

Министерство науки и высшего образования РФ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
Новосибирский государственный аграрный университет

На правах рукописи

Петров Алексей Фёдорович

**ГЕНЕАЛОГИЧЕСКАЯ И ГЕНЕТИЧЕСКАЯ СТРУКТУРА
ЧЁРНО-ПЁСТРОГО СКОТА ЗАПАДНОЙ СИБИРИ**

**4.2.5. «Разведение, селекция,
генетика и биотехнология животных»**

Диссертация на соискание ученой степени
кандидата биологических наук

Научный руководитель:

доктор биологических наук, доцент,
Камалдинов Евгений Варисович

Новосибирск 2025

ОГЛАВЛЕНИЕ

Введение.....	5
1. Обзор литературы.....	14
1.1 Тенденции развития животноводства в России.....	14
1.2 Исторические аспекты формирования популяции чёрно-пёстрого скота в России и его характеристика.....	15
1.3 Влияние голштинизации.....	17
1.4 Роль генеалогической структуры в совершенствовании генетического потенциала крупного рогатого скота.....	18
1.5 Фенотипические дистанции.....	20
1.6 Генетические дистанции.....	21
1.7 Молекулярно-генетические методы в молочном скотоводстве.....	22
1.8 Микросателлитный анализ.....	24
1.9 Молекулярно-генетические маркеры SNP.....	26
1.10 Прогноз продуктивности животных на основании геномной оценки.....	28
1.11 Оценка генетического потенциала животных с использованием SNP.....	29
1.12 Значение информационных технологий в селекционном процессе молочного скотоводства.....	37
1.13 Особенности работы с большими массивами данных при проведении исследования.....	42
2. Материалы и методы исследований.....	44
2.1 Создание единой региональной базы верифицированных данных.....	44
2.2 Статистический анализ.....	50
2.3 Оценка генеалогической структуры.....	50
2.4 Оценка генетической структуры.....	52
2.4.1 Генетическая структура по микросателлитным локусам.....	52
2.4.2 Кластеризации предприятий по общим быкам, генетическому сходству по микросателлитам и вклад быков производителей в генеалогическую структуру популяции чёрно-пёстрого скота Западной Сибири.....	55

2.4.3	Генетическая структура по SNP.....	56
2.5	Валидация баз данных на примере кластеризации стран по признакам молочной продуктивности.....	58
2.5.1	Кластеризация стран по признакам молочной продуктивности.....	58
2.5.2	Моделирование удоя по исходным данным региональных предприятий.....	59
3.	Результаты исследований.....	62
3.1	Оценка генеалогической структуры.....	62
3.1.1	Оценка качества входных генеалогических данных региональных предприятий.....	62
3.1.2	Генеалогическая структура чёрно-пёстрого скота Западной Сибири...75	
3.1.3	Роль отцовских предков в формировании генеалогической структуры маточного поголовья Западной Сибири.....	79
3.2	Валидация данных первичного цифрового учёта на примере признаков молочной продуктивности предприятия.....	82
3.3	Валидация входных данных первичного цифрового учёта на примере признаков молочной продуктивности зарубежного происхождения.....	86
3.4	Роль паратипических факторов в изменчивости удоя региональной субпопуляции черно-пёстрого скота.....	90
3.5	Генетическая структура популяции по STR-локусам.....	95
3.6	Кластеризация предприятий по используемым быкам и микросателлитам.....	114
3.7	Генетическая структура популяции по SNP-локусам, ассоциированными с признаками молочной продуктивности.....	120
	Обсуждение результатов исследований.....	132
	Заключение.....	145
	Предложения производству.....	147
	Библиографический список.....	148
	Приложения.....	178

Приложение А: Анализ качества входных генеалогических баз данных племенных предприятий Западной Сибири.....	178
Приложение Б: Ранжирование предприятий по качеству генеалогической информации.....	182
Приложение В: SNP ассоциированные с признаками молочной продуктивности.....	186
Приложение Г: Оценка изменчивости и распределения входных данных без предварительной обработки.....	189
Приложение Д: Уникальные генотипы по локусам.....	195

ВВЕДЕНИЕ

В указе Президента Российской Федерации «О Стратегии научно-технологического развития РФ» обозначено 9 приоритетных направлений научно-технологического развития страны, среди которых необходимо выделить переход к высокопродуктивному и экологически чистому агрохозяйству, хранению и эффективной переработке сельскохозяйственной продукции, созданию безопасных и качественных продуктов питания [100]. Достижение поставленных задач невозможно без внедрения современных отечественных цифровых решений и новых наукоемких подходов в селекционно-племенной работе, производстве продуктов питания, совершенствовании геномных технологий [76, 89, 114, 119]. Важную роль при этом играет и системная оценка генетического разнообразия, необходимая для управления генетическими ресурсами и мониторинга [40, 84]. В будущем, все это может привести к снижению зависимости отрасли от импорта племенного материала и обретению технологического суверенитета страны.

