	6,35	5,14	0,00	0,00	0,00
SK125	6,72	6,38	5,53	6,43	6,32	5,55	6,58	6,35	5,54	0,00	0,00	0,00
SK126	7,52	7,04	5,58	7,08	6,76	5,34	7,30	6,90	5,46	0,00	0,00	0,00
SK127	8,45	6,03	4,91	8,31	6,28	4,91	8,38	6,15	4,91	0,00	0,00	0,00

SK128	7,87	7,04	6,06	8,11	7,26	6,32	7,99	7,15	6,19	0,00	0,00	0,00
SK129	7,22	6,53	5,30	7,56	6,85	5,57	7,39	6,69	5,44	0,00	0,00	0,00
SK130	7,92	6,43	5,09	8,02	6,45	4,96	7,97	6,44	5,02	0,00	0,00	0,00
SK131	7,32	6,49	5,52	7,24	6,43	5,32	7,28	6,46	5,42	0,00	0,00	0,00
SK132	6,45	6,07	4,89	6,81	6,19	4,71	6,63	6,13	4,80	0,00	0,00	0,00
SK133	7,77	6,83	5,31	8,10	6,84	5,47	7,93	6,84	5,39	0,00	0,00	0,00
SK134	8,21	7,07	5,93	8,25	6,98	5,89	8,23	7,03	5,91	0,00	0,00	0,00
SK135	7,17	6,36	5,41	7,02	6,15	5,05	7,09	6,25	5,23	0,00	0,00	0,00
SK136	7,52	6,77	5,62	7,58	6,34	5,30	7,55	6,55	5,46	0,00	0,00	0,00
SK137	7,21	6,39	5,19	6,85	6,30	5,34	7,03	6,34	5,27	0,00	0,00	0,00
SK138	7,54	6,65	5,30	7,27	6,26	5,15	7,41	6,45	5,22	0,00	0,00	0,00
SK139	7,93	6,34	5,33	8,19	6,18	5,19	8,06	6,26	5,26	0,00	0,00	0,00
SK140	7,55	6,69	5,65	7,86	6,70	5,55	7,70	6,69	5,60	0,00	0,00	0,00
SK141	6,96	6,36	5,25	7,10	6,46	5,35	7,03	6,41	5,30	0,00	0,00	0,00
SK142	7,24	6,06	5,19	7,45	6,42	5,33	7,35	6,24	5,26	0,00	0,00	0,00
SK143	7,73	6,60	5,36	7,44	6,48	5,32	7,58	6,54	5,34	0,00	0,00	0,00
SK144	6,91	6,53	5,35	7,02	6,65	5,49	6,97	6,59	5,42	0,00	0,00	0,00
SK145	7,28	6,32	5,20	7,42	6,68	5,74	7,35	6,50	5,47	0,00	0,00	0,00
SK146	7,50	6,41	5,37	7,30	6,18	5,28	7,40	6,29	5,32	0,00	0,00	0,00
SK147	7,04	6,41	5,22	7,19	6,61	5,34	7,12	6,51	5,28	0,00	0,00	0,00
SK148	7,21	6,58	5,71	7,40	6,73	5,65	7,30	6,66	5,68	0,00	0,00	0,00
SK149	7,20	6,16	5,37	6,79	6,14	5,25	6,99	6,15	5,31	0,00	0,00	0,00
SK150	7,02	6,44	5,10	7,31	6,57	5,64	7,17	6,50	5,37	0,00	0,00	0,00
SK151	7,96	7,00	5,56	8,07	7,08	5,74	8,02	7,04	5,65	0,00	0,00	0,00
SK152	6,76	6,10	4,95	7,68	6,74	5,41	7,22	6,42	5,18	0,00	0,00	0,00
SK153	6,85	6,29	5,37	7,43	6,95	5,63	7,14	6,62	5,50	0,00	0,00	0,00
SK154	7,38	6,66	5,52	7,43	6,26	4,94	7,41	6,46	5,23	0,00	0,00	0,00
SK155	6,62	6,16	5,07	6,90	6,19	5,19	6,76	6,18	5,13	0,00	0,00	0,00
SK156	7,92	6,96	6,08	7,82	6,56	5,73	7,87	6,76	5,90	0,00	0,00	0,00
SK157	7,67	7,05	6,24	7,48	6,75	5,91	7,57	6,90	6,08	0,00	0,00	0,00

