

Федеральное государственное автономное образовательное
учреждение высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

На правах рукописи

Новикова Серафима Валерьевна

**Изучение генетической адаптации в популяциях лиственницы
сибирской (*Larix sibirica* Ledeb.) с использованием данных
полногеномного генотипирования**

Специальность 1.5.7 — Генетика

Диссертация

на соискание ученой степени кандидата биологических наук

Научный руководитель:

кандидат биологических наук

Крутовский Константин Валерьевич

Красноярск 2025

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
1.ОБЗОР ЛИТЕРАТУРЫ	13
1.1. Лиственница сибирская. Краткая характеристика	13
1.2. Адаптивный генетический потенциал лесных популяций в условиях глобального изменения климата.....	15
1.3. Ландшафтная геномика. Поиск ассоциаций геном-окружающая среда (GEA).....	16
1.4. Дендрохронология, годовичные кольца и дендрофенотипические показатели	20
1.5. Дендрогеномика. Полногеномный поиск ассоциаций (GWAS).....	20
1.6. ddRAD секвенирование как метод сокращения размерности генома для целей полногеномного генотипирования	23
2.МАТЕРИАЛЫ И МЕТОДЫ	26
2.1. Сбор образцов.....	26
2.2. Климат и биоклиматические переменные	30
2.3. Дендрохронологические данные	35
2.4. Выделение ДНК, подготовка ddRADseq библиотек и секвенирование	37
2.5. Обработка результатов секвенирования и генотипирование	39
2.6. Расчет основных популяционно-генетических параметров	40
2.7. Поиск корреляции между генетическими, географическими и экологическими координатами.....	41
2.8. Поиск корреляции между индивидуальной гетерозиготностью и дендрофенотипами.....	41
2.9. Поиск генетических маркеров, изменчивость которых, возможно, находится под влиянием отбора	42
2.10. Поиск генетических маркеров адаптации, связанных с переменными окружающей среды	43
2.11. Полногеномный поиск ассоциаций между генотипом и дендрофенотипом.....	45
2.12. Функциональная аннотация SNP-маркеров.....	47
3.РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ.....	48

3.1. Обработка результатов секвенирования, генотипирование	48
3.1.1. Популяции, отобранные для изучения адаптации к условиям высокогорья	48
3.1.2. Популяции, отобранные для изучения адаптации к широтному градиенту климатических условий	49
3.1.3. Популяции, отобранные для дендрохронологического анализа и изучения механизмов адаптации к условиям засухи	50
3.2. Структура популяции, основные популяционно-генетические статистики	50
3.2.1. Популяционно-генетические статистики	51
3.2.2. Иерархический анализ молекулярной дисперсии (AMOVA), индексы фиксации (F_{ST})	52
3.2.3. Изоляция расстоянием (IBD)	54
3.2.4. Анализ главных компонент (PCA)	56
3.2.5. Биогеографический анализ происхождения (Admixture)	60
3.3. Адаптация к условиям высокогорья	64
3.3.1. Поиск SNP-аутлайеров	64
3.3.2. Биоклиматические переменные	66
3.3.3. Поиск ассоциаций между генотипом и окружающей средой (GEA анализ)	69
3.3.4. SNP-маркеры, полученные в рамках исследования адаптации к условиям высокогорья	71
3.4. Адаптация к широтному градиенту климатических условий	77
3.4.1. Поиск SNP-аутлайеров	77
3.4.2. Биоклиматические переменные	78
3.4.3. Поиск ассоциаций между генотипом и средой (GEA анализ)	80
3.4.4. SNP-маркеры, полученные в рамках исследования адаптации к широтному градиенту климатических условий	81
3.5. Адаптация к условиям засухи	85
3.5.1. Дендрохронологические переменные, отражающие индивидуальный ответ дерева на условия засухи	85
3.5.2. Полногеномный поиск ассоциаций с использованием индивидуальных дендрофенотипов (GWAS)	89

3.5.3. SNP-маркеры, полученные в рамках изучения механизмов адаптации к условиям засухи	91
3.5.4. Корреляция между индивидуальной гетерозиготностью и дендрофенотипами (Heterozygosity-fitness correlations, HFCs)	95
ЗАКЛЮЧЕНИЕ	100
ВЫВОДЫ	103
СПИСОК ЛИТЕРАТУРЫ.....	104
ПРИЛОЖЕНИЕ. Аннотация SNP-маркеров, полученных в рамках исследования адаптации к условиям высокогорья	120

ВВЕДЕНИЕ

Актуальность темы исследования

Климатические модели прогнозируют увеличение продолжительности и интенсивности засушливых периодов [1]. В условиях изменяющегося климата в лесных экосистемах наблюдаются изменения в распределении границ лесов: верхняя граница высокогорных лесов смещается вверх, а северная граница леса расширяется в сторону более северных широт [2]. Для изучения процессов, происходящих в таких экосистемах, особое внимание следует уделить хвойным, так как они играют ключевую роль в структуре бореальных лесов [3].

Одним из эдификатор бореальных лесов Сибири является лиственница сибирская (*Larix sibirica* Ledeb.). Она обладает обширным ареалом с широким диапазоном климатических условий и экологической неоднородностью, что дает основание для выделения климатипов и географических рас вида, отчасти определяемых широтой произрастания. Высота произрастания, из-за влияния неблагоприятных факторов, которые действуют как основные силы в отборе наиболее приспособленных генотипов и стимулируют локальную адаптацию и дифференциацию, приводит к разделению *L. sibirica* еще и по высотным экотипам.

Методы секвенирования с сокращенной представленностью генома (Reduced Representation Sequencing, RRS) широко используются в популяционной генетике благодаря своей эффективности и доступности [4]. Секвенирование участков, ассоциированных с сайтами рестрикции (Restriction-site Associated DNA sequencing, RADseq), позволяет получить данные о тысячах молекулярно-генетических маркеров – однонуклеотидных полиморфизмов (single nucleotide polymorphism, SNPs) у большого числа особей, не ресеквенируя весь геном. Совместное использование данных генотипирования популяций из контрастных условий произрастания с биоклиматическими показателями позволяет провести поиск генетических основ адаптации и связать генетическую изменчивость с изменчивостью

конкретных адаптивных дендрофенотипов – характеристик дерева, основанных на годовичных кольцах прироста древесины [5].

Степень разработанности темы

Исследования подобного рода активно развиваются в последние годы, в том числе благодаря внедрению методов полногеномного генотипирования. Генетический аспект в исследовании процессов, протекающих в популяциях видов сибирских бореальных лесов, ранее был представлен сравнительно небольшим набором молекулярных маркеров [6–9].

С увеличением доступности полногеномных данных для организмов с мегагеномами, а также с появлением ddRADseq и других методов полногеномного генотипирования, для хвойных стали возможны полногеномный поиск ассоциаций (genome-wide association studies, GWAS) с использованием индивидуальных дендрофенотипов и поиск ассоциаций между генотипом и параметрами окружающей среды (Genome-Environment Association, GEA). Чаще всего, для хвойных видов исследования сфокусированы на их засухоустойчивости и холодоустойчивости [10–14], а в качестве переменных окружающей среды используется температура, количество осадков, высотность [15–17].

Цели и задачи исследования

Основной целью работы является поиск генетических основ адаптации популяций лиственницы сибирской к различным средовым факторам в контрастных экологических условиях с использованием биоклиматических и дендрофенотипических характеристик.

Достижение поставленной цели связано с реализацией следующих задач:

1. Прогенотипировать популяции *L. sibirica* из контрастных условий произрастания различных частей ареала, рассчитать основные популяционно-генетические параметры и определить пространственно-генетическую структуру исследуемых популяций.

2. Провести биоинформатический ассоциативно-корреляционный анализ для обнаружения связи между изменчивостью молекулярно-генетических маркеров и изменчивостью средовых климатических факторов.

3. Выполнить полногеномный поиск ассоциаций с использованием индивидуальных адаптивных дендрофенотипов – индексов устойчивости к засухе.

4. Проанализировать взаимосвязь между дендрофенотипами и уровнем индивидуальной генетической изменчивости (гетерозиготностью).

Научная новизна

Впервые пространственно-генетическая структура популяций *L. sibirica* была оценена с помощью данных полногеномного генотипирования методом ddRADseq. Была продемонстрирована пригодность данного метода для изучения мегагеномов хвойных. Полученный набор данных секвенирования для 488 деревьев из 37 точек сбора в дальнейшем может использоваться для различных исследовательских целей, таких как оценка генетического разнообразия, изучение структуры популяций, выявления гибридизации и интрогрессии, изучения адаптации и отбора.

Впервые выявлены маркеры и участки генома *L. sibirica*, изменчивость которых статистически достоверно связана с экологическими градиентами, ассоциированными как с высотой произрастания, так и с широтностью. На основании большого количества генетических маркеров показано наличие корреляции между генотипической изменчивостью дерева и характером его индивидуального ответа на засуху, определяемого на основании дендрохронологических индексов.

Выявлена взаимосвязь уровня индивидуальной гетерозиготности со стабильностью прироста древесины у лиственницы сибирской, что позволяет рассматривать индивидуальную гетерозиготность как потенциальный показатель приспособленности особей в условиях засушливого и нестабильного климата.

Теоретическая и практическая значимость исследований.

Результаты работы позволят глубже понять генетические механизмы, лежащие в основе адаптации *L. sibirica* к различным климатическим условиям. Данное исследование дало возможность выявить гены, потенциально участвующие в локальной адаптации. Обнаруженные SNP с выраженной адаптивной изменчивостью после функциональной проверки могут стать перспективными генетическими маркерами для разработки SNP-чипов с целью отслеживания адаптивной генетической изменчивости в популяциях лиственницы, выявления популяций с высоким генетическим потенциалом и осуществления экологического мониторинга. Предложенный подход, объединяющий генетические и дендрофенотипические данные, может лечь в основу научно обоснованного управления природными ресурсами, а также для установления географического происхождения древесины в рамках противодействия незаконным рубкам и обороту посадочного материала.

Основные положения, выносимые на защиту

1. Изученные популяции *L. sibirica* из контрастных частей ареала демонстрируют разделение на уральскую, алтайскую и сибирскую группы. Частоты генотипов коррелируют с географическим положением и условиями окружающей среды, что указывает на наличие пространственной и экологической изоляции, связанной с широтно-долготной дифференциацией и адаптацией к высотным условиям.
2. Участки генома *L. sibirica* содержат однонуклеотидные полиморфизмы, изменчивость которых статистически достоверно связана с экологическими градиентами, и однонуклеотидные полиморфизмы, изменчивость которых ассоциирована с характером индивидуального ответа дерева на засуху.
3. Уровень индивидуальной гетерозиготности в популяциях *L. sibirica* достоверно коррелирует со стабильностью прироста древесины, что позволяет рассматривать индивидуальную гетерозиготность как

потенциальный показатель приспособленности особей к условиям засушливого и нестабильного климата.

Апробация результатов

Результаты работы представлены на 7 международных и всероссийских научных конференциях:

1. Международная конференция «Лесные экосистемы бореальной зоны: биосферная роль, биоразнообразие, экологические риски», г. Красноярск, 16–20 сентября 2024 г.

2. Всероссийская научная конференция с международным участием «Продуктивность лесов в условиях меняющегося климата», посвященная 100-летию со дня рождения Н. И. Казиминова, и 8-е Международное совещание «Сохранение и рациональное использование лесных генетических ресурсов», г. Петрозаводск, 2–7 сентября 2024 г.

3. Международный Конгресс «VIII Съезд Вавиловского общества генетиков и селекционеров, посвященный 300-летию российской науки и высшей школы», г. Саратов, 14–19 июня 2024 г.

4. VII Международная научная конференция «Генетика, геномика, биоинформатика и биотехнология растений» (PLANTGEN 2023), г. Казань, 10–15 июля 2023 г.

5. VI Всероссийская научная конференция с международным участием «Устойчивость растений и микроорганизмов к неблагоприятным факторам среды» Иркутск, пос. Большое Голоустное 3–7 июля 2023 г.

6. Научная конференция с международным участием «Генетические процессы в популяциях», посвященная 50-летию юбилею лаборатории популяционной генетики им. Ю. П. Алтухова ИОГен РАН и 85-летию со дня рождения академика Юрия Петровича Алтухова, г. Москва, 11–14 октября 2022 г.

7. 6-я Международная конференция-совещание «Сохранение лесных генетических ресурсов» г. Щучинск, Казахстан, 16-20 сентября 2019 г.

Публикации

По теме диссертации опубликовано 4 работы в журналах, рекомендованных ВАК РФ, в том числе 4 в журналах, индексируемых в базе WoS и Scopus:

1. **Novikova, S.V.** Study of the Genetic Mechanisms of Siberian Stone Pine (*Pinus sibirica* Du Tour) Adaptation to the Climatic and Pest Outbreak Stresses Using Dendrogenomic Approach / S.V. Novikova, N.V. Oreshkova, V.V. Sharov, D.A. Kuzmin, D.A. Demidko, E.M. Bisirova, D.F. Zhirnova, L.V. Belokopytova, E.A. Babushkina, K.V. Krutovsky // Int. J. Mol. Sci. – 2024. – Vol. 25, № 21. – P. 11767.

2. **Novikova, S.V.** Study of the Genetic Adaptation Mechanisms of Siberian Larch (*Larix sibirica* Ledeb.) Regarding Climatic Stresses Based on Dendrogenomic Analysis / S.V. Novikova, N.V. Oreshkova, V.V. Sharov, D.F. Zhirnova, L.V. Belokopytova, E.A. Babushkina, K.V. Krutovsky // Forests. – 2023. – Vol. 14, № 12. – P. 2358.

3. **Новикова, С.В.** Генетическая структура и географическая дифференциация популяций лиственницы сибирской (*Larix sibirica* Ledeb.) на основе генотипирования генома путем секвенирования / С.В. Новикова, Н.В. Орешкова, В.В. Шаров, В.Л. Семериков, К.В. Крутовский // Сибирский экологический журнал. – 2023. – Т. 30, № 5. – С. 675-691.

4. **Novikova, S.V.** Genetic Adaptation of Siberian Larch (*Larix sibirica* Ledeb.) to High Altitudes / S.V. Novikova, V.V. Sharov, N.V. Oreshkova, E.P. Simonov, K.V. Krutovsky // Int. J. Mol. Sci. – 2023. – Vol. 24, №5. – P. 4530.

Личный вклад

Постановка цели и задач, выбор методов, анализ и обсуждение результатов исследования, и написание статей были проведены совместно с научным руководителем Крутовским К.В.

Автором работы выполнены частично или при участии коллег, отмеченных в разделе благодарности: полевые выезды и сбор образцов хвои, выделение ДНК и подготовка ddRAD библиотек для последующего

секвенирования. Сбор кернов, дендрохронологический анализ и расчет индивидуальных дендрофенотипов были организованы сотрудниками лаборатории дендроэкологии и экологического мониторинга Хакасского технического института, филиала СФУ. Секвенирование образцов осуществлялось сторонней компанией на коммерческой основе. Первичная обработка «сырых» данных секвенирования, их очистка, демультиплексинг и выравнивание на геном проводилось сотрудниками лаборатории геномных исследований и биотехнологии ФИЦ КНЦ СО РАН.

В ходе исследования автором лично проведены: генотипирование секвенированных образцов, вычисление популяционно-генетических параметров, корреляционный и ассоциативный анализ геномных, географических, биоклиматических и дендрохронологических показателей, а также аннотирование полученных наборов SNP.

Благодарности

Автор искренне признателен научному руководителю, заведующему лабораторией лесной геномики, профессору кафедры геномики и биоинформатики СФУ, к.б.н. Крутовскому К.В. за инициирование исследования, научное руководство и всестороннюю поддержку в ходе выполнения работы.

Автор выражает благодарность заведующей кафедрой геномики и биоинформатики д.б.н. Ямских И.Е. за ценные комментарии и поддержку, а также заведующему лабораторией молекулярной экологии растений Института экологии растений и животных УрО РАН, д.б.н. Семерикову В.Л. за руководство проектом, сбор образцов и выделение ДНК части описанных в работе популяций, сотрудникам Хакасского технического института, филиала СФУ: чл.-корр. Бабушкиной Е.А., к.б.н. Жирновой Д.Ф., к.б.н. Белокопытовой Л.В. и академику Ваганову Е.А. за проведение сборов, расчетов и анализа дендрохронологических данных, используемых в работе.

Особую благодарность автор выражает заведующей лабораторией геномных исследований и биотехнологии ФИЦ КНЦ СО РАН, к.б.н.

Орешковой Н.В. и к.б.н. Симонову Е.П. за содействие в реализации молекулярно-генетических этапов исследования, а также заведующему кафедрой высокопроизводительных вычислений, к.т.н. Кузьмину Д.А. и научному сотруднику ФИЦ КНЦ СО РАН Шарову В.В. за помощь в проведении вычислений и обработке данных секвенирования.

Также автор признательна к.б.н. Бондар Е.И., Мирошниковой К.А., Бирюкову В.В., Акуловой В.С. и Тараненко Е.А., а также всем членам лаборатории лесной геномики за поддержку и ценные советы во время работы над диссертацией.

Исследование проведено на базе лаборатории лесной геномики СФУ под руководством профессора К. В. Крутовского в рамках трех проектов:

- «Изучение генетической адаптации в популяциях лиственницы сибирской с использованием данных полногеномного генотипирования», финансируемого Российским Фондом Фундаментальных Исследований (РФФИ) в рамках Программы инициативных проектов (грант № 19-04-00964, руководитель К. В. Крутовский).
- «Изучение геномной и эпигеномной изменчивости, связанной с адаптацией древесных растений к гетерогенной среде», поддержанного РФФИ (грант № 18-29-13044, руководитель В. Л. Семериков).
- «Изучение генетической адаптации деревьев к стрессовым средовым факторам на основе полногеномного и дендрохронологического анализа в контексте глобального изменения климата», поддержанного Российским Научным Фондом (грант № 19-14-00120, руководитель К. В. Крутовский).

Также часть работы была проделана в рамках двух базовых проектов лаборатории геномных исследований и биотехнологии ФИЦ КНЦ СО РАН под руководством к.б.н. Н. В. Орешковой ФИЦ КНЦ СО РАН: FWES-2022-0003 (2022-2024) «Биотехнологические и молекулярно-биологические основы разработки биологических средств защиты растений нового поколения» и FWES-2025-0015 (2025-2027) «Геномные, популяционные и метагеномные исследования лесных биоценозов. Таксономический и филогенетический анализ растительных сообществ лесных экосистем».

1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Лиственница сибирская. Краткая характеристика

Лиственница сибирская (*Larix sibirica* Ledeb., 1833) – хвойное дерево из рода *Larix* семейства Сосновые (Pinaceae), занимающее обширные ареалы в бореальных и высокогорных регионах Северной Евразии. Являясь доминирующей хвойной породой на территории России, она имеет ключевое значение, как в функционировании природных экосистем, так и в экономике районов ее произрастания [18–20].

Лиственница сибирская — крупное дерево, достигающее высоты 30–40 метров, с диаметром ствола 80—100 см. Кора молодых деревьев гладкая, светло-соломенного цвета, с возрастом становится толстой и глубоко бороздчатой. Хвоя мягкая, узколинейная, длиной 1,3–4,5 см, собрана в пучки по 30–40 штук. Шишки мелкие (2–4 см), яйцевидной формы, созревают в течение одного сезона. Семена мелкие, с крылышками, что способствует их распространению ветром. Как и у всех представителей рода, хвоя лиственницы сибирской ежегодно опадает на зиму [21].

L. sibirica занимает обширный ареал, охватывающий Западную и Восточную Сибирь, Алтай, Саяны, Урал и частично Монголию и Китай (Рисунок 1). Она доминирует в таежных лесах, образуя как чистые лиственничники, так и смешанные леса с сосной, кедром, елью и пихтой. Вид предпочитает хорошо дренированные почвы, но способен расти на бедных, каменистых и заболоченных участках. *L. sibirica* устойчива к низким температурам, выдерживая морозы до -60°C , а корневая система адаптирована к вечной мерзлоте, что позволяет ей доминировать в экстремальных условиях северных лесов.

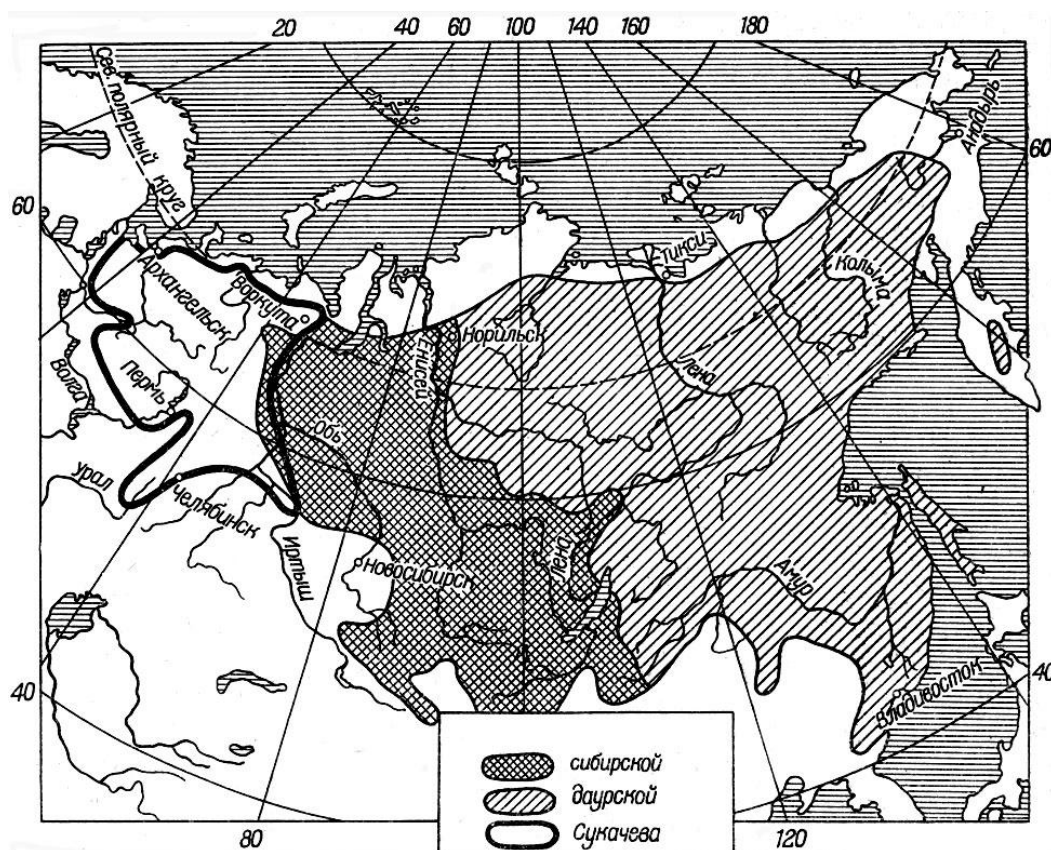


Рисунок 1. Ареал *L. sibirica* по кн. И.Ю. Коропачинский "Древесные растения Сибири"

[22]

Лиственница сибирская обладает высокой фитоценотической пластичностью, что позволяет ей успешно конкурировать с другими древесными видами. Обширный ареал способствует формированию географических экотипов, различающихся по морфологическим характеристикам [19,23], что указывает на высокую фенотипическую пластичность лиственницы и ее способность адаптироваться к разнообразным условиям среды [24]. Выделяют несколько географических рас, климатипов и экотипов лиственницы сибирской [25,26], среди которых две расы - *rossica* (север европейской части России к западу от Урала) и *obensis* (бассейн реки Оби, за исключением Алтая) – были объединены Н.В. Дылисом (1947) [27] в качестве самостоятельного вида – лиственницы Сукачева (*Larix sukaczewii* N. Dylis) (Рисунок 1) [28].

Как и у большинства хвойных, размер генома *L. sibirica* достаточно большой – 12 млрд пар нуклеотидных оснований, что существенно

превышает размеры генома резуховидки Таля (*Arabidopsis thaliana*), первого расшифрованного генома растений (114,5 млн п.н.о.), хотя и не является выдающимся по сравнению с другими хвойными (например, размер генома *P. taeda* более 20 млрд п.н.о, *P. sylvestris* – 22 млрд п.н.о., *P. lambertiana* 31 млрд п.н.о.)[29]. Существует ряд трудностей и ограничений, возникающих при работе с большими и сложными геномами. На текущий момент полные геномные сборки среди представителей рода *Larix* имеются лишь для двух видов – *L. kaempferi* и *L. sibirica* [30–32]. Завершенная в 2023 году аннотация генома лиственницы сибирской [30] значительно облегчила проведение генетических исследований и заложила важную основу для реализации масштабных ассоциативных исследований, а также для разработки молекулярных маркеров для целей геномной селекции.

1.2. Адаптивный генетический потенциал лесных популяций в условиях глобального изменения климата

Климатические прогнозы указывают на рост продолжительности засушливых сезонов и усиление засух в ряде регионов [33,34]. Это связано с повышением температур, которое ведет к увеличению испарения и, как результат, к более интенсивному водному стрессу для растительности [35,36]. В настоящее время, в связи с изменяющимися климатическими условиями, в лесных экосистемах наблюдаются процессы смещения границы леса, когда верхняя граница высокогорных лесов сдвигается выше, а леса распространяются севернее. Степень адаптации популяций будет зависеть от фенотипической изменчивости, силы отбора, плодовитости, межвидовой конкуренции и биотических взаимодействий [37]. Популяции умеренных и бореальных экосистем демонстрируют клинальную изменчивость (постепенное изменение фенотипических признаков) вдоль температурных градиентов, что указывает на существенную локальную адаптацию. Хвойные, такие как ель, лиственница, пихта и сосна, являющиеся ключевыми системообразующими лесными древесными растениями бореальных экосистем, и при этом обладающие высокой фенотипической пластичностью

и генетической изменчивостью, представляют большой интерес для научного сообщества в рамках изучения адаптивного потенциала лесных популяций к этим процессам [38,39].

1.3. Ландшафтная геномика. Поиск ассоциаций геном-окружающая среда (GEA)

Понимание генетических механизмов, лежащих в основе формирования адаптаций к разнообразным условиям окружающей среды является фундаментальным вопросом эволюционной биологии. Было показано, как давление отбора, вызванное гетерогенной средой, может способствовать локальной адаптации, отдавая предпочтение различным аллелям в зависимости от географической локации [40,41], точно так же, как экологические барьеры для потока генов могут усилить генетическую дифференциацию между ранее связанными популяциями из-за генетического дрейфа [42]. В некоторых случаях локальная адаптация из-за естественного отбора и уменьшения потока генов может в конечном итоге привести к видообразованию, что ставит генетическую изменчивость на фундаментальный уровень иерархии биоразнообразия [43,44]. Тем не менее, внимание к неоднородности окружающей среды в генетических исследованиях возросло только в последнее 20 лет [45,46]. Это связано с появлением и широким внедрением методов высокопроизводительного секвенирования, которые позволяют генерировать и генотипировать десятки и сотни тысяч генетических маркеров для любого организма, в том числе для таких, которые обладают большими и сложными геномами [38,47,48]. За последнее время с помощью этих методов было исследовано множество видов животных и растений [13,49–52].

Ранее понятие «ландшафтная генетика» было официально представлено в основополагающей статье Манела и соавт. [53], а затем расширено до ландшафтной геномики [54,55]. Ландшафтную генетику можно определить как «исследования, сочетающие популяционную генетику, ландшафтную экологию и пространственные аналитические

методы для количественного определения влияния ландшафтного состава, конфигурации и качества окружения на микроэволюционные процессы, такие как поток генов, дрейф и отбор с использованием нейтральных и адаптивных генетических данных» [56].

Возможность генотипирования большого числа маркеров и генов трансформировало ландшафтную генетику в ландшафтную геномику [14,57]. Активное развитие ландшафтной геномики во многом обусловлено растущим вниманием к экологическим и эволюционным последствиям глобальных антропогенных изменений, включая климатические трансформации, сокращение и фрагментацию природных ареалов. Особую актуальность приобретает необходимость прогнозирования и минимизации последствий этих изменений, что становится ключевой исследовательской задачей в условиях антропоцена, характеризующегося беспрецедентным по масштабам антропогенным воздействием на окружающую среду и сопутствующим сокращением биоразнообразия [58]. Ландшафтная геномика обладает огромным потенциалом для внесения вклада в такие исследования, поэтому неудивительно, что количество исследований по этой теме с 2003 года растет в геометрической прогрессии [59–63].

В настоящее время существуют два основных методологических подхода в ландшафтной геномике. Первый подход предполагает выявление локусов с аномально высокими показателями генетической дифференциации, т.н. F_{ST} -outlier тест. Второй подход, поиск ассоциаций геном-окружающая среда (genotype-environment association, GEA), основан на установлении ассоциаций между генетическими маркерами и параметрами окружающей среды.

Согласно F_{ST} -outlier тесту, действие направленного отбора в различных экологических условиях приводит к увеличению генетической дивергенции между популяциями, превышающей уровень, который может быть объяснен нейтральными эволюционными процессами. В то же время стабилизирующий отбор поддерживает генетическую дифференциацию на

уровне, статистически более низком по сравнению с ожидаемым при случайном распределении [64].

Поиск ассоциаций между генотипом и параметрами окружающей среды строится на предположении, что частоты аллелей локусов, ассоциированных с адаптивными признаками или участвующих в адаптации к конкретным экологическим факторам, должны демонстрировать значимую корреляцию с вариабельностью соответствующих экологических параметров и выраженные различия между географически обособленными популяциями [65].

Методы поиска ассоциаций на уровне геном-окружающая среда позволяют устанавливать явные связи с экологическим статусом индивидуальных деревьев, включая в анализ в качестве предикторов большое количество соответствующих переменных. Использование предикторов также может повысить мощность и обнаружить даже те селективные события, которые не вызывают высокой генетической дифференциации среди популяций [66–68].

С момента появления GEA одномерные статистические методы (без учета связей с другими переменными) доминировали в этом направлении [69]. Эти методы одновременно проверяют один локус и одну прогностическую переменную и включают обобщенные линейные модели [54,70], варианты линейных моделей смешанных эффектов [71,72] и непараметрические подходы [73]. Хотя эти методы работают хорошо, они могут давать повышенный уровень ложноположительных результатов при отсутствии поправки на множественные сравнения, что имеет повышенное значение при использовании больших наборов геномных данных. Поправки, такие как Бонферрони, могут быть чрезмерно консервативными (потенциально удаляя истинные положительные результаты), в то время как альтернативные методы поправки, такие как частота ложного обнаружения (false discovery rate, FDR), основаны на предположении о нормальном распределении p -значений, что может нарушаться для наборов эмпирических

данных [74]. Эти проблемы не должны препятствовать использованию одномерных методов, но поправки следует выбирать осторожно. Также другие подходы могут лучше подходить для геномных данных высокой размерности.

Появившиеся позднее многомерные подходы, одновременно анализирующие множество локусов, хорошо подходят для наборов данных, включающих сотни образцов, отобранных по многим тысячам генетических маркеров. На данный момент считается, что по сравнению с одномерными методами, эти подходы позволяют более эффективно обнаруживать мультилокусный отбор, поскольку они учитывают, как группы маркеров коварируют в ответ на переменные окружающей среды [68]. Это важно учитывать, так как что ожидается, что многие адаптивные процессы приведут к слабым, мультилокусным молекулярным сигналам из-за отбора на постоянной генетической изменчивости, который еще не привел к фиксации аллелей [75]. Поэтому выявление соответствующих закономерностей (например, скоординированных сдвигов частот аллелей во многих локусах), лежащих в основе этих адаптивных процессов, необходимо как для улучшения нашего понимания генетической основы локальной адаптации, так и для продвижения приложений этих данных, например, для сохранения эволюционного потенциал видов [76–78].

Исследования, основанные на геномном сканировании с применением описанных выше подходов, позволили идентифицировать участки геномов, находящихся под действием отбора и продемонстрировать адаптивную дивергенцию популяций. Эти работы получили свидетельства как в пользу конвергентного, так и параллельного характера генетических изменений в процессе адаптации к условиям окружающей среды. Также ряду исследователей удалось определить конкретные факторы среды, селективно действующие на локусы с резко выделяющимися значениями генетической дифференциации и связать её с функциональной ролью этих локусов [79,80].

1.4. Дендрохронология, годовые кольца и дендрофенотипические показатели

Методы дендрохронологического анализа, основанные на изучении изменчивости и точной датировке годовых колец древесины, дают возможность оценивать воздействие факторов окружающей среды на ростовые процессы древесных растений [81–83]. Динамика прироста годовых колец содержит обширную информацию об индивидуальной реакции деревьев на биотические и абиотические факторы, такие как колебания температуры, осадки и динамика численности вредителей [84–86], а также фиксируют значительные изменения среды, включая засухи, которые существенно влияют на прирост [87–89]. Корреляция динамики годового прироста с климатическими параметрами дает возможность идентифицировать различные дендрофенотипы и установить их взаимосвязь с генетической изменчивостью.

Ширина годового кольца (tree ring width, TRW) представляет собой ключевой показатель дендрофенотипической изменчивости. На основе TRW вычисляются производные характеристики: прирост базальной площади (BAI), стандартизированная ширина колец (RWI), средние ($AvTRW$) и дисперсия ($VarTRW$) показателей кольцевого прироста. Другие важные индивидуальные дендрофенотипы, предложенные Ллорет и соавт. [90] – это индексы сопротивления R_t , устойчивости R_s и восстановления R_c . Эти показатели отражают индивидуальную реакцию древесных растений на биотические и абиотические стрессовые факторы и широко применяются для оценки устойчивости радиального прироста к воздействию засухи и другим неблагоприятным условиям, включая инвазии фитофагов.

1.5. Дендрогеномика. Полногеномный поиск ассоциаций (GWAS)

Масштабное генотипирование множества деревьев в популяциях по обширному набору генов и генетических маркеров дает возможность установить взаимосвязь между генетической изменчивостью и вариабельностью специфических дендрофенотипов. Данный

междисциплинарный метод объединяет подходы дендрохронологии, дендрозкологии, дендроклиматологии, генетики и геномики, будучи впервые примененным в исследованиях [91–93], и получающим все более широкое распространение в последние годы [94–97]. Развивающееся направление, основанное на этом подходе, получило название "дендрогеномика", сам термин был впервые введен в публикации [98].

В настоящее время исследования в данной области направлены на решение нескольких важных проблем, связанных с динамикой границы леса, пространственно-временной структурой популяций, ростом и репродукцией деревьев, механизмами устойчивости к биотическим и абиотическим стрессам.

Для горной тсуги (*Tsuga mertensiana* Bong. Carr) на выборке из 163 деревьев, растущих на разных высотах на полуострове Кенай, Аляска, было проведено исследование с использованием данных о возрасте, ширине и дисперсии годовых колец с использованием 4665 SNP-маркеров, полученных в ходе полногеномного генотипирования [92,93]. С использованием 353 высокоинформативных SNP-маркеров (отобранных из 171019 SNP) была детально исследована возрастная, генетическая, родственная и репродуктивная структура древостоя на границе леса. Результаты показали: средняя дистанция распространения семян составила лишь 73 м; степень родства между деревьями оказалась невысокой; признаки клонального размножения отсутствовали. Достоверной взаимосвязи между индивидуальной гетерозиготностью и радиальным приростом деревьев не установлено, однако выявлена статистически значимая положительная корреляция между средней шириной годового кольца и изменчивостью этого признака в популяции. Это может указывать на нестабильность роста быстрорастущих особей, независимо от степени их гетерозиготности, в экстремальных условиях альпийского экотона.

Анализ ассоциаций выполнили с применением индивидуальных дендрофенотипов - индексов устойчивости (R_t , R_s и R_c), характеризующих

реакцию радиального прироста на засуху, и 130 SNP-маркеров, ассоциированных с устойчивостью к стрессам у 193 особей пихты европейской (*Abies alba* Mill.), переживших экстремальную засуху 1976 года [94]. Выявлена связь изменчивости 15 генов (включая гены фотосинтеза и засухоустойчивости) с вариабельностью дендрофенотипических показателей.

Дополнительно провели сопоставление дендрофенотипов (Rt, Rs, Rc и RRs), отражающих индивидуальную реакцию прироста на засуху, и характеристик древесины с результатами генотипирования 147 деревьев из 11 географических популяций в испытательных культурах ели европейской (*Picea abies* L.), созданных за пределами естественного ареала на высоте 250 м н.у.м. в северо-восточной Австрии [12]. Обнаружены достоверные ассоциации 29 SNP-маркеров с признаками засухоустойчивости, климатической чувствительности роста и свойствами древесины, объяснявшие 11-43% фенотипической изменчивости.

Для *P. abies*, комбинируя климатические и дендрофенотипические и данные с анализом 14 микросателлитных локусов, исследовали влияние дендрофенотипических, пространственных и климатических факторов на радиальный прирост и репродуктивный успех [99]. Установлено положительное влияние прироста и средней температуры предшествующего вегетационного сезона на материнскую репродуктивную эффективность. Особи, сохранявшие высокие темпы роста несмотря на ограничения, обладали наибольшей приспособленностью.