Так, в последние годы наблюдается зависимость отрасли молочного скотоводства от деятельности западных компаний, осуществляющих отбор и подбор животных на племенных предприятиях. При таком подходе практически не учитываются хозяйственные условия организаций в разных природно-климатических условиях отдельно взятых регионов и страны в целом. Нет централизованных источников, содержащих всю доступную информацию о генеалогической и генетической структуре популяций скота, инструментов, позволяющих судить о достоверности данных первичного зоотехнического учёта и, как следствие, вести селекционно-племенную работу на предприятиях по выверенным данным [47]. Все это негативно сказывается на качестве генеалогической и генетической структур маточного поголовья и снижает эффективность отечественного племенного животноводства. Этот факт подтверждается практически отсутствием быков отечественной селекции, полученных во втором и последующих поколениях быков в Российской Федерации, которые были бы сопоставимы по качеству с зарубежными.

В селекционных программах, реализуемых зарубежными компаниями, часто не уделяется должного внимания оценке по качеству потомства быков, в которых отдаётся предпочтение иностранным производителям, прошедшим геномную оценку за рубежом с учётом только иностранного маточного поголовья [43, 103, 142]. Это объясняется отсутствием в зарубежных базах данных сведений об отечественном маточном поголовье и быках отечественной селекции, не включённых в международные информационные системы.

Обозначенная проблема усугубляется в регионах с разнообразными природно-климатическими и хозяйственными условиями Западной Сибири, что представляет собой серьёзный вызов для отечественного племенного скотоводства. Это предъявляет особые требования к достоверности генеалогических данных и сопряжённой с ними генетической структурой трансконтинентальных пород. Голштинская порода занимает лидирующие позиции в молочном скотоводстве и определяет подавляющий уровень валового производства молочного сырья, где потенциал дальнейшего роста немыслим без дальнейшей цифровой трансформации отрасли.

К сожалению, в настоящее время генеалогические базы данных на уровне предприятия или региона разобщены и содержат значительное количество ошибок, наиболее распространёнными среди которых являются дублирующиеся записи о животных и заикливания в их родословных. Эти недостатки серьёзно препятствуют точному определению генеалогической и генетической структур, что в свою очередь, снижает точность и достоверность популяционно-генетических исследований.

Таким образом, ведение первичного зоотехнического учёта с применением отечественных цифровых технологий и формирование единой региональной электронной базы данных фенотипов и генотипов сельскохозяйственных животных в качестве надёжного и верифицированного источника данных о генеалогической и генетической структуре в том числе для разрабатываемой федеральной государственной информационно-аналитической системы племенных ресурсов (ФГИАС ПР), является ключевым шагом. Это позволит кардинально улучшить качество ге-

неалогической информации и понимание генетической структуры региональной популяции скота на новом уровне. Именно это понимание создаст прочный фундамент для развития отечественной цифровой крупномасштабной селекции, обеспечит её эффективность и способствует обретению технологического суверенитета в профильной области.

Степень разработанности темы исследований. Проблема повышения эффективности молочного скотоводства [2, 13, 34, 67, 154, 178, 210], и в частности чёрно-пёстрого скота [30, 57, 129], является предметом активных исследований во всем мире. Накоплен значительный объем знаний по различным аспектам этой проблемы, от традиционной селекции до современных геномных технологий [181, 219].

Существует немало работ, посвящённых оценке влияния голштинизации, роли генеалогической структуры в совершенствовании генетического потенциала крупного рогатого скота и методам оценки племенной ценности на основе анализа родословных. Значительный прогресс, достигнут в области молекулярно-генетических методов. Так, микросателлитный профиль [44, 157, 259] и SNP-маркеры [64, 154, 178, 253] широко используются для изучения генетической структуры популяций, определения родства и идентификации генов, ассоциированных с хозяйственно-полезными признаками. Но, несмотря на значительный прогресс в каждом из перечисленных направлений, существует потребность в комплексном подходе, интегрирующем традиционные методы селекции с современными геномными технологиями и информационными решениями. Особую актуальность приобретают исследования, направленные на адаптацию существующих методик к специфическим условиям различных регионов и пород, а также на развитие методов верификации и валидации данных первичного зоотехнического учёта.

Создание единой региональной базы данных фенотипов и генотипов чёрно-пёстрого скота по популяциям Сибири также является актуальной, так как впервые позволит провести комплексный анализ генетической и генеалогической структуры популяции в целом на основе верифицированных данных. Это станет основой для построения перспективных математических моделей для прогноза

уровней селекционных признаков животных, основанных на выверенных данных с учётом широкого спектра региональных и производственных факторов.