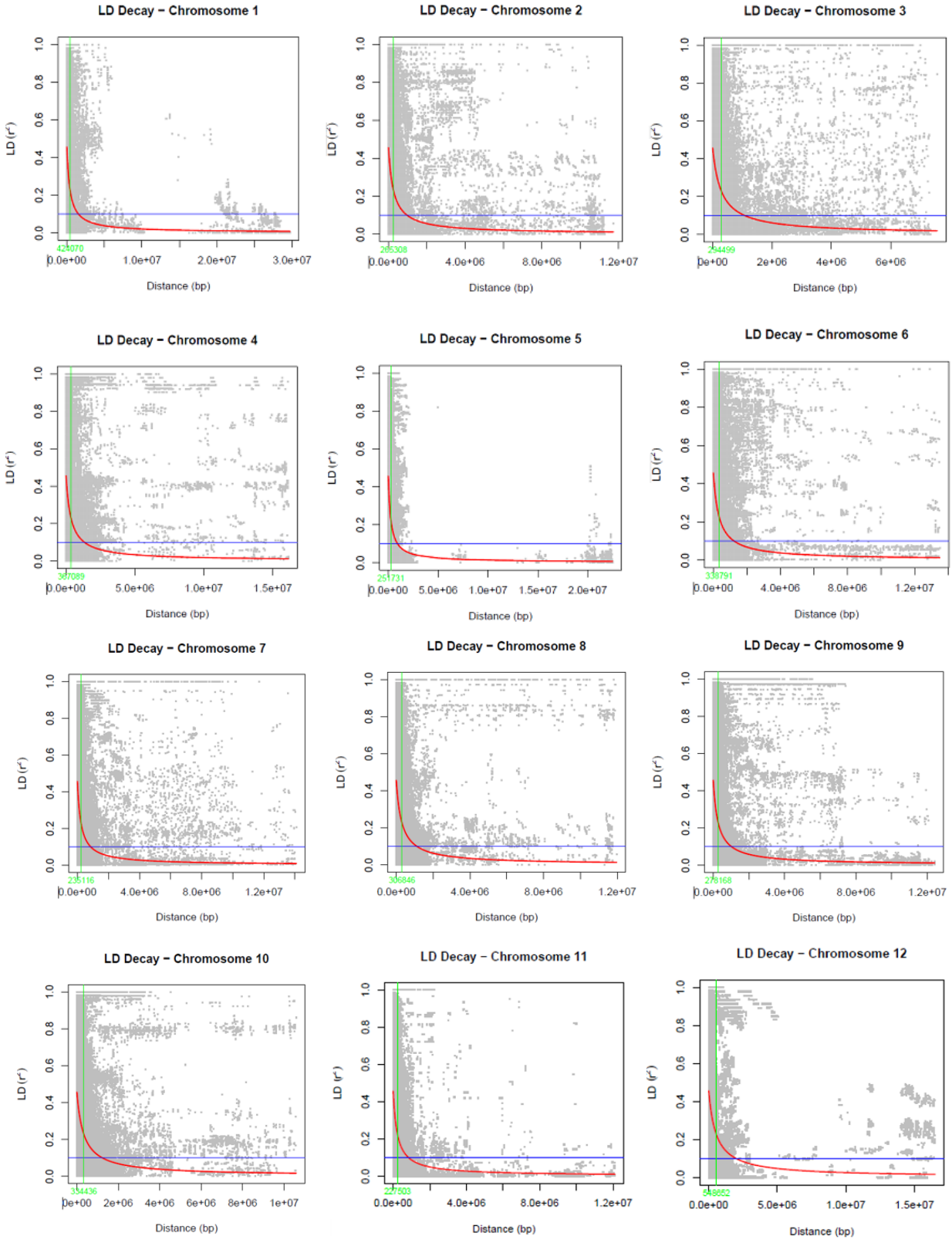
SK158	7,21	6,59	5,66	7,90	6,73	5,67	7,56	6,66	5,67	0,00	0,00	0,00
SK159	8,58	6,95	5,96	8,08	6,81	5,64	8,33	6,88	5,80	0,00	0,00	0,00
SK160	7,28	6,60	5,64	7,28	6,58	5,62	7,28	6,59	5,63	0,00	0,00	0,00
SK161	7,02	6,34	5,38	7,20	6,53	5,65	7,11	6,43	5,52	0,00	0,00	0,00
SK162	7,41	6,86	6,29	6,65	6,52	5,87	7,03	6,69	6,08	0,00	0,00	0,00
SK163	8,68	7,13	5,73	7,87	6,41	5,16	8,28	6,77	5,44	0,00	0,00	0,00
SK164	7,83	6,94	5,60	8,17	6,89	5,68	8,00	6,91	5,64	0,00	0,00	0,00
SK165	8,11	6,77	5,37	7,57	6,36	5,11	7,84	6,57	5,24	0,00	0,00	0,00
SK166	8,20	6,69	5,59	8,79	6,38	5,51	8,50	6,53	5,55	0,00	0,00	0,00
SK167	7,45	6,79	5,76	7,71	6,29	5,25	7,58	6,54	5,50	0,00	0,00	0,00
SK168	7,04	6,19	4,81	7,46	6,76	5,33	7,25	6,48	5,07	0,00	0,00	0,00
SK169	7,15	5,81	4,60	7,68	6,29	4,68	7,42	6,05	4,64	0,00	0,00	0,00
SK170	7,17	6,31	5,20	7,53	6,66	5,47	7,35	6,49	5,34	0,00	0,00	0,00
SK171	NA	NA	NA	7,86	6,56	4,89	7,86	6,56	4,89	0,00	0,00	0,00
SK172	NA	NA	NA	6,87	6,58	5,27	6,87	6,58	5,27	0,00	0,00	0,00
SK173	7,72	6,14	4,92	7,08	6,35	5,35	7,40	6,24	5,14	0,00	0,00	0,00
SK174	7,33	6,55	5,02	7,50	6,70	5,12	7,42	6,62	5,07	0,00	0,00	0,00
SK175	8,09	6,64	4,81	8,57	7,30	5,59	8,33	6,97	5,20	0,00	0,00	0,00
SK176	7,37	6,49	5,45	7,85	6,80	5,68	7,61	6,65	5,57	0,00	0,00	0,00
SK177	7,60	6,73	5,17	7,60	6,84	5,60	7,60	6,78	5,39	0,00	0,00	0,00
SK178	7,25	6,77	5,47	7,36	6,75	5,42	7,30	6,76	5,44	0,00	0,00	0,00
SK179	7,48	6,00	4,61	8,08	6,76	5,24	7,78	6,38	4,93	0,00	0,00	0,00
SK180	6,98	6,26	4,85	8,02	7,10	6,04	7,50	6,68	5,44	0,00	0,00	0,00
SK181	7,15	6,38	5,06	7,66	6,69	5,66	7,40	6,53	5,36	0,00	0,00	0,00
SK182	7,34	6,73	5,45	7,52	6,63	5,39	7,43	6,68	5,42	0,00	0,00	0,00
SK183	7,67	6,30	4,94	7,13	6,16	4,92	7,40	6,23	4,93	0,00	0,00	0,00
SK184	8,05	6,82	6,14	7,87	7,07	6,27	7,96	6,95	6,21	0,00	0,00	0,00
SK185	7,56	6,88	6,19	8,01	6,66	5,86	7,79	6,77	6,02	0,00	0,00	0,00
SK186	7,74	7,05	6,18	7,80	6,90	6,09	7,77	6,97	6,14	0,00	0,00	0,00
SK187	7,37	6,76	5,80	7,77	7,19	6,16	7,57	6,98	5,98	0,00	0,00	0,00