Для ели белой (*Picea glauca*), на 1473 деревьях, представляющих 43 популяции [100], была проанализирована связь климатических (градиенты температуры, осадков и засушливости), и 11 дендрометрических и связанных с засухой признаков с 6153 SNP-маркерами в 2606 генах-кандидатах, а также с данными транскрипции 37491 гена. В исследовании выявили 285 генов, демонстрирующих связь с фенотипическими характеристиками и климатическими параметрами. Из них 110 генов, дифференциально экспрессируемых при стрессе дефицита влаги. Построенная сеть

взаимосвязей "фенотип-генотип-среда" содержала 8 генов, участвующих в адаптации *P. glauca* к засухе, экспрессия четырех из этих генов оказалась чувствительной к условиям водного дефицита.

1.6. ddRAD секвенирование как метод сокращения размерности генома для целей полногеномного генотипирования

Метод RADSeq (Restriction site Associated DNA Sequencing), позволяет секвенировать только участки ДНК, прилегающие к сайтам рестрикции – нуклеотидным мотивам, специфично распознаваемым ферментами рестриктазами. Технология RADSeq была разработана в 2008 году и первоначально выступала как метод быстрого поиска SNP и генотипирования [4]. На данный момент RADSeq является одним из ключевых подходов высокопроизводительного секвенирования с редуцированной представленностью генома. Данная методика находит широкое применение для генотипирования популяций с использованием технологий секвенирования нового поколения [101]. В основе метода лежит отбор коротких фрагментов ДНК, примыкающих к специфическим сайтам узнавания рестриктазами, разрезающими нуклеотидную последовательность в местах узнавания, расположенных практически равномерно по всему геному, что позволяет проводить его репрезентативный анализ и создавать библиотеки для высокопроизводительного секвенирования [102].

Изначально метод создавался для внутривидового геномного анализа, однако современные исследования в различных областях – филогенетике, популяционной, экологической и эволюционной геномике – продемонстрировали потенциальную применимость RADSeq и для межвидовых исследований [103,104].

RADSeq объединяет два ключевых метода молекулярной биологии, обеспечивающих его широкое использование для различных исследовательских целей: 1) применение фрагментации ДНК специфическими рестрикционными ферментами и 2) возможность маркировки образцов индивидуальными баркодами для последующей

идентификации после секвенирования. Данный подход позволяет параллельно анализировать до 150-200 образцов, что существенно снижает стоимость процедуры секвенирования.

Метод RADSeq обеспечивает массовое генотипирование многочисленных образцов по сотням тысяч SNP-маркеров, равномерно распределенных по изучаемому геному. Данная технология применима для популяционно-генетических исследований видов, не обладающих референсным геномом или имеющих ограниченные геномные ресурсы (**Рисунок 2**).

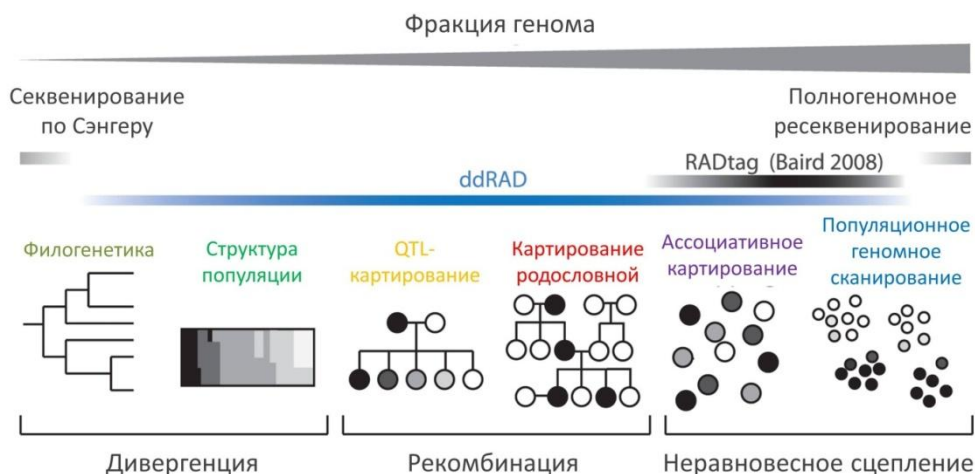


Рисунок 2 Применение ddRAD секвенирования для генотипирования позволяет оптимизировать количество генетических маркеров для конкретного экспериментального подхода [102].

В ddRADseq (double digest), используется два рестрикционных фермента (один с часто встречающимся мотивом сайта рестрикции и второй с редким), что позволяет осуществлять точный выбор размера итоговой библиотеки и достигать лучшей воспроизводимости (**Рисунок 3**). Предпочтительно также, чтобы одна из рестриктаз была метилчувствительная, чтобы избежать малоинформативные и обеднённые генами районы, которые обычно представлены высокоповторяющейся и высокометилированной ДНК.

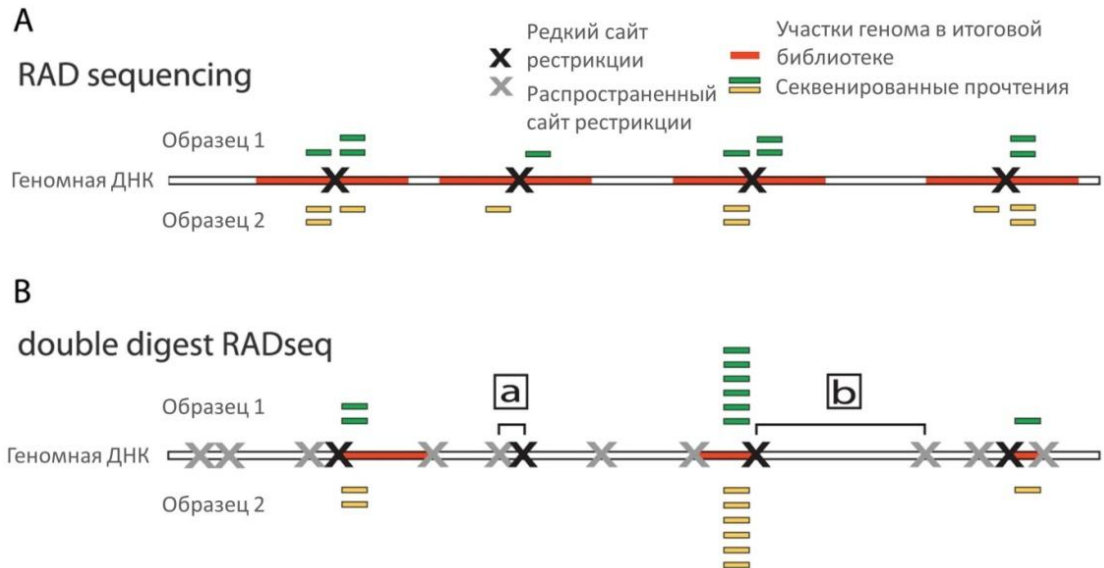


Рисунок 3. Секвенирование RAD (**A**). Обработка одной рестриктазой в сочетании со вторичной случайной фрагментацией. Секвенирование ddRADseq (**B**). Обработка двумя рестриктазами с последующим точным отбором по размеру, исключающим области, фланкированные либо [a] очень близкими, либо [b] очень удаленными сайтами рестрикции [102].

2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. Сбор образцов

В работе были использованы данные трех независимых проектов, в рамках которых проводились сборы образцов (июнь и июль 2019 – 2021 г.) на территории 7 регионов России: Красноярский край, Свердловская область, Республики Хакасия, Алтай, Тыва и Башкортостан, Ямало-Ненецкий автономный округ. Суммарно было отобрано, секвенировано и генотипировано 492 образца лиственницы сибирской из 37 точек сбора. Из них 231 дерево из 24 выборок было отобрано для исследования генетических механизмов адаптации к высотной поясности, 125 деревьев из 8 популяций отобрано для исследования влияния широтной поясности произрастания, 136 деревьев из 5 выборок были отобраны для дендрохронологического анализа устойчивости к засухе (Таблица 1, Рисунок 4).



Рисунок 4 Расположение точек сбора образцов *L. sibirica*. Круглыми маркерами на карте обозначены точки сбора образцов для изучения адаптации к высокогорным условиям, ромбами обозначены участки, где были собраны образцы для изучения изменчивости в широтном градиенте климатических условий, косыми крестиками обозначены точки сбора, в которых для образцов были дополнительно получены керны и рассчитаны дендрохронологические показатели устойчивости к условиям засухи.

Отбор образцов *L. sibirica* в Алтае-Саянском регионе (выборки 1-24, см. **Таблицу 1**) осуществлялся вдоль семи высотных трансект (А, С, D, Е, F, G, К), при этом каждая выборка собиралась на высотах, приблизительно соответствующих 500, 1000, 1500 или 2000 м.н.у.м. (что отражено в суффиксах _500, _1000, _1500 и _2000 в названиях выборок).

Также хвоя для экстракции ДНК отбиралась с 125 деревьев *L. sibirica* из восьми выборок (№25-32, см. **Таблицу 1**) вдоль двух географических трансект, характеризующихся градиентом биоклиматических параметров, обусловленных широтной поясностью. С учетом биогеографического распределения, сбор материала проводился: 1) в ареале западной расы *L. sibirica* – западная (уральская) трансекта (WT: выборки LAB, IVD, SEV, SIB); 2) в пределах типовой восточной расы - восточная (сибирская) трансекта (ET: выборки NU, NBSK, KRAS, TUR) (**Таблица 1, Рисунок 1**).

Сбор образцов из пяти выборок (с 33 по 37, см. **Таблицу 1**) был проведен в предгорьях Батеневского хребта – восточного отрога Кузнецкого Алатау, Республика Хакасия. Образцы древесины и хвои *L. sibirica* принадлежат выборкам, расположенным с интервалом 30-50 км в лесостепном поясе вдоль предгорий Батеневского хребта (рисунок 4). Участки выбраны на хорошо освещенных склонах, ориентированных от юго-запада до юго-востока.

Таблица 1 Расположение точек сбора образцов *L. sibirica*

№	Выборка	Кол-во образцов	Регион	Координаты и в.н.у.м
Биоклиматические переменные				
1	A_h_500	10	Саянский перевал, южнее г. Абаза, граница Респ. Хакасия и Респ. Тыва	52,487968 N 90,087579 E 548
2	A_h_1000	10		51,932477 N 89,827601 E 1021
3	A_h_1500	10		51,819086 N 89,798667 E 1522
4	A_h_2000	8		51,71465625 N 89,87788125 E

№	Выборка	Кол-во образцов	Регион	Координаты и в.н.у.м
				2086
5	C_h_500	9	Вблизи г. Туран, граница Респ. Хакасия и Респ. Тыва	52,03813 N 94,41143 E 676
6	C_h_1000	10		52,227013 N 93,686452 E 1496
7	C_h_1500	10		52,296698 N 93,68363 E 1496
8	G_h_500	10	Вблизи с. Приисковый, Республика Хакасия	54,82633219 N 89,12211998 E 471
9	G_h_1000	9		54,64112847 N 88,72485354 E 854
10	G_h_1500	10		54,63334245 N 88,6891428 E 1300
11	D_h_500	10	Семинский перевал, Чуйский тракт, Шебалинский район, Республика Алтай	51,4850364 N 85,5575046 E 580
12	D_h_1000	8		51,22784775 N 85,6391965 E 992
13	D_h_1500	10		51,0661722 N 85,5899068 E 1526
14	D_h_2000	10		51,0599355 N 85,6916137 E 2049
15	E_h_1000	10	Онгудайский район, Республика Алтай	50,37591557 N 87,01755721 E 922
16	E_h_1500	9		50,31867622 N 86,96392227 E 1446
17	E_h_2000	9		50,2958231 N 86,97921488 E 2082
18	F_h_500	9	Улаганский район, Республика Алтай	51,00016428 N 88,04201381 E 520
19	F_h_1000	10		50,91642577 N 88,20840134 E 1033
20	F_h_1500	10		50,75531748 N 88,08690841 E

№	Выборка	Кол-во образцов	Регион	Координаты и в.н.у.м
				1431
21	F_h_2000	10		50,49987377 N 87,65713938 E 2027
22	K_h_1000	10	Западная часть Каа-Хемского района Республики Тыва	51,78529563 N 95,40050714 E 1030
23	K_h_1500	10		51,92241823 N 95,46238891 E 1480
24	K_h_2000	10		52,00558523 N 95,57869646 E 1976
25	SIB	15	Хребет Ирэндик, вблизи г. Сибай (Республика Башкортостан)	52,78988889 N 58,27030555 E 574
26	SEV	16	п. Северка, окрестности г. Екатеринбурга (Свердловская область)	56,86845147 N 60,31872144 E 289
27	IVD	16	Окрестности г. Ивдель, Свердловская область	60,71000174 N 60,37234722 E 180
28	LAB	14	Окрестности г. Лабытнанги, Ямало-Ненецкий автономный округ	66,67899206 N 66,41047936 E 83
29	TUR	16	г. Туран в 65 километрах к северо-западу от г. Кызыла, южные отроги Западного Саяна, Республика Тыва	51,93063368 N 94,32375 E 1128
30	KRAS	16	Окрестности г. Красноярска, Красноярский край	55,95346875 N 92,15703299 E 12
31	NBSK	16	Окрестности г. Ноябрьска, Ямало-Ненецкий автономный округ	63,11023438 N 75,36250347 E 146
32	NU	16	Окрестности г. Н. Уренгой, Ямало-Ненецкий автономный округ	66,09730903 N 76,71616146 E 22
Дендрохронологические данные				
33	TUI	29	Окрестности с. Туим, Республика Хакасия	54,40455279 N 89,95829959 E 550-600
34	SON	29	Окрестности с. Сон, Республика Хакасия, на северо-восточной оконечности Батеневского кряжа	54,364 N 90,366 E 530-600
35	KAM	28	Окрестности д. Камызяк, Республика Хакасия	53,89924782 N 90,53280257 E 700-770

№	Выборка	Кол-во образцов	Регион	Координаты и в.н.у.м
36	BOG	23	Окрестности с. Боград, Республика Хакасия, на северо-восточной оконечности Батеневского кряжа	54,2655 N 90,6875 E 550-620
37	BID	27	Окрестности с. Вершино-Биджа, Республика Хакасия, на северо-восточной оконечности Батеневского кряжа	54,00268119 N 90,98333389 E 640-670

2.2. Климат и биоклиматические переменные

Юг Красноярского края и республика Хакасия. Большая часть исследования была проведена на популяциях, произрастающих в предгорьях Батеневского хребта, восточного отрога Кузнецкого Алатау. Данная горная система характеризуется относительно малыми перепадами высот, на южных и юго-восточных склонах, отличающихся большей засушливостью, наблюдаются участки редколесий, перемежающиеся степными сообществами. При этом значительную часть площади занимают смешанные лесные массивы с доминированием *L. sibirica*, сосны обыкновенной (*Pinus sylvestris* L.) и березы повислой (*Betula pendula* Roth.). Леса в исследуемом районе входят в состав государственного лесного фонда. Архивные записи Государственного комитета по лесу Республики Хакасия были использованы в качестве источника данных о нарушениях леса за последние десятилетия [105].

Территория характеризуется резко континентальным климатом с выраженными сезонными и суточными колебаниями температур, сочетающим жаркие летние периоды с морозными малоснежными зимами, продолжающимися с ноября по март [106]. Среднегодовая температура составляет 0°C, а годовой объем осадков варьирует от 470 до 560 мм, причем около 80% из них выпадает в период с положительными температурами, достигая максимума в июле. Безморозный период, когда минимальные суточные температуры остаются выше нуля, длится в среднем с мая по

сентябрь. В этот же промежуток времени среднесуточные температуры превышают отметку $+5^{\circ}\text{C}$. Кроме того, в течение этого сезона отмечаются характерные для региона значимые отрицательные корреляции между ежемесячными значениями температур и осадков. При этом минимальная температура не демонстрирует зависимости от количества выпадающих осадков [107]. Таким образом, климат в вегетационный период варьируется между двумя состояниями: жарким и сухим или прохладным и влажным, т.е. экстремальными и благоприятными для роста деревьев условиями.

В качестве основного источника климатических данных для популяций Батеневского хребта использовано пространственное распределенное поле ежемесячных рядов минимальной, средней, максимальной температуры и осадков из Climate Research Unit Time-Series [108,109]. В связи с изменением методики метеорологических наблюдений в России с 1936 года, а также значительным увеличением числа опорных метеостанций, климатические данные были ограничены периодом их наибольшей достоверности – с 1936 по 2019 год. Пространственные климатические градиенты в районе исследования определяли по исходному полю с разрешением $0,5 \times 0,5^{\circ}$ географической сетки, а для дендроклиматического анализа и выявления региональных климатических трендов использовали ряды, автоматически усредненные по всей территории Батеневского кряжа ($53,5\text{--}54,5^{\circ}\text{N}$ $89,5\text{--}91,5^{\circ}\text{E}$). Также для уточнения климатического режима отдельных лет использовали суточные данные температур и осадков расположенной в северной части района исследования метеостанции Ширы ($54^{\circ}30' \text{ N } 89^{\circ}56' \text{ E}$).

Республика Тыва (выборки А, С, К, TUR). Климат Тывы резко континентальный, рельеф горно-котловинный. Лето в горах умеренно теплое, а в котловинах — жаркое. Среднегодовая температура воздуха составляет $-5,5^{\circ}\text{C}$, при этом абсолютный минимум достигает -59°C , а максимум — $+38^{\circ}\text{C}$. Годовое количество осадков варьируется от 200 мм в котловинах до 1000 мм в горных районах. Продолжительность вегетационного периода — 150–160 дней. Для региона характерны поздние весенние заморозки, которые

могут наблюдаться примерно до 10 июня, а также ранние осенние заморозки, начинающиеся уже в конце августа. Безморозный период длится от 60 до 100 дней. Средняя температура июля в долинах составляет около 20 °С, тогда как в горных районах на высоте более 1000 метров она опускается до 12 °С.

Республика Алтай (выборки D, E, F). Климат резко континентальный, с коротким жарким летом и продолжительной морозной зимой. Среднегодовая температура воздуха в долинах колеблется от 0 до +5 °С, что является одним из самых высоких показателей для регионов Сибири. В горных районах среднегодовая температура опускается до -7 °С. Вегетационный период продолжается около 75-120 дней.

Республика Башкортостан (выборка SIB). Климат региона характеризуется как континентальный, с умеренно теплым или жарким летом и холодной зимой. Летняя жара усиливается из-за близости к засушливым степям Оренбургской и Челябинской областей, а также Казахстана, расположенных на юге республики. Уральские хребты, протянувшиеся с севера на юг, создают резкие различия в климате между западными и восточными склонами. Основная часть осадков (40–60 %) выпадает в течение трех летних месяцев. Рельеф оказывает значительное влияние на разнообразие климатических условий. Среднегодовая температура в регионе составляет +0,3 °С в горных районах и +2,8 °С на равнинах. Количество дней с положительной температурой воздуха варьируется от 200 до 205 на равнинах и от 188 до 193 в горах.

Свердловская область (выборки SEV, IVD). Регион расположен в зоне резко континентального климата. Климатические условия региона характеризуются суровыми продолжительными зимами. На равнинных территориях Зауралья средние январские температуры колеблются от -20°С в северных районах до -15°С на юге, достигая -17°С на юго-востоке. Летний период умеренно теплый, с повышением температур до жарких показателей в юго-восточных районах. Июльские температурные показатели составляют 16°С на севере и 19°С на юго-востоке. Вегетационный период длится 120-130

дней. Суммы осадков демонстрируют пространственную изменчивость: от 449 мм в восточных районах до 655 мм в западных, а также от 467 мм в южных до 518 мм в северных частях региона в год.

Ямало-Ненецкий автономный округ (NBSK, LAB, NU). Регион относится к районам Крайнего Севера. Климат преимущественно субарктический, а на крайнем севере переходит в арктический. Зима длится от 180 дней на западе до 230 дней на востоке. В летние месяцы возможны заморозки, а зимой — оттепели. Округ подвержен частому вторжению атлантических и арктических воздушных масс, что приводит к постоянной изменчивости погоды. Зимой и осенью преобладают ветры с южной составляющей, а летом — северные и северо-восточные, вызванные вторжением холодного арктического воздуха на нагретый материк, где в это время наблюдается пониженное атмосферное давление.

Вся территория округа находится в зоне избыточного увлажнения. Годовое количество осадков варьируется от 400 до 700 мм. Минимум осадков наблюдается в феврале, а максимум — в августе и сентябре. Не менее 30 % осадков выпадает в виде снега, присутствует многолетняя мерзлота. Климат округа демонстрирует тенденцию к потеплению. В 2020 году в Нарьян-Маре впервые с начала метеорологических наблюдений была зафиксирована положительная среднегодовая температура (+0,6 °C).

Биоклиматические переменные. На основании измеряемых на местах сбора для выборки в целом и для каждого дерева индивидуально географических координат и высоты произрастания, для каждого дерева при помощи пакетов R raster и sp были получены значения климатических переменных из базы данных WorldClim [110] в разрешении 2,5 минуты (примерно 2,6 км). Биоклиматические показатели вычисляются на основе месячных данных о температуре и количестве осадков для получения биологически релевантных параметров. Эти показатели представляют собой усредненные значения за период 1970-2000 гг. и широко применяются в модельных исследованиях. Биоклиматические переменные характеризуют:

годовые тенденции (среднегодовая температура, суммарные годовые осадки), сезонную динамику (годовая амплитуда температур и количества осадков), экстремальные и лимитирующие факторы среды (температуры наиболее холодного и теплого месяцев, количество осадков в наиболее влажный и сухой кварталы). Из 19 доступных в базе WorldClim, в работе были использованы 9 биоклиматических переменных:

1. ***Среднегодовая температура (Temp)***, полученная на основе данных о месячной температуре, рассчитываемая путём усреднения среднемесячных температур за весь год

2. ***Изотермальность (Isothermality)***, соотношение между средним суточным диапазоном температуры (Mean Diurnal Range) и годовым диапазоном температуры (Temperature Annual Range). Она выражается в процентах и отражает, насколько суточные колебания температуры сравнимы с годовыми. Высокая изотермальность (близкая к 100%) указывает на то, что суточные колебания температуры сравнимы с годовыми. Это характерно для регионов с умеренным климатом, где сезонные изменения температуры незначительны, например, в некоторых прибрежных или тропических зонах. Низкая изотермальность (близкая к 0%) означает, что суточные колебания температуры малы по сравнению с годовыми. Это типично для континентальных регионов с резкими сезонными перепадами температур.

3. ***Температурная сезонность (Temperature Seasonality, TempSeas)***, степень изменчивости температуры в течение года. Она рассчитывается как стандартное отклонение месячных значений температуры, умноженное на 100. Высокие значения указывают на значительные сезонные колебания температуры, что характерно для регионов с резко континентальным климатом. Низкие значения свидетельствуют о малой изменчивости температуры в течение года, что типично для регионов, где температура остается относительно постоянной круглый год (например, экваториальные зоны).

4. **Средняя температура самого холодного квартала (*MTofCQ - mean temperature of coldest quarter*)** определяется для самого холодного трёхмесячного периода в течение года. Переменная помогает оценить, насколько холодным является самый зимний период года в конкретном регионе.

5. **Средняя температура самого теплого квартала (*MTofWQ – mean temperature of warmest quarter*)**, определяется для самого тёплого трёхмесячного периода в течение года. Данный показатель позволяет количественно оценить степень прогрева в наиболее тёплый сезон для конкретной территории.

6. **Годовые осадки (*Annual Precipitation, Prcp*)**, сумма всех оценок количества осадков за 12 месяцев.

7. **Уровень инсоляции (кДж м⁻² сут⁻¹, *Srad – Solar Radiation*)**, относится к количеству солнечной энергии, получаемой на определенной площади поверхности за 24-часовой период. Зависит от широты, сезона, времени суток, облачности и атмосферных условий.

8. **Осадки самого влажного месяца (*PofWM - precipitation of wettest month*)**, определяется самый влажный месяц года и подсчитывается сумма осадков за этот период.

9. **Средняя скорость ветра (м с⁻¹, *Wind*)**.

Для всех анализируемых переменных с использованием пакета *vegan* в R [111] выполнены анализ главных компонент (PCA) и оценка факторов инфляции дисперсии (VIF) предикторных переменных с установленным пороговым значением $VIF \leq 10$.

2.3. Дендрохронологические данные

Образцы древесины (керны) взрослых неповрежденных живых деревьев *L. sibirica* брали на уровне груди. Сбор и обработка образцов проводились стандартными методиками дендрохронологии. Измеренные на установке LINTAB (Rinntech-Metriwerk GmbH & Co., Гейдельберг, Германия) с помощью программы TSAP-Win [113] индивидуальные ряды ширины

годовых колец (TRW) перекрестно датировали, датировку проверили в программе COFESHA [114]. В качестве дендрофенотипов использовали среднее значение ширины годовых колец, обобщенных raw хронологий (avTRW) и дисперсию (varTRW).

Из климатических экстремумов, значимо подавляющих прирост *L. sibirica* в региональном масштабе, были выбраны пять засух (1951, 1963-65, 1974-76, 1999, 2015 г.), произошедших с интервалом более 6 лет. Для оценки воздействия стрессовых факторов на радиальный прирост *L. sibirica* применялись разработанные Ллорет и соавт. [90] индексы сопротивления (resistance, $R_t = G_d / G_{prev}$), восстановления (recovery, $R_c = G_{post} / G_d$) и устойчивости (resilience, $R_s = G_{post} / G_{prev}$), G_{prev} - средний прирост в период за 3 года до засухи, G_d - средний прирост в период засухи, G_{post} - средний прирост в течение 3 лет после засухи, рассчитанные по стандартизированным хронологиям. Дополнительно анализировался индекс относительной устойчивости $RR_s = R_s - R_t$. В качестве фенотипических характеристик учитывались: средняя длина 15 хвоинок (avLn) и дисперсия (varLn), возраст дерева на момент отбора проб (Age). Тренды радиального прироста (trendTRW) для отдельных деревьев определялись как наклоны линейной регрессии стандартизированных (с устранением долгосрочных трендов, связанных с размером и возрастом дерева, динамикой древостоя) временных рядов за 30-летний интервал (1990-2019 гг.). Всего 27 дендрофенотипических признаков были рассмотрены в работе (Таблица 2). Фенотипические данные, не соответствующие нормальному распределению по результатам теста Шапиро-Уилка [115] были нормализованы с помощью алгоритма рангового обратного нормального преобразования (rank-based inverse normal transformation, INT).

Таблица 2. Описание дендрофенотипических переменных, рассчитанных для образцов *L. sibirica*.

Дендрофенотип	Обозначение				
Возраст дерева на момент сбора	Age				
Средняя длина хвои	avLn				
Дисперсия длины хвои	varLn				
Среднее значение ширины годичных колец	avTRW				
Дисперсия ширины годичных колец	varTRW				
Тенденции радиального прироста (1990-2019)	trendTRW				
Индекс:	Период засухи:				
	1951	1963-1965	1974-1976	1999	2015
сопротивления	Rt1	Rt2	Rt3	Rt4	Rt5
восстановления	Rc1	Rc2	Rc3	Rc4	Rc5
устойчивости	Rs1	Rs2	Rs3	Rs4	Rs5
относительной устойчивости	RRs1	RRs2	RRs3	RRs4	RRs5

2.4. Выделение ДНК, подготовка ddRADseq библиотек и секвенирование

Из хвои собранных образцов *L. sibirica* методом СТАВ [116] была выделена ДНК. Этот процесс включает в себя термостатирование растертой хвои с буфером на основе цетилтриметиламмонийбромид (СТАВ), несколько этапов центрифугирования и обработки хлороформом с последующим добавлением холодного изопропанола для осаждения ДНК. После отмывки осадка 80% этанолом, ДНК растворяли в деионизированной воде.

Концентрацию выделенной ДНК проверяли при помощи флуориметра Qubit 2.0, набор Qubit dsDNA BR (Thermo Fisher Scientific, Уолтем, США). Оценку чистоты и качества проводили с использованием Implen NanoPhotometer P330 (Implen, Мюнхен, Германия). Для работы были отобраны образцы высококачественной ДНК с A260/230 около 1,8 и концентрацией 20-150 нг/мкл, что соответствует протоколу приготовления ddRADseq библиотек.

ddRADseq библиотеки готовили на основании модифицированной версии протокола [102]. Производили обработку рестрикционными ферментами *EcoRI* и *MseI* [117], подобранными *in silico*, используя программу ddRADseqTools [118]. После обработки рестриктазами для каждого образца было проведено лигирование адаптеров, содержащих участок, уникальный для каждого образца (индивидуального дерева) для его последующей идентификации (баркод). После очистки реакционной смеси от несвязавшихся адаптеров с использованием магнитных частиц Agencourt AMPure XP (Beckman Coulter, Бреа, США) проводили ПЦР-амплификацию лигированных продуктов с применением высокоточной полимеразы Q5 Hot Start HF (New England BioLabs, Ипсвич, США) с промежуточным контролем качества библиотек методом гель-электрофореза для каждого образца. Амплифицированные ПЦР-продукты объединяли в пулы по 60-80 образцов. Для подготовки к секвенированию из каждого пула выделяли фрагменты размером 300-700 п.н.о. путем эксцизии соответствующей полосы после электрофореза в 2,5%-ном агарозном геле. Экстракцию ДНК из геля выполняли с применением набора QIAquick Gel Extraction Kit (Qiagen, Хильден, Германия). Контроль качества созданных пулов ddRADseq библиотек выполнялся методом капиллярного электрофореза на системе Bioanalyzer 2100 Agilent Technologies с применением набора High Sensitivity DNA (Agilent Technologies, Санта-Клара, США). Количественное определение концентрации пулов проводили с использованием флуориметра Qubit 2.0. Финальный контроль качества подготовленных библиотек перед секвенированием выполнялся на системе Agilent 2200 TapeStation System (Agilent Technologies, Санта-Клара, США).

Секвенирование ddRADseq библиотек выполняли на высокопроизводительной платформе NovaSeq 6000 (Illumina, Сан-Диего, США) с применением проточной ячейки S1. Длина прочтений составила 100 п.н.о., количество прочтений варьировало в диапазоне 400–500 млн. на дорожку.

2.5. Обработка результатов секвенирования и генотипирование

Полученные данные проходили несколько этапов начальной обработки. Секвенированные прочтения подвергались фильтрации и обрезке с установленными параметрами качества: MINQUAL=23 (минимальный показатель качества чтения) и MINLEN=40 (минимальная длина чтения после обработки), что позволяло сохранить информативные последовательности.

Каждая последовательность анализировалась на предмет наличия сайтов рестрикции *EcoRI*. На основании индивидуальных для каждого образца уникальных баркодированных последовательностей *EcoRI*-адаптеров, , выполнялось демультиплексирование - разделение данных по отдельным образцам в библиотеке. Все процедуры обработки осуществлялись с использованием программы `process_radtags` из пакета Stacks [119] при этом статистика по каждому этапу фильтрации собиралась программой FastQC [120], а отчеты визуализировались в MultiQC [121]. Очищенные прочтения каждого образца выравнивали с помощью программы Bowtie2 версии 2.3 [122] на референсный геном *L. sibirica* [31].

Поиск локусов, содержащих однонуклеотидные полиморфные позиции (SNP calling) осуществлялся с использованием программ Ustacks и Populations ПО Stacks, со следующими параметрами фильтрации:

- локус присутствует как минимум у 80% всех генотипируемых образцов;
- локус присутствует в каждой выборке;
- локус присутствует как минимум у 60-80% образцов в каждой выборке (пороговое значение подбиралось индивидуально под разные наборы образцов);
- наблюдаемая гетерозиготность по нуклеотидной позиции не должна превышать 0,6-0,8 (по всем генотипируемым образцам);
- минимальное покрытие минорного (самого редкого) аллеля – не менее 3-5;

минимальная частота минорного аллеля использовалась для фильтрации в значениях 0,01, 0,05 и 0,1 в зависимости от набора данных и поставленных задач.

2.6. Расчет основных популяционно-генетических параметров

Для анализа генетической изменчивости использовались следующие пакеты R: *adegenet* [123], *poppr* [124], *vcfR* [125], а также утилита Populations ПО Stacks. Для каждой трансекты были рассчитаны: частота наиболее часто встречающегося аллеля (P), доля гетерозиготных особей в данной популяции (H_O), гетерозиготность ожидаемая при равновесии Харди-Вайнберга (H_E), оценка нуклеотидного разнообразия (P_i), индекс фиксации (прокси коэффициента инбридинга особи относительно субпопуляции) (F_{IS}) согласно Хартлу и Кларку [126].

Определение показателей межпопуляционной дифференциации (F_{ST}) и иерархический анализ молекулярной дисперсии для оценки генетической дифференциации популяций (Analysis of Molecular Variance, AMOVA) выполнены с использованием программного обеспечения Arlequin v.3.5.2 [127].

Дополнительно был применен метод главных компонент (PCA) с реализацией в пакете R *vegan* [111], результаты которого для анализа популяционной структуры были визуализированы средствами пакета *ggplot2* R [128].

Исследование генетической структуры также проводилось с применением алгоритма *Admixture*, реализованного в программе *AdmixPipe* [129], позволяющего оценивать максимальное правдоподобие выделения генетических кластеров на основе данных генотипирования. Анализ основан на модели, которая предполагает, что геномы особей в популяции могут быть представлены как смесь генетических компонентов от нескольких предковых популяций. Программа оценивает пропорции этих компонентов для каждой особи, что позволяет визуализировать генетическую структуру в виде диаграмм. Оценка правдоподобия для значений K в диапазоне от 1 до 25, с

выполнением 20 повторностей для каждого значения K позволила выбрать оптимальное число генетических кластеров (K). Наилучшее значение K было выбрано на основе анализа ошибки кросс-валидации и показателя ΔK для каждой группы популяций.

2.7. Поиск корреляции между генетическими, географическими и экологическими координатами

Матрицы попарных географических расстояний и экологических параметров рассчитывались с применением евклидовой метрики в R-пакете *vegan*. Матрица попарных генетических дистанций для выполнения теста Мантелла, основанная на информации о SNP-маркерах, была сформирована в программе TASSEL [130] и вычислялась по формуле $1-IBS$, где IBS представляет собой показатель генетического сходства на основе идентичности генотипов (Identity by State). Тестирование по методу Мантелла проводилось в том же пакете с 999 пермутациями и уровнем значимости $p=0,001$.

2.8. Поиск корреляции между индивидуальной гетерозиготностью и дендрофенотипами

Наличие корреляции между индивидуальной гетерозиготностью и фенотипом дерева может указывать на несколько важных процессов, происходящих в популяции. Положительная корреляция между гетерозиготностью и дендрофенотипом может указывать на гетерозис, т.е. в таком случае гетерозиготные особи демонстрируют повышенную приспособленность, жизнеспособность или продуктивность по сравнению с гомозиготными.

Отрицательная корреляция между гетерозиготностью и дендрофенотипом может свидетельствовать об инбредной депрессии. Это происходит, когда снижение гетерозиготности (например, из-за инбридинга) приводит к ухудшению фенотипических признаков, таких как выживаемость, плодовитость или устойчивость к стрессу. Корреляция также может указывать на то, что изучаемый дендрофенотип находится под влиянием

конкретных генов, связанных с локусами, где наблюдается высокая гетерозиготность. Это может быть связано со сверхдоминированием (когда гетерозиготы имеют преимущество перед гомозиготами) или эпистазом (неаллельным взаимодействием между генами).

Для каждого образца был посчитан индивидуальный уровень гетерозиготности: Наблюдаемую индивидуальную гетерозиготность рассчитывали как количество полиморфных сайтов в гетерозиготном состоянии, деленное на общее количество полиморфных сайтов, генотипированных для данного образца. Анализ корреляций между индивидуальной гетерозиготностью и дендрофенотипами проводили с использованием непараметрического теста, ранговой корреляции Спирмена и параметрического, критерия корреляции Пирсона с помощью пакета R stats [131].

2.9. Поиск генетических маркеров, изменчивость которых, возможно, находится под влиянием отбора

Выявление генетических маркеров аутлайеров с аномально высокими показателями дифференциации и изменчивости, необъяснимыми нейтральными эволюционными процессами, выполнялось с применением трех популяционно-генетических методов, реализованных в соответствующих программах: Arlequin, BayeScan [132] и PCAdapt [133].

Для анализа адаптации к широтному климатическому градиенту вычисления в BayeScan, Arlequin и PCAdapt проводились независимо для восточной (сибирской, ET) и западной (уральской, WT) трансект. Контроль уровня ложного обнаружения (FDR) осуществлялся через q-значения с порогом $\alpha = 0,05$ [134]. В результате для каждой трансекты были отобраны SNP, достоверно идентифицированные всеми тремя программами.