Цель и задачи исследования. Цель – оценка генеалогической и генетической структуры чёрно-пёстрого скота Западной Сибири по верифицированным данным первичного зоотехнического учёта для разработки научно-обоснованных основ региональных программ крупномасштабной селекции. В соответствии с поставленной целью в работе были обозначены следующие задачи:

- разработать и внедрить комплексную методологию верификации и валидации данных первичного зоотехнического учёта в молочном скотоводстве Западной Сибири, включающую математическую модель прогнозирования удоя и обеспечивающую формирование достоверной и научно-обоснованной базы для генеалогического и генетического анализа;
- выполнить комплексную оценку несмещённой генеалогической структуры популяции чёрно-пёстрого скота на региональном уровне, включая уровень её дискретности, для определения оптимальных стратегий управления генетическими ресурсами;
- провести изучение несмещённой генетической структуры популяции чёрно-пёстрого скота по STR-локусам с использованием больших массивов верифицированных данных для оценки генетического разнообразия и выявления особенностей популяции скота Западной Сибири;
- проанализировать несмещённую генетическую структуру популяции по SNP-локусам, ассоциированным с признаками молочной продуктивности для совершенствования методов оценки племенных качеств молочного скота в рамках региональных программ крупномасштабной селекции, в том числе геномной селекции.

Научная новизна исследований. Впервые для чёрно-пёстрого скота Западной Сибири в комплексе проанализированы генеалогическая и генетическая структуры популяции на региональном уровне с использованием уточнённых (верифицированных) данных первичного зоотехнического учёта, что обеспечивает целостное представление о генетических ресурсах региона для разработки про-

грамм крупномасштабной селекции. Разработана и внедрена комплексная методология верификации и валидации данных, обеспечивающая получение достоверной основы для последующего генетического анализа и принятия обоснованных селекционных решений. Создана региональная база данных верифицированных генотипов и фенотипов чёрно-пёстрого скота, решающая проблему цифровой идентификации животных для развертывания региональных селекционных программ. Продемонстрировано повышение точности математических моделей прогнозирования удоя при использовании уточнённых данных, что обосновывает их практическое применение для повышения эффективности селекционных программ в молочном скотоводстве Западной Сибири.

Теоретическая значимость заключается в расширении и углублении знаний о несмещенных генеалогической и генетической структурах поголовья чёрно-пёстрого скота Западной Сибири, а также об изменчивости признаков молочной продуктивности. Разработка и апробация комплексной методологии валидации и верификации больших объемов входных данных первичного зоотехнического учёта, основанной на программных алгоритмах, математическом моделировании и многомерном статистическом анализе, вносит вклад в развитие теории и практики управления качеством данных в животноводстве. Полученные результаты и разработанная методология формируют научно-теоретическую базу для дальнейших исследований в области геномной селекции молочного скота. Это важно для разработки региональных программ, направленных на повышение продуктивности и адаптивности скота в Западной Сибири и их дальнейшего масштабирования в РФ и странах ЕАЭС.

Практическая значимость работы определяется созданием прикладных инструментов и методических решений для существенного повышения эффективности селекционно-племенной работы с чёрно-пёстрым скотом Сибири. В практику племенных и товарных предприятий, а также региональных информационно-селекционных центров, предлагается к внедрению разработанное прикладное программное обеспечение, верифицированная региональная база данных и математическая модель изменчивости уровня удоя. Интеграция данных разработок в селек-

ционный процесс обеспечит повышение точности оценки племенной ценности животных, целенаправленную оптимизацию и совершенствование региональных селекционных программ.

Методология и методы исследования

Методологической основой послужили работы отечественных и зарубежных исследователей в релевантных областях науки. Генеалогическая и генетическая структура изучалась на основе записей о животных в базах данных предприятий Сибирского федерального округа (СФО). Сведения о фенотипах и генотипах верифицировались согласно принципам многомерной статистики, биоинформатики и программирования. Данные для изучения генеалогической структуры были дополнены информацией из зарубежных открытых источников. В основе генетических исследований были использованы микросателлитные профили (STR-локусы) и однонуклеотидный полиморфизм (SNP) черно-пёстрого скота Западной Сибири.

Основные положения, выносимые на защиту

- Впервые предложенный комплекс методов верификации данных первичного зоотехнического учёта объединяет программные алгоритмы цифровой идентификации (с привлечением всех доступных источников), методы многомерной статистики и многоуровневую визуализацию генеалогических деревьев с использованием массивов больших данных. Данный комплекс обеспечивает формирование достоверной основы для популяционно-генетических исследований молочного скота в Сибирском федеральном округе (СФО), качество которой превосходит традиционные подходы, опирающиеся на не верифицированные данные.
- Комплексная оценка несмещённой генеалогической структуры популяции чёрно-пёстрого скота Западной Сибири, проведённая на региональном уровне, позволила не только выявить, но и количественно оценить вклад ведущих быков-производителей в формирование регионального маточного поголовья и его генеалогическую дискретизацию, что является важным при оптимизации селекционных программ.