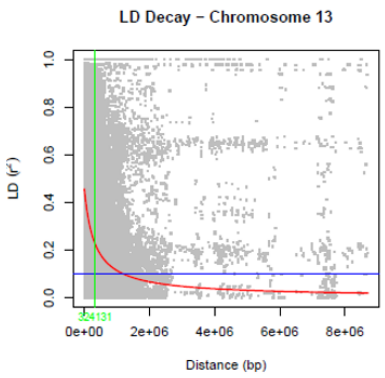
SK188	7,61	6,85	6,32	8,14	6,81	6,38	7,88	6,83	6,35	0,00	0,00	0,00
SK189	NA	NA	NA	7,82	6,89	6,05	7,82	6,89	6,05	0,00	0,00	0,00
SK190	7,18	6,47	5,82	7,47	6,73	5,99	7,33	6,60	5,90	5,00	0,00	2,50
SK191	7,84	6,84	5,77	7,63	6,83	5,64	7,73	6,84	5,70	0,00	0,00	0,00
SK192	7,88	6,29	5,35	6,87	6,43	5,54	7,37	6,36	5,45	0,00	0,00	0,00
SK193	7,77	7,20	6,04	7,64	7,02	6,00	7,71	7,11	6,02	0,00	0,00	0,00
SK194	8,01	6,65	5,93	7,79	6,33	5,35	7,90	6,49	5,64	0,00	0,00	0,00
SK195	NA	NA	NA	8,82	7,32	6,64	8,82	7,32	6,64	0,00	0,00	0,00
SK196	7,93	6,67	5,66	7,78	6,75	5,97	7,86	6,71	5,82	0,00	0,00	0,00
SK197	7,30	6,74	5,84	7,34	6,65	5,78	7,32	6,69	5,81	0,00	0,00	0,00
SK198	7,44	6,49	5,39	7,83	6,92	5,70	7,63	6,70	5,55	0,00	0,00	0,00
SK199	8,26	6,83	6,05	8,10	6,50	5,77	8,18	6,66	5,91	0,00	0,00	0,00
SK200	8,08	6,72	5,93	7,73	6,93	5,87	7,91	6,82	5,90	0,00	0,00	0,00
SK201	8,48	7,92	6,37	8,52	8,25	7,06	8,50	8,09	6,71	0,00	0,00	0,00
SK202	6,94	6,61	5,57	7,75	6,97	5,80	7,35	6,79	5,69	0,00	0,00	0,00
SK203	7,69	6,46	5,28	7,77	6,20	5,40	7,73	6,33	5,34	0,00	0,00	0,00
SK204	7,80	6,68	5,46	7,99	6,67	5,67	7,90	6,67	5,56	0,00	5,00	2,50
SK205	7,40	6,23	5,11	7,77	6,37	5,18	7,58	6,30	5,15	0,00	0,00	0,00
SK206	8,28	5,98	4,90	7,48	5,95	5,00	7,88	5,97	4,95	0,00	5,00	2,50
SK207	8,44	6,17	5,00	8,66	6,49	5,59	8,55	6,33	5,29	0,00	0,00	0,00
SK208	7,00	6,18	5,13	7,20	6,24	5,27	7,10	6,21	5,20	0,00	0,00	0,00
SK209	7,15	6,43	5,52	6,70	5,98	5,10	6,93	6,20	5,31	0,00	0,00	0,00
SK210	8,57	6,57	4,74	8,69	7,17	5,83	8,63	6,87	5,29	0,00	0,00	0,00
SK211	7,68	6,67	5,28	7,46	5,90	5,27	7,57	6,28	5,28	0,00	0,00	0,00
SK212	8,61	7,56	5,63	8,74	7,72	5,68	8,68	7,64	5,65	5,00	5,00	5,00
SK213	8,17	6,54	5,42	7,48	5,96	4,95	7,83	6,25	5,19	0,00	0,00	0,00
SK214	7,37	6,42	5,29	6,96	6,08	5,24	7,16	6,25	5,26	0,00	0,00	0,00
SK215	7,22	6,75	5,71	7,14	6,57	5,58	7,18	6,66	5,65	0,00	0,00	0,00
SK216	7,98	6,78	5,77	7,90	6,62	5,45	7,94	6,70	5,61	0,00	0,00	0,00
SK217	7,74	6,52	5,50	7,58	6,00	4,97	7,66	6,26	5,24	0,00	0,00	0,00