Для обнаружения SNP-маркеров, находящихся под действием отбора, применялся R пакет PCAdapt, позволяющий выполнять поиск маркеро-кандидатов на основе анализа главных компонент. В ходе работы программа выполняет две последовательные задачи: сначала на центрированной и

масштабированной матрице генотипов выполняется анализ методом главных компонент, затем происходит вычисление тестовых статистик и p -значений на основе корреляций между SNP-маркерами и первыми K главными компонентами.

2.10. Поиск генетических маркеров адаптации, связанных с переменными окружающей среды

Анализ ассоциаций "генотип-среда" (Genotype-Environment Association analysis, GEA) представляет собой широко применяемый в популяционной генетике подход, направленный на выявление генетических маркеров или локусов, ассоциированных с адаптацией организмов к конкретным экологическим факторам, включая температурные условия, влажность, почвенные характеристики, высотную поясность и т.п. Этот подход позволяет выявить генетические основы адаптации без необходимости проведения экспериментов, полезен для изучения видов с длительным жизненным циклом. Однако существует риск ложноположительных результатов из-за влияния на частоты аллелей популяционной структуры и других спутывающих факторов (т.н. confounding factors).

Идентификация локусов, чья изменчивость потенциально связана с адаптацией к экологическим факторам, проводилась с использованием трех взаимодополняющих популяционно-генетических методов, реализованных в программах BayeScEnv [67], LFMM2 [135] и RDA [136] с целью снижения вероятности обнаружения ложных ассоциаций.

Анализ адаптации к широтному климатическому градиенту выполнялся отдельно для восточной (сибирской, ET) и западной (уральской, WT) трансект. В BayeScEnv и LFMM2 каждая климатическая переменная анализировалась независимо, тогда как в RDA проводилось одновременное тестирование пяти отобранных климатических параметров – среднегодовая температура, осадки, скорость ветра, влажность самого засушливого месяца и уровень инсоляции (Temp, Prec, Wind, PofDM и Srad соответственно).

При исследовании адаптации к высокогорным условиям рассматривались климатические переменные Temp, Iso, TempSeas, MTofWQ, MTofCQ, Prec (среднегодовая температура, изотермальность, температурная сезонность, средняя температура самого теплого и самого холодного кварталов, осадки) и синтетический предиктор PC1, полученный на их основе.

Контроль ложных открытий в BayeScEnv и LFMM2 осуществлялся через фильтрацию по q -значениям (порог $\alpha = 0,05$), тогда как для RDA использовался FDR-контроль с двусторонним p -значением 0,0027 (3 стандартных отклонения). В результате были отобраны SNP, достоверно идентифицированные всеми тремя аналитическими подходами.

Метод смешанных моделей со скрытым фактором (*LFMM*, Latent Factor Mixed Models) оценивает искажающие факторы (confounders) при выявлении генетических полиморфизмов, демонстрирующих корреляцию с градиентами окружающей среды или другими переменными. Искажающие факторы, или скрытые факторы, затем включаются в статистическую модель для проверки связи между генотипами и интересующей переменной. Переменные окружающей среды вводятся в модель как постоянная величина и используются независимо друг от друга, в то время как структура популяции моделируется с учётом её возможной подразделенности на K генетических кластеров, чтобы избежать ложно-положительных ассоциаций между SNP-маркерами и факторами среды. В основе метода лежат регуляризованные оценки методом наименьших квадратов для смешанных моделей со скрытым фактором. Так как метод крайне чувствителен к наличию пропущенных значений в данных, предварительно была проведена импутация на основании оценки индивидуального вклада («примеси», admixture) каждого потенциального кластера в генотип дерева. Также на этой стадии было выбрано наиболее вероятное число таких кластеров. Далее такие кластеры оценивались в модели как скрытые факторы.

Метод, реализованный в BayeScEnv, направлен на обнаружение локальной адаптации, связанной с переменными окружающей среды, с использованием генотипических данных средней и высокой плотности. Чтобы учесть структуру популяции, применяется F -модель, которая лежит в основе известного программного обеспечения BayeScan, с которым BayeScEnv разделяет большую часть своего кода. Метод основан на полиномиальной модели Дирихле. Отбор генетических маркеров вводится путем разложения коэффициентов F_{ST} на компоненты, специфичные для популяции, общие для всех локусов, и компоненты, специфичные для локуса, общие для всех популяций, с использованием логистической регрессии. Для каждого локуса МСМС (Марковские цепи Монте-Карло) с обратимым скачком исследует модели с отбором и без него и оценивает их относительные апостериорные вероятности. Модели BayeScEnv запускались отдельно для каждой средовой переменной с использованием одних и тех же параметров: 20 пилотных прогонов с 5000 итераций; интервал прореживания 10 запусков; 5000 выведенных итераций и длина прожига 30000 запусков. Все остальные параметры были сохранены по умолчанию.

2.11. Полногеномный поиск ассоциаций между генотипом и дендрофенотипом

Полногеномный поиск ассоциаций (Genome-Wide Association Study, GWAS) позволяет идентифицировать генетические полиморфизмы (например, SNP-маркеры), которые связаны с определенными фенотипическими признаками, такими как урожайность, устойчивость к болезням и засухе, так как является эффективным и широко используемым подходом для изучения ассоциаций между генотипом и фенотипом.

Однако такие аспекты, как стратификация (подразделенность) популяции, вызванная селективно-нейтральными факторами (такими как генетический дрейф и изоляция популяций), точность фенотипирования, сложная архитектура признаков и неравновесие по сцеплению могут привести к ложноположительным результатам. Для контроля этого

необходимы правильные статистические методы (например, смешанные модели). Кроме того, из-за большого количества протестированных SNP требуется строгая коррекция с учётом множественного тестирования.

Анализ ассоциаций "генотип-дендрофенотип" выполняли с применением обобщенных линейных моделей (Generalized linear model, GLM) и смешанных линейных моделей (Mixed linear model, MLM) в программном пакете TASSEL, где каждый SNP-локус рассматривался как независимая переменная, а дендрофенотипический показатель - как зависимая переменная. В MLM-анализ включались матрица родства K (kinship), вычисленная методом Centered IBS, и матрица популяционной структуры Q , полученная с использованием алгоритма sNMF (основанного на оценке энтропийного критерия и качестве подгонки модели методом кросс-валидации) для оптимального числа кластеров K [137]. Для выявления устойчивых ассоциаций, проявляющихся в различные засушливые периоды, проводили отбор общих SNP-локусов для соответствующих дендрофенотипов, удовлетворяющих критерию статистической значимости $p\text{-value} < 0.01$.

Ассоциации генотип-фенотип оценивали также с помощью Байесовской разреженной линейной смешанной модели (Bayesian sparse linear mixed model, BSLMM), реализованной в программном пакете GEMMA v.0.97 [138]. BSLMM представляет собой полигенную модель, учитывающую как вклад отдельных SNP с выраженным эффектом, так и совокупное влияние множества SNP с малым воздействием на фенотипическую изменчивость. Данная модель включает два типа предикторов фенотипа: сильные эффекты от небольшого количества индивидуальных SNP и полигенный эффект, обусловленный комбинацией множества слабовлияющих SNP. Для идентификации значимых ассоциаций "генотип-фенотип" по каждому признаку анализировали апостериорную вероятность включения (Posterior Inclusion Probability, PIP), отбирая маркеры с $PIP > 0,25$.

2.12. Функциональная аннотация SNP-маркеров

Для анализа геномных участков, содержащих маркеры генетической изменчивости, а также SNP-маркеры, ассоциированные с локальной адаптацией, была использована аннотация геномной сборки *L. sibirica*. Аннотация основана на установлении гомологов в неизбыточной (non-redundant, «nr») базе NCBI GenBank [139], отфильтрованной с использованием таксономического идентификатора для видов клады высшие растения (Embryophyta). Поиск белковых доменов производился с помощью программы InterProScan на сервере EMBL-EBI [140]. Термины геномной онтологии были определены с использованием платформы OmixBox, в частности встроенного ПО Blast2GO [141]. Аннотирование ассоциированных с локальной адаптацией SNP-маркеров выполнялось с применением инструмента аннотации однонуклеотидных полиморфизмов SNPdat [142]. Дополнительно, ввиду ограничений полноты доступной геномной аннотации, для SNP, демонстрирующих потенциально адаптивную изменчивость (из консенсусного списка), по их геномным координатам были извлечены фланкирующие нуклеотидные последовательности (± 1000 п.н.о.), после чего проведен поиск гомологичных последовательностей в базе данных NCBI BLAST.

3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3.1. Обработка результатов секвенирования, генотипирование

В рамках исследования для решения различных задач: изучения адаптации к условиям высокогорья, к широтному градиенту климатических условий и к условиям засухи, были использованы три неперекрывающихся выборки образцов, что обусловлено спецификой каждой из частей работы. Построение локусов и определение полиморфных позиций проводилось отдельно для каждой части исследования с учетом целей и задач, что позволило оптимизировать процесс и минимизировать потенциальные ошибки. Дополнительно, для изучения структуры популяции, был получен набор однонуклеотидных полиморфизмов, общий для всех задействованных в работе 488 образцов *L. sibirica*, включающий в себя 2938 SNP-маркеров.

3.1.1. Популяции, отобранные для изучения адаптации к условиям высокогорья

По результатам высокопроизводительного секвенирования для 250 образцов Алтае-Саянского региона было получено более 3,2 млрд прочтений длиной 100 п.н.о. После проведения первичной обработки и фильтрации по качеству было отобрано 3,1 млрд прочтений, в среднем 13,5 млн прочтений на образец. В среднем 97% прочтений каждого образца были успешно картированы на референсный геном *L. sibirica*. Результаты картирования представлены в Таблице 3.

Таблица 3 Результаты картирования секвенированных прочтений на референсный геном *L. sibirica*

Трансекта	Среднее кол-во ($\pm$ CO) прочтений на образец, млн	Не картированные прочтения, млн	Картированы уникально, млн	Картированные в несколько мест в геноме, млн	Доля успешно картированных прочтений, %
A	16,5 $\pm$ 9,9	0,690	1,378	14,432	96,22%
C	14,7 $\pm$ 13,2	0,423	1,313	13,032	97,22%
D	9,1 $\pm$ 3,4	0,176	0,698	8,241	98,08%
E	16,9 $\pm$ 10,5	0,948	1,469	14,512	94,59%

Трансекта	Среднее кол-во ($\pm$ CO) прочтений на образец, млн	Не картированные прочтения, млн	Картированы уникально, млн	Картированные в несколько мест в геноме, млн	Доля успешно картированных прочтений, %
F	6,4 $\pm$ 2,3	0,148	0,487	5,835	97,60%
G	13,1 $\pm$ 5,1	0,257	1,116	11,775	98,02%
K	12,9 $\pm$ 6,02	0,374	1,092	11,493	97,22%

Из последующего анализа ввиду низкого уровня картирования на референсный геном и недостаточного количества прочтений (≤ 1 млн) полностью исключили 19 секвенированных образцов. При генотипировании 231 образца для исследования адаптации к условиям высокогорья был сформирован набор из 20 тысяч локусов, содержащих 25143 биаллельных SNP.

Дополнительно сформировали набор данных из 761 предположительно селективно-нейтрального SNP, отобранный из некодирующих областей генома для оценки потенциальной генетической структуры популяции. Также набор данных из 550 предположительно адаптивных SNP был получен в ходе поиска маркеров-аутлайеров, изменчивость которых коррелировала с биоклиматическими переменными.

3.1.2. Популяции, отобранные для изучения адаптации к широтному градиенту климатических условий

В результате секвенирования для 125 деревьев, отобранных для изучения адаптации к широтному градиенту климатических условий и принадлежащих западной (уральской, WT) и восточной (сибирской, ET) трансектам, было суммарно получено 1,7 млрд. нуклеотидных прочтений. Около 1,2 млрд. прочтений длиной от 32 до 92 п.н.о., в среднем по 9,6 млн $\pm$ 0,4 млн прочтений на образец были отобраны после проведения первичной обработки и фильтрации по качеству. Для каждого образца в среднем по 97,86% полученных прочтений были успешно картированы на референсный геном *L. sibirica*.

При поиске полиморфных позиций (SNP) было отобрано 87 тыс. локусов, соответствующих критериям фильтрации и содержащих 47 929 биаллельных SNP-маркеров для 125 образцов. Доля отсутствующих данных по частотам аллелей на образец составила в среднем $5,98\% \pm 0,41\%$. Пропущенные частоты были сгенерированы с помощью процедуры импутации (k-nearest neighbor genotype imputation method, LD-kNNi, метод импутации генотипа k-ближайших соседей, LD-kNNi).

3.1.3. Популяции, отобранные для дендрохронологического анализа и изучения механизмов адаптации к условиям засухи

Для 150 образцов *L. sibirica*, отобранных для изучения механизмов адаптации к условиям засухи, в результате секвенирования было получено 1,8 миллиарда прочтений. После первоначальной обработки и фильтрации, четырнадцать деревьев были исключены из-за плохого качества секвенирования, а для анализа было отобрано около 1,4 миллиарда прочтений длиной от 32 до 92 п.н.о., в среднем по 10 млн $\pm 0,4$ млн прочтений на образец. В среднем 97% прочтений для каждого образца были успешно сопоставлены с референсным геномом *L. sibirica*. Всего было отобрано 11 550 SNP, которые соответствовали критериям фильтрации, описанным в главе 2.5. Для пар SNP с неравновесным сцеплением (linkage disequilibrium, $LD\ r^2 > 0,8$) для дальнейшего анализа был отобран только один SNP. Окончательный набор данных содержал 9742 биаллельных SNP, генотипированных в 136 деревьях *L. sibirica*.

3.2. Структура популяции, основные популяционно-генетические статистики

Анализ генетической структуры популяций *L. sibirica* проводился на всех 37 задействованных в исследовании выборках, охватывающих различные климатические зоны из отдаленных частей ареала с контрастными экологическими условиями. Полногеномное генотипирование 488 деревьев по большому количеству генетических маркеров дает возможность выявить закономерности дифференциации, уровень генетического полиморфизма и

проанализировать пространственную генетическую структуру, включая выявление изолированных групп и поток генов между ними. Для изучения структуры популяции был использован набор однонуклеотидных полиморфизмов, общий для всех 488 образцов *L. sibirica* и включающий в себя 2938 SNP-маркеров.

3.2.1. Популяционно-генетические статистики

Был проведен расчет основных популяционно-генетических статистик, таких как частота наиболее часто встречающегося аллеля (P), доля гетерозиготных особей в данной популяции (H_O), гетерозиготность ожидаемая при равновесии Харди-Вайнберга (H_E), оценка нуклеотидного разнообразия (P_i), коэффициент инбридинга особи относительно субпопуляции (F_{IS}) (Таблица 4).

Таблица 4 Значения параметров генетической изменчивости $\pm$ SE по 2938 SNP для каждой из 37 точек сбора лиственницы сибирской.

Pop ID	P	SE	H _O	SE	H _E	SE	P _i	SE	F _{IS}	SE
SEV	0,856	0,003	0,214	0,004	0,198	0,003	0,205	0,004	-0,012	0,026
SIB	0,860	0,003	0,207	0,004	0,191	0,004	0,198	0,004	-0,018	0,024
IVD	0,858	0,003	0,210	0,004	0,196	0,003	0,203	0,004	-0,014	0,030
LAB	0,824	0,003	0,243	0,004	0,248	0,003	0,258	0,003	0,042	0,023
NBSK	0,848	0,002	0,236	0,004	0,225	0,003	0,233	0,003	0,003	0,025
NU	0,854	0,002	0,234	0,003	0,221	0,003	0,229	0,003	-0,001	0,027
D_h_500	0,844	0,002	0,254	0,004	0,230	0,003	0,243	0,003	-0,017	0,015
D_h_1500	0,840	0,003	0,242	0,004	0,231	0,003	0,245	0,003	0,011	0,016
D_h_1000	0,848	0,003	0,247	0,004	0,222	0,003	0,237	0,003	-0,019	0,012
D_h_2000	0,839	0,002	0,266	0,004	0,237	0,003	0,250	0,003	-0,033	0,016
E_h_1500	0,838	0,003	0,247	0,004	0,234	0,003	0,249	0,003	0,015	0,013
E_h_2000	0,846	0,003	0,218	0,004	0,222	0,003	0,238	0,003	0,057	0,017
E_h_1000	0,840	0,002	0,213	0,004	0,233	0,003	0,247	0,003	0,112	0,018
F_h_2000	0,840	0,002	0,286	0,004	0,233	0,003	0,247	0,003	-0,087	0,019
F_h_500	0,837	0,002	0,297	0,004	0,239	0,003	0,255	0,003	-0,089	0,017
F_h_1500	0,852	0,002	0,248	0,004	0,220	0,003	0,234	0,003	-0,026	0,021
F_h_1000	0,838	0,002	0,263	0,004	0,236	0,003	0,250	0,003	-0,028	0,017
G_h_1500	0,843	0,002	0,228	0,004	0,232	0,003	0,245	0,003	0,066	0,016
G_h_1000	0,838	0,002	0,227	0,004	0,236	0,003	0,253	0,003	0,073	0,015
G_h_500	0,837	0,002	0,248	0,004	0,240	0,003	0,254	0,003	0,024	0,014
A_h_1500	0,840	0,002	0,232	0,004	0,239	0,003	0,254	0,003	0,075	0,018
A_h_1000	0,848	0,003	0,230	0,004	0,224	0,003	0,242	0,003	0,035	0,015
A_h_2000	0,846	0,002	0,239	0,004	0,228	0,003	0,245	0,003	0,023	0,012
TUI	0,842	0,002	0,238	0,003	0,243	0,002	0,248	0,002	0,049	0,030

Pop ID	P	SE	H _O	SE	H _E	SE	P _i	SE	F _{IS}	SE
A_h_500	0,854	0,003	0,167	0,005	0,218	0,004	0,235	0,005	0,196	0,025
SON	0,872	0,002	0,160	0,003	0,198	0,003	0,203	0,003	0,187	0,061
KAM	0,865	0,002	0,184	0,003	0,208	0,003	0,213	0,003	0,129	0,061
BOG	0,837	0,002	0,258	0,003	0,248	0,002	0,254	0,003	0,005	0,024
BID	0,838	0,002	0,244	0,003	0,248	0,002	0,253	0,003	0,042	0,025
KRAS	0,851	0,002	0,239	0,003	0,228	0,003	0,236	0,003	0,005	0,026
C_h_1500	0,842	0,002	0,248	0,004	0,234	0,003	0,250	0,003	0,011	0,023
C_h_1000	0,852	0,003	0,152	0,003	0,217	0,003	0,233	0,004	0,225	0,022
TUR	0,846	0,002	0,230	0,003	0,234	0,003	0,242	0,003	0,042	0,026
C_h_500	0,847	0,002	0,202	0,004	0,225	0,003	0,242	0,003	0,108	0,018
K_h_1000	0,840	0,002	0,240	0,004	0,236	0,003	0,249	0,003	0,043	0,017
K_h_1500	0,837	0,002	0,238	0,004	0,243	0,003	0,257	0,003	0,068	0,013
K_h_2000	0,834	0,003	0,206	0,004	0,241	0,003	0,258	0,003	0,134	0,022

Примечание: Точки сбора приведены в географическом порядке, с запада на восток.

3.2.2. Иерархический анализ молекулярной дисперсии (AMOVA), индексы фиксации (F_{ST})

Иерархический анализ молекулярной дисперсии (AMOVA) применялся для оценки генетической дифференциации популяций и распределения генетического разнообразия на разных уровнях структуры: внутри популяций, между популяциями и между группами популяций). В рамках AMOVA проводилось разделение общей генетической изменчивости на эти структурные иерархические компоненты, для каждого из них рассчитывались соответствующие индексы F_{ST} . Образцы разделены на три группы так, что выборки западной (уральской, WT) трансекты сформировали первую группу, выборки Республики Алтай сформировали вторую, а все остальные – третью (Таблица 5).

Таблица 5 Результаты AMOVA

Источник изменчивости	Сумма квадратов	Дисперсионная составляющая	Доля в общей изменчивости, %	<i>F-index</i>
Между группами	21701,197	45,969	15,015	F_{ST} : 0,178 F_{SC} : 0,033 F_{CT} : 0,150
Между популяциями внутри групп	10465,567	8,495	2,775	
Внутри популяций	207000,959	251,695	82,211	
Сумма	239167,723	306,159		

Индексы фиксации Райта оценивали размеры эффектов по доле изменчивости между группами ($F_{CT} = 0,150$), между популяциями внутри групп ($F_{SC} = 0,033$), и между всеми популяциями ($F_{ST} = 0,178$), все из которых были достоверными и имели уровень значимости $p \ll 0,001$.

Также оценку степени генетической дифференциации по параметру F_{ST} проводили попарно между всеми 37 выборками, среднее значение $F_{ST} = 0,284$, с минимальным значением 0,003 между выборками SON и KAM, и максимальным 0,686 между IVD и NBSK; все полученные значения статистически значимы ($p \ll 0,01$) (Рисунок 5).

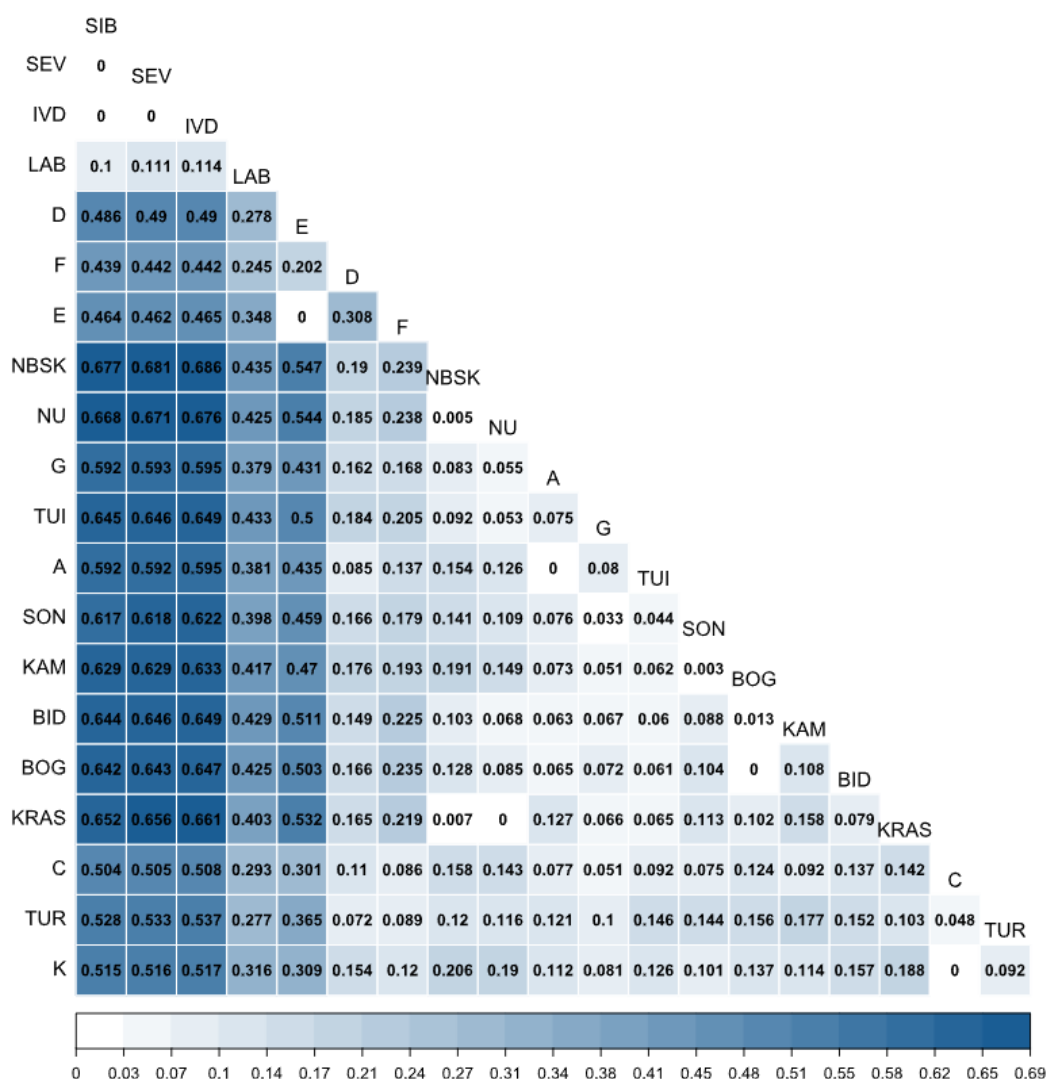


Рисунок 5. Попарные индексы фиксации F_{ST} , рассчитанные между 37 выборками на основе всех 2938 SNP.

3.2.3. Изоляция расстоянием (IBD)

Географически ограниченное распространение может формировать генетическую структуру популяции и приводить к корреляции между генетической и географической дистанцией, называемой изоляцией по расстоянию (Isolation by distance, IBD). Для обнаружения такой корреляции был проведен тест Мантелы с использованием всех 2938 SNP, и координат 37 точек сбора образцов *L. sibirica*. Его результаты выявили умеренную положительную и высокодостоверную линейную связь ($r = 0,4302$, $p\text{-value} = 0,003$) между географическим и генетическим расстоянием (**Рисунок 6**).

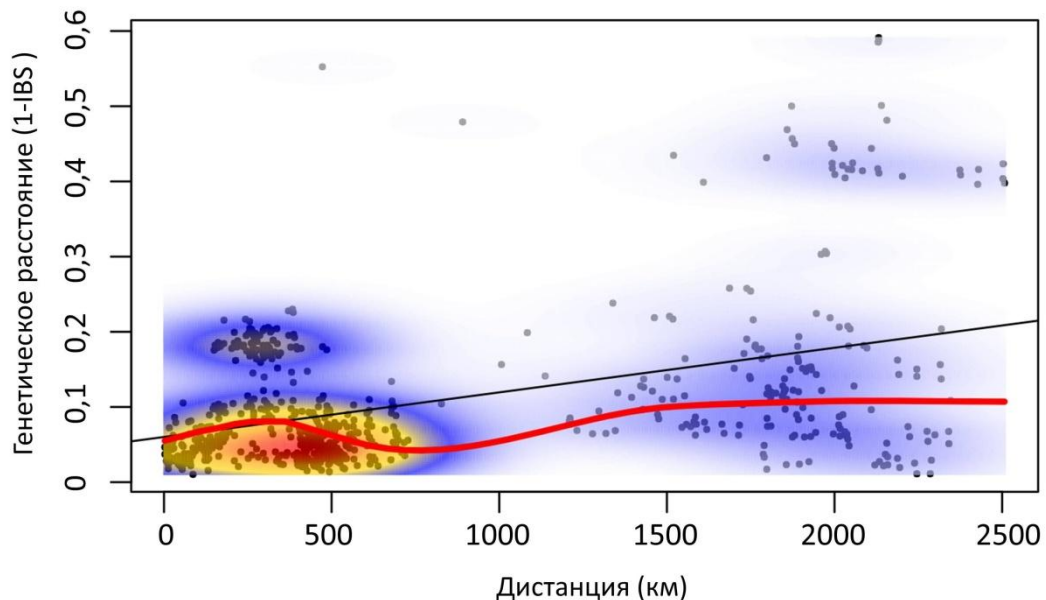


Рисунок 6 Результаты теста Мантелы, демонстрирующие линейную корреляцию между генетической и географической дистанцией в изученных популяциях *L. sibirica* при анализе 2938 SNP-маркеров ($r = 0,4302$, $p\text{-value} = 0,003$). Цветовая шкала отражает градиент плотности корреляционных значений (красный – максимальная плотность, синий – минимальная), где красная кривая соответствует сглаженному локальному среднему, а черная линия - линейной регрессионной зависимости.

Также тест Мантелы был проведен отдельно для образцов, отобранных для изучения адаптации к условиям высокогорья ($r = 0,206$, $p\text{-value} = 0,001$) и для образцов, отобранных для изучения адаптации к широтному градиенту климатических условий, отдельно для каждой трансекты (**Рисунок 7**).

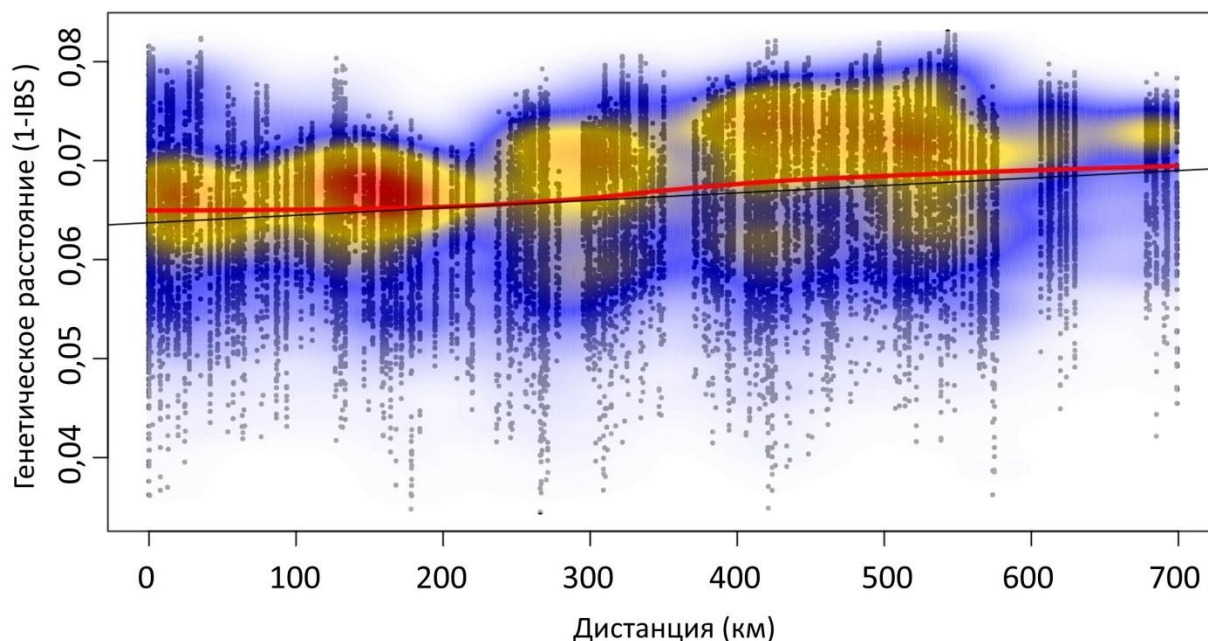


Рисунок 7 Результаты теста Мантела демонстрируют наличие линейной зависимости между генетическими и географическими расстояниями в популяциях *L. sibirica*, отобранных для исследования адаптации к высокогорным условиям (Алтае-Саянский регион), при анализе 25 143 SNP-маркеров. Цветовая шкала отражает градиент плотности корреляционных значений (красный – максимальная плотность, синий – минимальная), где красная кривая соответствует сглаженному локальному среднему, а черная линия – линейной регрессионной зависимости.

Для выявления признаков изоляции популяций по условиям среды (Isolation by Environment, IBE), возникающей под воздействием экологических факторов на генетическую дифференциацию, для выборок из западной (WT) и восточной (ET) трансект был проведен Мантел-тест. Результаты подтвердили статистически значимую линейную корреляцию генетического расстояния с географическим и экологическим, рассчитанным на основе биоклиматических переменных (**Таблица 6**).

Таблица 6 Коэффициенты корреляции (r) Мантел-теста для выборок западной и восточной трансект

Группа	Коэфф.корреляции r между генетической и географической м.	Коэфф. корреляции r между генетической матрицей и матрицей расстояний клим. параметров	Коэфф. корреляции r между географической матрицей и матрицей расстояний клим. параметров
Все образцы	0,469	0,227	0,695
Трансекта ET	0,078	0,071*	0,947
Трансекта WT	0,397	0,386	0,905

Примечание: для всех r p -value = 0,001, * p -value = 0,01

3.2.4. Анализ главных компонент (PCA)

Метод анализа главных компонент (PCA, Principal component analysis) помогает выявить закономерности в распределении генетических вариаций между популяциями и определить, насколько они дифференцированы, так как дает возможность редуцировать многомерные генетические данные до небольшого числа основных компонент, пригодных для визуализации. Это помогает обнаружить скрытые структуры в данных, такие как наличие подгрупп внутри популяций, которые могут быть связаны с географической изоляцией, экологическими условиями или историческими событиями. В генетических исследованиях, особенно в GWAS (Genome-Wide Association Studies), PCA используется для учета стратификации популяций, чтобы избежать ложных ассоциаций между генетическими маркерами и признаками.

Анализ главных компонент был проведен с использованием набора 2938 SNP для 488 деревьев из 37 точек сбора образцов лиственницы сибирской (**Рисунок 8**). В результате проведенного анализа было выявлено, что выборки западной, уральской трансекты (SIB, SEV, IVD и LAB) образуют отдельный кластер в плоскости первой компоненты, образцы Алтайских выборок (трансекты D, E, F) в плоскости третьей компоненты визуально отделяются от остальных выборок.

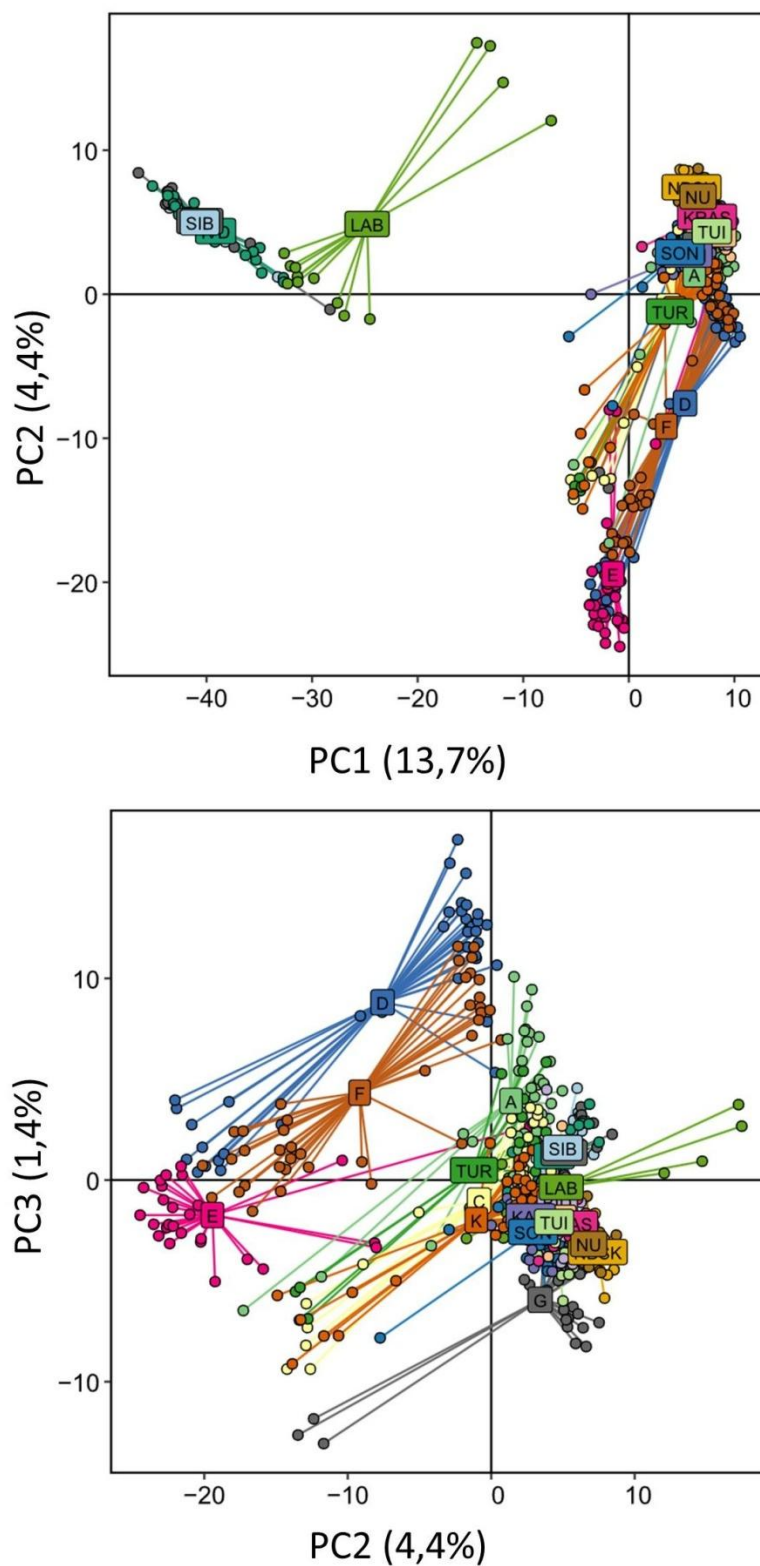


Рисунок 8 Результаты анализа главных компонент, цветом обозначены точки сбора, центроиды выборок подписаны. 24 выборки по высотной поясности укрупнены до 7 трансект.