- Изучение несмещённой генетической структуры популяции чёрно-пёстрого скота Западной Сибири по STR-локусам на основе верифицированных данных показало сохранение высокого уровня генетического разнообразия, несмотря на интенсивную голштинизацию, с выраженным уровнем полиморфизма по локусам TGLA122, TGLA227 и TGLA53, что свидетельствует об особенностях её генетической структуры.
- Анализ несмещённой генетической структуры популяции по SNP-локусам, ассоциированными с признаками молочной продуктивности, позволил для региональной популяции выявить высокое значение генетического разнообразия Нея по полиморфизму rs109421300, расположенному рядом с геном DGAT1. Данный ген связан с выходом жира, молока, процентами жира и белка, что открывает перспективы для использования данного SNP в геномной селекции черно-пёстрого скота Западной Сибири.

Степень достоверности и апробация результатов исследований.

Проводимые исследования были основаны на использовании большого массива данных первичного электронного зоотехнического учёта (более 6 млн. записей о происхождении животных и показателях молочной продуктивности из региональных и международных баз данных), которые включали региональные микросателлитные профили и данные однонуклеотидных полиморфизмов (SNP) с использованием чипов плотностью 50К. Все полученные массивы данных подвергались многосторонней верификации и валидации. Верификация проводилась путём сопоставления полученных результатов с данными из открытых источников (CDN, CDCB, базы данных ВНИИплем в открытом доступе), а валидация – посредством статистического анализа и многомерного моделирования с последующей интерпретацией результатов. Использование такого подхода позволило уточнить генеалогическую и генетическую структуру чёрно-пёстрого скота Западной Сибири.

Представленные результаты диссертационного исследования были доложены и обсуждены на следующих конференциях и мероприятиях: круглый стол "Проекты для рынка Фуднет" (в рамках цикла мероприятий "Что такое рынки

НТИ") (6 октября 2022), "SibAgroTech 2022" - I Научно-практическая конференция (Новосибирск 2022), «Биотехнология в растениеводстве, животноводстве и сельскохозяйственной микробиологии XXIII». Материалы 23-ей всероссийской молодежной научной конференции (Москва 2023), Всероссийская научно-практическая конференция «Инновационные методы и технологии повышения продуктивности в молочном животноводстве» (Барнаул 2024), Сборник трудов научно-практической конференции преподавателей, аспирантов, магистрантов и студентов Новосибирского ГАУ Выпуск 9 (Новосибирск 2024), Conference Proceedings: Global Food Forum 2021. Dialogue without borders (Бишкек, 2022).

Публикации результатов исследования. По теме диссертационного исследования было опубликовано 7 работ. В рецензируемых научных журналах, рекомендованных ВАК - 5 статей в журналах «Вестник НГАУ» и «Животноводство и кормопроизводство»; работа в издании «Veterinary World» из списка «Белого списка»; получено свидетельство о регистрации базы данных.

Структура и объем работы. Диссертация представлена на 196 страницах машинописного текста, содержит 14 таблиц, 30 рисунков и 5 приложений. В библиографический список включено 260 источников, из которых отечественных 137 и зарубежных 123.

Текст диссертации включает введение, обзор литературы, материалы и методы исследования, результаты исследований, обсуждения результатов, предложения, библиографический список и приложения.

Личный вклад автора. Автор принимал непосредственное участие в разработке целей и задач, осуществлял отбор животных для геномной оценки, сбор биологического материала в ходе выездных работ, обработку и статистический анализ полученных результатов и их публикацию. Написал программное обеспечение, которое позволило объединить локализованные массивы с данными о животных из разных сельскохозяйственных предприятий в единую базу данных фенотипов и генотипов западно-сибирского чёрно-пёстрого скота с уточнением родословной по иностранным источникам и алгоритмам верификации. Построил

модели по признакам молочной продуктивности на основе уточнённых данных и доложил результаты исследования на научных конференциях.

Благодарности. Автор выражает благодарность научному руководителю доктору биологических наук, доценту Камалдинову Евгению Варисовичу за помощь, поддержку и руководство на протяжении всего исследования. Гарту Владимиру Владимировичу, доктору сельскохозяйственных наук, доценту и Куликовой Светлане Геннадьевне, доктору биологических наук, профессору за их опыт и неоценимые советы в ходе реализации поставленных задач. Автор благодарит АО «Новосибирскагроплем» – предприятие по племенной работе в Новосибирской области, ЗАО племзавод «ИРМЕНЬ», СПК Береговой, ООО "Мираторг-Генетика – за всестороннее сотрудничество на протяжении всего исследования. Благодарность выражается коллективу лаборатории прикладной биоинформатики ФГБОУ ВО Новосибирский ГАУ за возможность использовать вычислительные мощности серверного оборудования для вычислений необходимых при выполнении поставленных задач. Искренняя признательность также выражается коллективу кафедры ветеринарной генетики и биотехнологии за ценные советы и содействие в ходе выполнения работы.