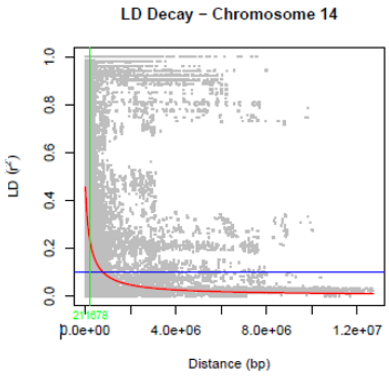
SK218	7,07	6,67	5,46	7,00	6,58	5,43	7,04	6,62	5,44	0,00	0,00	0,00
SK219	7,54	6,70	5,55	7,37	6,11	5,26	7,45	6,40	5,41	0,00	0,00	0,00
SK220	7,09	6,12	5,22	7,29	6,18	5,44	7,19	6,15	5,33	0,00	0,00	0,00
SK221	7,34	6,17	5,31	7,50	6,45	5,35	7,42	6,31	5,33	0,00	0,00	0,00
SK222	7,54	6,55	5,45	6,99	5,95	4,78	7,26	6,25	5,12	0,00	0,00	0,00
SK223	7,51	5,91	5,04	8,04	5,71	4,93	7,78	5,81	4,99	0,00	0,00	0,00
SK224	7,80	6,68	5,53	7,16	5,66	5,06	7,48	6,17	5,29	0,00	0,00	0,00
SK225	7,66	6,70	5,72	7,26	6,33	5,47	7,46	6,52	5,59	5,00	5,00	5,00
SK226	6,89	6,49	5,32	8,17	6,85	5,69	7,53	6,67	5,51	0,00	0,00	0,00
SK227	8,08	6,63	5,30	8,13	6,02	5,13	8,10	6,32	5,21	0,00	5,00	2,50
SK228	7,25	6,02	5,18	7,31	6,24	5,38	7,28	6,13	5,28	0,00	0,00	0,00
SK229	7,81	5,95	5,35	7,66	6,19	5,54	7,73	6,07	5,45	0,00	0,00	0,00
SK230	7,32	6,19	5,20	7,16	6,33	5,31	7,24	6,26	5,25	0,00	0,00	0,00
SK231	7,54	6,48	5,60	7,52	6,60	5,50	7,53	6,54	5,55	0,00	0,00	0,00
SK232	8,24	6,85	5,24	8,04	6,67	5,22	8,14	6,76	5,23	0,00	0,00	0,00
SK233	8,41	6,61	5,66	7,88	6,18	5,21	8,14	6,39	5,43	0,00	0,00	0,00
SK234	6,91	6,34	5,21	7,00	6,04	5,13	6,96	6,19	5,17	0,00	0,00	0,00
SK235	7,44	6,29	5,26	7,47	6,46	5,40	7,45	6,37	5,33	0,00	0,00	0,00
SK236	7,22	6,46	5,56	7,90	7,03	6,15	7,56	6,75	5,86	0,00	0,00	0,00
SK237	7,95	6,30	5,35	7,75	6,64	5,69	7,85	6,47	5,52	0,00	0,00	0,00
SK238	8,44	6,75	5,42	8,39	6,88	5,51	8,41	6,81	5,47	0,00	0,00	0,00
SK239	7,52	6,54	5,56	7,57	6,66	5,49	7,55	6,60	5,52	0,00	0,00	0,00
SK240	7,76	6,60	5,56	7,46	6,85	5,57	7,61	6,72	5,57	0,00	0,00	0,00
SK241	6,97	5,69	4,69	7,77	6,35	5,28	7,37	6,02	4,98	0,00	0,00	0,00
SK242	8,28	6,81	5,67	8,53	6,58	5,01	8,40	6,69	5,34	0,00	0,00	0,00
SK243	7,59	6,37	5,14	7,23	5,65	4,76	7,41	6,01	4,95	0,00	0,00	0,00
SK244	7,74	6,80	5,67	7,87	6,90	5,91	7,80	6,85	5,79	0,00	0,00	0,00
SK245	6,60	6,14	4,71	6,78	6,37	4,81	6,69	6,25	4,76	0,00	0,00	0,00
SK246	6,96	6,37	5,02	7,08	6,43	5,18	7,02	6,40	5,10	0,00	0,00	0,00
SK247	8,22	6,74	5,46	8,17	6,37	5,41	8,20	6,55	5,43	-	-	-

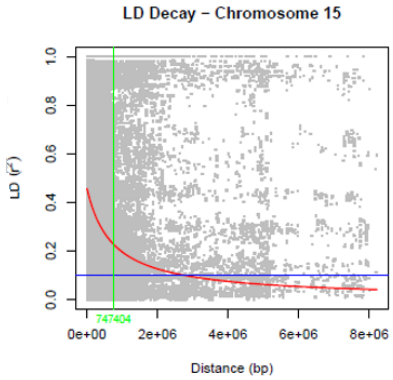
SK248	7,06	6,83	5,73	7,21	6,81	5,90	7,13	6,82	5,81	0,00	0,00	0,00
SK249	7,30	6,14	4,89	7,47	6,31	5,19	7,39	6,23	5,04	0,00	0,00	0,00
SK250	7,66	6,58	5,07	7,69	6,38	5,31	7,67	6,48	5,19	0,00	0,00	0,00
SK251	7,54	6,57	5,41	7,70	6,80	5,67	7,62	6,69	5,54	-	-	-
SK252	7,15	6,06	4,77	7,22	5,68	4,48	7,19	5,87	4,62	-	-	-