PCA и DAPC (дискриминантный анализ главных компонент) также были выполнены отдельно для популяций, отобранных в рамках трех

независимых частей исследования. На данных, отобранных для изучения адаптации к высокогорью (231 дерева) было выявлено, что исследуемые выборки не образуют ярко выраженных однозначных групп, однако можно обозначить наличие нескольких условных кластеров: выборки из трансект А, С и К (Тыва) образуют один кластер, выборки трансект D, E и F (Алтай) образуют второй (отделяются как и при добавлении остальных образцов, на наборе SNP-маркеров для 488 деревьев), трансекта G (Хакасия) по результатам DAPC занимает обособленное положение (**Рисунок 9**).

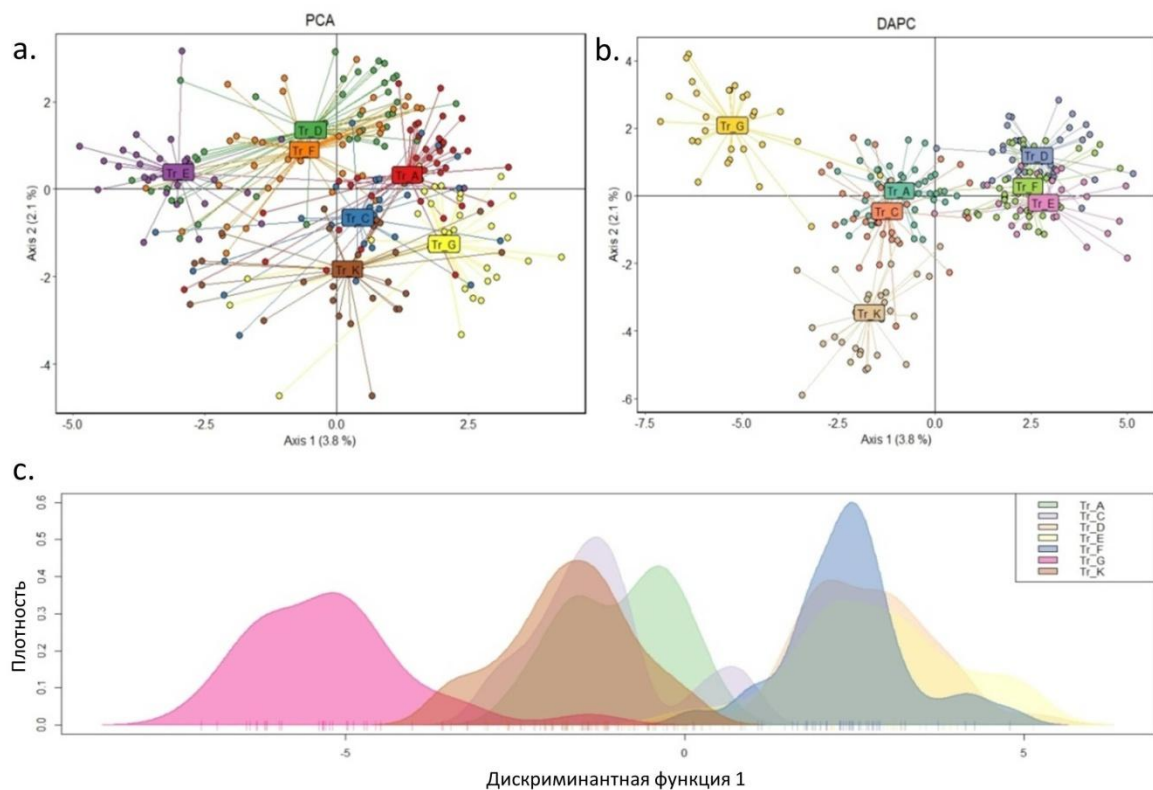


Рисунок 9 Результаты анализа главных компонент (PCA, **a**) и дискриминантного анализа главных компонент (DAPC, **b**), а также график плотности выборки вдоль первой дискриминантной функции (**c**) для выборок, отобранных для изучения адаптации к высокогорью (231 дерева)

Популяции, отобранные в рамках изучения адаптации к широтному градиенту климатических условий дифференцируются лучше. В результате проведенного анализа было выявлено, что исследуемые выборки образуют 3 четких кластера: четыре выборки, принадлежащие восточной (сибирской) трансекте ET формируют один кластер, выборки западной (уральской) SEV,

IVD и LAB – второй, отдельное положение занимает выборка SIB (Башкортостан, **Рисунок 10**).

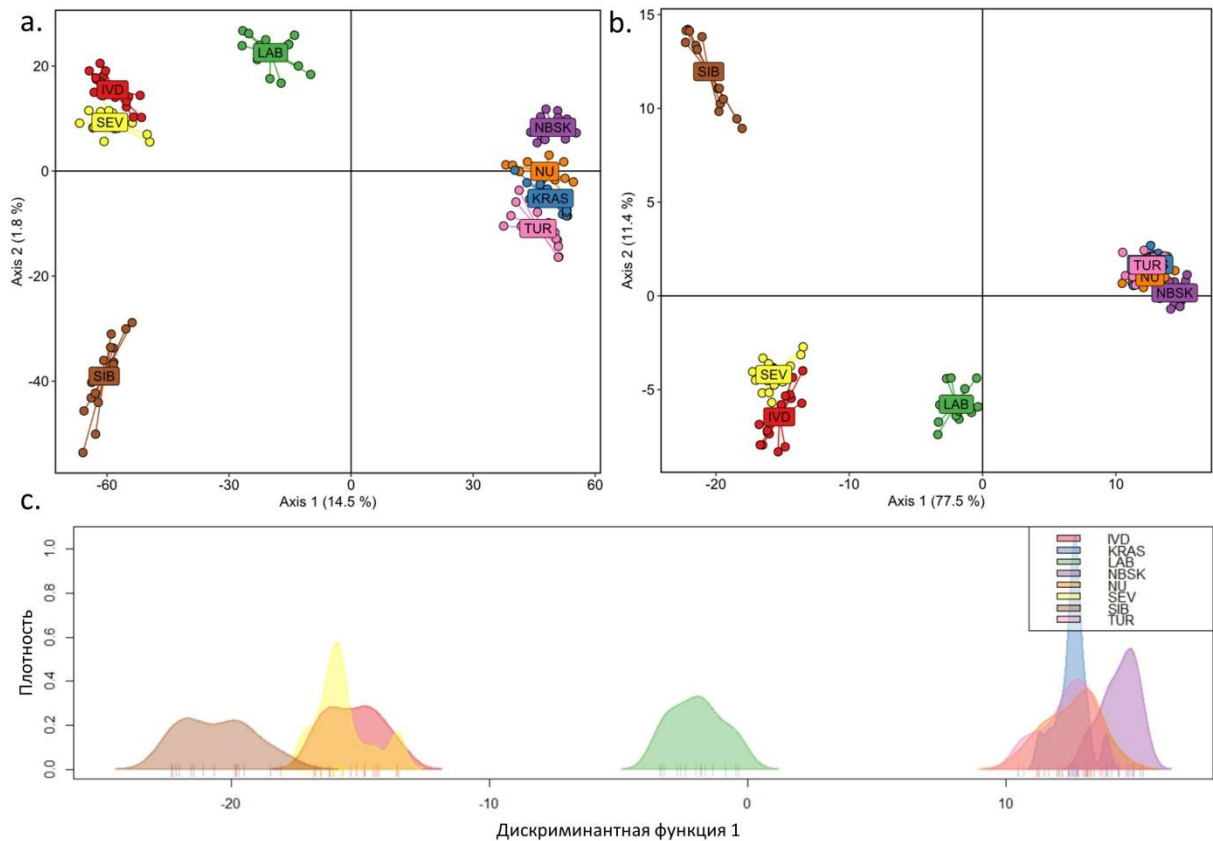


Рисунок 10. Результаты анализа главных компонент (**a**) и дискриминантного анализа главных компонент (**b**), график плотности образцов вдоль первой дискриминантной функции (**c**) для 125 образцов, отобранных для изучения адаптации к широтному градиенту климатических условий

Анализ 9 742 SNP-маркеров у 136 деревьев, отобранных для дендрохронологических исследований и изучения адаптационных механизмов к засухе, выявил отсутствие четко дифференцированных кластеров. Однако предварительно можно выделить два потенциальных кластера: первый объединяет образцы из выборок BID, BOG и TUI, тогда как второй формируют образцы SON и KAM (**Рисунок 11**).

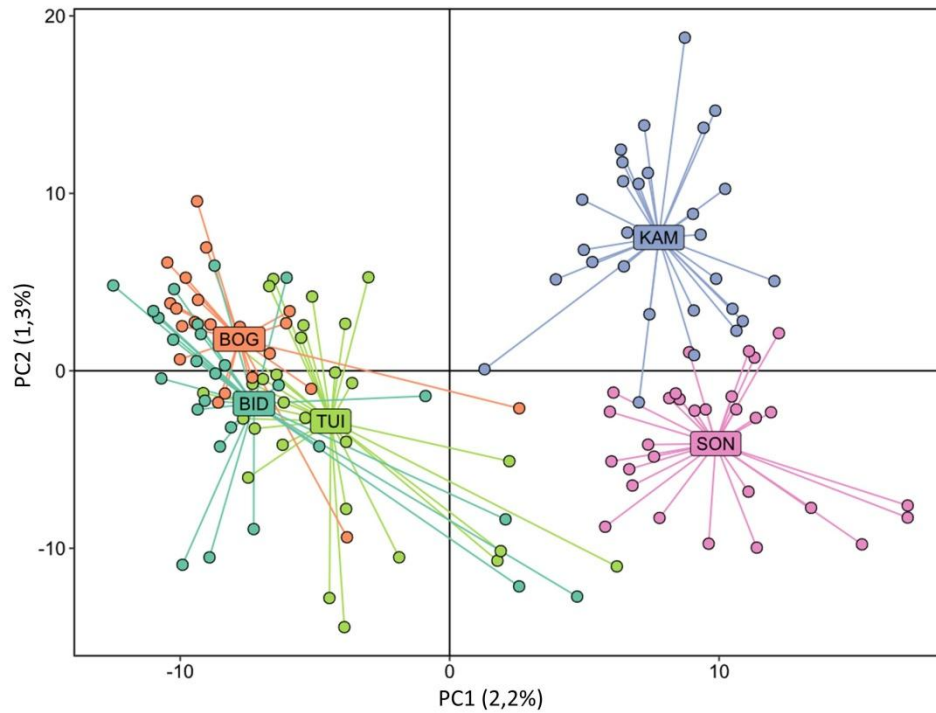


Рисунок 11 Результаты анализа главных компонент 136 деревьев *L. sibirica*, произрастающих в районе Батеневского края, отобранных для изучения механизмов адаптации к условиям засухи

3.2.5. Биogeографический анализ происхождения (Admixture)

Анализ Admixture является важным инструментом в популяционной генетике и экологии и используется для решения ряда задач, связанных с изучением генетической структуры, истории и адаптации видов. Admixture помогает выявить генетически однородные группы (кластеры) внутри популяций и проанализировать смешанность состава (анг. Admixture – примесь) набора индивидов на основе данных о генотипах и тем самым строить предположения о происхождении популяции. Это особенно полезно для видов с широким ареалом распространения, где популяции могут быть изолированы географически или экологически. Admixture позволяет определить степень смешения и выявить гибридные особи при наличии гибридизации между близкородственными видами или популяциями. Также анализ помогает восстановить исторические события, такие как миграции, расширение ареалов или изоляция популяций, вызванная изменениями климата или ландшафта. В GWAS-исследованиях Admixture и подобные алгоритмы используются для учета стратификации (подразделенности)

популяций, чтобы избежать ложных ассоциаций между генетическими маркерами и признаками. Это особенно важно в исследованиях, где популяции имеют сложную генетическую историю.

Для понимания популяционной структуры, в качестве одного из этапов GWAS был проведен поиск наиболее вероятного числа кластеров (субпопуляций) K , в диапазоне K от 1 до 25, с 20 повторностями для каждого K , и были получены оценки вклада каждого кластера в индивидуальный генотип деревьев (Q-значения). Значение K , при котором ΔK достигает максимума, считается наиболее вероятным числом генетических кластеров. Метод ΔK показал, что более вероятными значениями являются K , равные 2, 3, 4, 5, 8, 12, 15 (**Рисунок 12**). Вклад каждого генетического кластера в индивидуальные генотипы отдельных деревьев при $K = 3, 5, 15$ согласно Admixture показаны на **Рисунке 13**. При $K=3$ наблюдается разделение выборок на 3 группы: уральскую (выборки SIB, SEV, IVD и LAB), алтайскую (выборки из трансект D, E и F) и сибирскую (выборки Красноярского края, Республик Тыва и Хакасия).

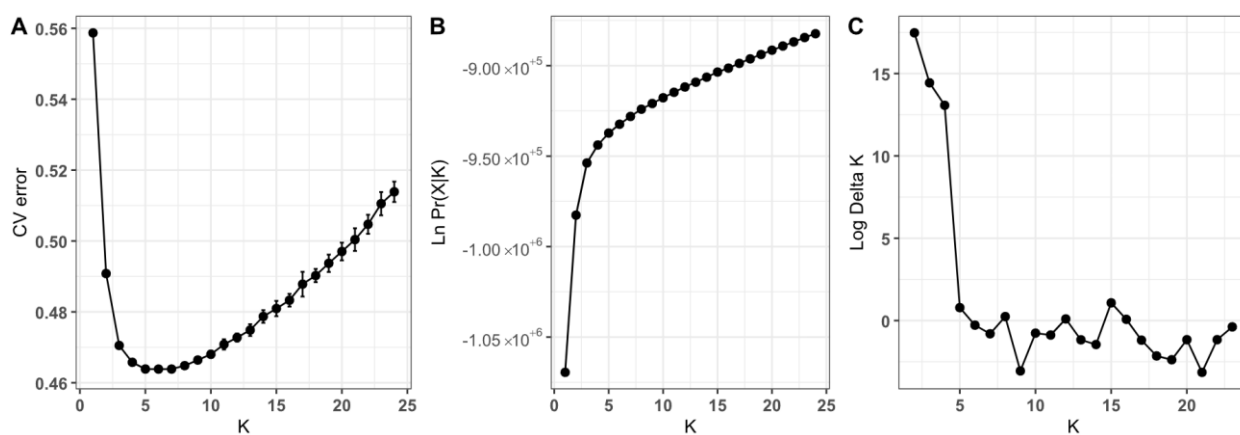


Рисунок 12 Анализ вероятного числа генетических кластеров K . Ошибка предсказания перекрестной проверки (**A**), логарифм правдоподобия (апостериорной вероятности, **B**) и значения логарифма ΔK (**C**) для каждого K на основе всех 2938 SNP для 488 деревьев из 37 точек сбора.

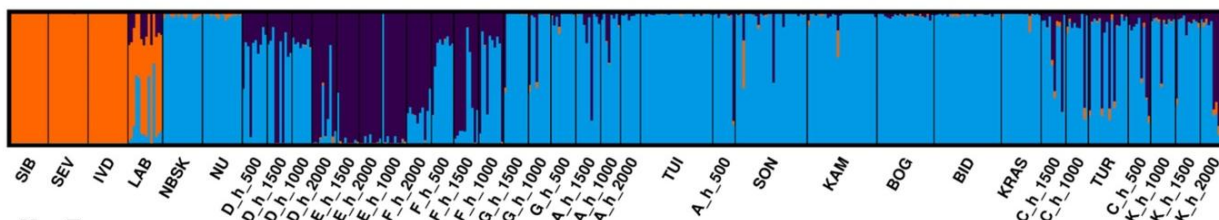
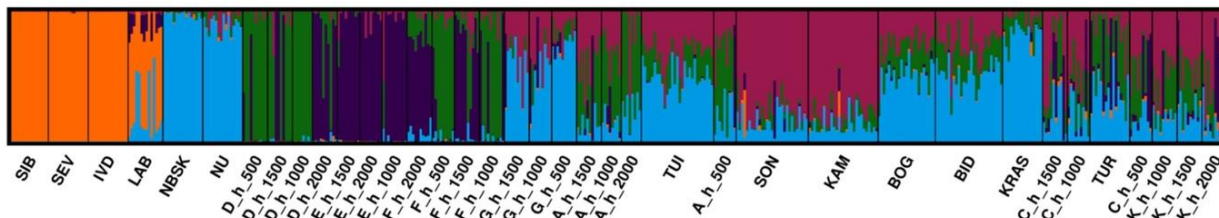
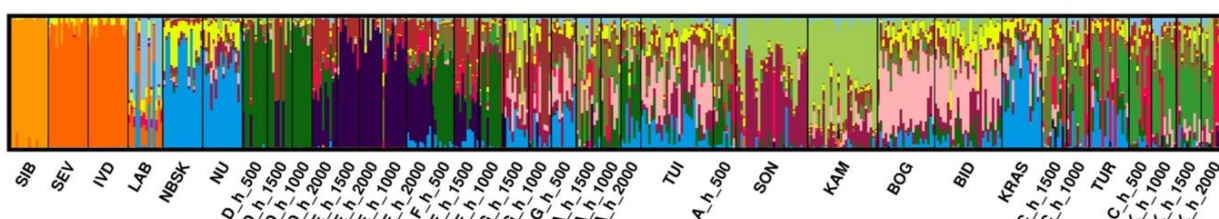
$K = 3$  **$K = 5$**  **$K = 15$** 

Рисунок 13. Вклад («admixture») каждого из генетических кластеров (при $K = 3, 5, 15$) в индивидуальные деревья (Q -значения), обозначенный разными цветами. Образцы и выборки расположены на графике в географическом порядке, с запада на восток.

Дополнительно, если рассмотреть отдельно каждый из наборов данных, использованных в работе, видно, что в популяциях, отобранных для изучения адаптации к высокогорью, деревья, собранные в Горном Алтае также группируются в отдельный кластер и генетически в основном отличаются от деревьев, собранных в других регионах (**Рисунок 14**).

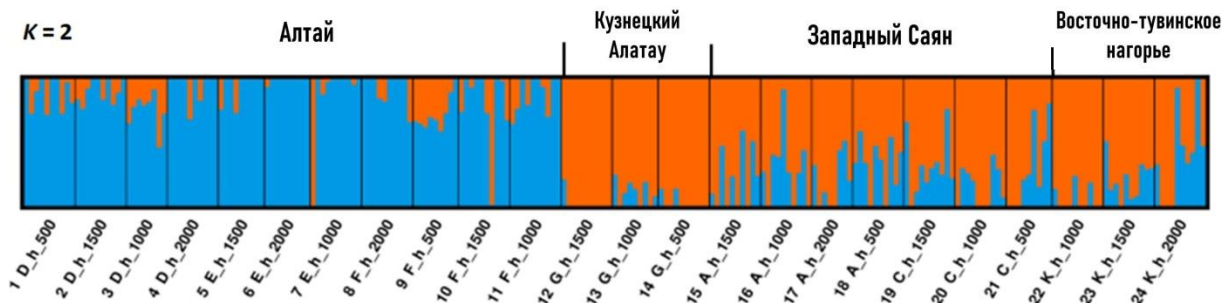


Рисунок 14 Примесь каждого из двух кластеров ($K = 2$), обозначенных двумя разными цветами (значения Q), соответственно, в отдельных деревьях *L. sibirica*, представляющих 24 выборки, собранных на разной высоте в четырех географических регионах и основанных на наборе из нейтральных SNP.

Для 125 деревьев из 8 популяций, отобранных для исследования влияния широтной поясности произрастания, распределение каждого из двух генетических кластеров в индивидуальных генотипах показано на **Рисунке 15**. В целом разделение на группы происходит согласно принадлежности изучаемых образцов к географической трансекте (при $K=2$), однако принадлежащие западной (уральской, WT) трансекте образцы выборки LAB (окрестности г. Лабытнанги, ЯНАО) занимают промежуточное положение (**Рисунок 15**).

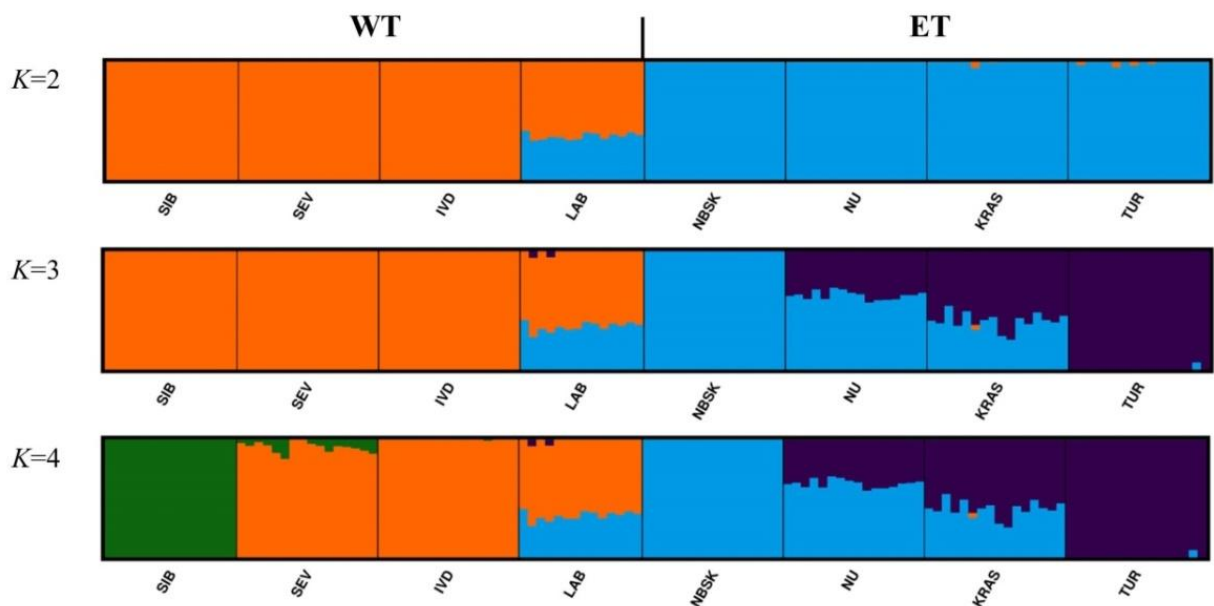


Рисунок 15. Вклад («admixture») каждого из генетических кластеров (при $K = 2, 3$ и 4) в генотип индивидуальных деревьев, обозначенный разными цветами для образцов, отобранных для изучения адаптации к широтному градиенту климатических условий.

Следует отдельно отметить некоторую генетическую обособленность популяции SIB в Республике Башкортостан от остальных популяций *L. sibirica*. Вероятно, это отражает более выраженную дифференциацию популяций Южного Урала по сравнению с более северными, за счет более длительного стабильного существования в этой части Урала, который рассматривается в качестве плейстоценового рефугиума для многих видов деревьев, включая *L. sibirica*. [143–145].

Ранее были получены данные об интрогрессивной гибридизации между *L. sibirica* и *L. sukaczewii* на севере – в Приполярном Урале [146–148]. Наши

данные по SNP-маркерам это подтверждают – популяция LAB (г. Лабытнанги, Ямало-Ненецкий автономный округ) занимает промежуточное положение при кластеризации (**Рисунок 13 и 15**), что соответствует также её географическому расположению на границе между ареалами лиственницы сибирской и Сукачева (или границы между западной и восточной расами *L. sibirica*), где лиственница сибирская заходит с востока на ареал лиственницы Сукачева близко к Полярному Уралу.

Результаты анализа программы Admixture для образцов Батеневского кряжа представлены на **Рисунке 16**. Алгоритм sNMF (sparse nonnegative matrix factorization), применявшийся для вычисления матрицы Q-значений, продемонстрировал схожие результаты. Деревья из выборок SON и KAM формируют обособленную группу и демонстрируют генетические отличия от деревьев, отобранных на остальных участках (TUI, BOG и BID) при количестве генетических кластеров $K=2$.

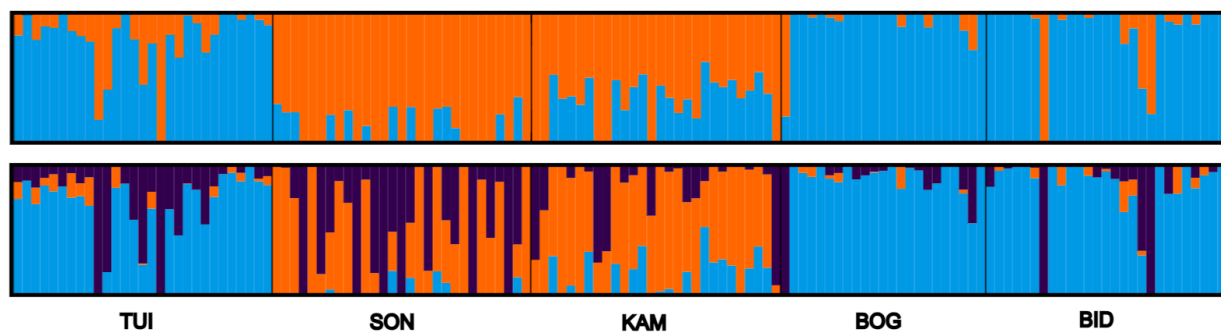


Рисунок 16 Распределение каждого из двух ($K = 2$, сверху) и трех ($K = 3$, внизу) генетических кластеров в генотипах отдельных деревьев *L. sibirica* Батеневского кряжа (TUI, SON, KAM, BOG и BID).

3.3. Адаптация к условиям высокогорья

3.3.1. Поиск SNP-аутлайеров

SNP-маркеры, уровень изменчивости и дифференциации которых не может быть объяснен селективно-нейтральными процессами, скорее всего, находятся под влиянием отбора и вовлечены в локальную биологическую адаптацию. В данной работе были использованы несколько наиболее эффективных популяционно-генетических подходов для поиска таких

кандидатных адаптивных SNP, реализованных в трех программах: PCAdapt, Arlequin и Bayescan.

С помощью программы PCAdapt на 24 выборках, собранных на разных высотах, был выполнен поиск маркёров-кандидатов на основе анализа главных компонент. Для того чтобы выбрать количество K главных компонент, сначала PCA анализ был проведен на большом их количестве, $K = 20$, построен график (**Рисунок 17**) убывания собственных значений. На основании этих данных было выбрано количество K главных компонент 3.

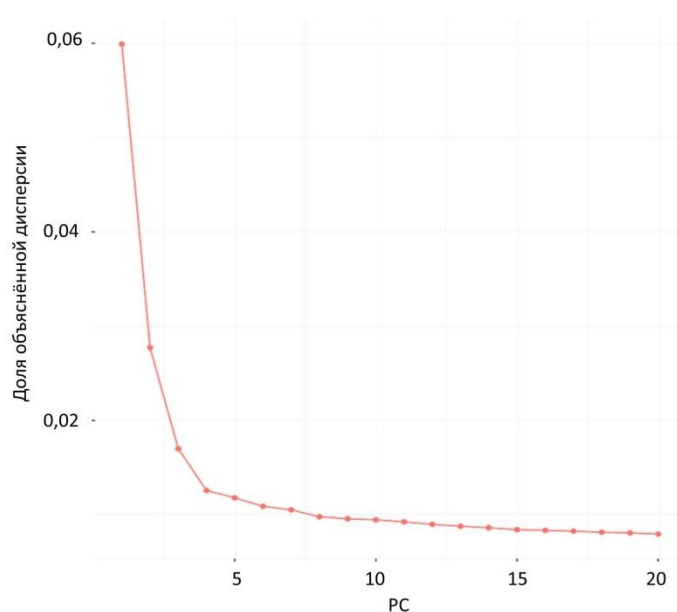


Рисунок 17 Доля дисперсии, объясненной каждой компонентой от 1 до 20 в PCAdapt.

На **Рисунке 18** представлено распределение статистической оценки значимости для каждого SNP-маркера (т.н., Манхэттенский график, англ. Manhattan plot), представленных как отрицательный десятичный логарифм p -значений по оси Y для каждого SNP-маркера по оси X . Изменчивость 212 выделенных синим цветом маркёров достоверно отличалась от ожидаемой для селективно-нейтральных маркёров.

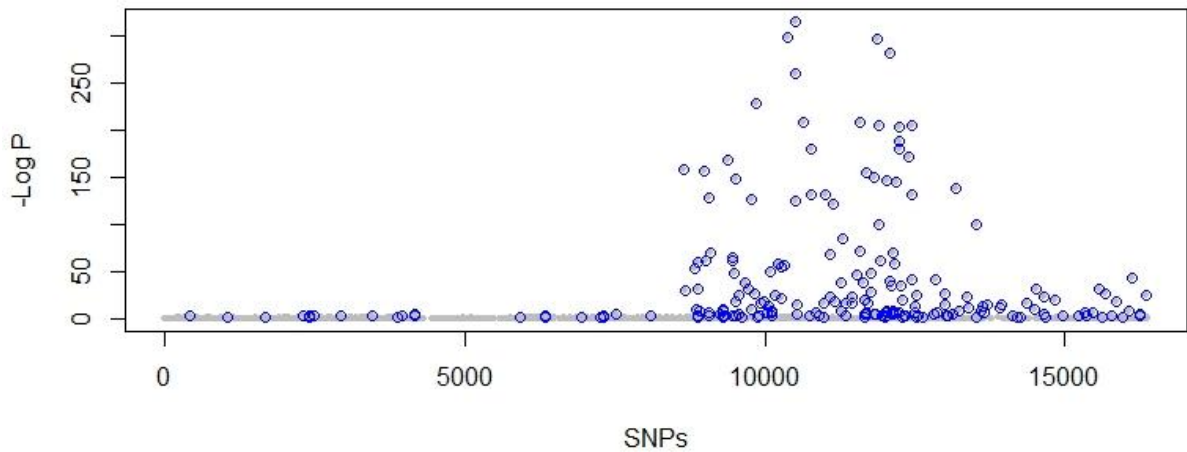


Рисунок 18 Манхеттенский график маркеров на основании их p -значений, полученных в PCAadapt.

Полученный набор из 212 маркеров-аутлайеров, которые сильно дифференцированы между популяциями, затем сравнивали с наборами маркеров, изменчивость которых коррелировала с факторами окружающей среды.

3.3.2. Биоклиматические переменные

Были отобраны биоклиматические переменные, отражающие ключевые аспекты климата, такие как температура, осадки, сезонность и др. Их использование позволяет выявить взаимосвязь между генетическими особенностями популяций и их экологическими нишами. Использование этих биоклиматических данных помогает определить, какие факторы среды оказывают наибольшее влияние на генетическую изменчивость и адаптацию видов.

Чтобы проследить взаимосвязь между 11 климатическими переменными и высотой произрастания *L. sibirica* на 24 участках сбора, был выполнен анализ главных компонент. Первая компонента (PC1) аккумулировала 53,6% вариации, связанной, в первую очередь, с показателями средней температуры самого холодного квартала (MTofCQ) и среднегодовых осадков (Prec) – более высокая температура и низкое количество осадков соответствуют положительным значениям главной компоненты PC1. Вторая компонента (PC2) объяснила 32,2% вариации, большая высота произрастания (Altitude) и более низкая средняя температура

самого теплого квартала (MTofWQ) соответствовали положительным значениям PC2 (Рисунок 19).

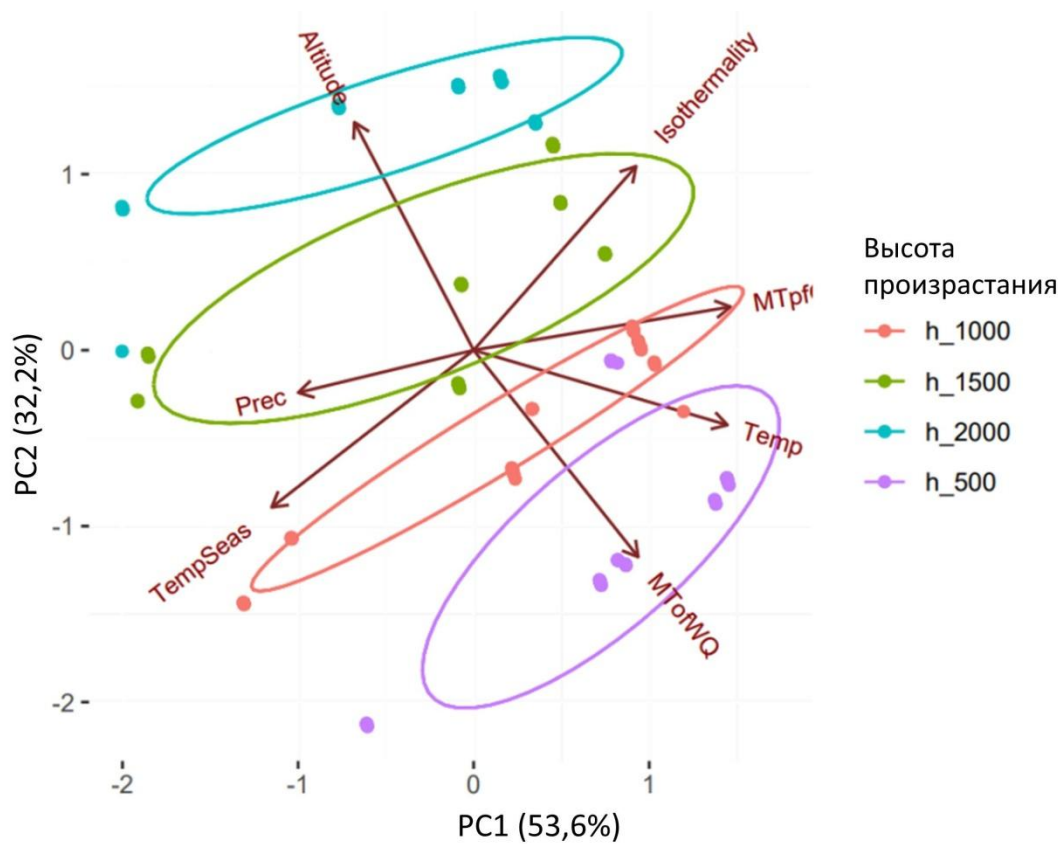


Рисунок 19 Распределение семи средовых факторов в пространстве главных компонент (PCA), примерная высота произрастания обозначена цветом.

Также чтобы проследить взаимосвязь между переменными были рассчитаны коэффициенты корреляции Пирсона, корреляционная матрица была визуализирована (Рисунок 20). Показатель среднегодовой температуры коррелировал с показателями средней температуры самого теплого и самого холодного квартала – с коэффициентами корреляции 0,81 и 0,89 соответственно. Выявлена сильная обратная зависимость между высотой произрастания (Altitude) и средней температурой самого теплого квартала (MTofWQ) ($r = -0,89$); температурной сезонностью (TempSeas) и средней температурой самого холодного квартала (MTofCQ) ($r = -0,88$).

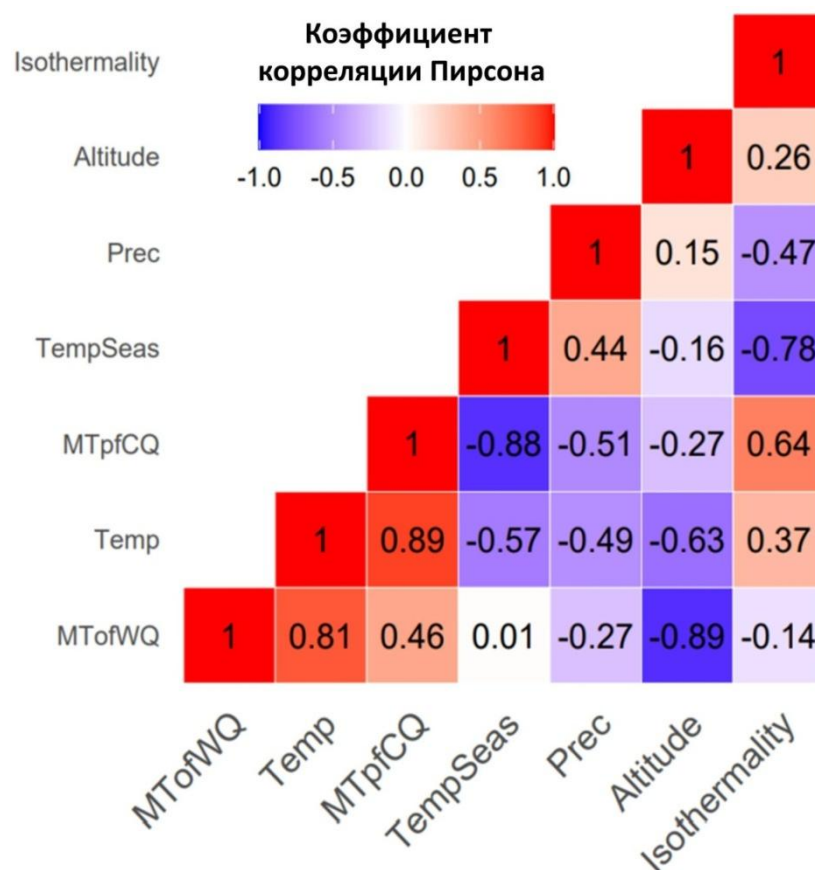


Рисунок 20 Матрица попарных корреляций факторов окружающей среды с указанием коэффициентов корреляции. Обозначения: среднегодовая температура (Temp), изотермальность (Isothermality), температурная сезонность (TempSeas), средняя температура самого теплого квартала (MTofWQ), средняя температура самого холодного квартала (MTpfCQ), годовые осадки (Prec).