1. ОБЗОР ЛИТЕРАТУРЫ

1.1 Тенденции развития животноводства в России

Разнообразие пород крупного рогатого скота и их генетический потенциал являются ключевыми элементами устойчивого развития молочного скотоводства в России и обеспечивают продовольственную безопасность в стране. Наблюдаемый в последние несколько десятилетий ориентир производства в сторону увеличения показателей молочной продуктивности негативно отразился на генетическом разнообразии пород скота молочного направления продуктивности [5, 40, 84, 95, 124] и привёл к планомерному увеличению генетического сходства существующих популяций [83]. Наряду с сокращением генетических ресурсов крупного рогатого скота в отрасли наблюдается критическая зависимость от импорта племенного материала [45], технологий и оборудования. Чтобы эффективно реагировать на современную конъюнктуру рынка [1, 3] важно изучать генетический потенциал, как локальных пород [120], так и иметь представление о генофонде распространённых в нашей стране трансконтинентальных и аборигенных пород крупного рогатого скота на региональном и федеральном уровнях [93].

Согласно ФГБНУ ВНИИплем на период 1 января 2024 года общее поголовье крупного рогатого скота в России составило 17 млн. 67,9 тыс. гол. Из них на долю сельскохозяйственных организаций приходилось 7 млн. 801,9 тыс. гол. В хозяйствах населения сосредоточилось порядка 6 млн. 374,6 тыс. гол. и 2 млн. 891,4 тыс. гол. в крестьянских (фермерских) хозяйствах и у индивидуальных предпринимателей [67]. В соответствии с новой методикой принятой Евразийской экономической комиссией (ЕЭК) по определению породности [99], в результате породной инвентаризации скота, было установлено, что только за один 2023 год численность голштинской породы выросла на 10,05% и составила 64,37% от общей численности скота молочного направления в России. В период с 2010 по 2023 гг. поголовье голштинской породы выросло более чем на 59,72% и данная порода получила самое большое распространение по всей территории Российской Федерации [67]. Рост численности голштинской породы и рост долей кровности

по ней [32, 60] вызывает особую озабоченность в реалиях современного животноводства у ведущих специалистов отрасли.

Одним из факторов, который сдерживает развитие отечественных популяций скота, является недостаточное количество как самих быков-производителей отечественной селекции, так и маточного поголовья с высоким генетическим потенциалом. Влияние голштинской крови на продуктивные качества, а также конституционные и экстерьерные особенности чёрно-пёстрого скота изучается в различных регионах России [68, 77, 78], но необходимость продолжения такой работы с вовлечением отечественных пород продолжает вызывать озабоченность.

1.2 Исторические аспекты формирования популяции чёрно-пёстрого скота в России и его характеристика

В Россию завоз чёрно-пёстрого скота из Голландии начался при Петре I и продолжался до первой половины 20 столетия. Проводились скрещивания голландского скота с местным, а полученные помеси разводились «в себе». Так, в Поволжье велись попытки выведения новой породы скота, на основе голландского чёрно-пёстрого, ещё до 30-х годов 20-го века, но, несмотря на ведение племенной книги, порода так и не была зарегистрирована. Известно, что в племенную книгу голландского чёрно-пёстрого скота Поволжья в период с 1916 по 1925 г. было внесено порядка 87% общего числа поголовья этого скота, что составляло порядка 1500 особей [12].

После Октябрьской революции ведением племенной книги по голландскому чёрно-пёстрому скоту в Поволжье занялся Энгельсовский госплемрассадник, который в 1955 году был переименован в «Краснокутский» и ещё ряд племенных предприятий Саратовской области. Вместе с тем, наряду с совершенствованием показателей продуктивности, росла и численность чёрно-пёстрого скота относительно других популяций этого вида сельскохозяйственных животных как в этом, так и в других регионах нашей страны.

Согласно Герчикову Н.П. [24], в некоторых областях к 50-м годам чёрно-пёстрый скот составлял около 75,5 % от общего поголовья скота. В это же время

путём скрещивания аборигенных пород с остфризским и голландским скотом создавались новые типы и породные группы чёрно-пёстрого скота [102, 123, 135], которые в 1959 г. были объединены в одну чёрно-пёструю породу [9, 126]. Было принято решение о разведении чёрно-пёстрого скота на всей территории Советского союза. Для этой цели во многих областях создавались племенные рассадники, племенные фермы и племенные совхозы. На территории Новосибирской области был основан Барабинский племенной рассадник [10].