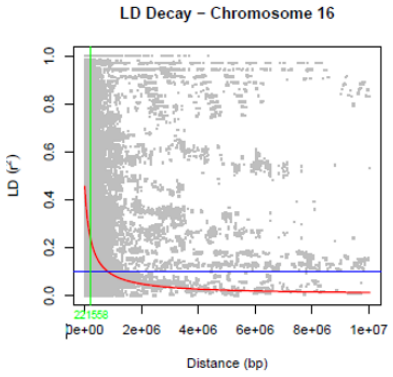
ПРИЛОЖЕНИЕ 3

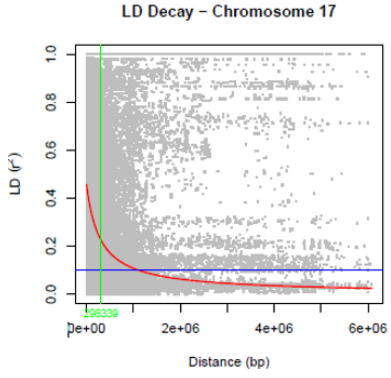


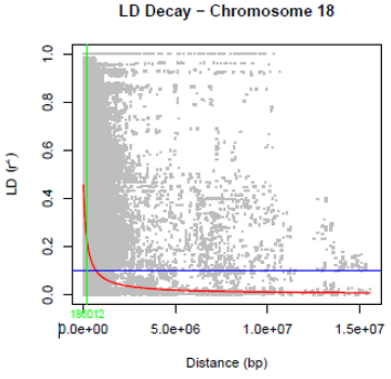

LD Decay – Chromosome 13
LD (r²)
Distance (bp)
0e+00
2e+06
4e+06
6e+06
8e+06

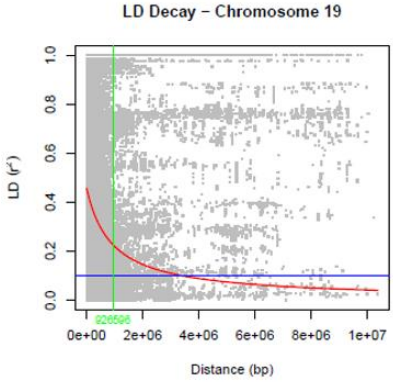

LD Decay – Chromosome 14
LD (r²)
Distance (bp)
0.0e+00
4.0e+06
8.0e+06
1.2e+07

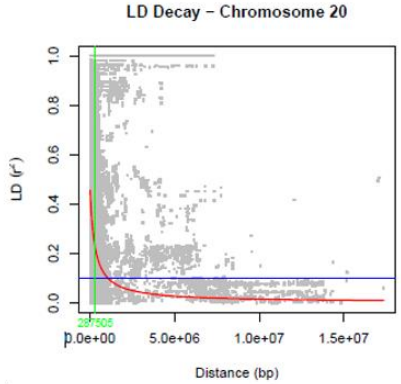

LD Decay – Chromosome 15
LD (r²)
Distance (bp)
0e+00
2e+06
4e+06
6e+06
8e+06


LD Decay – Chromosome 16
LD (r²)
Distance (bp)
0e+00
2e+06
4e+06
6e+06
8e+06
1e+07