Образцы из 24 выборок были разделены на четыре группы, основываясь на показателях высоты произрастания, 500, 1000, 1500 или 2000 метров н.у.м. соответственно. Для каждой группы были посчитаны средние значения климатических переменных (**Таблица 7**). Далее все значения были стандартизированы для анализа с помощью базовой функции R scale, путем вычитания среднего арифметического и деления на стандартное отклонение.

Таблица 7. Средние значения и стандартное отклонение (SD) показателей факторов окружающей среды по четырем высотным группам

Группа	Высота		Temp		Isothermality		TempSeas	
	Среднее	SD	Среднее	SD	Среднее	SD	Среднее	SD
h_500	555,00	68,035	-0,034	0,155	0,239	0,011	133,747	14,908
h_1000	986,31	71,993	-0,229	0,194	0,239	0,012	134,773	17,894
h_1500	1455,32	76,218	-0,402	0,249	0,242	0,012	130,262	15,502
h_2000	2038,46	46,335	-0,520	0,248	0,246	0,005	125,604	16,209

Продолжение таблицы 7

Группа	MTofWQ		MTofCQ		Prec	
	Среднее	SD	Среднее	SD	Среднее	SD
h_500	1,609	0,119	-1,824	0,359	4,554	0,551
h_1000	1,418	0,045	-2,032	0,470	4,941	1,323
h_1500	1,206	0,112	-2,129	0,484	5,328	1,643
h_2000	1,039	0,108	-2,171	0,494	5,032	1,161

3.3.3. Поиск ассоциаций между генотипом и окружающей средой (GEA анализ)

Анализ с использованием программы LFMM2 выявил 40 SNP-маркеров, изменчивость которых коррелировала с составным предиктором PC1, отражающим преимущественно вариации биоклиматических показателей – среднегодового количества осадков (Prec) и средней температуры самого холодного квартала (MTofCQ), а также 49 SNP, ассоциированных с высотой (Alt), включая 3 общих SNP для обоих предикторов (PC1 и Alt). Применение программы BayeScEnv позволило идентифицировать 94 SNP, демонстрирующих связь с предиктором PC1, из которых 41 SNP также показал корреляцию с высотным градиентом (ALT). Связь экологических факторов с первыми тремя осями анализа избыточности (RDA) отображена в **Таблице 8**.

Таблица 8. Корреляция факторов окружающей среды с первыми тремя осями RDA.

Фактор окружающей среды	RDA1	RDA2	RDA3
Alt	-0,015	0,259	0,155
Temp	0,051	-0,284	0,031
Isothermality	0,039	0,032	0,099
TempSeas	-0,102	0,056	-0,079
MTofWQ	-0,014	-0,310	-0,030
MTofCQ	0,087	-0,193	0,061
Prec	-0,003	0,170	-0,113

Собственные значения первых трех осей RDA и проекции SNP на них представлены на **Рисунке 21**. Всего было обнаружено 158 значимых SNP, включая 46, которые были значимы по нескольким осям RDA (двустороннее p -значение = 0,0027). Alt, Prec и Isothermality наиболее тесно связаны с найденными SNP-маркерами (**Рисунок 21**).

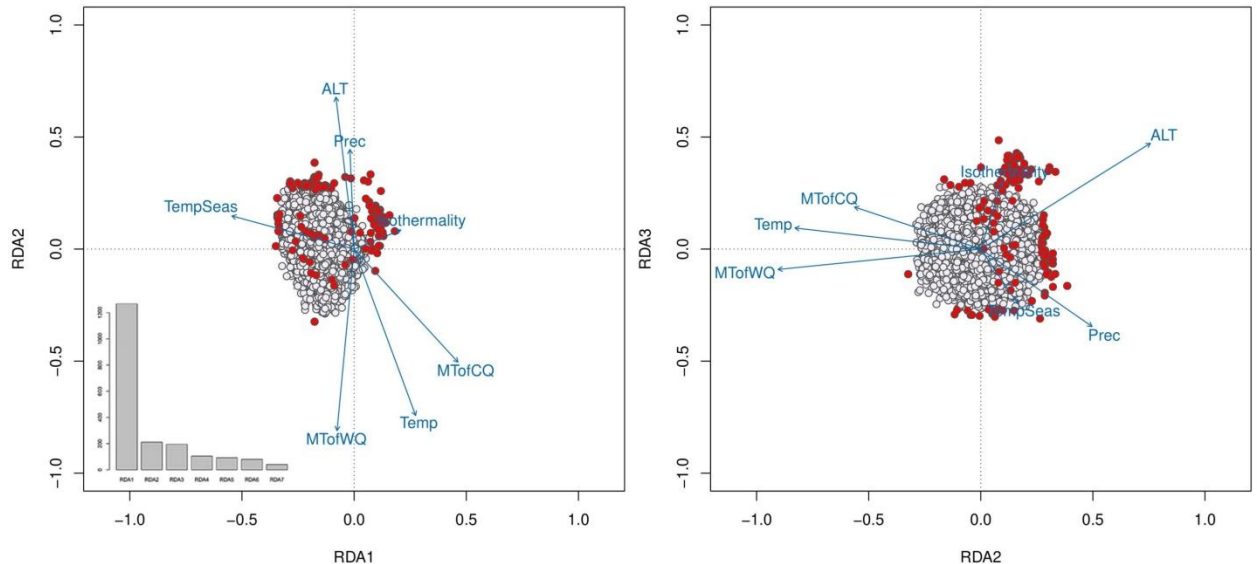


Рисунок 21. Собственные значения первых трех осей RDA и проекции SNP на RDA1 и RDA2 (левый график) и RDA2 и RDA3 (правый график). Значимые SNP выделены красным.

Всего при использовании четырех методов было выявлено 550 уникальных значимых SNP-аутлайеров, в том числе 49, которые коррелировали с факторами окружающей среды (результаты LFMM, BayeScEnv и RDA), и 46 для всех четырех методов GEA, включая PCAdapt (**Рисунок 22**). Среди всех найденных 550 SNP-аутлайеров 67 SNP коррелировали с высотой на основе LFMM или BayeScEnv, а среди 46 SNP из консенсусного списка таких SNP оказалось подавляющее большинство (43).

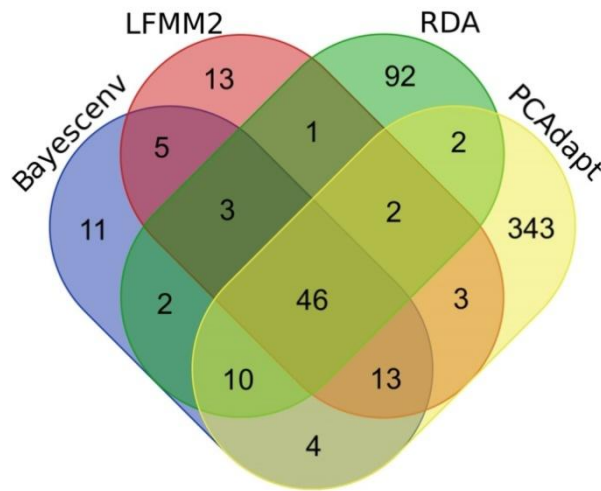


Рисунок 22. Диаграмма Венна, обобщающая результаты поиска значимых SNP-маркеров с использованием LFMM, BayeScEnv, RDA и PCAdapt.

3.3.4. SNP-маркеры, полученные в рамках исследования адаптации к условиям высокогорья

Из 550 значимых SNP-маркеров, 61 располагался в 49 скаффолдах, включающих в себя аннотированные гены: 41 SNP располагался в межгенных областях, в том числе 18 на расстоянии менее 10 тыс. п.н.о. от ближайшего гена, 20 SNP в кодирующих областях генов. Из ассоциированных с высотой произрастания (Alt) 67 SNP, четыре располагались в трех скаффолдах, содержащих аннотированные гены: три из них в межгенных районах, в том числе один SNP на расстоянии менее 10 тысяч п.н.о. от генов, и один SNP, LS.31130.28092 в кодирующей области гена *LS_31130-0.0* (Приложение).

Клеточные компоненты, где локализовано функционирование продуктов этих генов, включают широкий спектр мембранных комплексов (Рисунок 23).

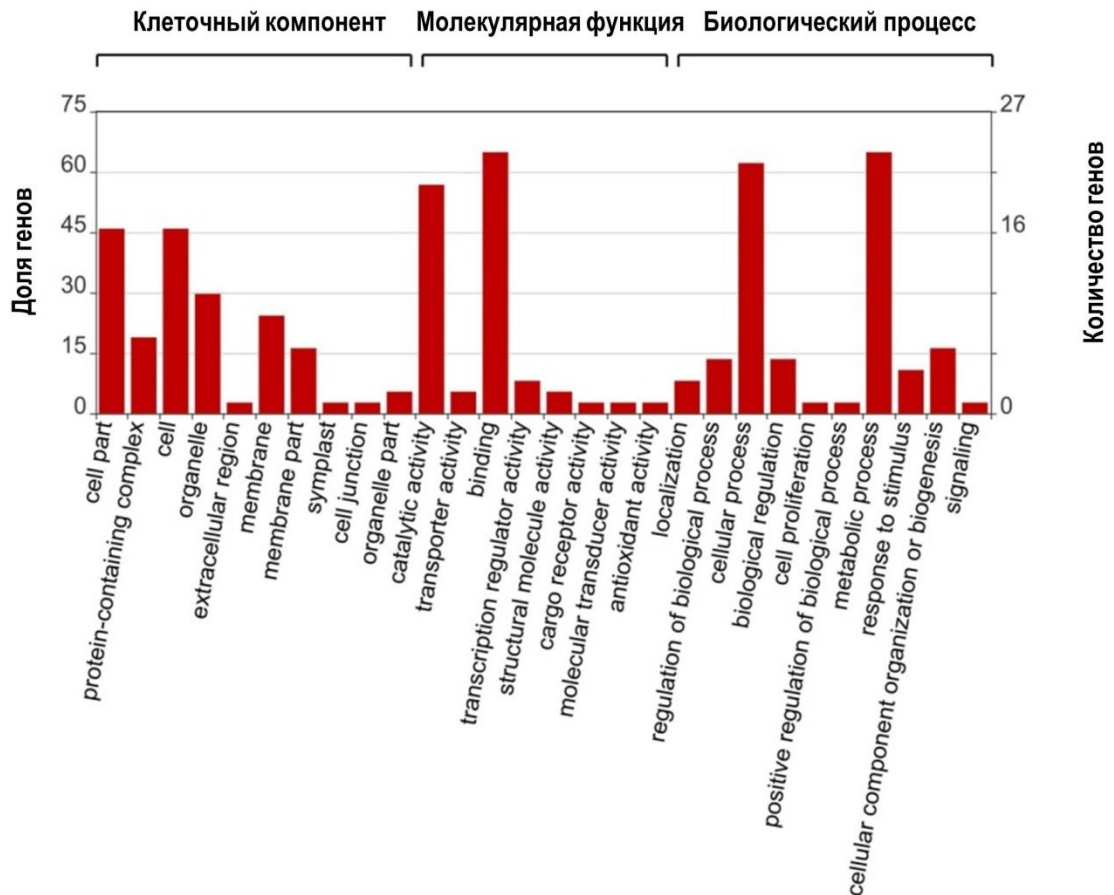


Рисунок 23 GO (Gene Ontology) аннотация генов в скаффолдах, содержащих найденные SNP-маркеры, предположительно участвующих в адаптации к условиям высокогорья.

Продукты генов, содержащих найденные SNP-маркеры, согласно терминам генной аннотации, преимущественно участвуют в ключевых биологических процессах клеточного метаболизма, включая метаболизм макромолекул (белков, углеводов, азотистых оснований), биосинтез органических соединений (ароматических и азотистых веществ, фосфорсодержащих соединений), регуляторные процессы экспрессии генов, ассоциированные с указанными метаболическими путями. **(Приложение).**

Наибольший интерес для последующего изучения представляют 16 несинонимичных однонуклеотидных замен, приводящих к изменениям в кодируемой аминокислотной последовательности, что может в свою очередь повлиять на функционирование соответствующих белковых продуктов.

Поиск BLAST в базе данных NCBI GenBank последовательностей, гомологичных скаффолдам со значимыми SNP, но без аннотированных

генов, выявил большое количество сходных последовательностей, представляющих некоторые участки митохондриального и хлоропластного геномов голосеменных растений. Сравнение данных последовательностей с геномами органелл *L. sibirica* позволило выявить 85 SNP митохондриального происхождения, все из которых локализовались в некодирующих участках митохондриального генома. Часть этих маркеров потенциально может относиться к регуляторным элементам и нуждается в дополнительном углубленном изучении в рамках специального исследования.

Суммарно, 20 из найденных маркеров располагались в экзонных, 41 в межгенных и 489 в негенных регионах (**Приложение**). Из 550 SNP, 67 маркеров ассоциировались с высотой над уровнем моря согласно результатам как минимум одного из двух методов (LFMM и BayeScEnv). Девять маркеров локализовались в шести генах, причем лишь один - несинонимичный SNP LS.31130.28092 – был идентифицирован всеми четырьмя методами как связанный с высотой произрастания. Данный маркер расположен в гене, кодирующем мембранный белок с неидентифицированной функцией.

Краткая характеристика белковых продуктов генов, содержащих значимые SNP в кодирующих областях, позволяет предположить их функциональную роль в адаптации к условиям окружающей среды.

Синонимичный SNP LS.9849.59355 и несинонимичный SNP LS.9849.59395, выявленные как маркеры-аутлайеры методом PCAdapt, не были отобраны тремя методами GEA-анализ (RDA, LFMM и BayeScEnv). Оба маркера расположены в гене *LS_9849-0.1*, кодирующем белок позднего эмбриогенеза (LEA) D-34. Гены LEA экспрессируются в различных органах растений (семенах, проростках, корнях) на всех стадиях развития. При стрессе растения аккумулируют повышенные количества LEA-белков, которым приписывают многофункциональную роль: защита клеточных структур от дегидратации, предотвращение стресс-индуцированного повреждения белков, связывание ионов и рефолдинг денатурированных

белков. Предполагается, что LEA-белки могут выполнять шапероноподобные функции, предотвращая клеточные повреждения [149].

Несинонимичный SNP LS.36255.2146 был ассоциирован с синтетическим предиктором PC1 и локализован в гене *LS_36255-0.2*, кодирующем нуклеоредоксин 1 (NRX1). В растительных клетках оксидоредуктаза NRX1 обеспечивает защиту антиоксидантных ферментов (включая каталазу) от окислительного повреждения, индуцированного активными формами кислорода. Современные исследования [150], свидетельствуют о ключевой роли NRX1 в прямой регуляции детоксикационной способности клеток по отношению к пероксиду водорода, что способствует защите растений от окислительного стресса, вызванного факторами окружающей среды.

Два несинонимичных SNP (LS.73031.9433 и LS.73031.9438), ассоциированных со всеми исследуемыми экологическими предикторами, локализованы в гене *LS_73031-0.1*, кодирующем F-box белок, гомологичный At1g67340. F-box домен представляет собой консервативный мотив длиной около 50 аминокислот, функционирующий как модуль белок-белковых взаимодействий. Первоначально F-box белки были идентифицированы как компоненты SCF убиквитинлигазных комплексов, где они опосредуют связывание субстратов для последующего убиквитин-зависимого протеолиза. Однако в последнее время обнаружено участие этих белков в не-SCF комплексах, выполняющих различные клеточные функции [151].

Несинонимичный SNP LS.118661.524, идентифицированный как аутлайер методом PCAdapt (но не отобранный тремя GEA-методами), расположен в гене *LS_118661-0.1*, кодирующем EXORDIUM-подобный белок 3 (EXO). Данный белок рассматривается как потенциальный посредник брассиностероид-индуцированного роста [152] и необходим для процессов клеточного деления в листовых тканях.

Ген EXO необходим для деления клеток в листьях. Паттерны экспрессии генов указывают на то, что EXO опосредует BR-индуцированный

рост листьев. Считается, что EXO участвует в сигнальном процессе, который координирует реакции BR с сигналами окружающей среды или развития.

Несинонимичный SNP LS.3984510.7337 был обнаружен как аутлайер в PCAdapt, но не был выбран тремя методами GEA. Он расположен в гене *LS_3984510-0.1*, который кодирует консервативную субъединицу 1-подобного олигомерного комплекса Гольджи (COG1). COG1 поддерживает правильную структуру и функцию комплекса Гольджи во время ретроградного транспорта везикул. У резуховидки Таля (*Arabidopsis thaliana*) комплекс COG1 функционирует при росте клеток, размножении и других процессах, включая прямое взаимодействие с компонентами системы секреции. Недавние эксперименты выявили защитную роль комплекса COG1 у растений, в том числе при взаимодействии растений с широким спектром патогенных организмов [153].

Несинонимичный SNP LS.4015301.864 был связан со всеми предикторами окружающей среды и расположен в гене *LS_4015301-0.0*, который кодирует изоформу X2 ацил-КоА-редуктазы жирных кислот 4 (FAR4). Данный белок катализирует процесс восстановления насыщенных (но не ненасыщенных) C16 и C18 ацил-КоА производных до соответствующих жирных спиртов. В современных исследованиях на резуховидке Таля [154] FAR4 (вместе с FAR1 и FAR5) был идентифицирован как ключевой фермент биосинтеза первичных жирных спиртов, входящих в состав суберина - защитного биополиэфира, образованного феруловой кислотой, глицерином и алифатическими компонентами.

Несинонимичный SNP LS.4023983.5884, определенный как аутлайер в PCAdapt (но не выявленный тремя GEA-методами), локализован в гене *LS_4023983-0.1*, кодирующем адаптерный белок AP2A1 (субъединицу 1 AP2 комплекса). AP2 комплекс играет центральную роль в клатрин-опосредованном эндоцитозе, обеспечивая взаимодействие между транспортными белками, мембранными липидами и клатрином. Экспериментальные данные демонстрируют участие AP2 у резуховидки Таля

в процессе эндоцитоза рецептора BRI1 (BRASSINOSTEROID INSENSITIVE1) [155] – ключевого компонента сигнального каскада brassиностероидов. Эти фитогормоны обладают выраженной рострегулирующей активностью и участвуют в контроле многочисленных физиологических процессов, включая устойчивость к абиотическим стрессам, регуляцию сроков цветения, фертильность и развитие пыльцы.

Несинонимичный SNP LS.4033175.6319 отмечен как аутлайер PCAdapt, но не был выбран тремя методами GEA. Он расположен в гене *LS_4033175-0.1*, кодирующем целлюлозосинтазоподобный белок E6 (CSL). Представитель подсемейства ферментов, близкородственных целлюлозосинтазам, которые у некоторых видов растений участвуют в биосинтезе целлюлозы и различных полимеров β -гликанов [156].

Синонимичный SNP LS.4078980.3679 отмечен как аутлайер PCAdapt. Он локализован в гене *LS_4078980-0.1*, кодирующем АТФазу, ремоделирующую хроматин ISWI, которая регулирует транскрипцию кодирующих и не кодирующих РНК путем мобилизации нуклеосом и контроля длины линкерной ДНК, разделяющей нуклеосомы [157].

Несинонимичный SNP LS.5133697.5690 отмечен как аутлайер PCAdapt, но не обнаружен методами GEA. Маркер локализован в гене *LS_5133697-0.1*, кодирующем фосфолипазу A1 глицеролипидов PLIP2. Современные исследования демонстрируют, что гиперэкспрессия PLIP2 существенно ингибирует ростовые процессы растений и вызывает аккумуляцию биологически активных форм жасмоната и родственных оксипинов [158]. У *Arabidopsis thaliana* PLIP2 опосредует взаимосвязь между АВА-зависимым ответом на абиотический стресс и оксипиновой сигнальной системой.

Несинонимичный SNP LS.5135911.1098, идентифицированный как аутлайер методом PCAdapt (но не выявленный GEA-методами), расположен в гене *LS.5135911-0.1*, кодирующем растительный внутриклеточный Ras-родственный белок 1 LRR-типа (PIRL1). Белки PIRL структурно отличаются от более крупных и хорошо изученных классов растительных LRR-белков.

Исследования мутантных форм с Т-ДНК вставками подтвердили значимую роль PIRL1 в процессах раннего развития пыльцы [159].

Несинонимичный SNP LS.5188799.2951 отмечен как аутлайер PCAadapt, но не обнаружен методами GEA. Он расположен в гене *LS_5188799-0.1*, который кодирует олигопептидный транспортер 7 (OPT7). Транспорт пептидов включает транслокацию пептидов (длинных остатков 2–6 аминокислот) через клеточную мембрану энергозависимым образом. Идентификация нескольких OPT у *A. thaliana* позволяет предположить, что они могут играть разные функциональные роли [160].

Ранее С. Чжэн и соавт. [161] показали наличие высотной адаптации у популяции тополя тибетского (*Populus szechuanica* var. *tibetica*). Обнаружено два участка генома, один из которых (четыре гена, хромосома 15) связан с высотной изменчивостью, а другой (10 генов, хромосома 6) – с реакцией на количество солнечного излучения. Один из генов, найденных в этом исследовании, был ортологичен *At3g47110*, обнаруженному у *A. Thaliana*. Белок LRR, кодируемый этим геном, взаимодействует с FERRIC REDUCTASE DEFECTIVE3 (FRD3), который опосредованно участвует в стабильном развитии микроспор при росте пыльцевых трубок. Другой ген кодирует транскрипционный фактор MADS-box 47, который участвует в формировании органов цветка, частично за счет подавления сигнального пути брассиностероидов. Фосфолипид гидропероксидглутатионпероксидаза 1 (GPX) представляет собой группу белков, которые защищают клетки от окислительного повреждения, вызванного активными формами кислорода (АФК).

3.4. Адаптация к широтному градиенту климатических условий

3.4.1. Поиск SNP-аутлайеров

Также в PCAadapt был проведен поиск SNP-аутлайеров для 125 деревьев из 8 выборок, отобранных для изучения адаптации к широтному градиенту климатических условий. На начальном этапе был выполнен анализ PCA на большом количестве главных компонент ($K = 15$) и построен график

убывания значений дисперсии, объяснённой каждой компонентой от 1 до 15, чтобы выбрать наиболее значимые компоненты. На основании этих данных для обеих трансе́кт было выбрано три главных компоненты. Для восточной трансе́кты (ET) было найдено 168 SNP-аутлайеров, а для западной трансе́кты (WT) 551 SNP-аутлайер. Из них 38 SNP-аутлайеров были общими для западной и восточной трансе́кт.

С помощью программы Arlequin для восточной трансе́кты (ET) был найден 1971 SNP-аутлайер, а для западной трансе́кты (WT) – 3885 SNP-аутлайер. Из них 263 маркера-аутлайнера были общими для обеих трансе́кт.

Аналогично, с помощью программы Bayescan для восточной трансе́кты (ET) было найдено 104 SNP-аутлайнера, для западной трансе́кты (WT) – 286 аутлайеров. Из них 13 маркеров были общими для обеих трансе́кт.

После сопоставления результатов программ PCAdapt, Arlequin и Bayescan для каждой трансе́кты были сформированы консенсусные наборы маркеров-аутлайеров, обнаруженных всеми тремя программами: 36 высокодостоверных SNP-аутлайеров для трансе́кты ET, и 225 для трансе́кты WT. Сравнение данных наборов показало наличие 9 SNP-аутлайеров общих для обеих трансе́кт.

3.4.2. Биоклиматические переменные

Для деревьев из восьми выборок, принадлежащих к западной и восточной трансе́ктам, из базы данных WordClim были получены данные по 11 биоклиматическим переменным. На основании значений VIF и показателей корреляции были отобраны 5 переменных: температура, осадки, осадки самого влажного месяца, инсоляция и скорость ветра (Temp, Precipitation, PofWM (Precipitation of Wettest Month), Srad, Wind, **Рисунок 24**). Значения VIF для пяти отобранных переменных варьировали от 1,684 до 8,229 и не превышали порогового значения, что указывает на то, что взаимозависимость среди этих предикторов не повлияет на результаты дальнейшего анализа [162].

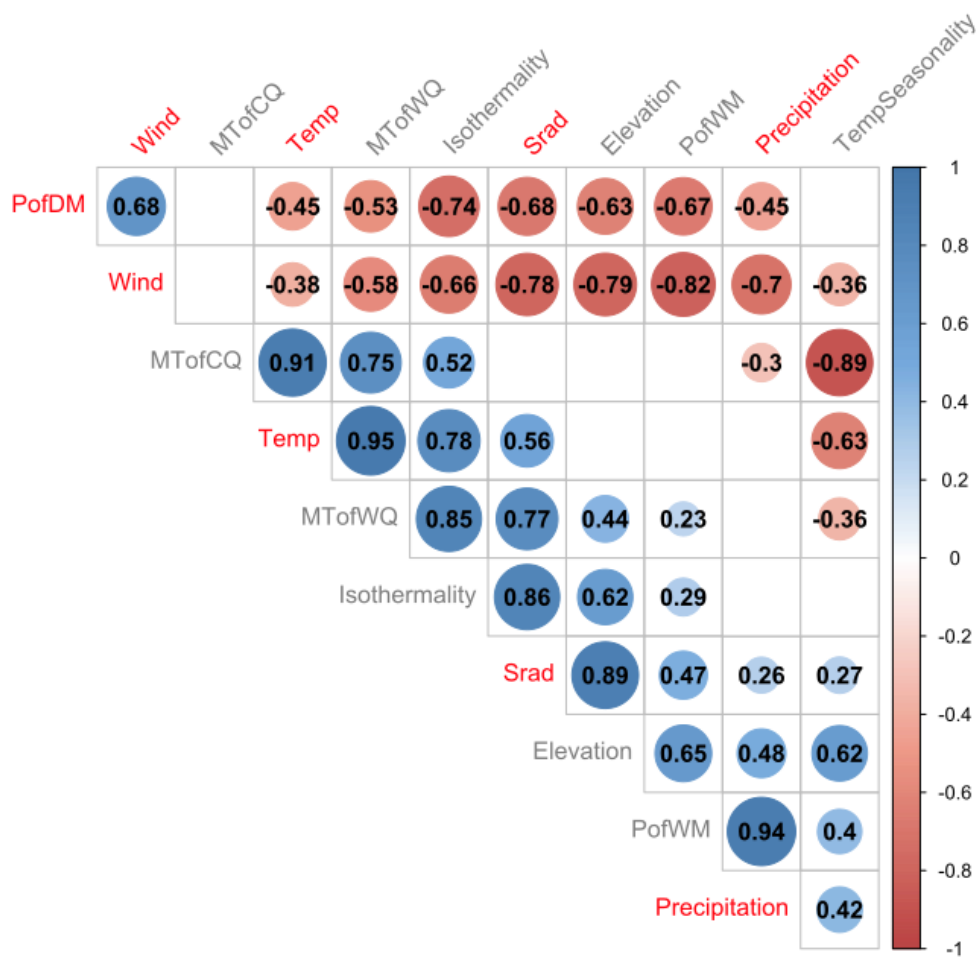


Рисунок 24 Матрица корреляций факторов окружающей среды с указанием коэффициентов корреляции. Красным цветом выделены биоклиматические переменные, отобранные для анализа

После визуализации методом главных компонент взаимосвязей между биоклиматическими переменными в местах сбора образцов была обнаружено, что первая компонента описывает 61,1% вариации, связанной, в первую очередь, с показателями скорости ветра (Wind) и солнечного излучения (Srad) – более высокие показатели Wind и уменьшение Srad соответствуют положительным значениям главной компоненты PC1. Вторая компонента описывает 23,3% вариации, высокие значения годовой температуры (Temp) и низкие показатели влажности (Prec) здесь соответствуют положительным значениям PC2 (**Рисунок 25**). Для каждой выборки также были посчитаны средние значения климатических переменных.

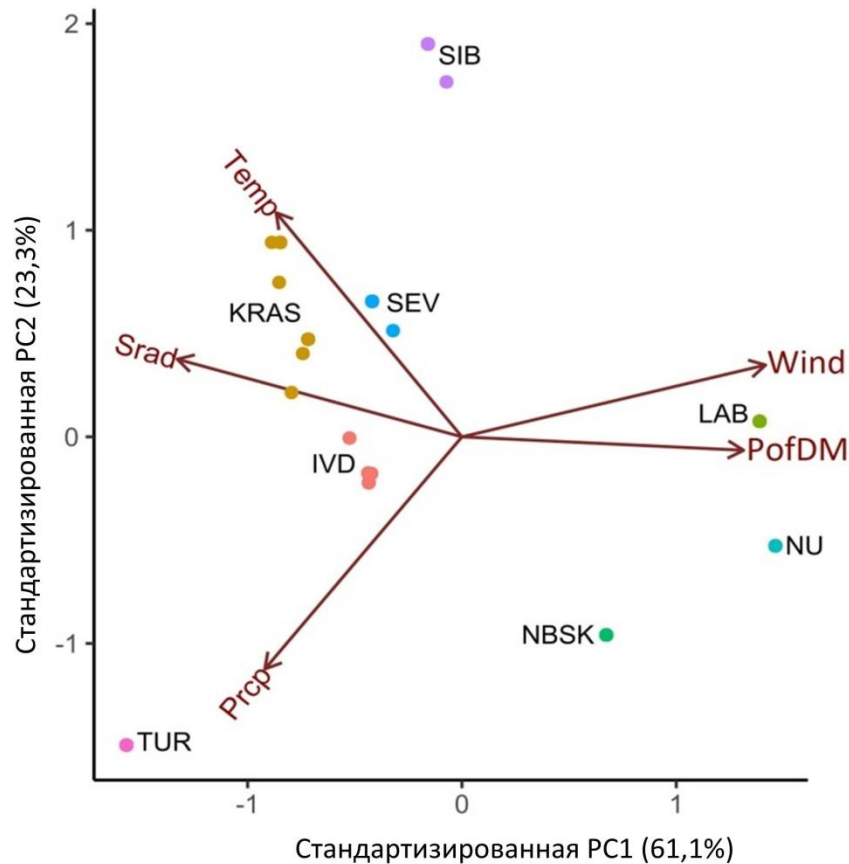


Рисунок 25. Распределение средовых факторов в пространстве главных компонент (PCA) в популяциях западной и восточной трансект *L. sibirica*

3.4.3. Поиск ассоциаций между генотипом и средой (GEA анализ)

Программа LFMM2 выявила для восточной трансекты 2167 маркеров, чья изменчивость коррелировала с изменчивостью одной или нескольких из пяти биоклиматических переменных, включая 357 SNP-маркеров (17%), коррелировавших с изменчивостью всех пяти средовых факторов. Для западной трансекты найдено 3725 таких SNP, из них 2113 (около 57%) оказались общими для всех пяти используемых климатических переменных. Всего 196 проассоциированных SNPs были общими для обеих трансект.

Также с помощью программы BayeScEnv для восточной трансекты было обнаружено 154 SNPs, связанных с изменчивостью одной или нескольких из пяти биоклиматических переменных, из них 32 коррелировали с изменчивостью всех пяти. Для данных западной трансекты было обнаружено 424 таких SNPs, из которых 84 коррелировали с изменчивостью

всех пяти средовых факторов. Всего программой BayeScEnv было выявлено 32 SNPs, общих для обеих трансе́кт.

По результатам RDA анализа было выявлено 887 значимых адаптивных SNPs для восточной трансе́кты и 769 SNPs для западной трансе́кты, 17 из них оказались общими для двух трансе́кт.

Было найдено пять высокодостоверных SNPs для восточной трансе́кты, обнаруженных всеми тремя программами и коррелирующих хотя бы с одним средовым фактором, и аналогично также пять SNPs для западной трансе́кты. Маркеров, обнаруженных всеми тремя методами и общих для обеих трансе́кт, найдено не было. Из маркеров, связанных с переменными окружающей среды в обеих трансе́ктах, два SNPs были найдены как RDA, так и LFMM2 (**Рисунок 26**).

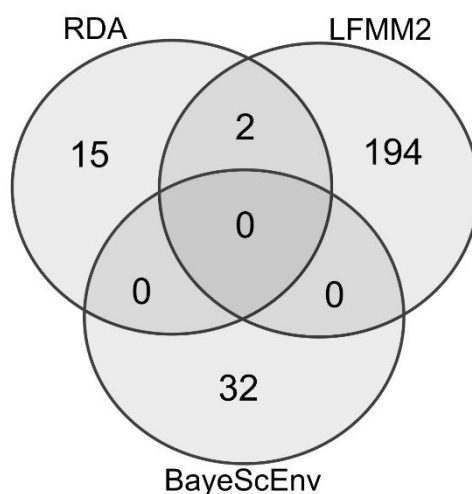


Рисунок 26 Диаграмма Венна результатов поиска значимых маркеров, связанных с биоклиматическими переменными в западной и восточной трансе́ктах методами LFMM2, BayeScEnv и RDA

3.4.4. SNP-маркеры, полученные в рамках исследования адаптации к широтному градиенту климатических условий

Для 9 общих SNP-аутлайеров, а также для 12 SNP, изменчивость которых достоверно коррелировала с изменчивостью средовых факторов, был проведен анализ геномных участков, содержащих эти SNP-маркеры. Для них были извлечены последовательности интервалом ± 10000 п.н.о. Процедура выравнивания этих последовательностей на базы NCBI GenBank

показала наличие в базе гомологичных участков для 18 SNP-маркеров. Пять SNP-маркеров расположены в межгенных областях, а три SNP оказались расположены в скаффолдах, не содержащих гены – для них не найдены гомологичные последовательности в NCBI GenBank. Краткие результаты представлены в **Таблице 9**.

Таблица 9 Результаты аннотации 19 значимых SNPs, связанных с адаптацией к широтному градиенту климатических условий

SNP	Локализация и аннотация (позиция в NCBI GenBank нуклеотидной последовательности)	Коррелирующий средовой фактор в трансектах WT и/или ET
LS.4447298.1210	9375 п.н.о. от начала митохондриального гена, кодирующего рибосомальный белок S11 (2195303 п.н.о., MT797188.1)	аутлайер в WT и ET
LS.4806638.6052	2818 п.н.о. от конца митохондриального гена, кодирующего рибосомальный белок S7 и на расстоянии гены 30124 и 30704 п.н.о. перед генами рибосомальных белков L5 и S14, соответственно (889123 п.н.о., MT797189.1)	
LS.4766550.9636	интрон 1 митохондриального гена NAD1 (1651865 п.н.о., MT797188.1)	
LS.4800638.59265	55579 п.н.о. после митохондриального гена NAD2 (548253 п.н.о., MT797190.1)	
LS.5154937.1260 LS.5119432.22927 LS.5156723.4944 LS.4695581.16666 LS.4400751.1818	интрон 1 митохондриального гена рибосомального белка L2 (позиция в MT797187.1: 806147, 1077727, 1146853, 2293241, 2322428, соответственно)	все аутлайеры в WT и ET, кроме LS.5156723.4944, который аутлайер только в ET, а также его изменчивость коррелировала с Prcp (ET), Srad (ET) и PofDM (ET)
LS.4586690.9126	интрон 4 митохондриального гена NAD2 (1957472 п.н.о., MT797188.1)	аутлайер в WT и ET
LS.28474.5984	11 тысяч п.н.о. перед ядерным геном, кодирующем белок гомологичный связанному с микротрубочками futsch-подобному белку <i>Arabidopsis thaliana</i> (NP_201061.1)	Temp (ET)
LS.4137682.158	в ядерном гене, кодирующем неизвестный белок <i>Picea sitchensis</i> (ABR16307.1), гомологичный транспозону <i>Prunus dulcis</i> (BBG96345.1)	Temp (ET)

SNP	Локализация и аннотация (позиция в NCBI GenBank нуклеотидной последовательности)	Коррелирующий средовой фактор в трансектах WT и/или ET
LS.132659.5239 LS.132659.5292	после ядерного гена, кодирующего неизвестный белок <i>Shorea leprosula</i> (GKV31644.1), гомологичный транспозону <i>Glycine soja</i> (RZC29890.1)	аутлайер в WT; PofDM (WT)
LS.1251904.144	в ядерном гене, кодирующем неизвестный белок <i>Picea sitchensis</i> (QHR90813.1), гомологичный zinc finger domain-containing protein бактерии <i>Stenotrophomonas</i> sp. (MBD3828844.1), хотя встречается и в других хвойных видах, что может быть связано с бактериальным загрязнением или какой-то эндогенной бактерией	аутлайер в WT; Temp (WT)
LS.3903199.5631	после ядерного гена, кодирующего неизвестный белок <i>Shorea leprosula</i> (GKV44695.1), гомологичный ретровирусному полипротеину Pol транспозона TNT 1-94 <i>Cajanus cajan</i> (KYP38095.1)	Temp (WT)
LS.3983248.4557	в участке, гомологичном гену, кодирующему неизвестную mRNA <i>Picea glauca</i> (BT118153.1)	PofDM, Prcp, Srad, Temp (WT и ET), Wind (WT)
LS.2939623.746	в участке, гомологичном гену, кодирующему бактериальный белок retropepsin-like aspartic protease (WP_160144215.1) и неизвестный гипотетический белок <i>Picea sitchensis</i> (QHR91704.1)	Prpcp, Srad, Temp (WT и ET), PofDM (ET)
LS.7403.79464	контиги этих трёх SNP-маркеров не содержат гены, и в NCBI GenBank для них не найдены гомологичные последовательности	PofDM (WT)
LS.3963342.9449		Temp (ET)
LS.4465.9359		Temp (ET)

По результатам выравнивания нуклеотидных последовательностей, содержащих SNP-аутлайеры на митохондриальный геном *L. sibirica* [163] было установлено, что 10 найденных маркеров расположены в межгенном пространстве митохондриального генома: интронах генов NAD1 и NAD2, в интронах генов рибосомальных протеинов L2, L5, S7, S11, S14 рядом с генами COX2, NAD4 и ATP4. Из них пять SNP-маркеров (LS.5154937.1260, LS.5119432.22927, LS.5156723.4944, LS.4695581.16666 и LS.4400751.1818) находятся в интроне 1 митохондриального гена рибосомального белка L2, т.е. сцеплены. Четыре из них были выявлены как аутлайеры в западной и восточной трансектах, но LS.5156723.4944, находящийся на расстоянии 69126 п.н.о. от ближайшего SNP-маркера, показал отличное от четырех других маркеров поведение, что может быть объяснено большим размером интрона и удалённостью от других SNP-маркеров, и идентифицирован как

аутлайер только для восточной трансекты. Его изменчивость коррелировала переменными, отражающими суммарное количество осадков (Prec), количество осадков самого засушливого месяца (PofDM) и уровень солнечного излучения (Srad).