Разнообразие исторических и природно-климатических условий, различия в кормовой базе и условиях содержания от региона к региону привели к разнообразным результатам в селекции чёрно-пёстрой породы. Но, несмотря на это, породе свойственен ряд характерных особенностей, среди которых наличие средней длины шеи, удлинённой головы, ровные спина, поясница и холка, глубокая грудная клетка, крестец, конституция крепкая, плотная, реже грубая или нежная. Конечности крепкие, правильно поставленные. Вымя хорошо прикреплено к брюшной стенке и характеризуется выраженной чашеобразной, округлой формой [94, 96]. Передние соски поставлены нормально, задние часто сближены. Скот имеет высокие показатели молочной продуктивности, по сравнению с другими породами [16, 52–54], приспособляемость к разнообразным условиям содержания в промышленных комплексах и удовлетворительными мясными качествами [31, 36, 61].

Изучению становления чёрно-пёстрой породы в России посвящено множество работ. Так, Костомахин Н.М. [58] выделил следующие периоды:

I период (1930 – 1940 гг.). Распространение поглотительного скрещивания местных пород с чёрно-пёстрым скотом, который завозился из Европы.

II период (1940 – 1945 гг.). Ввоз в страну быков немецкого и шведского происхождения для использования на маточном поголовье, которое было получено путём скрещивания с производителями, ранее импортированными из Германии и Прибалтики. Начало разведения полученных помесей «в себе».

III период (1945 – 1975 гг.). Поступление производителей из Сибири и Прибалтики, позднее животных из Швеции, в центральные области страны, постра-

давших от военных действий. Выделение из большого массива отечественного чёрно-пёстрого скота нескольких популяций (среднерусский, сибирский, уральский и др.) в общую чёрно-пёструю породу и утверждение её в 1959 г. Третий период также характеризовался импортом животных для пополнения поголовья чёрно-пёстрой породы из стран Западной и Восточной Европы (Швеции, Нидерландов, Дании, Германии и Польши).

IV период (с 1975 г. по настоящее время). Начало ввоза в страну племенного материала из США и Канады. Продолжение ввоза маточного поголовья и семенного материала из стран Западной Европы. Повсеместная голштинизация чёрно-пёстрого скота.

Как было отмечено ранее, в настоящее время интенсификация молочного скотоводства в России продолжается путём введения в эксплуатацию высокопродуктивных пород. Среди таких пород особое место занимает голштинский скот американского и канадского происхождения. Использование генофонда этой породы обусловлено самыми высокими показателями молочной продуктивности, а также рядом технологических качеств (высокая степень адаптации к различным климатическим условиям и т.д.). Существенное влияние на развитие молочного скотоводства оказывают улучшение уровня кормления и содержания животных, использование быков-производителей, оцениваемых по качеству потомства и ряд других экономических и организационных мероприятий. Выполнение обозначенных выше условий позволяет добиться максимальной реализации заложенного генетического потенциала [62, 144, 187, 214].

1.3 Влияние голштинизации

Как отмечалось ранее, в последние годы вектор отбора чёрно-пёстрого скота направлен на увеличение показателей молочной продуктивности. Одним из движущих элементов, позволяющих достичь высоких показателей, является прилитие крови голштинской породы, посредством массового использования семени голштинских быков иностранной селекции с целью совершенствования местного маточного поголовья. Наряду с увеличением показателей удоя [12, 34, 129], на-

блюдается негативное влияние на продуктивное долголетие скота и период его хозяйственного использования [30, 101, 112, 117, 136].

Проводимое в последние десятилетия поглотительное скрещивание местного маточного поголовья быками голштинской породы, привело к росту генетического сходства и уменьшению генетической изменчивости популяций чёрно-пёстрого скота в нашей стране [69, 72]. Такая тенденция продолжает сохраняться и вызывает беспокойство. Опасность потери генетического разнообразия может привести к планомерному увеличению инбридинга и росту частоты проявления ряда моногенных заболеваний. В работе Н.А. Мартынова [69] отмечается, что вместе со снижением генетической изменчивости в популяциях, возрастает вероятность случайного изменения аллельных частот. Увеличить своё распространение могут аллели, ассоциированные с аутосомно-рецессивными заболеваниями. Практика животноводства показывает, что в отсутствие контроля частотный спектр рецессивных аллелей, ассоциированных с заболеваниями, может превышать 10%. Примером этого является цитруллинемия, которая описана в статьях Робинсоном Дж. с соавторами [236, 237], первый голштинский гаплотип фертильности (НН1) в канадской популяции [159] или дефицит уридинмонофосфатсинтазы (DUMPS) [198]. Повышение уровня генетического груза оказывает значительное влияние на рентабельность животноводства, поэтому исследования и контроль распространения за генетическими нарушениями в популяциях крупного рогатого скота должен являться неотъемлемой частью селекционных процессов [41].

1.4 Роль генеалогической структуры в совершенствовании генетического потенциала крупного рогатого скота

Комплекс взаимосвязанных мероприятий, направленных на улучшение племенных и продуктивных качеств животных находится в основе селекционно-племенной работы. Селекция направлена на получение животных, обладающих ценными племенными достоинствами. Отличительной особенностью таких животных является не только экстерьер, интерьер, конституция и продуктивность, но и

способность передавать ценные качества своему потомству, тем самым оказывая улучшающее влияние на стадо или породу в целом. Отбор и подбор позволяют сформировать однородную по типу телосложения и направлению продуктивности группу животных.