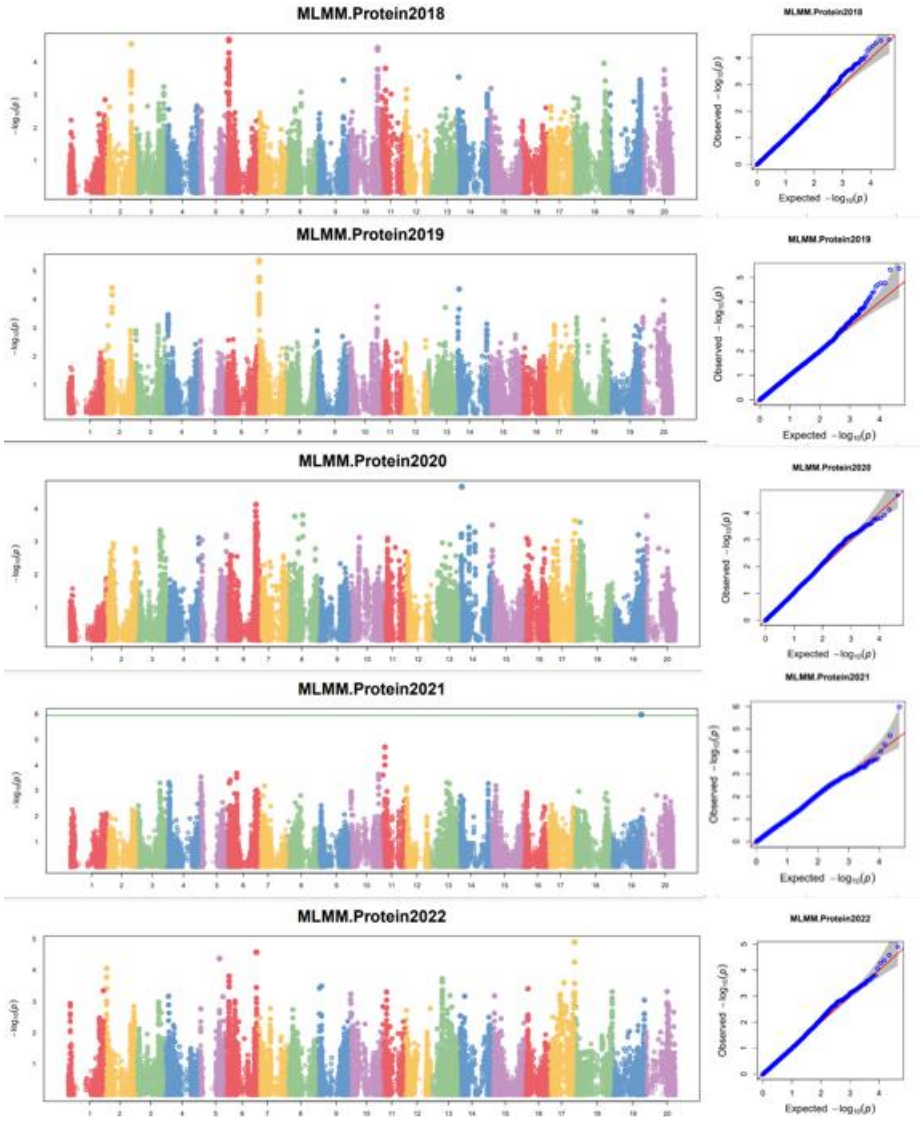

LD Decay – Chromosome 17
LD (r²)
Distance (bp)
2e+06
4e+06
6e+06

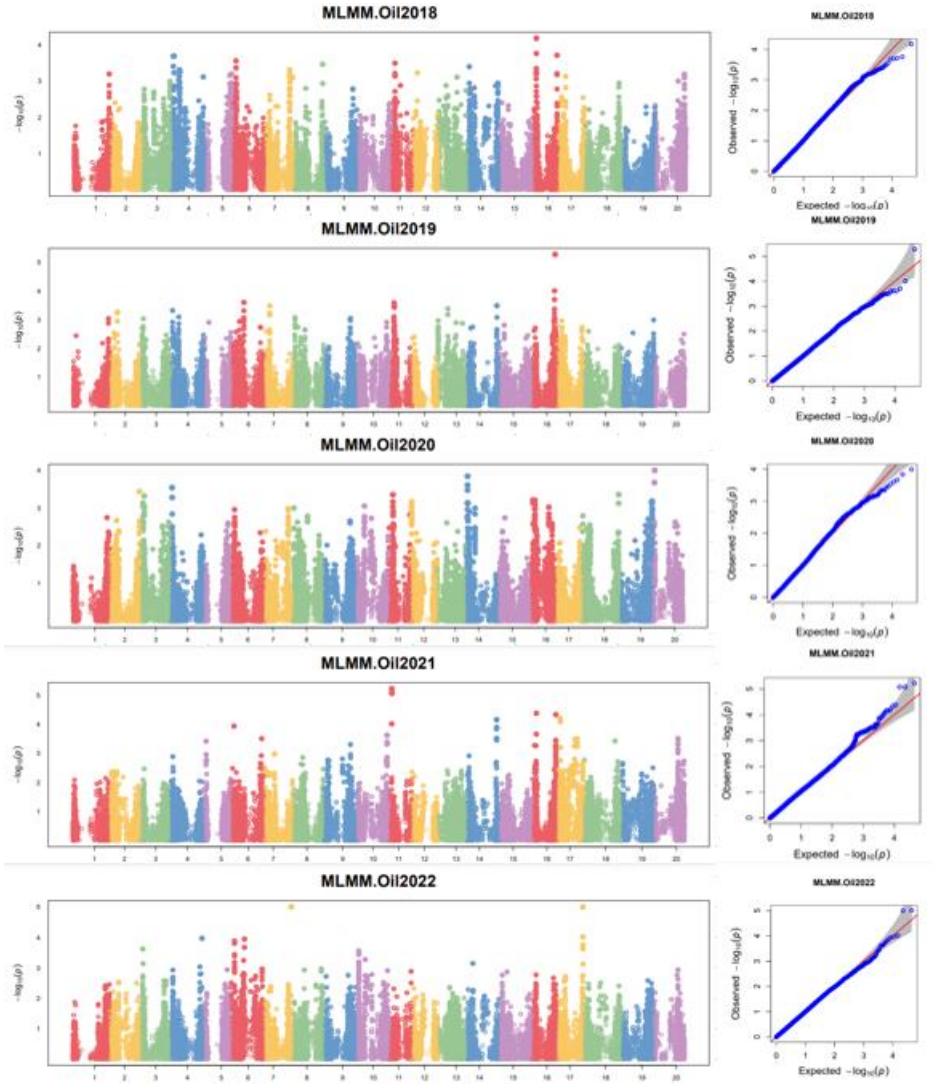

LD Decay – Chromosome 18
LD (r²)
Distance (bp)
0.0e+00
5.0e+06
1.0e+07
1.5e+07

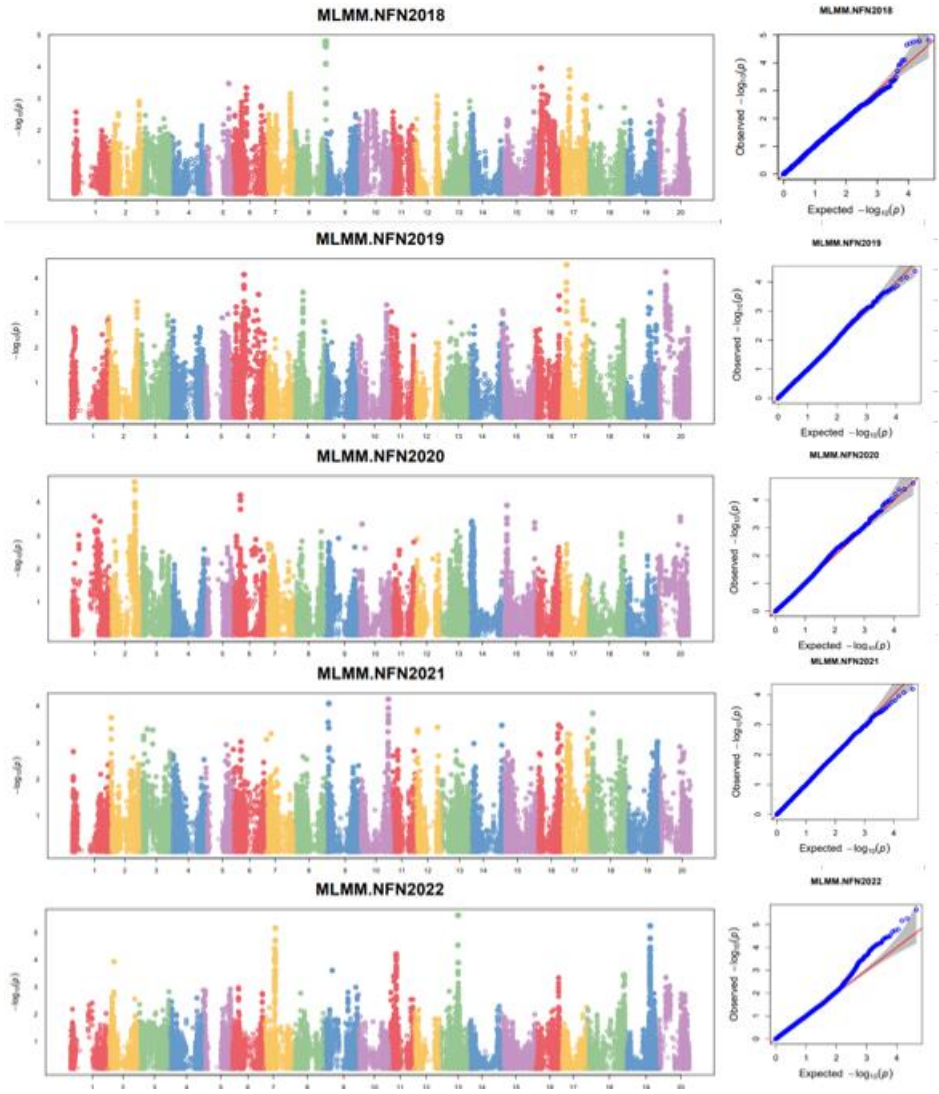

LD Decay – Chromosome 19
LD (r²)
Distance (bp)
0e+00
2e+06
4e+06
6e+06
8e+06
1e+07