На основании данных геномной аннотации было установлено, что маркер LS.28474.5984, изменчивость которого коррелирует с изменчивостью показателя Temp для восточной трансекты, расположен на расстоянии 11 тысяч п.н.о. от гена, кодирующего связанный с микротрубочками futsch-подобный белок, вовлеченный в регуляцию процесса метаболизма азотистых соединений, а также процессы роста и развития тканей растения [164].

Для последовательностей, содержащих SNP-маркеры LS.4137682.158, LS.132659.5239, LS.132659.5292, LS.1251904.144 и LS.3903199.5631, ассоциированные с климатической переменной Temp для западной трансекты, были найдены гомологичные участки в неохарактеризованных белках *Picea sitchensis* и *Shorea leprosula*.

Для маркеров LS.7403.79464, LS.3963342.9449 и LS.4465.9359, коррелирующих со средовыми факторами, поиск гомологов в базах NCBI GenBank не дал достоверных результатов.

Локализация большинства из обнаруженных адаптивных маркеров в митохондриальном геноме представляет интерес. Известно, что митохондриальные гены играют решающую роль в локальной адаптации растений [165, 166], а найденные маркеры-кандидаты вероятно могут находиться в регуляторных областях и влиять на экспрессию данных генов.

Наличием среди обнаруженных маркеров значительного числа митохондриальных маркеров можно также объяснить большую дифференциацию популяций по этим наборам SNP, в сравнении с ядерными маркерами – аллозимными и микросателлитными локусами, описанными в ранних исследованиях [167]. В значительной мере такая существенная дифференциация митохондриальных маркеров может быть вызвана материнским характером наследования митохондриального генома,

следствием чего является пониженная скорость пространственного перемещения митохондриальных генов, переносимых с семенами, по сравнению с ядерными генами, переносимыми как семенами так и пыльцой.

Анализ митохондриальных SNP-маркеров одновременно и совместно с ядерными маркерами может привести к ошибочному выделению их как аутлайеров, только на основании их повышенной дифференциации по сравнению ядерными SNP-маркерами. В то же время, связь изменчивости митохондриальных SNP-маркеров с климатическими переменными может быть опосредована высокой корреляцией последних с географическими факторами.

Из маркеров, связанных с переменными окружающей среды в обеих трансектах, два SNPs были найдены как RDA, так и LFMM2. Небольшое число общих для двух трансект адаптивных маркеров, ассоциированных с изменчивостью факторов окружающей среды и подтверждённых более чем одним методом может быть вызвано либо относительно небольшой выборкой, недостаточной статистически для обнаружения искомых ассоциаций, или же небольшим вкладом наблюдаемой генетической изменчивости в локальную адаптацию *L. sibirica* в изученном градиенте климатических условий. Необходимо исследовать также эпигенетические механизмы адаптации, так как эпигенетические модификации (метилование ДНК, модификации гистонов) могут регулировать экспрессию генов в ответ на стресс (засуху, морозы, патогены), обеспечивая пластичность фенотипа.

3.5. Адаптация к условиям засухи

3.5.1. Дендрохронологические переменные, отражающие индивидуальный ответ дерева на условия засухи

Для того чтобы понять, как различаются значения фенотипических признаков в популяциях, были рассчитаны средние значения и построены диаграммы размаха для каждого признака (**Рисунок 27**).

Возраст. Средний возраст исследованных деревьев из популяции Батеневского кряжа составил 124 года, от 34 до 319 лет. Самый высокий средний возраст был в популяции КАМ, $153,82 \pm 13,93$; самый низкий средний возраст был в популяции SON, $102,83 \pm 5,56$. Только между этими популяциями тест Вилкоксона показал наличие значимых отличий ($p\text{-value} = 0,019$).

avLn. Средняя длина хвои для всех деревьев составила $22,79 \pm 0,34$ см. Среди популяций: в популяции КАМ $26,49 \pm 0,92$, BOGR $23,48 \pm 0,50$, BID $22,94 \pm 0,52$, SON $21,13 \pm 0,49$, TUIM $20,11 \pm 0,50$, при попарном сравнении популяции значительно различались ($p\text{-значение} < 0,05$), за исключением пар BID – BOGR и SON – TUIM.

varLn. Дисперсия длины хвои по всем деревьям в среднем составила $14,42 \pm 0,84$. Среднее для популяций: КАМ $21,52 \pm 3,05$, BOGR $10,54 \pm 0,92$, BID $14,32 \pm 1,63$, SON $12,66 \pm 0,99$, TUIM $12,24 \pm 0,94$. Только для популяций BOGR и КАМ тест Вилкоксона показал существенную разницу ($p\text{-значение} = 0,048$).

avTRW. Средняя ширина годичных колец в популяции SON составила $1,10 \pm 0,07$, в популяции BID — $0,85 \pm 0,09$, КАМ $0,83 \pm 0,11$, BOGR $0,78 \pm 0,05$, TUIM $0,77 \pm 0,05$. Популяция SON значительно отличалась от других популяций.

varTRW. Наибольшая дисперсия данных ширины годичных колец также наблюдалась в популяции SON и была равна $0,97 \pm 0,19$, среднее значение varTRW в популяциях КАМ составило $0,63 \pm 0,16$, TUIM $0,43 \pm 0,06$, BID $0,36 \pm 0,07$, BOGR $0,33 \pm 0,04$. Популяция SON значительно отличалась от других популяций.

trendTRW. Средние значения trendTRW были отрицательными во всех популяциях, кроме BOGR. КАМ $-0,008 \pm 0,002$, BOGR $0,001 \pm 0,003$, BID $-0,004 \pm 0,002$, SON $-0,022 \pm 0,003$, TUIM $-0,012 \pm 0,002$. Популяция SON значительно отличалась от всех других популяций, а популяция TUIM также значительно отличалась от популяций BOGR и BID.

Rt. Для всех изученных периодов засухи самый высокий средний индекс устойчивости был получен для периода засухи 1963-1965 гг. и был равен $0,91 \pm 0,03$. Для всех периодов засухи, кроме 1963–1965 гг., самые высокие средние значения индекса устойчивости наблюдались в популяции SON ($1,25 \pm 0,10$, $0,95 \pm 0,04$, $0,94 \pm 0,06$, $0,70 \pm 0,07$).

Rc. Самое высокое среднее значение индекса восстановления наблюдалось для засухи 1999 г. ($8,31 \pm 3,09$). Для засух 1951 и 1974-1976 гг. самые высокие значения Rc наблюдались в популяции КАМ ($4,48 \pm 1,31$ и $1,40 \pm 0,07$), для засух 1963–1965 и 1999 гг. в выборке BOGR наблюдались самые высокие значения Rc ($1,33 \pm 0,06$ и $24,74 \pm 15,38$ соответственно), а для засухи 2015 г. максимальное среднее значение индекса восстановления (Rc5) наблюдалось в выборке TUIM ($6,29 \pm 1,33$). Самые низкие значения Rc (значительно отличающиеся от других популяций) для всех периодов засухи продемонстрировала выборка SON.

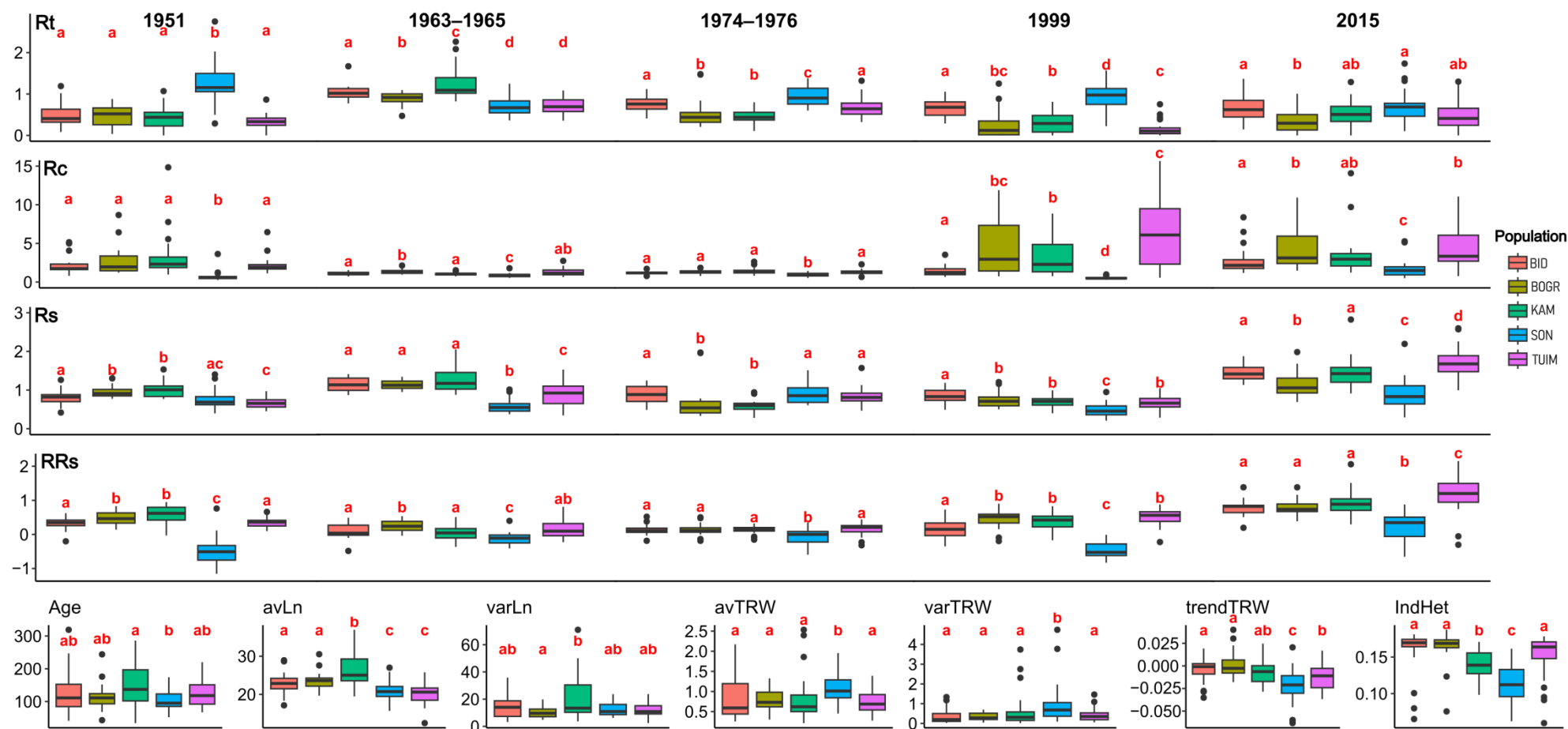


Рисунок 27 Диаграмма распределения значений 26 дендрофенотипических признаков и индивидуальной гетерозиготности (IndHet) в популяциях Батеневского кряжа.

Rs. Среднее значение индекса устойчивости для 136 деревьев варьировало от $1,33 \pm 0,04$ (засуха 2015 г.) до $0,68 \pm 0,02$ по данным засухи 1999 г. Для засух 1951 и 1963-1965 гг. самые высокие средние значения индекса устойчивости наблюдались у деревьев популяции КАМ ($1,01 \pm 0,04$ и $1,30 \pm 0,07$); для засух 1974-1976, 1999 и 2015 гг. самые высокие средние значения индекса устойчивости наблюдались у деревьев популяций SON, BID и TUIM ($0,89 \pm 0,04$, $0,85 \pm 0,03$ и $1,68 \pm 0,07$ соответственно). Для трех из пяти засух, 1963-1965, 1999 и 2015 годов, самые низкие значения индекса устойчивости наблюдались у популяции SON ($0,60 \pm 0,03$, $0,48 \pm 0,03$, $0,90 \pm 0,07$), что существенно отличалось от других популяций.

RRs. Среднее значение индекса относительной устойчивости для всех деревьев варьировало от $0,78 \pm 0,04$, исходя из данных о засухе 2015 года, до $0,09 \pm 0,02$, исходя из данных о засухе периода 1963–1965 годов. Для засухи 1951 года самые высокие значения относительной устойчивости наблюдались у деревьев популяции КАМ ($0,59 \pm 0,04$); за период 1963–1965 гг. самые высокие значения RR наблюдались в популяции BOGR ($0,25 \pm 0,03$). Согласно данным за последние три засухи, самые высокие значения RR наблюдались в популяции TUIM ($0,16 \pm 0,03$, $0,50 \pm 0,04$ и $1,19 \pm 0,09$). Самые низкие значения индекса относительной устойчивости наблюдались в популяции SON за все пять периодов засухи ($-0,49 \pm 0,09$, $-0,11 \pm 0,03$, $-0,05 \pm 0,04$, $-0,46 \pm 0,04$, $0,20 \pm 0,08$).

3.5.2. Полногеномный поиск ассоциаций с использованием индивидуальных дендрофенотипов (GWAS)

При поиске ассоциаций генотип-фенотип с помощью Байесовской разреженной линейной смешанной модели (BSLMM), доля вариативности дендрофенотипов, объясняемая генотипами (PVE) в среднем $0,78$, от минимальных значений по индексу сопротивления Rt5 ($0,39$) до максимальных по среднему приросту годичных колец avTRW ($0,98$). При этом доля генетической изменчивости, объясняемая вариантами с относительно большим эффектом (PGE) в среднем равнялась $0,44$ и

варьировала от 0,03 для индекса $Rc4$ до 0,96 у $avTRW$. Для 21 из 25 дендрофенотипов методом BSLMM были найдены ассоциированные SNP-маркеры со значениями апостериорных вероятностей включения $PIP > 25\%$. Всего методом BSLMM было отобрано 2523 предположительно адаптивных SNP-маркеров, 453 из которых ассоциированы с более чем одним дендрофенотипом.

При поиске ассоциаций методом смешанных линейных моделей (MLM), с учетом структуры популяций и матрицы попарного родства, суммарно было найдено 1380 предположительно адаптивных SNP для 25 дендрофенотипов, среди которых 565 ассоциированы с несколькими фенотипами (от 2 до 8) одновременно.

Подход на основе общих линейных моделей (GLM), в отличие от MLM, не учитывает популяционную структуру и степень родства между изучаемыми деревьями, что может приводить к ложноотрицательным или ложноположительным ассоциациям. С помощью метода GLM было выявлено 2031 предположительно адаптивный SNP-маркер, 1179 из которых (58%) были выявлены для более чем одного дендрофенотипа, а один SNP (LS.4109781.252) ассоциирован с 10 различными дендрофенотипами.

Результаты методов GLM и MLM показали высокий уровень перекрытия: 1276 SNP-маркеров, соответствующие 63% всех SNP, детектированных GLM и 92% SNP, детектированных при помощи MLM, были общими для обоих подходов. В целом, при сравнении всех трех использованных в работе методов поиска ассоциаций, 359 SNP-маркеров были отмечены как предположительно адаптивные по результатам BSLMM, MLM и GLM (**Рисунок 28**).

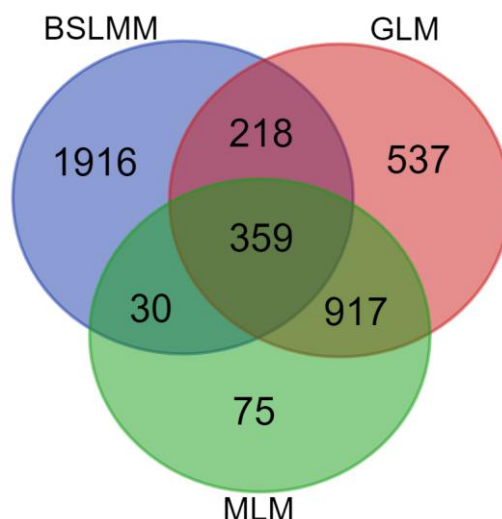


Рисунок 28 Диаграмма Венна, суммирующая результаты поиска SNP, связанных хотя бы с одним из 26 дендрофенотипов, с использованием трех методов: BSLMM, GLM и MLM.

3.5.3. SNP-маркеры, полученные в рамках изучения механизмов адаптации к условиям засухи

Для 371 SNP-маркера, обнаруженного всеми тремя методами ассоциативного анализа, мы проанализировали участки генома, в которых они локализованы. Только 26 SNP-маркеров принадлежат скаффолдам с аннотированными участками генома согласно геномной аннотации *L. sibirica*, тогда как остальные располагаются в неаннотированных скаффолдах, без доступной информации о структурных и функциональных элементах. Адаптивные SNP-маркеры в основном были расположены в межгенных регионах; только пять располагались в кодирующих областях генов, три из них являются несинонимичными нуклеотидными заменами и приводят к изменению кодируемой аминокислоты (**Таблица 10**).

Таблица 10 Результаты аннотации 26 потенциально адаптивных SNP, полученных в ходе полногеномного поиска ассоциаций с использованием индивидуальных дендрофенотипов

SNP	Дендрофенотип (метод GWAS)	Положение*	Белок/продукт гена	Функция	Источник
LS.3651056.309	Rt1 (GLM, MLM), avTRW (BSLMM)	экзон	цитохром P450 89A2	оксидоредуктазная активность	[168,169]
LS.3905950.5973	avTRW, varTRW (GLM, MLM), Rs4 (BSLMM)	экзон	ксилоглюкан эндотрансглюкозилаза / гидролаза белок В	строение клеточной стенки растущих тканей	[170]
LS.4033175.6261	Rs1 (GLM, MLM), RRs1 (BSLMM)	экзон	белок Е6, подобный целлюлозосинтазе	процесс биосинтеза целлюлозы	[171]

SNP	Дендрофенотип (метод GWAS)	Положе ние*	Белок/продукт гена	Функция	Источник
LS.4447596. 15042	Age (GLM, MLM), Rc1 (BSLMM)	экзон	нехарактеризованный белок LOC116133164	-	-
LS.4817075. 929	varLn (GLM, MLM), Rc5 (GLM, BSLMM), RRs5 (GLM)	экзон	белок Brevis radix-like 4	модулятор роста корней	[172]
LS.3903843. 2180	Rc2, Rc3, RRs2, RRs3 (MLM, GLM) Rs2 (GLM), Rs3 (BSLMM)	78	RING-H2 палец белок ATL3	убиквитинлигаза	[173,174]
LS.4185574. 3908	varLn (GLM, MLM), Rs3 (BSLMM)	325	альфа-маннозидаза 2	производство гликанов сложного типа	[175]
LS.7774.124 31	RRs4, Rs2 (GLM, MLM), RRs5 (GLM, BSLMM) Rc5, Rt5 (GLM), Age (BSLMM)	3951	О- фукозилтрансфераза 19-подобный	компонент сигнального пути гиббереллина	[176]
LS.10638.27 578	avTRW (GLM, MLM, BSLMM), Age, Rc4 (GLM)	5283	белок 1, содержащий мотивы GPALPP (белок липополисахарид- специфического ответа)	липополисахарид -специфический ответ	-
LS.56935.13 023	Rs4 (GLM, MLM, BSLMM)	5957	вероятная альдо- кеторедуктаза 4	деградация глифосата	[177]
LS.21791.50 056	RRs1, Rc5 (GLM, MLM), Rs3 (BSLMM)	6066	неизвестный	-	-
LS.61616.10 689	Rc1, Rt1 (GLM, MLM), Rs5 (BSLMM)	6109	белок Portal 56	-	-
LS.24961.61 616	RRs5 (GLM, MLM), Rt4 (BSLMM)	6200	неизвестный	-	-
LS.7196.672 2	RRs5, Rs5 (GLM, MLM), RRs2 (BSLMM)	8888	гипотетический белок теплового шока	-	-
LS.105004.9 682	avLn, Rt2 (GLM, MLM), Rc3 (GLM, BSLMM), RRs1, Rc1, Rt3 (GLM)	12733	фосфолипаза A1- Igamma3, хлоропластный	катализирует гидролиз фосфатидилхоли нае	[178]
LS.7184.303 5	Rs1 (GLM, MLM), RRs2 (BSLMM)	12939	неизвестный	-	-
LS.7252.104 584	Rc3 (GLM, MLM, BSLMM), RRs1, Rc1, (MLM, GLM), Rs1 (GLM)	15386	изоформа 2 белка семейства липоксигеназ, содержащая домен PLAT/LH2	защитные механизмы против патогенов	[179,180]
LS.14907.46	Rt2 (GLM, MLM,	15426	белок, подобный	может	[181,182]

SNP	Дендрофенотип (метод GWAS)	Положе ние*	Белок/продукт гена	Функция	Источник
110	BSLMM), Rs1 (GLM)		убиквитинлигазе E3 BIG BROTHER	ограничивать продолжительнос ть роста органов и их размер за счет активного разрушения критических стимуляторов роста.	
LS.15117.58 377	RRs2, Rc2 (GLM, MLM), Age (GLM), Rs4 (BSLMM)	34970	белок RAR1, содержащий домен, богатый цистеином и гистидином (CHRD1)	генно- опосредованная устойчивость к болезням	[183]
LS.30352.56 788	Rs4 (GLM, MLM), Rs3 (BSLMM)	54635	неизвестный	-	-
LS.16524.11 6919	Rs1 (GLM, MLM), Rs5 (BSLMM)	58871	неохарактеризованны й белок C1271.03с, содержащий домен FCP1	-	-
LS.12450.31 29	Rs2, avTRW, varTRW (GLM, MLM), avLn (BSLMM)	64166	белок RING-H2	убиквитинлигаза	[173,174]
LS.1920.812 78	Rc1 (GLM, MLM), Rt1 (GLM), Rt3 (BSLMM)	65622	белок, содержащий домен U-box, подобный белку 52	функционирует как E3- убиквитинлигаза.	[184]
LS.5783.879 94	Rt1 (GLM, MLM), Rc1 (GLM), Rs3 (BSLMM)	69130	АТФ-зависимая RECG-подобная ДНК- хеликаза	рекомбинация и репарация ДНК	[185]
LS.2549.428 67	Rc3 (GLM, MLM), RRs2 (BSLMM)	69752	гипотетический белок GW17_00026793	-	-
LS.16630.14 4091	varLn (GLM, MLM), avLn (BSLMM)	83423	аденилосукцинатлиаз а	синтез пуриновых нуклеотидов	[186]

Для всех фланкирующих последовательностей (± 1000 п.н.о. от SNP или максимально доступного интервала, если 2000 п.н.о. выходят за границы скаффолда), содержащих SNP, ассоциированные с дендрофенотипами, проводили поиск гомологичных последовательностей в базе данных NCBI BLAST. 57 содержащих SNP последовательностей показали высокую степень гомологии с задепонированными в базе последовательностями, причем большинство совпадений было выявлено с неизученными мРНК и гипотетическими белками *Picea glauca*, *Picea sitchensis* и *Pinus taeda*. Для 8

последовательностей были идентифицированы гомологи, содержащие аннотированные гены или другие структурные элементы растительных геномов. Например, для локуса SNP-маркера LS.4643215.2651, который связан с дендрофенотипом avLn по GLM и MLM, а также с индексом Rc3 по BSLMM, обнаружена гомология с мРНК, кодирующей мультипротеин-мостиковый фактор 1с (MBF1, XM_028068828) вигны (*Vigna unguiculata*). Это кофактор транскрипции, молекулярная функция которого заключается в формировании моста между факторами транскрипции и базальным механизмом транскрипции. Они участвуют в процессах развития и реакциях на стресс [187].

SNP-маркер LS.3970711.165, демонстрирующий ассоциации с дендрофенотипами Rc4 и RRs4 по результатам GLM/MLM и с RRs4 по результатам BSLMM, локализован в участке, гомологичном мРНК медь-метиламинооксидазоподобного белка (XM_022162594) *Helianthus annuus*, участвующего в защитных механизмах против грибных патогенов [188].

SNP LS.890128.210, ассоциированный с индексами Rs2 (MLM/GLM) и Rs3 (BSLMM), расположен в локусе, гомологичном MYB6-транскрипционному фактору *Pinus radiata* (KY196189), регулирующему биосинтез вторичных метаболитов [189].

SNP-маркер LS.3903843.2180, связанный с пятью дендрофенотипами (Rc2, Rc3, Rs2, RRs2, RRs3) согласно GLM и четырьмя (Rc2, Rc3, RRs2, RRs3) согласно MLM, и с Rs3 по BSLMM, находится на участке, гомологичном мРНК RING-H2 пальцевого белка типа ATL3 *Glycine soja* (XM_028350422), с дополнительной аннотацией для *L. sibirica* (Таблица 10).

SNP LS.4382060.16010, ассоциированный с индексами RRs1, Rs1, Rs3, Rs5, Rt3 (GLM), Rs3 (MLM) и Rc1 (BSLMM), локализован в митохондриальном гене, кодирующем рибосомальный белок S4 *Pinus strobus* (MN965308), участвующий в формировании устойчивости к заболеваниям [190].

SNP-маркер LS.29920.18803, связанный с дендрофенотипами avLn, Rs1, Rt4 (GLM/MLM), Rs2, Rt2 (GLM) и Rc1 (BSLMM), расположен в локусе, гомологичном полигалактуроназе *Papaver somniferum* (XM_026530000), участвующей в деградации пектина и ответах на биотический стресс [191].

Для двух последовательностей выявлены гомологи ретротранспозонной природы. Интервал, содержащий SNP-маркер LS.9095140.133 (ассоциированный с Rs4 (GLM/MLM) и avLn (BSLMM)), показал гомологию с РНК-направленной ДНК-полимеразой *Brassica oleracea* (XM_013761928). Последовательность SNP LS.4042787.3127 (связанная с переменными RRs1, Rc1, Rs2, Rt1, Rt3 (GLM), RRs1, Rc1, Rt1 (MLM) и Rs5 (BSLMM)) демонстрировала гомологию с ретротранспозонами *Pinus pinaster* (DQ394069) и *Pinus radiata* (AJ004945).

3.5.4. Корреляция между индивидуальной гетерозиготностью и дендрофенотипами (Heterozygosity-fitness correlations, HFCs)

Для 136 деревьев, отобранных для дендрохронологического анализа и изучения адаптации к условиям засухи, был посчитан индивидуальный уровень гетерозиготности (IndHet), в среднем он составил 0,144. Наименьшее среднее значение индивидуальной гетерозиготности было получено для выборки SON ($0,112 \pm 0,005$), наибольшее для выборки BOG ($0,164 \pm 0,005$). В ходе корреляционного анализа для каждого дендрофенотипа были рассчитаны коэффициенты корреляции Пирсона и Спирмена с индивидуальной гетерозиготностью (**Таблица 11**). Эти коэффициенты были рассчитаны как для нестандартизированных по хронологии данных, так и для индексов, рассчитанных после процедуры стандартизации.

Таблица 11. Коэффициент корреляции Пирсона (r) и коэффициент ранговой корреляции Спирмена (r_s) между дендрофенотипами и индивидуальной гетерозиготностью.

Дендрофенотип	r	P	r_s	P
Age	-0,006	0,9413	0,011	0,8979
avLn	0,047	0,5858	0,030	0,7325
varLnr	-0,055	0,5262	-0,150	0,0808
avTRW	-0,049	0,5718	-0,046	0,5911
varTRW	-0,118	0,1701	-0,164	0,0560
trendTRW	0,372	<0,0001	0,371	<0,0001
Rc1	0,322	0,0004	0,366	<0,0001
Rc2	0,345	<0,0001	0,351	<0,0001
Rc3	0,339	<0,0001	0,437	<0,0001
Rc4	0,307	0,0007	0,435	<0,0001
Rc5	0,318	0,0003	0,331	0,0002
Rs1	0,064	0,4918	-0,028	0,7615
Rs2	0,228	0,0111	0,278	0,0019
Rs3	0,074	0,4098	0,062	0,4898
Rs4	0,180	0,0370	0,311	0,0002
Rs5	0,142	0,1097	0,200	0,0237
Rt1	-0,432	<0,0001	-0,378	<0,0001
Rt2	-0,076	0,4014	0,021	0,8144
Rt3	-0,286	0,0011	-0,258	0,0034
Rt4	-0,342	<0,0001	-0,280	0,0010
Rt5	-0,203	0,0216	-0,215	0,0146
RRs1	0,471	<0,0001	0,316	0,0005
RRs2	0,378	<0,0001	0,353	<0,0001
RRs3	0,405	<0,0001	0,493	<0,0001
RRs4	0,423	<0,0001	0,401	<0,0001
RRs5	0,216	0,0144	0,258	0,0033

Примечание: жирным шрифтом выделены дендрофенотипы с p -значениями $< 0,05$.

Было выявлено наличие достоверной положительной связи между индексом восстановления (Rc), индексом относительной устойчивости (RRs) и индивидуальной гетерозиготностью по всем пяти периодам засухи (p -value $< 0,05$). Индекс сопротивления (Rt) показал достоверную отрицательную корреляцию с индивидуальной гетерозиготностью по четырем засухам, исключая период засухи 1963-1965 г. (Rt2). Индекс устойчивости (Rs) коррелировал согласно коэффициентам корреляции Пирсона и Спирмена с индивидуальной гетерозиготностью только по данным засух 1963-65 и 1999 г., но только на основе коэффициента ранговой корреляции Спирмена для данных засухи 2015 года.

Отсутствие значимой корреляции между индивидуальной гетерозиготностью и дендрофенотипами $avTRW$ и $varTRW$ согласуется с данными, полученными ранее Е.А. Бабушкиной и коллегами [91] для двух популяций лиственницы сибирской на основе 8 наиболее полиморфных микросателлитных локусов (SSR), хотя в случае $varTRW$ она была отрицательной согласно коэффициенту ранговой корреляции Спирмена ($rS = -0,164$) и почти значимой ($P = 0,056$), что может отчасти указывать на то, что повышенная гетерозиготность стабилизирует индивидуальное развитие, снижая дисперсию роста древесины. Не было обнаружено значимой корреляции между индивидуальной гетерозиготностью и возрастом, средней длиной хвои или дисперсией длины хвои. Наличие положительной корреляции между индивидуальной гетерозиготностью и индексами восстановления (Rc) и относительной устойчивости (RRs) для всех изученных периодов засухи и частично индексами устойчивости (Rs) для трех периодов засухи возможно означает, что различия в засухоустойчивости отдельных деревьев *L. sibirica* связаны с уровнем индивидуальной гетерозиготности, повышение которого улучшает адаптивные возможности организма (**Рисунок 29, a,b,d**).

Однако также была обнаружена значимая отрицательная связь между уровнем гетерозиготности и индексом устойчивости (Rt), который напрямую отражает эффективность роста *L. sibirica* в засушливые годы в четырех из пяти периодов (Рисунок 29, c). Возможно, один из адаптивных механизмов лиственницы заключается в резком снижении роста в периоды засухи как части стратегии выживания, которая более выражена у деревьев с более высоким уровнем индивидуальной гетерозиготности.

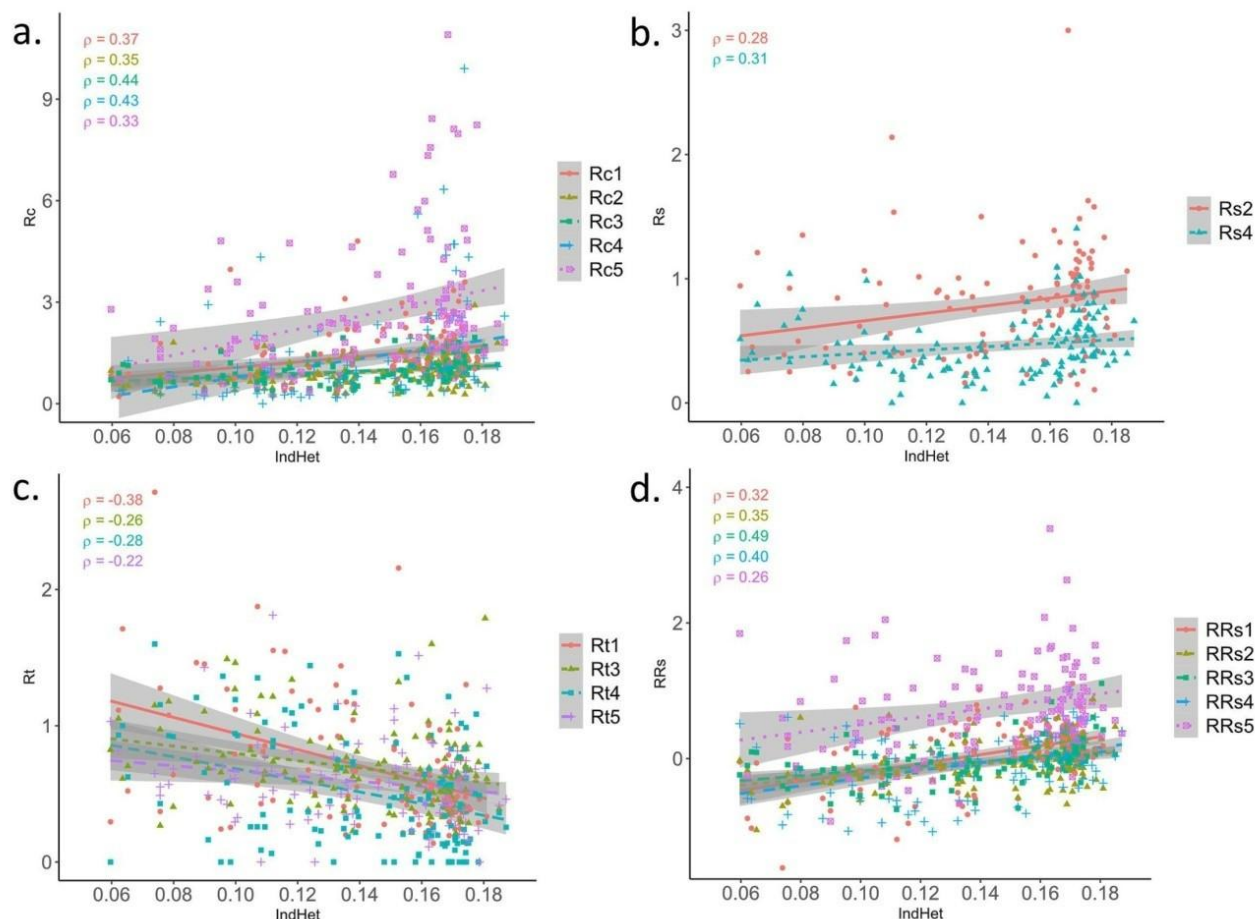


Рисунок 29. График, отображающий зависимость индексов восстановления (а), устойчивости (b), сопротивления (с) и относительной устойчивости (d) от уровня индивидуальной гетерозиготности (IndHet).

Средний радиальный прирост и средний тренд радиального прироста (trendTRW) представлены в **Таблице 12**. Положительные значения тренда означают, что скорость роста увеличивалась за последние 30 лет в период 1990-2019 гг. с учетом тренда размер-возраст, а отрицательные значения означают наоборот – уменьшалась. Если среднее значение индекса роста больше единицы, то деревья росли в среднем быстрее за последние 30 лет, чем за всю свою жизнь; если меньше единицы, то медленнее. Скорость роста немного снижалась в четырех из пяти популяций, несмотря на то, что в трех из пяти популяций деревья росли в среднем быстрее за последние 30 лет, чем за всю свою жизнь.