Ранее, при изучении крупного рогатого скота в России было принято разделять линии на заводские и генеалогические. В настоящее время генеалогические линии играют существенную роль при подборе родительских пар, где выдающийся бык производитель иностранной селекции с высоким генетическим потенциалом представляет особый интерес. В связи с активным использованием поглотительного скрещивания в течение последних десятилетий роль заводских линий значительно снизилась и представляет интерес преимущественно при разведении скота отечественных пород [104].

Бык-производитель является основным движущим структурным элементом в породе и выделяется среди своих сверстников своими потомками по уровню селекционных признаков. Если рассматривать породу в целом, то ветви, которые формируют её типичность немногочисленны и связаны кличками нескольких выдающихся предков. Изучению генеалогических линий и ветвей чёрно-пёстрого скота посвящено множество работ [14, 42, 57, 81, 106, 127, 130]. На основании этих работ можно сделать вывод, что в современной России широко распространены высокоценные производители, которые являются сыновьями и внуками лидеров голштинской породы Пикланда МГФ 393 и Гвидона МГФ 394 линии Монтвик Чифтейна; Сувенира МГФ 195 и Боншаса МГФ 361 линии Рефлекшн Соверинга. В последние годы замечено резкое сокращение (либо полное отсутствие) популяции животных от Шквала МГФ 386 и Стингера МГФ 422 линии Силинг Трайджун Рокит.

Направленная селекционно-племенная работа, с применением достаточного количества производителей, оказывает влияние на продолжительность хозяйственного использования и заводской линейной структуры. Подбор позволяет получить более ценных производителей согласно чётким критериям отбора.

1.5 Фенотипические дистанции

Исследования в биологии часто сопровождаются широким применением дискриминантного и кластерного анализа с вычислением фенотипических расстояний. Известны работы по изучению видов ив лесной зоны Западной Сибири [6, 20–22]. В представленных работах применяется дискриминантный анализ и рассчитываются межвидовые фенотипические дистанции по комплексу морфометрических признаков листа для изучения ареала распространения популяций ив. Существуют работы, в которых для оценки популяционной структуры ели применяются дистанции Махаланобиса. В работах Попова П.П. [91, 92] межвидовой ареал распространения ели сибирской и европейской исследуется с применением квадратов дистанций Махаланобиса (SMD) крайних вариантов популяций (в пределах от 0 до 100). На основании полученных результатов строятся кластеры, которые позволяют судить о распространении видов ели по различным признакам. Известны публикации в которых для изучения генетической и фенотипической дивергенции популяций растений также проводится по фенотипическим и генетическим дистанциям [29, 105, 132, 148].

В сельском хозяйстве дистанции применяются для изучения сортов при гибридизации. К примеру, Евклидовы расстояния используют для определения взаимосвязи между родительскими сортами при получении сорта гибрида. Такие метрики приведены в работах по изучению эфиопской горчицы [246], подсолнечника [158], проса [177] и др. видов растений сельскохозяйственного назначения. Широкое распространение в изучении популяций и генотипа культурных сортов растений и пород животных получили расстояния Махаланобиса, которые представляют собой меру расстояния между векторами случайных величин, обобщающую понятие Евклидово расстояние [226, 173, 204, 255]. Среди отечественных публикаций следует отметить работу Е.В. Камалдинова [48], где им был предложен метод расчёта межпородных фенотипических дистанций по комплексу биохимических показателей у популяций свиней в Западной Сибири. Автор отмечает, что для учёта сложных корреляций между изучаемыми признаками, недостаточно

использования распространённых одномерных статистических методов [15, 70, 85, 118], которые классически применяются при статистической обработке биохимических и любых других количественных признаков с высоким уровнем изменчивости. Чтобы избежать появления смещённой оценки, из-за ограниченного понимания изучаемого процесса, автором предлагается применять многомерные критерии, которые при наличии большого количества многомерных первичных данных позволяют исследователю делать выводы на качественно новом уровне [46, 51].

1.6 Генетические дистанции

Наряду с фенотипическими дистанциями широкое распространение в популяционных исследованиях получили генетические дистанции, которые также называются генетическими расстояниями. Генетические дистанции позволяют определить генетическое различие (дивергенцию) между видами, подвидами или популяциями одного вида [23]. Генетическое сходство тем больше, чем меньше генетическое расстояние. Так в простейшем случае разница в частотах определенного признака между двумя или более популяциями одного вида, могут быть представлены как генетические дистанции между ними.