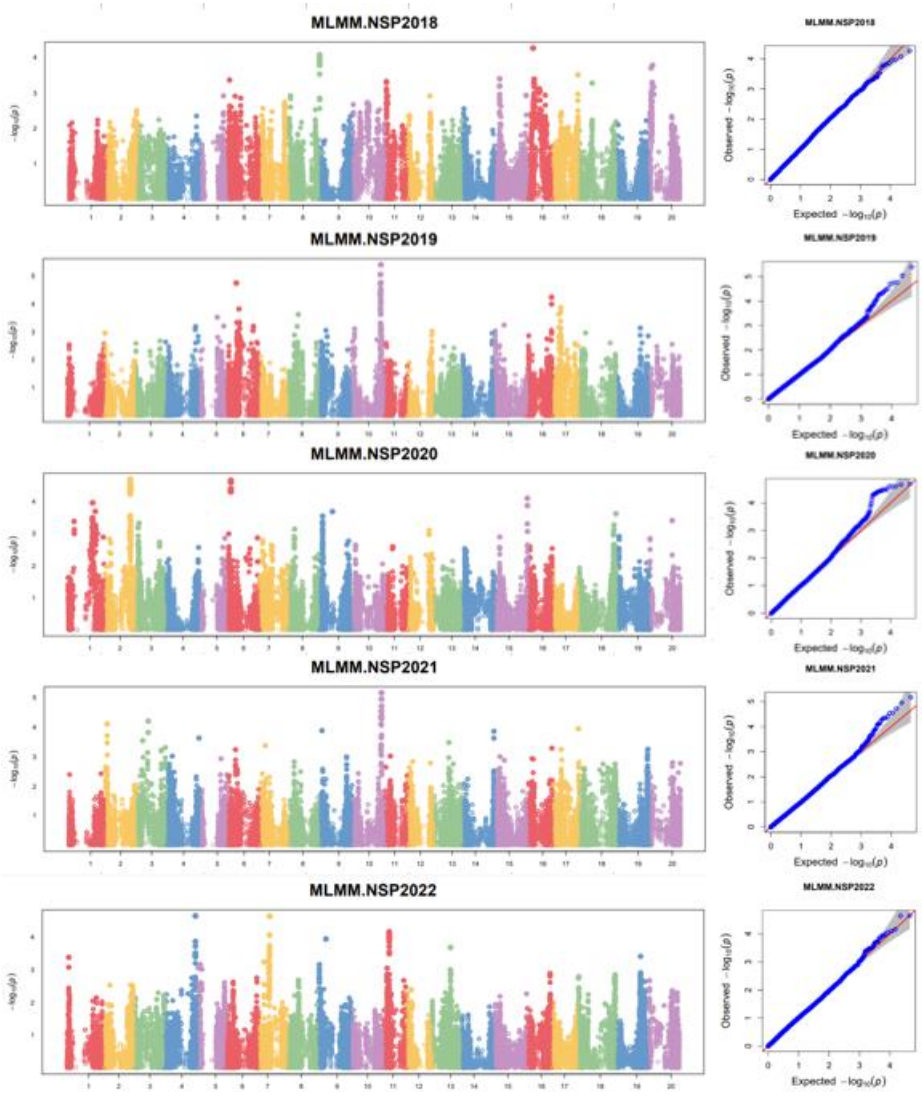

LD Decay – Chromosome 20
LD (r²)
Distance (bp)
0.0e+00
5.0e+06
1.0e+07
1.5e+07

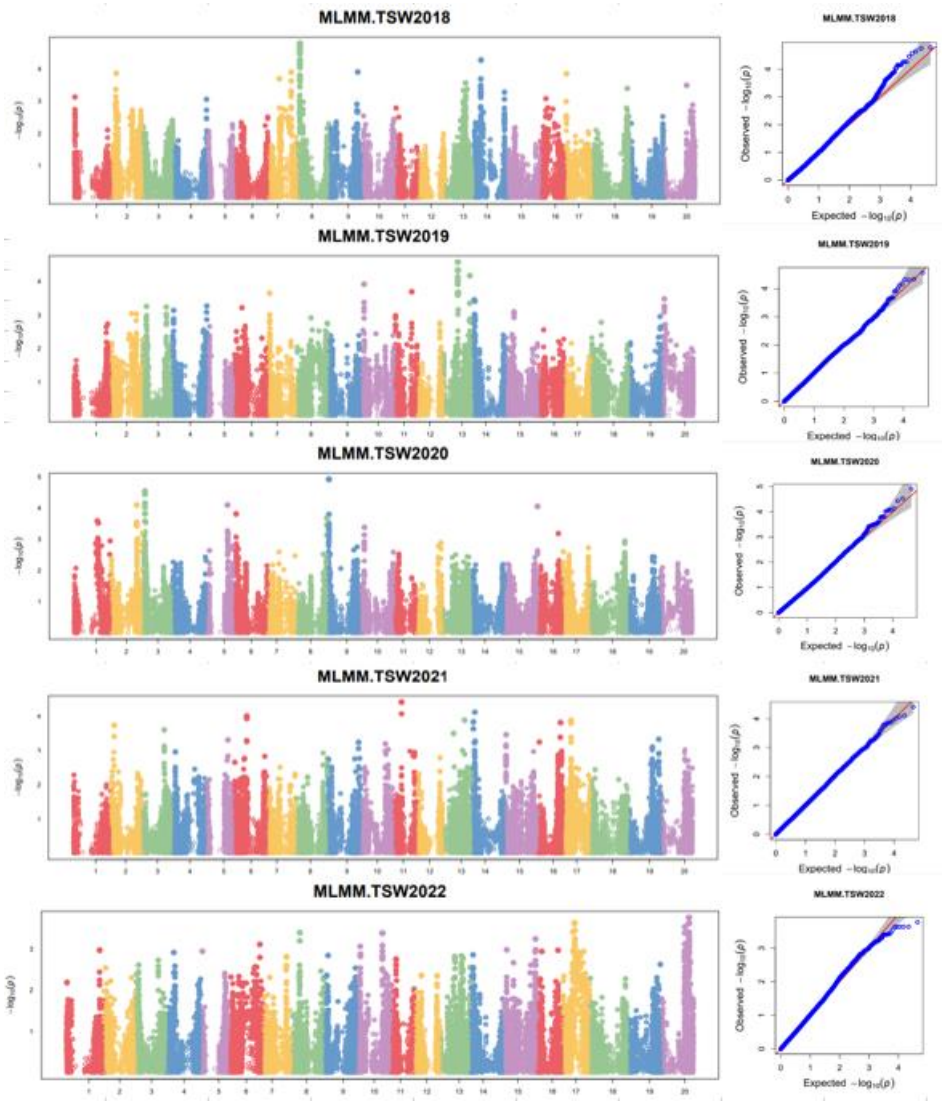
ПРИЛОЖЕНИЕ 4

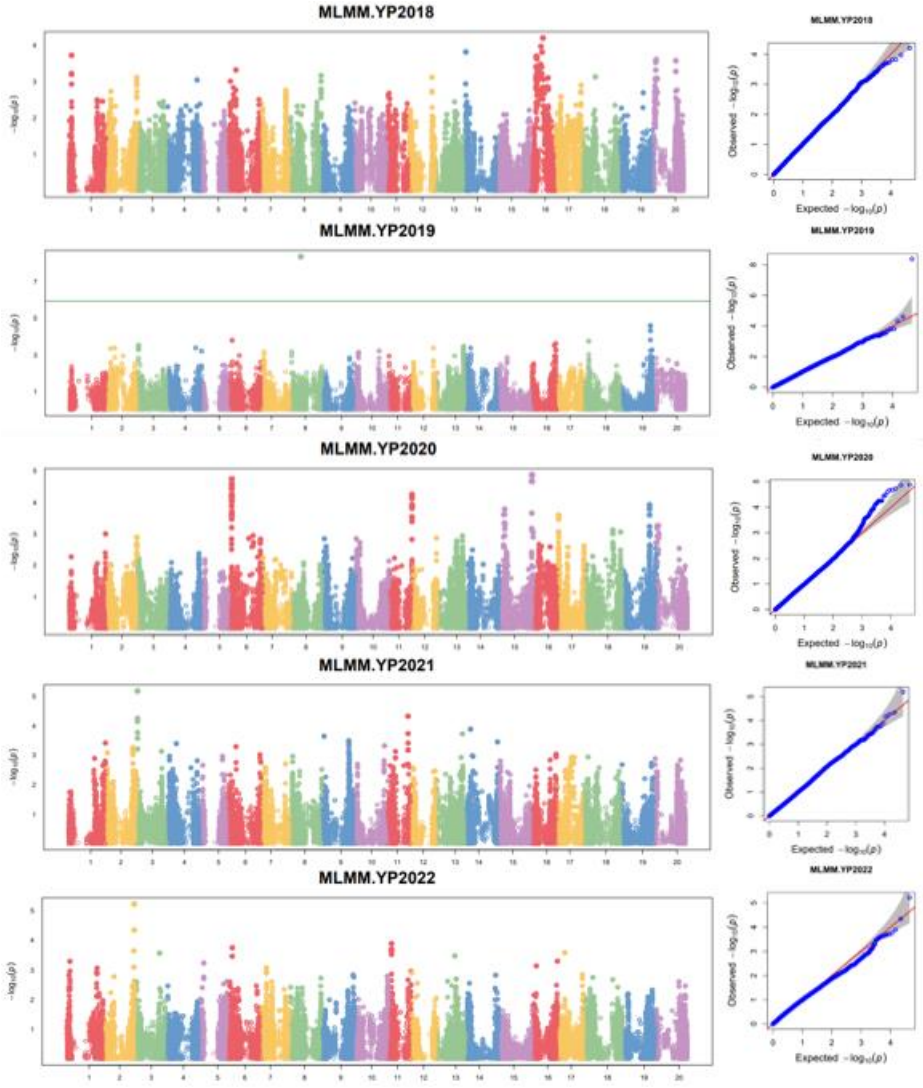



MLMM.Oil2018
MLMM.Oil2019
MLMM.Oil2020
MLMM.Oil2021
MLMM.Oil2022


MLMM.NFN2018
MLMM.NFN2018
MLMM.NFN2019
MLMM.NFN2019
MLMM.NFN2020
MLMM.NFN2020
MLMM.NFN2021
MLMM.NFN2021
MLMM.NFN2022
MLMM.NFN2022


MLMM.NSP2018
MLMM.NSP2018
MLMM.NSP2019
MLMM.NSP2019
MLMM.NSP2020
MLMM.NSP2020
MLMM.NSP2021
MLMM.NSP2021
MLMM.NSP2022
MLMM.NSP2022


MLMM.TSW2018
MLMM.TSW2018
MLMM.TSW2019
MLMM.TSW2019
MLMM.TSW2020
MLMM.TSW2020
MLMM.TSW2021
MLMM.TSW2021
MLMM.TSW2022
MLMM.TSW2022


MLMM.YP2018
MLMM.YP2019
MLMM.YP2020
MLMM.YP2021
MLMM.YP2022