Таблица 12 Показатели тренда и среднего радиального прироста за последние 30 лет в период 1990-2019 гг. в пяти выборках лиственницы сибирской Батеневского кряжа.

Выборка	BID	BOG	KAM	SON	TUIM
trendTRW	-0,002	0,002	-0,009	-0,017	-0,009
средний радиальный прирост	0,973	1,017	0,982	1,042	1,058

Была обнаружена значимая положительная корреляция между индивидуальной гетерозиготностью и трендом радиального прироста (trendTRW). Основываясь на этих данных, деревья с большим уровнем гетерозиготности демонстрируют меньшее снижение скорости роста (Рисунок 30).

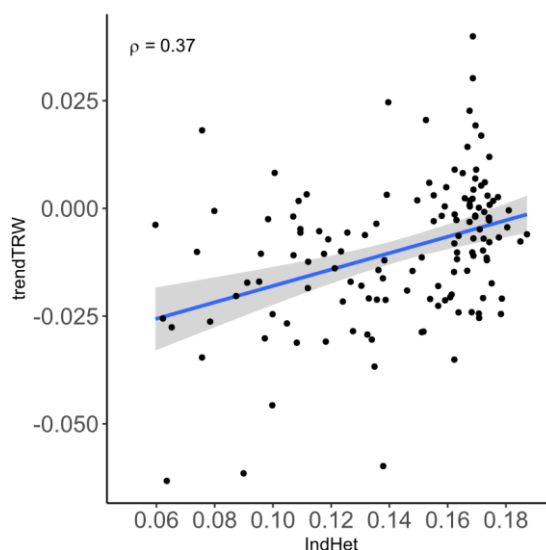


Рисунок 30 Зависимость между уровнем индивидуальной гетерозиготности (IndHet) и трендом радиального прироста (trendTRW).

Важно отметить, что в четырех из пяти популяций скорость роста несколько снижалась за последние 30 лет в период 1990-2019 гг. Вероятно, это связано с изменением климата, в частности, с увеличением дефицита воды, вызванного потеплением. Эти наблюдения подтверждаются недавними тенденциями к снижению урожайности зерновых культур, вызванными изменением климата, наблюдаемыми в степных территориях исследуемого региона [192]. Однако создание крупных водохранилищ до сих пор смягчало большинство тенденций к засухе в регионе [89] по сравнению с большей частью континентальной Азии, где в последнее время во многих местах на южных и нижних окраинах лесных массивов наблюдалось резкое снижение роста деревьев и гибель лесов [1,193].

ЗАКЛЮЧЕНИЕ

Данная работа демонстрирует преимущества методов секвенирования с уменьшенной репрезентативностью (англ. reduced-representation sequencing), в частности метода ddRADseq, для популяционных исследований на немодельных организмах, обладающих большими геномами. В ходе работы по большому количеству локусов было прогенотипировано 492 дерева *L. sibirica* из контрастных условий произрастания, при охвате 37 популяций из разных частей ареала, относительно быстро и экономически эффективно по сравнению с полногеномным ДНК-секвенированием. Протоколы ddRADseq позволили масштабировать пробоподготовку за счет индивидуального баркодирования и последующего демультимплексинга, что также выгодно отличает метод от более простых, но плохо масштабируемых подходов популяционной генетики, например анализа микросателлитных локусов.

Анализ пространственно-генетической структуры исследуемых природных популяций *L. sibirica*, в целом, подтвердил ранее полученные данные об относительно слабой селективно нейтральной структуре близко расположенных популяций хвойных, включая лиственницу [194,195], что объясняется в основном интенсивным потоком генов [196]. Однако по результатам анализа главных компонент (PCA), иерархического анализа молекулярной дисперсии (AMOVA) и анализа биогеографического происхождения индивидуальных деревьев (Admixture) наблюдается разделение популяций по меньшей мере на 3 кластера: уральский, алтайский и саянский с зоной гибридизации в крайне северных участках ареала западной и восточной рас лиственницы сибирской, что согласуется с ранее полученными данными [197].

Было идентифицировано несколько SNP-маркеров в генах-кандидатах, изменчивость которых была связана с высотностью и другими биоклиматическими переменными, это гены-гомологи *FAR4*, *AP2A1*, *CSL*, *LEA*, *NRX1*, *F-box*, *ISWI*, *PLIP2*, *EXO*, *COG1*, *PIRL1* и *OPT7*-подобные гены.

На основании этих результатов можно предположить, что *L. sibirica* адаптировалась к большим высотам отчасти за счет вспомогательных функций, связанных с воспроизводством в условиях абиотического стресса, таких как защита шаперонов от клеточного повреждения, поддержка и стимуляция роста клеток, сигнализация стресса, эпигенетическая регуляция через ремоделирование хроматина и т. д., хотя необходимо получить больше информации о том, как эти гены регулируют адаптацию. Учитывая, что большинство экологических факторов высокогорья оказывают стрессовое воздействие на *L. sibirica*, настоящее исследование дает возможность: идентифицировать ключевые гены и SNP-маркеры, имеющие значение для селекционной работы и создать фундамент для разработки SNP-чипа, который позволит отслеживать как нейтральную, так и адаптивную генетическую изменчивость в различных популяциях *L. sibirica*.

Использование шести различных методов поиска адаптивных маркеров-кандидатов позволило получить набор данных из 21 высокодостоверного SNP-маркера, изменчивость 12 из которых коррелировала с изменчивостью средовых климатических факторов, связанных с широтой произрастания. Аннотация геномных участков, содержащих эти 21 SNP показала, что большинство маркеров сосредоточены в межгенных областях митохондриального генома, преимущественно вблизи генов *NAD1*, *NAD2*, *NAD4*, *L2*, *L5*, *S7*, *S11*, *S14*, *COX2* и *ATP4*. Ассоциированный со средовой изменчивостью в восточной трансекте, маркер LS.28474_5984 расположен недалеко от гена *futsch*-подобного белка, участвующего в метаболизме азотистых соединений. Некоторые маркеры находятся, возможно в промоторном районе и могут влиять на экспрессию генов. Локализация большинства из обнаруженных адаптивных маркеров в митохондриальном геноме представляет интерес. Известно, что митохондриальные гены играют решающую роль в локальной адаптации растений [165,166], а найденные маркеры-кандидаты вероятно могут находиться в регуляторных областях и влиять на экспрессию данных генов.

Представленные данные могут служить научной основой для оптимизации природопользования, разработки методов рационального использования изучаемого вида, выявления популяций *L. sibirica* с хорошим генетическим потенциалом и проведения экологического мониторинга.

Поиск ассоциаций генотип-фенотип с использованием трех различных методов выявил 371 потенциально адаптивных SNP. Из них только 26 SNP были расположены в аннотированных регионах: 21 в межгенных регионах и пять в регионах, кодирующих гены.

Обнаружена значимая положительная корреляция между индивидуальной гетерозиготностью и индексами восстановления, устойчивости и относительной устойчивости, что позволяет предположить, что различия в засухоустойчивости отдельных деревьев *L. sibirica* связаны с уровнем индивидуальной гетерозиготности, повышение которого улучшает адаптивные возможности. Однако также была обнаружена значимая отрицательная связь между уровнем гетерозиготности и индексом сопротивления, который напрямую отражает эффективность роста *L. sibirica* в засушливые годы. Возможно, один из адаптивных механизмов лиственницы заключается в резком снижении роста в периоды засухи как части стратегии выживания, которая более выражена у деревьев с более высоким уровнем индивидуальной гетерозиготности. Является ли это следствием повышенного гомеостаза или гетерозиса и сверхдоминирования, определить сложно и требует дополнительных исследований.

ВЫВОДЫ

1. Впервые на основании данных полногеномного генотипирования установлена пространственная популяционно-генетическая структура изученных выборок *L. sibirica*. По результатам анализа показано разделение выборок как минимум на три группы: уральскую, алтайскую и сибирскую. Обнаружена значимая корреляция между географическим положением и генетической дифференциацией, выявлены сигналы изоляции вследствие гетерогенности окружающей среды и возможной локальной адаптации.

2. Выявлены участки генома, содержащие однонуклеотидные полиморфизмы, изменчивость которых статистически достоверно связана с вариабельностью экологических факторов, что указывает на процессы локальной адаптации в популяциях.

3. Гены-кандидаты, ассоциированные с адаптацией к условиям высокогорья и широтному градиенту, вовлечены в процессы стресс-ответа, фотоморфогенеза и регуляции роста, что позволяет рассматривать эти процессы как вовлеченные в механизмы локальной адаптации.

4. Найдены участки генома и гены-кандидаты, ответственные за устойчивость деревьев к повторяющимся климатическим стрессам в условиях юга Сибири на основании результатов ассоциативного анализа между генотипами и дендрофенотипами, отражающими индивидуальный ответ дерева на условия засухи, включая стандартные индексы сопротивления, восстановления и устойчивости.

5. Обнаруженная значимая положительная корреляция между индивидуальной гетерозиготностью и индексами восстановления и устойчивости позволяет предположить, что различия в засухоустойчивости отдельных деревьев *L. sibirica* популяций юга Сибири связаны с уровнем индивидуальной гетерозиготности, повышение которого улучшает адаптивные возможности.

СПИСОК ЛИТЕРАТУРЫ

1. Mukherjee, S. Climate Change and Drought: a Perspective on Drought Indices/ S. Mukherjee, A. Mishra, K.E. Trenberth // Curr Clim Change Rep. – 2018. – Vol. 4, № 2. – P. 145-163.
2. Grace, J. Impacts of Climate Change on the Tree Line / J. Grace // Annals of Botany. – 2002. – Vol. 90, № 4. – P. 537-544.
3. Лукина, Н.В. Функциональная классификация лесов: актуальность и подходы к разработке / Н.В. Лукина и др. // Лесоведение. – 2021. – № 6. – С. 566-580.
4. Baird, N.A. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers / N.A. Baird et al. // PLoS ONE. – 2008. – Vol. 3, № 10. – P. e3376.
5. Крутовский, К.В. Дендрогеномика – новая междисциплинарная область исследований адаптивного генетического потенциала лесных древесных популяций, интегрирующая дендрохронологию, дендрэкологию, дендроклиматологию и геномику / К.В. Крутовский // Генетика. – 2022. – Т. 58, № 11. – С. 1225-1239.
6. Орешкова, Н.В. Разработка микросателлитных маркеров лиственницы сибирской (*Larix sibirica* Ledeb.) на основе полногеномного *de novo* секвенирования / Н.В. Орешкова и др. // Генетика. – 2017. – Т. 53, № 11. – С. 1278-1284.
7. Орешкова, Н.В. Разработка ядерных микросателлитных маркеров с длинными (трех-, четырех-, пяти- и шестинуклеотидными) мотивами для трех видов лиственницы на основе полногеномного *de novo* секвенирования лиственницы сибирской (*Larix sibirica* Ledeb.) / Н.В. Орешкова и др. // Генетика. – 2019. – Т. 55, № 4. – С. 418-425.
8. Васильева, Ю.С. Анализ генетического разнообразия и структуры популяций западной расы лиственницы сибирской *Larix sibirica* Ledeb. Урала на основе полиморфизма межмикросателлитных маркеров / Ю.С. Васильева и др. // Бюллетень науки и практики. – 2018. – Т. 4, № 12. – С. 113-124.
9. Экарт, А.К. Изменчивость локуса *mt44* митохондриальной ДНК в популяциях ели сибирской / А.К. Экарт и др. // Генетика. – 2020. – Т. 56, № 7. – С. 842-847.
10. Eckert, A.J. Patterns of Population Structure and Environmental Associations to Aridity Across the Range of Loblolly Pine (*Pinus taeda* L., Pinaceae) / A.J. Eckert et al. // Genetics. – 2010. – Vol. 185, № 3. – P. 969-982.
11. Heer, K. Linking dendroecology and association genetics in natural populations: Stress responses archived in tree rings associate with SNP genotypes in silver fir

- (*Abies alba* Mill.) / K. Heer et al. // *Molecular Ecology*. – 2018. – Vol. 27, № 6. – P. 1428-1438.
12. Trujillo-Moya, C. Drought Sensitivity of Norway Spruce at the Species' Warmest Fringe: Quantitative and Molecular Analysis Reveals High Genetic Variation Among and Within Provenances / C. Trujillo-Moya et al. // *G3 Genes|Genomes|Genetics*. – 2018. – Vol. 8, № 4. – P. 1225-1245.
 13. De La Torre, A.R. Dissecting the Polygenic Basis of Cold Adaptation Using Genome-Wide Association of Traits and Environmental Data in Douglas-fir / A.R. De La Torre et al. // *Genes*. – 2021. – Vol. 12, № 1. – P. 110.
 14. Segura-Sanchez, C.A. Dendrogenomic resilience mechanisms of two endangered Mexican spruces / C.A. Segura-Sanchez et al. // *Plant Diversity*. – 2025. – P. S2468265925000976.
 15. Hornoy, B. Genetic Adaptation to Climate in White Spruce Involves Small to Moderate Allele Frequency Shifts in Functionally Diverse Genes / B. Hornoy et al. // *Genome Biology and Evolution*. – 2015. – Vol. 7, № 12. – P. 3269-3285.
 16. Yu, Y. Using landscape genomics to delineate seed and breeding zones for lodgepole pine / Y. Yu et al. // *New Phytologist*. – 2022. – Vol. 235, № 4. – P. 1653-1664.
 17. Shu, M. Identifying genetic variation associated with environmental gradients and drought-tolerance phenotypes in ponderosa pine / M. Shu, E.V. Moran // *Ecology and Evolution*. – 2023. – Vol. 13, № 10. – P. e10620.
 18. Бобров, Е.Г. История и систематика лиственницы / Е.Г. Бобров. – Ленинград: Наука, 1972. – 96 с.
 19. Дылис, Н.В. Лиственница / Н.В. Дылис. – Москва: Лесная промышленность, 1981. – 96 с.
 20. Абаимов, А.П. Биоразнообразие лиственниц Азиатской России / А.П. Абаимов и др. – Новосибирск: Акад. изд-во «ГЕО», 2010. – 160 с.
 21. Абаимов, В.Ф. Дендрология : учеб. пособие для студ. высш. учеб. заведений / В.Ф.Абаимов. – 3-е изд., перераб. – М. : Издательский центр «Академия», 2009. – 368 с.
 22. Коропачинский, И.Ю. Древесные растения Сибири. / И. Ю. Коропачинский. – Новосибирск : Наука, 1983. – 384 с.
 23. Онучин, В.С. О некоторых морфологических признаках Лиственницы сибирской, произрастающей в Туве / В.С. Онучин // *Лиственница (сборник статей)*. – 1962. – Т. 29. – С. 22-35
 24. Kruse, S. High gene flow and complex treeline dynamics of *Larix* Mill. stands on the Taymyr Peninsula (north-central Siberia) revealed by nuclear microsatellites / S. Kruse et al. // *Tree Genetics & Genomes*. – 2018. – Vol. 14 № 2.

25. Абаимов, А. Морфофизиологическая изменчивость лиственниц Сибири и Дальнего Востока / А. Абаимов и др. // Биоразнообразие лиственниц Азиатской России. Новосибирск: Акад. изд-во «ГЕО», 2010. – С. 21-33.
26. Милютин, Л.И. Биоразнообразие лиственниц России / Л.И. Милютин // Хвойные бореальной зоны. – 2003. – Т. 21, № 1. – С. 1-4.
27. Дылис, Н.В. Сибирская лиственница. Материалы к систематике, географии и истории / Н.В. Дылис – Москва: Изд-во МОИП, 1947. – 137 с.
28. Семериков, В.Л. Популяционная структура и молекулярная систематика видов *Larix* Mill.: дис. ... д-ра биол. наук: 03.00.05 / Владимир Леонидович Семериков. – Екатеринбург, 2007. – 260 с.
29. Gonzalez-Ibyea, D. Assessing the Gene Content of the Megagenome: Sugar Pine (*Pinus lambertiana*) / D. Gonzalez-Ibyea et al. // G3 Genes|Genomes|Genetics. Oxford University Press (OUP). – 2016. – Vol. 6, № 12. – P. 3787-3802.
30. Bondar, E.I. Annotation of Siberian Larch (*Larix sibirica* Ledeb.) Nuclear Genome—One of the Most Cold-Resistant Tree Species in the Only Deciduous GENUS in Pinaceae / E.I. Bondar et al. // Plants. – 2022. – Vol. 11, № 15. – P. 2062.
31. Kuzmin, D.A. Stepwise large genome assembly approach: a case of Siberian larch (*Larix sibirica* Ledeb) / D.A. Kuzmin et al. // BMC Bioinformatics. – 2019. – Vol. 20, № S1. – P. 37.
32. Sun, C. The *Larix kaempferi* genome reveals new insights into wood properties / C. Sun et al. // JIPB. – 2022. – Vol. 64, № 7. – P. 1364-1373.
33. Dai, A. Increasing drought under global warming in observations and models / A. Dai // Nature Clim Change. – 2013. – Vol. 3, № 1. – P. 52-58.
34. Seager, R. Model Projections of an Imminent Transition to a More Arid Climate in Southwestern North America / R. Seager et al. // Science. – 2007. – Vol. 316, № 5828. – P. 1181-1184.
35. Allen, C.D. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests / C.D. Allen et al. // Forest Ecology and Management. – 2010. – Vol. 259, № 4. – P. 660-684.
36. Serra-Maluquer, X. Changes in tree resistance, recovery and resilience across three successive extreme droughts in the northeast Iberian Peninsula / X. Serra-Maluquer, M. Mencuccini, J. Martínez-Vilalta // Oecologia. – 2018. – Vol. 187, № 1. – P. 343-354.
37. Aitken, S.N. Adaptation, migration or extirpation: climate change outcomes for tree populations / S.N. Aitken et al. // Evolutionary Applications. – 2008. – Vol. 1, № 1. – P. 95-111.
38. Leites, L. Forest tree species adaptation to climate across biomes: Building on the legacy of ecological genetics to anticipate responses to climate change / L. Leites,

- M. Benito Garzón // *Global Change Biology*. – 2023. – Vol. 29, № 17. – P. 4711-4730.
39. Тихонова, И.В. Генетическая изменчивость в популяциях *Pinus sylvestris*, *Picea obovata*, *Abies sibirica* и на вырубках в южной тайге Средней Сибири / И.В. Тихонова и др. // *Генетика*. – 2021. – Т. 57, № 3. – С. 296-310.
 40. Hedrick, P.W. Genetic Polymorphism in Heterogeneous Environments / P.W. Hedrick, M.E. Ginevan, E.P. Ewing // *Annu. Rev. Ecol. Syst.* – 1976. – Vol. 7, № 1. – P. 1-32.
 41. Richardson, D.M. Tree invasions: patterns, processes, challenges and opportunities / D.M. Richardson et al. // *Biol Invasions*. – 2014. – Vol. 16, № 3. – P. 473-481.
 42. Hitchings, S.P. Genetic substructuring as a result of barriers to gene flow in urban *Rana temporaria* (common frog) populations: implications for biodiversity conservation / S.P. Hitchings, T.J.C Beebee // *Heredity*. – 1997. – Vol. 79, № 2. – P. 117-127.
 43. Primack, R.B. *Essentials of conservation biology*. / R.B. Primack. – 6th ed. – Sunderland: Sinauer associates publ, 2014. – 564 p.
 44. Via, S. The Ecological Genetics of Speciation / S. Via // *The American Naturalist*. – 2002. – Vol. 159, № S3. – P. 1-7.
 45. Dyer, R.J. Is there such a thing as landscape genetics? / R.J. Dyer // *Mol Ecol*. – 2015. – Vol. 24, № 14. – P. 3518-3528.
 46. Storfer, A. Landscape genetics: where are we now?: TRENDS IN LANDSCAPE GENETICS / A. Storfer et al. // *Molecular Ecology*. – 2010. – Vol. 19, № 17. – P. 3496-3514.
 47. Helyar, S.J. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges: ANALYTICAL APPROACHES / S.J. Helyar et al. // *Molecular Ecology Resources*. – 2011. – Vol. 11. – P. 123-136.
 48. Stapley, J. Adaptation genomics: the next generation / J. Stapley et al. // *Trends in Ecology & Evolution*. – 2010. – Vol. 25, № 12. – P. 705-712.
 49. Castellana, S. Environmental genome-wide association studies across precipitation regimens reveal that the E3 ubiquitin ligase MBR1 regulates plant adaptation to rainy environments / S. Castellana et al. // *Plant Communications*. – 2024. № 5. – P. 101074.
 50. Etter, P.D. RAD Paired-End Sequencing for Local *De Novo* Assembly and SNP Discovery in Non-model Organisms / P.D. Etter, E. Johnson // *Data Production and Analysis in Population Genomics*. – 2012. – Vol. 888. – P. 135-151.
 51. Hale, M.C. Evaluating Adaptive Divergence Between Migratory and Nonmigratory Ecotypes of a Salmonid Fish, *Oncorhynchus mykiss* / M.C. Hale et al. // *G3 Genes|Genomes|Genetics*. – 2013. – Vol. 3, № 8. – P. 1273-1285.

52. Neale, D.B. A Genome Sequence for the Threatened Whitebark Pine / D.B. Neale et al. // *G3: Genes|Genomes|Genetics*. – 2024. – Vol. 14, № 5. – P. jkae061.
53. Manel, S. Landscape genetics: combining landscape ecology and population genetics / S. Manel et al. // *Trends in Ecology & Evolution*. – 2003. – Vol. 18, № 4. – P. 189-197.
54. Joost, S. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation / S. Joost et al. // *Mol Ecol*. – 2007. – Vol. 16, № 18. – P. 3955-3969.
55. Luikart, G. The power and promise of population genomics: from genotyping to genome typing / G. Luikart et al. // *Nat Rev Genet*. – 2003. – Vol. 4, № 12. – P. 981-994.
56. Balkenhol, N. Landscape genetics: concepts, methods, applications / N. Balkenhol et al. – Hoboken: Wiley Blackwell, 2016. – 264 p.
57. Balkenhol, N. Landscape Genomics: Understanding Relationships Between Environmental Heterogeneity and Genomic Characteristics of Populations/ N. Balkenhol et al. // *Population Genomics: Concepts, Approaches and Applications*. – Cham: Springer International Publishing, 2019. – P. 261-322.
58. Haddad, N.M. Habitat fragmentation and its lasting impact on Earth's ecosystems / N.M. Haddad et al.// *Sci. Adv*. – 2015. – Vol. 1, № 2. – P. e1500052.
59. Aitken, S.N. Conserving Evolutionary Potential: Combining Landscape Genomics with Established Methods to Inform Plant Conservation / S.N. Aitken, R. Jordan, H.R. Tumas // *Annual Review of Plant Biology*. Annual Reviews, – 2024. – Vol. 75, № 1. – P. 707-736.
60. Campbell, Q. Agricultural landscape genomics to increase crop resilience / Q. Campbell et al.// *Plant Communications*. – 2025. – Vol. 6, № 2. – P. 101260.
61. Dauphin, B. Re-thinking the environment in landscape genomics / B. Dauphin et al. // *Trends in Ecology & Evolution*. – 2023. – Vol. 38, № 3. – P. 261-274.
62. Grummer, J.A. Aquatic Landscape Genomics and Environmental Effects on Genetic Variation / J.A. Grummer et al. // *Trends in Ecology & Evolution*. – 2019. – Vol. 34, № 7. – P. 641-654.
63. Stölting, K.N. Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species / K.N. Stölting et al. // *Mol Ecol*. – 2013. – Vol. 22, № 3. – P. 842-855.
64. Storz, J.F. INVITED REVIEW: Using genome scans of DNA polymorphism to infer adaptive population divergence: GENOME SCANS AND ADAPTIVE POPULATION DIVERGENCE / J.F. Storz // *Molecular Ecology*. – 2005. – Vol. 14, № 3. – P. 671-688.
65. Coop, G. Using Environmental Correlations to Identify Loci Underlying Local Adaptation / G. Coop et al. // *Genetics*. – 2010. – Vol. 185, № 4. – P. 1411–1423.

66. De Mita, S. Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations / S. De Mita et al. // *Mol Ecol.* – 2013. – Vol. 22, № 5. – P. 1383-1399.
67. De Villemereuil, P. A new F_{ST} -based method to uncover local adaptation using environmental variables / P. De Villemereuil, O.E. Gaggiotti // *Methods Ecol Evol.* – 2015. – Vol. 6, № 11. – P. 1248-1258.
68. Rellstab, C. A practical guide to environmental association analysis in landscape genomics / C. Rellstab et al. // *Mol Ecol.* – 2015. – Vol. 24, № 17. – P. 4348-4370.
69. Grant, M.C. Genetic Differentiation Among Growth Forms of Engelmann Spruce and Subalpine Fir at Tree Line / M.C. Grant, J.B. Mitton // *Arctic and Alpine Research.* – 1977. – Vol. 9, № 3. – P. 259-263.
70. Stucki, S. High performance computation of landscape genomic models integrating local indices of spatial association / S. Stucki et al. // *Mol Ecol Resour.* – 2017. – Vol. 17, № 5. – P. 1072-1089.
71. Chang, C.-W. Physical geography, isolation by distance and environmental variables shape genomic variation of wild barley (*Hordeum vulgare* L. ssp. *spontaneum*) in the Southern Levant / C.-W. Chang et al. // *Heredity.* – 2022. – Vol. 128, № 2. – P. 107-119.
72. Yoder, P.S. A statistical filtering procedure to improve the accuracy of estimating population parameters in feed composition databases / P.S. Yoder, N.R. St-Pierre // *Journal of Dairy Science.* – 2014. – Vol. 97, № 9. – P. 5645-5656.
73. Hancock, A.M. Adaptation to Climate Across the *Arabidopsis thaliana* Genome / A.M. Hancock et al. // *Science.* – 2011. – Vol. 334, № 6052. – P. 83-86.
74. Benjamini, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing / Y. Benjamini, Y. Hochberg // *Journal of the Royal Statistical Society.* – 1995. – Vol. 57, № 1. – P. 289-300.
75. Yeaman, S. The genetic architecture of adaptation under migration-selection balance / S. Yeaman, M.C. Whitlock // *Evolution.* – 2011. – Vol. 65, № 7. – P. 1897-1911.
76. Harrison, P.A. Linkages between biodiversity attributes and ecosystem services: A systematic review / P.A. Harrison et al. // *Ecosystem Services.* – 2014. – Vol. 9. – P. 191–203.
77. Lasky, J.R. Genome-environment associations in sorghum landraces predict adaptive traits / J.R. Lasky et al. // *Sci. Adv.* – 2015. – Vol. 1, № 6. – P. e1400218.
78. Savolainen, O. Ecological genomics of local adaptation / O. Savolainen, M. Lascoux, J. Merilä // *Nat Rev Genet.* – 2013. – Vol. 14, № 11. – P. 807-820.

79. Bonin, A. Explorative Genome Scan to Detect Candidate Loci for Adaptation Along a Gradient of Altitude in the Common Frog (*Rana temporaria*) / A. Bonin et al. // Molecular biology and evolution. – 2006. – Vol. 23. – P. 773-783.
80. Nunes, V.L. Multiple approaches to detect outliers in a genome scan for selection in ocellated lizards (*Lacerta lepida*) along an environmental gradient: SELECTION IN OCELLATED LIZARDS / V.L. Nunes et al. // Molecular Ecology. – 2011. – Vol. 20, № 2. – P. 193-205.
81. Arzac, A. An Overview on Dendrochronology and Quantitative Wood Anatomy Studies of Conifers in Southern Siberia (Russia) / A. Arzac, M.V. Fonti, E.A. Vaganov // Progress in Botany. – 2021. – Vol. 83. – P. 161-181.
82. Krutovsky, K.V. Dendrogenomics Is a New Interdisciplinary Field of Research of the Adaptive Genetic Potential of Forest Tree Populations Integrating Dendrochronology, Dendroecology, Dendroclimatology, and Genomics / K.V. Krutovsky // Russian Journal of Genetics. – 2022. – Vol. 58, № 11. – P. 1273-1286.
83. Vaganov, E.A. Growth dynamics of conifer tree rings: images of past and future environments / E.A. Vaganov et al. – Berlin: Springer, 2006.
84. Fritts, H.C. Tree rings and climate / H.C. Fritts. – Caldwell: Blackburn Press, 2001. – 567 p.
85. Lloyd, A.H. Responses of the circumpolar boreal forest to 20th century climate variability / A.H. Lloyd, A.G. Bunn // Environ. Res. Lett. – 2007. – Vol. 2, № 4. – P. 045013.
86. Wang, T. Climatic signals in tree ring of *Picea schrenkiana* along an altitudinal gradient in the central Tianshan Mountains, northwestern China / T. Wang, H. Ren, K. Ma // Trees. – 2005. – Vol. 19, № 6. – P. 736-742.
87. Babushkina, E.A. Response of Four Tree Species to Changing Climate in a Moisture-Limited Area of South Siberia / E.A. Babushkina et al. // Forests. – 2019. – Vol. 10, № 11. – P. 999.
88. Cochran, P.H. Examples of mortality and reduced annual increments of white fir induced by drought, insects, and disease at different stand densities / P.H. Cochran // U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, – 1998. – 19 p.
89. Zhirnova, D.F. Climate change and tree growth in the Khakass-Minusinsk Depression (South Siberia) impacted by large water reservoirs / D.F. Zhirnova et al. // Scientific Reports. – 2021. – Vol. 11, № 1. – P. 14266.
90. Lloret, F. Components of tree resilience: effects of successive low-growth episodes in old ponderosa pine forests / F. Lloret, E.G. Keeling, A. Sala // Oikos. – 2011. – Vol. 120, № 12. – P. 1909-1920.

91. Babushkina, E.A. The effect of individual genetic heterozygosity on general homeostasis, heterosis and resilience in Siberian larch (*Larix sibirica* Ledeb.) using dendrochronology and microsatellite loci genotyping / E.A. Babushkina et al. // *Dendrochronologia*. – 2016. – Vol. 38. – P. 26-37.
92. Johnson, J.S. Seed dispersal at alpine treeline: an assessment of seed movement within the alpine treeline ecotone / J.S. Johnson et al. // *Ecosphere*. – 2017. – Vol. 8, № 1. – P. e01649.
93. Johnson, J.S. Growth and Its Relationship to Individual Genetic Diversity of Mountain Hemlock (*Tsuga mertensiana*) at Alpine Treeline in Alaska: Combining Dendrochronology and Genomics / J.S. Johnson et al. // *Forests*. – 2017. – Vol. 8, № 11. – P. 418.
94. Heer, K. Linking dendroecology and association genetics in natural populations: Stress responses archived in tree rings associate with SNP genotypes in silver fir (*Abies alba* Mill.) / K. Heer et al. // *Molecular Ecology*. – 2018. – Vol. 27, № 6. – P. 1428-1438.
95. Piotti, A. Influence of Spatiotemporal Dynamics on the Fine-Scale Spatial Genetic Structure of Differently Managed *Picea abies* Stands / A. Piotti et al. // *Forests*. – 2018. – Vol. 9, № 10. – P. 622.
96. Singh, P. Genetic architecture of disease resistance and tolerance in Douglas-fir trees / P. Singh et al. // *New Phytologist*. – 2024. – Vol. 243, № 2. – P. 705-719.
97. Zhou, J. Local diversity of drought resistance and resilience in *Populus tomentosa* correlates with the variation of DNA methylation / J. Zhou et al. // *Plant Cell & Environment*. – 2023. – Vol. 46, № 2. – P. 479-497.
98. Johnson, J.S. Advancing Biogeography Through Population Genomics / J.S. Johnson et al. // *Population Genomics*. – 2018. – P. 539-585.
99. Avanzi, C. Individual reproductive success in Norway spruce natural populations depends on growth rate, age and sensitivity to temperature / C. Avanzi et al. // *Heredity*. – 2020. – Vol. 124, № 6. – P. 685-698.
100. Depardieu, C. Connecting tree-ring phenotypes, genetic associations and transcriptomics to decipher the genomic architecture of drought adaptation in a widespread conifer / C. Depardieu et al. // *Molecular Ecology*. – 2021. – Vol. 30, № 16. – P. 3898-3917.
101. Davey, J.W. RADSeq: next-generation population genetics / J.W. Davey, M.L. Blaxter // *Briefings in Functional Genomics*. – 2010. – Vol. 9, № 5–6. – P. 416-423.
102. Peterson, B.K. Double Digest RADseq: An Inexpensive Method for *De Novo* SNP Discovery and Genotyping in Model and Non-Model Species / B.K. Peterson et al. // *PLoS ONE*. – 2012. – Vol. 7, № 5. – P. e37135.

103. Cariou, M. Is RAD-seq suitable for phylogenetic inference? *An in silico* assessment and optimization / M. Cariou, L. Duret, S. Charlat // *Ecology and Evolution*. – 2013. – Vol. 3, № 4. – P. 846-852.
104. Rubin, B.E.R. Inferring Phylogenies from RAD Sequence Data / B.E.R. Rubin, R.H. Ree, C.S. Moreau // *PLoS ONE*. – 2012. – Vol. 7, № 4. – P. e33394.
105. Zhirnova, D.F. To which side are the scales swinging? Growth stability of Siberian larch under permanent moisture deficit with periodic droughts / D.F. Zhirnova et al.// *Forest Ecology and Management*. – 2020. – Vol. 459. – P. 117841.
106. Алисов, Б.П. Климат России./ Б.П Алисов. – М.: Изд-во Моск. ун-та, 1956. 125 с.
107. Bazhenova, O.I. The structure of contemporary denudation in the steppes of the Minusinskaya depression / O.I. Bazhenova, E.M. Tyumentseva // *Geography and Natural Resources*. – 2010. – Vol. 31, № 4. – P. 362-369.
108. Harris, I. Updated high-resolution grids of monthly climatic observations – the CRU TS3 .10 Dataset / I. Harris et al.// *Intl Journal of Climatology*. – 2014. – Vol. 34, № 3. – P. 623-642.
109. Harris, I.C. CRU TS4.02: Climatic Research Unit (CRU) Time-Series (TS) version 4.02 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901- Dec. 2017)./ I.C. Harris, P.D. Jones // *Centre for Environmental Data Analysis (CEDA), University Of East Anglia Climatic Research Unit (CRU)*. – 2019.
110. Fick, S.E. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas / S.E. Fick, R.J. Hijmans // *Int. J. Climatol*. – 2017. Vol. 37, № 12. – P. 4302-4315.
111. Oksanen, J. vegan community ecology package version 2.6-2 / Oksanen J. et al. // *The Comprehensive R Archive Network*. – 2022.
112. Cook, E.R. *Methods of Dendrochronology* / E.R. Cook, L.A. Kairiukstis. – Dordrecht: Springer, 1990. – 351 p.
113. Rinn, F. *TSAP-Win: Time Series Analysis and Presentation for Dendrochronology and related Applications* / F. Rinn. – Heidelberg, Germany, Frank Rinn Distribution, 2003.
114. Holmes, R.L. Computer-assisted quality control in tree-ring dating and measurement / R.L. Holmes // *Tree-Ring Bulletin*. – 1983. – Vol. 43. – P. 69-78.
115. Shapiro, S.S. An analysis of variance test for normality (complete samples) / S.S. Shapiro, M.B. Wilk // *Biometrika*. Oxford University Press (OUP), – 1965. – Vol. 52, № 3–4. – P. 591-611.

116. Devey, M.E. A genetic linkage map for *Pinus radiata* based on RFLP, RAPD, and microsatellite markers / M.E. Devey et al.// Theoretical and Applied Genetics. – 1996. – Vol. 92, № 6. – P. 673-679.
117. Parchman, T.L. Genome-wide association genetics of an adaptive trait in lodgepole pine: ASSOCIATION MAPPING OF SEROTINY / T.L. Parchman et al. // Molecular Ecology. – 2012. – Vol. 21, № 12. – P. 2991-3005.
118. Mora-Márquez, F. DDRADSEQTOOLS: a software package for in silico simulation and testing of double-digest RADseq experiments / F. Mora-Márquez et al. // Mol Ecol Resour. – 2017. – Vol. 17, № 2. – P. 230-246.
119. Catchen, J. Stacks: an analysis tool set for population genomics / J. Catchen et al.// Mol Ecol. – 2013. – Vol. 22, № 11. – P. 3124-3140.
120. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data и [Electronic resource] / S. Andrews. – Babraham Bioinformatics. – 2010. – Mode of access: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
121. Ewels, P. MultiQC: summarize analysis results for multiple tools and samples in a single report / P. Ewels et al.// Bioinformatics. – 2016. – Vol. 32, № 19. – P. 3047-3048.
122. Langmead, B. Fast gapped-read alignment with Bowtie 2 / B. Langmead, S.L. Salzberg // Nat Methods. – 2012. – Vol. 9, № 4. – P. 357-359.
123. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers / T. Jombart // Bioinformatics. – 2008. – Vol. 24, № 11. – P. 1403-1405.
124. Kamvar, Z.N. Poppr : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction / Z.N. Kamvar, J.F. Tabima, N.J. Grünwald // PeerJ. – 2014. – Vol. 2. – P. e281.
125. Knaus, B.J. VCFR : a package to manipulate and visualize variant call format data in R / B.J. Knaus, N.J. Grünwald // Mol Ecol Resour. – 2017. – Vol. 17, № 1. – P. 44-53.
126. Hartl, D.L. Principles of population genetics. 4th ed. / D.L. Hartl, A.G. Clark – Sunderland: Sinauer associates, 2007. – 672 p.
127. Excoffier, L. Arlequin suite ver. 3.5: a new series of programs to perform population genetics analyses under Linux and Windows / L. Excoffier, H.E.L. Lischer // Molecular Ecology Resources. – 2010. – Vol. 10, № 3. – P. 564-567.
128. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. 2nd ed. / H. Wickham. – Cham: Springer International Publishing, 2016. – 195 p.
129. Musmann, S.M. AdmixPipe: population analyses in Admixture for non-model organisms / S.M. Musmann et al. // BMC Bioinformatics. – 2020. – Vol. 21, № 1. – P. 337.

130. Bradbury, P.J. TASSEL: software for association mapping of complex traits in diverse samples / P.J. Bradbury et al. // *Bioinformatics*. – 2007. – Vol. 23, № 19. – P. 2633-2635.
131. R Core Team. R: A Language and Environment for Statistical Computing. [Electronic resource] / R Core Team. – Vienna: R Foundation for Statistical Computing, 2013. – Mode of access: <http://www.R-project.org/>.
132. Foll, M. A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective / M. Foll, O. Gaggiotti // *Genetics*. – 2008. – Vol. 180, № 2. – P. 977-993.
133. Privé, F. Performing Highly Efficient Genome Scans for Local Adaptation with R Package pcadapt Version 4 / F. Privé et al.// *Molecular Biology and Evolution*. – 2020. – Vol. 37, № 7. – P. 2153-2154.
134. Storey, J.D. Statistical significance for genomewide studies / J.D. Storey, R. Tibshirani // *Proc. Natl. Acad. Sci.* – 2003. – Vol. 100, № 16. – P. 9440-9445.
135. Caye, K. LFMM 2: Fast and Accurate Inference of Gene-Environment Associations in Genome-Wide Studies / K. Caye et al.// *Molecular Biology and Evolution*. – 2019. – Vol. 36, № 4. – P. 852-860.
136. Capblancq, T. Redundancy analysis: A Swiss Army Knife for landscape genomics / T. Capblancq, B.R. Forester // *Methods Ecol Evol*. – 2021. – Vol. 12, № 12. – P. 2298-2309.
137. Frichot, E. LEA: An R package for landscape and ecological association studies / E. Frichot, O. François // *Methods Ecol Evol*. – 2015. – Vol. 6, № 8. – P. 925-929.
138. Zhou, X. Polygenic Modeling with Bayesian Sparse Linear Mixed Models / X. Zhou, P. Carbonetto, M. Stephens // *PLOS Genetics*. – 2013. – Vol. 9, № 2. – P. 1-14.
139. Clark, K. GenBank / K. Clark et al.// *Nucleic Acids Res.* – 2016. – Vol. 44, № D1. – P. D67-D72.
140. Blum, M. The InterPro protein families and domains database: 20 years on / M. Blum et al.// *Nucleic Acids Res.* – 2021. – Vol. 49, № D1. – P. D344-D354.
141. Gotz, S. High-throughput functional annotation and data mining with the Blast2GO suite / S. Gotz et al.// *Nucleic Acids Research*. – 2008. – Vol. 36, № 10. – P. 3420-3435.
142. Doran, A.G. Snpdat: Easy and rapid annotation of results from de novo snp discovery projects for model and non-model organisms / A.G. Doran, C.J. Creevey // *BMC Bioinformatics*. – 2013. – Vol. 14, № 1. – P. 45.
143. Araki, N.H.T. Phylogeography of *Larix sukaczewii* Dyl. and *Larix sibirica* L. inferred from nucleotide variation of nuclear genes / N.H.T. Araki et al.// *Tree Genetics & Genomes*. – 2008. – Vol. 4, № 4. – P. 611-623.

144. Semerikov, V.L. Mitochondrial DNA in Siberian conifers indicates multiple postglacial colonization centers / V.L. Semerikov et al. // Can. J. For. Res. – 2019. – Vol. 49, № 8. – P. 875-883.
145. Semerikov, V.L. Southern montane populations did not contribute to the recolonization of West Siberian Plain by Siberian larch (*Larix sibirica*): a range-wide analysis of cytoplasmic markers / V.L. Semerikov et al.// Molecular Ecology. – 2013. – Vol. 22, № 19. – P. 4958–4971.
146. Milyutin, L.I. Larch and larch forests of Siberia / L.I. Milyutin, K.D. Vishnevetskaia // U.S. Forest Service. – 1992. – P. 50-53.
147. Putenikhin, V.P. Phenotypic diversity, introgressive hybridization and microevolution of *Larix sukaczewii* Dyl. in the Urals. / V.P. Putenikhin // Gap. INRA, Unité d'Amélioration. – 2002. – P. 25-31.
148. Путенихин, В.П. Микроэволюционные аспекты внутривидовой дифференциации лиственницы Сукачева на Урале / В.П. Путенихин // Хвойные бореальной зоны. – 2003. – Т. 1. – С. 21-27.
149. Gao, J. Functional characterization of the late embryogenesis abundant (LEA) protein gene family from *Pinus tabulaeformis* (Pinaceae) in *Escherichia coli* / J. Gao, T. Lan // Sci Rep. – 2016. – Vol. 6, № 1. – P. 19467.
150. Kneeshaw, S. Nucleoredoxin guards against oxidative stress by protecting antioxidant enzymes / S. Kneeshaw et al.// Proc Natl Acad Sci U S A. – 2017. – Vol. 114, № 31. – P. 8414-8419.
151. Kipreos, E.T. The F-box protein family / E.T. Kipreos, M. Pagano // Genome Biology. – 2000. – Vol. 1, № 5. – P. reviews3002.1.
152. Schröder, F. The extracellular EXO protein mediates cell expansion in *Arabidopsis* leaves / F. Schröder et al.// BMC Plant Biology. – 2009. – Vol. 9, № 1. – P. 20.
153. Klink, V.P. The conserved oligomeric Golgi (COG) complex, a window into plant-pathogen interactions / V.P. Klink et al. // Journal of Plant Interactions. Taylor & Francis, – 2022. – Vol. 17, № 1. – P. 344-360.
154. Domergue, F. Three *Arabidopsis* fatty acyl-coenzyme A reductases, FAR1, FAR4, and FAR5, generate primary fatty alcohols associated with suberin depositio / F. Domergue et al.// Plant Physiol. – 2010. – Vol. 153, № 4. – P. 1539-1554.
155. Di Rubbo, S. The clathrin adaptor complex AP-2 mediates endocytosis of brassinosteroid insensitive1 in *Arabidopsis* / S. Di Rubbo et al. // Plant Cell. – 2013. – Vol. 25, № 8. – P. 2986-2997.
156. Wang, X. AtCSLD3, a cellulose synthase-like gene important for root hair growth in *arabidopsis* / X. Wang et al.// Plant Physiol. – 2001. – Vol. 126, № 2. – P. 575-586.

157. Bartholomew, B. ISWI chromatin remodeling: one primary actor or a coordinated effort? / B. Bartholomew // *Curr Opin Struct Biol.* – 2014. – Vol. 24. – P. 150-155.
158. Wang, K. Two Absciscic Acid-Responsive Plastid Lipase Genes Involved in Jasmonic Acid Biosynthesis in *Arabidopsis thaliana* / K. Wang et al. // *The Plant Cell.* – 2018. – Vol. 30, № 5. – P. 1006-1022.
159. Forsthoefel, N.R. The Arabidopsis Plant Intracellular Ras-group LRR (PIRL) Family and the Value of Reverse Genetic Analysis for Identifying Genes that Function in Gametophyte Development / N.R. Forsthoefel et al.// *Plants (Basel).* – 2013. – Vol. 2, № 3. – P. 507-520.
160. Koh, S. et al. An oligopeptide transporter gene family in Arabidopsis / S. Koh et al.// *Plant Physiol.* – 2002. – Vol. 128, № 1. – P. 21-29.
161. Zheng, C. Genetic adaptation of Tibetan poplar (*Populus szechuanica* var. *tibetica*) to high altitudes on the Qinghai-Tibetan Plateau / C. Zheng et al. // *Ecol Evol.* – 2020. – Vol. 10, № 20. – P. 10974-10985.
162. Zuur, A.F. A protocol for data exploration to avoid common statistical problems: Data exploration / A.F. Zuur, E.N. Ieno, C.S. Elphick // *Methods in Ecology and Evolution.* – 2010. – Vol. 1, № 1. – P. 3-14.
163. Putintseva, Y.A. Siberian larch (*Larix sibirica* Ledeb.) mitochondrial genome assembled using both short and long nucleotide sequence reads is currently the largest known mitogenome / Y.A. Putintseva et al.// *BMC Genomics.* – 2020. – Vol. 21, № 1. – P. 654.
164. Wang, Y. Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in Arabidopsis. / Y. Wang et al. // *Plant Physiol.* – 2008. – Vol. 148, № 3. – P. 1201-1211.
165. Bock, D.G. On the adaptive value of cytoplasmic genomes in plants / D.G. Bock, R.L. Andrew, L.H. Rieseberg // *Molecular Ecology.* – 2014. – Vol. 23, № 20. – P. 4899-4911.
166. Sloan, D.B. Using plants to elucidate the mechanisms of cytonuclear co-evolution / D.B. Sloan // *New Phytologist.* – 2015. – Vol. 205, № 3. – P. 1040-1046.
167. Орешкова, Н.В. Генетическое разнообразие, популяционная структура и дифференциация лиственниц сибирской, Гмелина и Каяндера по данным SSR-маркеров / Н.В. Орешкова, М.М. Белоконь, С. Жамъянсурен // *Генетика.* – 2013. – Т. 49, № 2. – С. 204-213.
168. Duan, F. Overexpression of SoCYP85A1, a Spinach Cytochrome p450 Gene in Transgenic Tobacco Enhances Root Development and Drought Stress Tolerance / F. Duan et al.// *Front. Plant Sci.* – 2017. – Vol. 8. – P. 1909.

169. Rao, M.J. CsCYT75B1, a Citrus CYTOCHROME P450 Gene, Is Involved in Accumulation of Antioxidant Flavonoids and Induces Drought Tolerance in Transgenic Arabidopsis / M.J. Rao et al. // *Antioxidants*. – 2020. – Vol. 9, № 2. – P. 161.
170. Xu, W. Arabidopsis TCH4, regulated by hormones and the environment, encodes a xyloglucan endotransglycosylase. / W. Xu et al. // *Plant Cell*. – 1995. – Vol. 7, № 10. – P. 1555-1567.
171. Zhu, J. A cellulose synthase-like protein is required for osmotic stress tolerance in Arabidopsis: SOS6 is important for osmotic stress tolerance in plants / J. Zhu et al.// *The Plant Journal*. – 2010. – Vol.63, № 1. – P. 128-40.
172. Jinu, J. Dehydration stress influences the expression of brevis radix gene family members in sorghum (*Sorghum bicolor*) / J. Jinu et al. // *Proc.Indian Natl. Sci. Acad.* – 2022. – Vol. 88, № 3. – P. 324-335.
173. Kim, S.J. Suppression of Arabidopsis RING E3 ubiquitin ligase AtATL78 increases tolerance to cold stress and decreases tolerance to drought stress / S.J. Kim, W.T. Kim // *FEBS Letters*. – 2013. – Vol. 587, № 16. – P. 2584-2590.
174. Suh, J.Y. Arabidopsis Tóxicos en Levadura 78 (AtATL78) mediates ABA-dependent ROS signaling in response to drought stress / J.Y. Suh et al. // *Biochemical and Biophysical Research Communications*. – 2016. – Vol. 469, № 1. – P. 8-14.
175. Wang, W. Overexpression of wheat α -mannosidase gene TaMP impairs salt tolerance in transgenic *Brachypodium distachyon* / W. Wang et al. // *Plant Cell Rep.* – 2020. – Vol. 39, № 5. – P. 653-667.
176. Liang, L. O-fucosylation of CPN20 by SPINDLY Derepresses Absciscic Acid Signaling During Seed Germination and Seedling Development / L. Liang et al.// *Front. Plant Sci.* – 2021. – Vol. 12. – P. 724144.
177. Simpson, P.J. Characterization of Two Novel Aldo–Keto Reductases from Arabidopsis: Expression Patterns, Broad Substrate Specificity, and an Open Active-Site Structure Suggest a Role in Toxicant Metabolism Following Stress / P.J. Simpson et al.// *Journal of Molecular Biology*. – 2009. – Vol. 392, № 2. – P. 465-480.
178. Seo, Y.S. Enzymatic characterization of class I DAD1-like acylhydrolase members targeted to chloroplast in *Arabidopsis* / Y.S. Seo et al. // *FEBS Letters*. – 2009. – Vol. 583, № 13. – P. 2301-2307.
179. Bateman, A. The PLAT domain: a new piece in the PKD1 puzzle / A. Bateman, R. Sandford // *Current Biology*. – 1999. – Vol. 9, № 16. – P. R588-S2.
180. Oliw, E.H. Plant and fungal lipoxygenases / E.H. Oliw // *Prostaglandins & Other Lipid Mediators*. – 2002. – Vol. 68–69. – P. 313-323.

181. Al-Saharin, R. Plant E3 Ligases and Their Role in Abiotic Stress Response / R. Al-Saharin, H. Hellmann, S. Mooney // Cells. – 2022. – Vol. 11, № 5. – P. 890.
182. Disch, S. The E3 Ubiquitin Ligase BIG BROTHER Controls Arabidopsis Organ Size in a Dosage-Dependent Manner / S. Disch et al.// Current Biology. – 2006. – Vol. 16, № 3. – P. 272-279.
183. Yuan, C. Genome-Wide Identification and Characterization of HSP90-RAR1-SGT1-Complex Members From Arachis Genomes and Their Responses to Biotic and Abiotic Stresses / C. Yuan et al.// Front. Genet. – 2021. – Vol. 12. – P. 689669.
184. Song, J. The U-box family genes in *Medicago truncatula*: Key elements in response to salt, cold, and drought stresses / J. Song et al. // PLoS ONE. – 2017. – Vol. 12, № 8. – P. e0182402.
185. Mohapatra, M.D. Helicase: A genetic tool for providing stress tolerance in plants / M.D. Mohapatra et al. // Plant Stress. – 2023. – Vol. 9. – P. 100171.
186. Zhang, Z. Adenylsuccinate Synthetase MoADE12 Plays Important Roles in the Development and Pathogenicity of the Rice Blast Fungus / Z. Zhang et al. // JoF. – 2022. – Vol. 8, № 8. – P. 780.
187. Jaimes-Miranda, F. The plant MBF1 protein family: a bridge between stress and transcription / F. Jaimes-Miranda, R.A. Chávez Montes // Journal of Experimental Botany. – 2020. – Vol. 71, № 6. – P. 1782-1791.
188. Rea, G. Copper Amine Oxidase Expression in Defense Responses to Wounding and *Ascochyta rabiei* Invasion / G. Rea et al.// Plant Physiol. – 2002. – Vol. 128, № 3. – P. 865-875.
189. Ambawat, S. MYB transcription factor genes as regulators for plant responses: an overview / S. Ambawat et al. // Physiol Mol Biol Plants. – 2013. – Vol. 19, № 3. – P. 307-321.
190. Gassmann, W. The Arabidopsis RPS4 bacterial-resistance gene is a member of the TIR-NBS-LRR family of disease-resistance genes / W. Gassmann, M.E. Hinsch, B.J. Staskawicz // Plant J. – 1999. – Vol. 20, № 3. – P. 265-277.
191. Liu, H. Overexpression of stress-inducible OsBURP16 , the β subunit of polygalacturonase 1, decreases pectin content and cell adhesion and increases abiotic stress sensitivity in rice / H. Liu et al. // Plant Cell & Environment. – 2014. – Vol. 37, № 5. – P. 1144-1158.
192. Babushkina, E.A. Climatically driven yield variability of major crops in Khakassia (South Siberia) / E.A. Babushkina et al.// Int J Biometeorol. – 2018. – Vol. 62, № 6. – P. 939-948.
193. Xu, H. Decreased vegetation growth in response to summer drought in Central Asia from 2000 to 2012 / H. Xu, X. Wang, X. Zhang // International

- Journal of Applied Earth Observation and Geoinformation. – 2016. – Vol. 52. – P. 390-402.
194. Bagnoli, F. Neutral patterns of genetic variation and applications to conservation in conifer species / F. Bagnoli et al. // Genetics, Genomics and Breeding of Conifers. – New-York: CRC Press and Edenbridge Science Publishers, 2011. – P. 141-195.
195. Neale, D.B. et al. Neutral Genetic Variation // The Conifers: Genomes, Variation and Evolution. – New York: Springer, 2019. – P. 181-224.
196. Krutovsky, K.V. Gene Flow, Spatial Structure, Local Adaptation, and Assisted Migration in Trees / K.V. Krutovsky et al. // Genomics of Tree Crops. – New York: Springer, 2012. – P. 71-116.
197. Semerikov, V.L. Intra- and interspecific allozyme variability in Eurasian *Larix* Mill. species / V.L. Semerikov, L.F. Semerikov, M. Lascoux // Heredity. – 1999. – Vol. 82, № 2. – P. 193-204.

ПРИЛОЖЕНИЕ

Аннотация SNP-маркеров, полученных в рамках исследования адаптации к условиям высокогорья

Скафф олд	Пози ция	Регион	Расстоя ние до ближай шего элемента	Идентиф икатор гена	Термин генной онтологии	Описание	Идентифик атор генной онтологии	Напр авле нность	Кодон	A/K	Синон имичн ость
LS_9849	59355	Экзон		LS_9849-0.1	-	Белок позднего эмбриогенеза D-34	-	+	AA[G/A]	[L/L]	Y
LS_9849	59395	Экзон		LS_9849-0.1	-	Белок позднего эмбриогенеза D-34	-	+	[G/A]TC	[V/I]	N
LS_3113 0	28092	Экзон		LS_31130-0.0	интегрированный компонент мембраны	неизвестный	GO:0016021	+	T[T/C]G	[L/S]	N
LS_3113 0	28096	Экзон		LS_31130-0.0	интегрированный компонент мембраны	неизвестный	GO:0016021	+	[G/A]GT	[G/S]	N
LS_3625 5	2146	Экзон		LS_36255-0.2	-	вероятный нуклеоредоксин 1	-	+	[T/G]CA	[S/A]	N
LS_6383 2	12967	Экзон		LS_63832-0.2	-	-	-	+	TG[T/C]	[C/C]	Y
LS_7303 1	9433	Экзон		LS_73031-0.1	связывание белков	F-box белок, подобный At1g67340	GO:0005515	-	[C/T]AG	[Q/-]	N
LS_7303 1	9438	Экзон		LS_73031-0.1	связывание с белками	F-box белок, подобный At1g67340	GO:0005515	-	T[C/G]C	[S/C]	N
LS_1186 61	524	Экзон		LS_118661-0.1	мембрана	белок подобный EXORDIUM 3	GO:0016020	+	G[C/G]C	[A/G]	N
LS_3984 510	7337	Экзон		LS_3984510-0.1	-	консервативный олигомерный комплекс Гольджи 1	-	-	C[G/A]T	[R/H]	N
LS_4015 301	864	Экзон		LS_4015301-0.0	жирно-ацил-КоА-редуктазная активность (алкоголизация), алкогольобразующая жирно-ацил-КоА-редуктазная активность, процесс	вероятная жирная ацил-КоА-редуктаза 4 изоформа X2	GO:0080019, GO:0102965, GO:0006629, GO:0055114	-	[A/G]CG	[T/A]	N

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин геномной онтологии	Описание	Идентификатор геномной онтологии	Направление	Кодон	A/K	Синонимичность
					метаболизма липидов, окислительно-восстановительный процесс						
LS_4023983	5884	Экзон		LS_4023983-0.1	клатрин-зависимый эндоцитоз, активность адаптера клатрина, адаптерный комплекс AP-2, внутриклеточный транспорт белков	Субъединица комплекса AP-2 альфа-1-подобная	GO:0072583, GO:0035615, GO:0030122, GO:0006886	-	[A/G]CA	[T/A]	N
LS_4033175	6319	Экзон		LS_4033175-0.1	неотъемлемый компонент мембраны, процесс биосинтеза целлюлозы, активность целлюлозосинтазы (УДФ-образующей)	белок E6, подобный целлюлозосинтазе	GO:0016021, GO:0030244, GO:0016760	-	C[G/A]T	[R/H]	N
LS_4078980	3679	Экзон		LS_4078980-0.1	ядро, связывание АТФ, связывание ДНК, активность ДНК зависимой АТФазы, АТФ-зависимое ремоделирование хроматина, связывание нуклеосом, позиционирование нуклеосом	Комплекс ремоделирования хроматина ISWI АТФаза CHR11	GO:0005634, GO:0005524, GO:0003677, GO:0008094, GO:0043044, GO:0031491, GO:0016584	-	TT[A/G]	[L/L]	Y
LS_4839632	266	Экзон		LS_4839632-0.0	оксидоредуктазная активность, действующая на парных донорах, с включением или восстановлением молекулярного кислорода, связывание ионов железа, окислительно-	-	GO:0016705, GO:0005506, GO:0055114, GO:0020037	+	AC[G/A]	[T/T]	Y

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин геномной онтологии	Описание	Идентификатор геномной онтологии	Направление	Кодон	A/K	Синонимичность
					восстановительный процесс, связывание гема						
LS_4839632	287	Экзон		LS_4839632-0.0	оксидоредуктазная активность, действующая на парных донорах, с включением или восстановлением молекулярного кислорода, связывание ионов железа, окислительно-восстановительный процесс, связывание гема	-	GO:0016705, GO:0005506, GO:0055114, GO:0020037	+	CA[C/G]	[H/Q]	N
LS_4851389	400	Экзон		LS_4851389-0.0	-	-	-	-	T[T/C]T	[F/S]	N
LS_5133697	5690	Экзон		LS_5133697-0.1	активность фосфолипазы A1, процесс метаболизма липидов	фосфолипаза A1 PLIP2, хлоропластная	GO:0008970, GO:0006629	-	[C/T]CA	[P/S]	N
LS_5135911	1098	Экзон		LS_5135911-0.1	связывание с белками	растительный внутриклеточный белок LRR 1, связанный с группой Ras	GO:0005515	+	C[C/T]G	[P/L]	N
LS_5188799	2951	Экзон		LS_5188799-0.1	неотъемлемый компонент плазматической мембраны, трансмембранный транспорт олигопептидов, активность трансмембранного транспортера олигопептидов, транспорт белков	Oligopeptide transporter 7	GO:0005887, GO:0035672, GO:0035673, GO:0015031	-	G[T/G]C	[V/G]	N

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин генной онтологии	Описание	Идентификатор генной онтологии	Направленность	Кодон	А/К	Синонимичность
LS_5177694	1419	Межгенный	472	LS_5177694-0.4	-	фактор транскрипции R2R3-MYB MYB4	-	-	-	-	-
LS_4742473	2991	Межгенный	1102	LS_4742473-0.1	эндоплазматический ретикулум, аппарат Гольджи, плазматическая мембрана, активность киназы, фосфорилирование	протеинкиназа	GO:0005783, GO:0005794, GO:0005886, GO:0016301, GO:0016310	-	-	-	-
LS_7946	29200	Межгенный	1228	LS_7946-0.3	связывание ДНК, регуляция транскрипции	домен C35	GO:0003677, GO:0006355	+	-	-	-
LS_4969215	6675	Межгенный	1421	LS_4969215-0.0	связывание ДНК, активность ДНК-связывающего фактора транскрипции, регуляция транскрипции	фактор транскрипции 2	GO:0003677, GO:0003700, GO:0006355	+	-	-	-
LS_4156437	11306	Межгенный	1431	LS_4156437-0.2	активность ДНК-связывающего фактора транскрипции, регуляция транскрипции, ДНК-матрица, связывание ДНК	фактор транскрипции теплового стресса, подобный фактору В-4	GO:0003700, GO:0006355, GO:0043565	-	-	-	-

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин генной онтологии	Описание	Идентификатор генной онтологии	Направленность	Кодон	А/К	Синонимичность
LS_4054	83319	Межгенный	2644	LS_4054-0.1	активность киназы гептапептидного повтора CTD R-полимеразы II, передача сигнала, позитивная регуляция пролиферации клеточной популяции, комплекс циклин-зависимой протеинкиназы голофермента, активность циклин-зависимой протеинсерин/треониновой киназы, фосфорилирование белков, ядро, связывание АТФ, цитоплазма, переход G1/S митотического клеточного цикла, регуляция экспрессии генов, связывание циклина, регуляция перехода G2/M митотического клеточного цикла	циклин-зависимая киназа B1-2-подобная	GO:0008353, GO:0007165, GO:0008284, GO:0000307, GO:0004693, GO:0006468, GO:0005634, GO:0005524, GO:0005737, GO:0000082, GO:0010468, GO:0030332, GO:0010389	+	-	-	-
LS_4806638	6052	Межгенный	2818	LS_4806638-0.0	митохондрия, рибосома, трансляция	рибосомальный белок S7	GO:0005739, GO:0005840, GO:0006412	+	-	-	-
LS_12180	48651	Межгенный	2929	LS_12180-0.1	-	-	-	-	-	-	-
LS_12180	48668	Межгенный	2946	LS_12180-0.1	-	-	-	-	-	-	-
LS_4093180	3370	Межгенный	3467	LS_4093180-0.1	мембрана, трансферазная активность, перенос гликозильных групп	галактан бета-1, 4-галактозилтрансфераза GALS1	GO:0016020, GO:0016757	+	-	-	-

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин геномной онтологии	Описание	Идентификатор геномной онтологии	Направленность	Кодон	A/K	Синонимичность
LS_17999	23782	Межгенный	4288	LS_17999-0.2	активность пероксидазы, реакция на окислительный стресс, окислительно-восстановительный процесс, клеточный процесс, связывание гема	катионная пероксидаза 1	GO:0004601, GO:0006979, GO:0055114, GO:0009987, GO:0020037	+	-	-	-
LS_128740	280	Межгенный	4996	LS_128740-0.0	цитоплазма, связывание ионов железа, оксигеназная активность инозитола, катаболический процесс инозитола, процесс биосинтеза L-аскорбиновой кислоты, окислительно-восстановительный процесс	инозитолоксигеназа 1 изоформа X2	GO:0005737, GO:0005506, GO:0050113, GO:0019310, GO:0019853, GO:0055114	+	-	-	-
LS_22886	63167	Межгенный	5324	LS_22886-0.1	неотъемлемый компонент мембраны, активность эндопептидазы серинового типа, связывание с белками, протеолиз	белок, содержащий убиквитин-ассоциированный домен	GO:0016021, GO:0004252, GO:0005515, GO:0006508	+	-	-	-
LS_4265687	3712	Межгенный	5896	LS_4265687-0.0	активность трансмембранного переносчика гема, митохондрия, неотъемлемый компонент мембраны, сборка цитохромного комплекса, транспорт гема, связывание гема	FN биогенез цитохрома с	GO:0015232, GO:0005739, GO:0016021, GO:0017004, GO:0015886, GO:0020037	+	-	-	-
LS_4447298	1393	Межгенный	9192	LS_4447298-0.0	митохондрия, рибосома, структурный компонент рибосомы, трансляция	рибосомальный белок S11	GO:0005739, GO:0005840, GO:0003735, GO:0006412	+	-	-	-
LS_4447298	1210	Межгенный	9375	LS_4447298-0.0	митохондрия, рибосома, структурный компонент рибосомы, трансляция	рибосомальный белок S11	GO:0005739, GO:0005840, GO:0003735,	+	-	-	-

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин геномной онтологии	Описание	Идентификатор геномной онтологии	Направление	Кодон	A/K	Синонимичность
							GO:0006412				
LS_16638	4951	Межгенный	9836	LS_16638-0.0	процесс катаболизма сахарозы, активность гликопептид-альфа-N-ацетилгалактозаминидазы, цитозоль, активность сахарозо-альфа-глюкозидазы	нейтральная/щелочная инвертаза	GO:0005987, GO:0033926, GO:0005829, GO:0004575	+	-	-	-
LS_16638	4912	Межгенный	9875	LS_16638-0.0	процесс катаболизма сахарозы, активность гликопептид-альфа-N-ацетилгалактозаминидазы, цитозоль, активность сахарозо-альфа-глюкозидазы	нейтральная/щелочная инвертаза	GO:0005987, GO:0033926, GO:0005829, GO:0004575	+	-	-	-
LS_8111	16237	Межгенный	13455	LS_8111-0.3	внутриклеточная анатомическая структура, клеточный цикл	белок семейства циклинов d2	GO:0005622, GO:0007049	+	-	-	-
LS_59565	919	Межгенный	13842	LS_59565-0.1	связывание нуклеиновых кислот, связывание ионов металлов	изоформа X1, подобная белку 53, содержащему домен цинкового пальца CCHH	GO:0003676, GO:0046872	+	-	-	-
LS_27094	20280	Межгенный	14606	LS_27094-0.0	процесс метаболизма углеводов, связывание крахмала, трансферазная активность	4-альфа-глюканотрансфераза DPE2	GO:0005975, GO:2001070, GO:0016740	+	-	-	-
LS_27055	33480	Межгенный	14698	LS_27055-0.1	организация клеточной стенки растительного типа, внеклеточная область	экспансин-A11	GO:0009664, GO:0005576	-	-	-	-
LS_27055	33471	Межгенный	14707	LS_27055-0.1	организация клеточной стенки растительного типа, внеклеточная область	экспансин-A11	GO:0009664, GO:0005576	-	-	-	-
LS_4365068	58380	Межгенный	14743	LS_4365068-0.2	митохондрия, рибосома, структурный компонент	рибосомальный белок S14	GO:0005739, GO:0005840,	+	-	-	-

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин генной онтологии	Описание	Идентификатор генной онтологии	Направленность	Кодон	А/К	Синонимичность
					рибосомы, трансляция		GO:0003735, GO:0006412				
LS_4365068	58331	Межгенный	14792	LS_4365068-0.2	митохондрия, рибосома, структурный компонент рибосомы, трансляция	рибосомальный белок S14	GO:0005739, GO:0005840, GO:0003735, GO:0006412	+	-	-	-
LS_594	14080	Межгенный	14819	LS_594-0.1	связывание нуклеотидов, хеликазная активность	DEADNAbox АТФ-зависимая R-хеликаза 31	GO:0000166, GO:0004386	+	-	-	-
LS_4628586	8618	Межгенный	14825	LS_4628586-0.1	-	-	-	+	-	-	-
LS_1819	118155	Межгенный	17413	LS_1819-0.6	ядро, связывание АТФ, связывание РНК, процессинг 3'-конца РНК, полиаденилирование РНК, процессинг мРНК, активность полинуклеотидаденилаттрансферазы	ядерный белок, подобный полиА-полимеразе 1	GO:0005634, GO:0005524, GO:0003723, GO:0031123, GO:0043631, GO:0006397, GO:0004652	+	-	-	-
LS_1819	118200	Межгенный	17458	LS_1819-0.6	ядро, связывание АТФ, связывание РНК, процессинг 3'-конца РНК, полиаденилирование РНК, процессинг мРНК, активность полинуклеотидаденилаттрансферазы	ядерный белок, подобный полиА-полимеразе 1	GO:0005634, GO:0005524, GO:0003723, GO:0031123, GO:0043631, GO:0006397, GO:0004652	+	-	-	-
LS_6815	43529	Межгенный	17478	LS_6815-0.0	связывание ионов магния, активность альфа-бисаболенсинтазы, активность терпенсинтазы	Цепь А, альфа-бисаболенсинтаза	GO:0000287, GO:0052681, GO:0010333	-	-	-	-

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин геномной онтологии	Описание	Идентификатор геномной онтологии	Направленность	Кодон	A/K	Синонимичность
LS_19054	39882	Межгенный	17695	LS_19054-0.1	связывание АТФ, фосфорилирование белков, активность протеинсерин/треониновой киназы	кокус устойчивости к болезни листовая ржавчина 10, рецептор-подобный протеинкиназе, подобный изоформе 1.2 ХЗ	GO:0005524, GO:0006468, GO:0004674	+	-	-	-
LS_15190	24325	Межгенный	20037	LS_15190-0.1	-	Белок декапирование 5-подобной Х1	-	+	-	-	-
LS_14769	40453	Межгенный	21146	LS_14769-0.2	внутриклеточная анатомическая структура, активность протеинкиназы, связывание АТФ, фосфорилирование белков	казеинкиназа I-подобная	GO:0005622, GO:0004672, GO:0005524, GO:0006468	+	-	-	-
LS_3317	43859	Межгенный	22739	LS_3317-0.4	активность АТФазы, когезия сестринских хроматид, ядро, связывание АТФ, плазмодесма, комплекс Smc5-Smc6, репарация двухцепочечных разрывов посредством гомологичной рекомбинации	структурное поддержание хромосомного белка 5 изоформы Х1	GO:0016887, GO:0007062, GO:0005634, GO:0005524, GO:0009506, GO:0030915, GO:0000724	+	-	-	-
LS_36744	8129	Межгенный	41475	LS_36744-0.1	связывание АТФ, раскручивание дуплекса ДНК, репарация межцепочечных сшивков, репарация двухцепочечных разрывов посредством синтез-зависимого отжига цепи, активность хеликазы четырехстороннего соединения, активность 3'-5' ДНК-хеликазы, связывание ДНК четырехстороннего	АТФ-зависимая ДНК-хеликаза mph1 изоформа Х1	GO:0005524, GO:0032508, GO:0036297, GO:0045003, GO:0009378, GO:0043138, GO:0000400, GO:0071932	+	-	-	-

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин генной онтологии	Описание	Идентификатор генной онтологии	Направленность	Кодон	А/К	Синонимичность
					соединения, реверс репликативной вилки						
LS_36744	8112	Межгенный	41492	LS_36744-0.1	связывание АТФ, раскручивание дуплекса ДНК, репарация межцепочечных сшивок, репарация двухцепочечных разрывов посредством синтез-зависимого отжига цепи, активность хеликазы четырехстороннего соединения, активность 3'-5' ДНК-хеликазы, связывание ДНК четырехстороннего соединения, реверс репликативной вилки	АТФ-зависимая ДНК-хеликаза mph1 изоформа X1	GO:0005524, GO:0032508, GO:0036297, GO:0045003, GO:0009378, GO:0043138, GO:0000400, GO:0071932	+	-	-	-
LS_20619	23839	Межгенный	43483	LS_20619-0.1	-	-	-	-	-	-	-

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин геномной онтологии	Описание	Идентификатор геномной онтологии	Направленность	Кодон	A/K	Синонимичность
LS_29436	10106	Межгенный	45794	LS_29436-0.2	комплекс эукариотического фактора инициации трансляции 3, эукариотический преинициаторный комплекс 43S, эукариотический преинициаторный комплекс 48S, активность фактора инициации трансляции, активность металлопептидазы, связывание фактора инициации трансляции, изопептидазная активность, образование цитоплазматического комплекса инициации трансляции, протеолиз	Субъединица F эукариотического фактора инициации трансляции 3	GO:0005852, GO:0016282, GO:0033290, GO:0003743, GO:0008237, GO:0031369, GO:0070122, GO:0001732, GO:0006508	+	-	-	-
LS_29436	10041	Межгенный	45859	LS_29436-0.2	комплекс эукариотического фактора инициации трансляции 3, эукариотический преинициаторный комплекс 43S, эукариотический преинициаторный комплекс 48S, активность фактора инициации трансляции, активность металлопептидазы, связывание фактора инициации трансляции, изопептидазная активность, образование цитоплазматического	Субъединица F эукариотического фактора инициации трансляции 3	GO:0005852, GO:0016282, GO:0033290, GO:0003743, GO:0008237, GO:0031369, GO:0070122, GO:0001732, GO:0006508	+	-	-	-

Скаффолд	Позиция	Регион	Расстояние до ближайшего элемента	Идентификатор гена	Термин генной онтологии	Описание	Идентификатор генной онтологии	Направленность	Кодон	А/К	Синонимичность
					комплекса инициации трансляции, протеолиз						
LS_22903	58023	Межгенный	54536	LS_22903-0.1	-	-	-	-	-	-	-
LS_10624	85795	Межгенный	65297	LS_10624-0.4	ядро, последовательность-специфическое связывание ДНК, активность ДНК-связывающего фактора транскрипции, специфическая для R-полимеразы II, регуляция транскрипции R-полимеразой II	гомеобокс-лейциновый белок-молния, подобный NAT4	GO:0005634, GO:0043565, GO:0000981, GO:0006357	-	-	-	-