Для расчёта генетических расстояний могут применяться различные методы. Одним из самых распространённых методов расчёта генетических дистанций является метод Нея [65, 220], который учитывает частоты аллелей в популяциях и сравнивает их друг с другом. Полученное значение отражает глубину генетического расстояний между сравниваемыми популяциями [258]. Ещё одним методом расчёта генетических дистанций между популяциями на основе частот аллелей является расстояние Эдвардса [157, 161], которое вычисляется как угловое значение между векторами, характеризующими частоты аллелей в популяциях в многомерном пространстве. Чем меньше вычисленное угловое расстояние, тем больше генетическое сходство между изучаемыми популяциями и, напротив, большее значение углового расстояния указывает на большую генетическую дивергенцию между ними. В 1983 году Рейнольдс с соавторами [234] предложили

меру генетического расстояния, которая была основана на генетической ковариации между популяциями и вычислялась как мера различия в частотах аллелей между популяциями, но скорректированная с учётом случайного генетического дрейфа. Данный метод расчёта генетического расстояния получил название расстояния Рейнольдса и, соответственно, был охарактеризован как способ вычисления генетического расстояния на основе различий в частотах аллелей, корректируя эти различия на генетический дрейф. Практическим инструментом для оценки генетической дифференциации между популяциями может быть и генетическое расстояние Роджерса [238], которое основано на простой, но эффективной идее вычисления разницы в частотах аллелей посредством измерения евклидовых расстояний между точками в многомерном пространстве, представляющими собой изучаемые популяции. Альберто Провести [230] предложил метод расчёта меры генетического расстояния, но уже не основанного на евклидовой метрике. Основная идея расстояния Провести заключается в том, чтобы измерить суммарную разницу в частотах аллелей между двумя популяциями. Внимание акцентируется на абсолютных значениях разностей частот, а не на квадратах разностей (как в евклидовом пространстве), что позволяет делать расчёт менее чувствительным к экстремальным значениям и, в некоторых случаях, более устойчивым к влиянию редких аллелей.

1.7 Молекулярно-генетические методы в молочном скотоводстве

Разведение животных является сложным процессом, включающим в себя множество аспектов, в том числе и молекулярно-генетические технологии и методики их применения [88]. До недавнего времени отбор осуществлялся только с использованием фенотипических признаков, которые можно рассматривать как конечный результат реализации генотипа в определённых условиях среды. По таким показателям оценивали эффективность селекции и принимали решения о дальнейших мероприятиях, направленных на повышение продуктивных и других качеств потомства животных. Идеальным сценарием можно рассматривать возможность предсказания генетического потенциала на основе имеющейся наслед-

ственной информации, с учётом многочисленных фенотипических данных. Современные генетические технологии позволяют более точно судить о племенном потенциале в отношении признаков отбора, репродуктивных качествах, показателях здоровья и генеалогическом разнообразии с привлечением геномных технологий. Такие технологии позволяют проводить раннюю оценку и отбор животных с высоким генетическим потенциалом продуктивности [89, 137].

Фенотипические показатели всегда играли и продолжают играть ключевую роль в селекционно-племенной работе, позволяя осуществлять постоянный мониторинг и обновлять существующие биометрические модели для сохранения консолидированности популяций в соответствии с целями отбора. Следует отметить наличие большого числа работ по фенотипам, в которых авторы исследуют связи между линейными признаками экстерьера и продуктивных качеств животных [18; 19, 66, 218].

Сбалансированное использование информации о геноме и данных о фенотипе позволяет значительно увеличить точность прогноза.

За последние 2-3 десятилетия методы молекулярной генетики позволили значительно повысить эффективность селекции [40, 96, 111]. Перед практической селекцией поставлена задача ранней оценки генетического потенциала животных по ряду хозяйственно-полезных признаков. Для поиска генов, которые отвечают за такие признаки и по которым следует вести отбор, стали применять ДНК-технологии с маркировкой локусов количественных признаков, а также для реализации отбора с использованием маркеров. Такой подход существенно повысил эффективность процесса селекции [7, 17, 63, 116, 121, 140].

Прогресс развития ДНК-технологий продолжается. В настоящее время ДНК-маркеры уже принято относить к третьему поколению. Белковые маркеры из-за стремительного развития ДНК-технологий постепенно вытесняются из популяционной генетики. Основной акцент делается на исследовании полиморфизма на геномном уровне, а не на только уровне продуктов экспрессии гена [184, 232].

Известны работы, где исследовалась связь генетических маркеров с показателями, отвечающими за молочную продуктивность чёрно-пёстрого скота. Так, в работе О.В. Костюниной, А.М. Бакай, Г.А. Бушовой с коллегами [59] установлен полиморфизм генов κ-казеина (CSN3), α-лактальбумина (LALBA), диацил О-ацилтрансферазы (DGAT) и гормона роста (BGH). Особи с генотипом КК/LL по DGAT1/BGH обладали наименьшим уровнем удоя, белка и жира за лактацию. Напротив, положительная тенденция по содержанию жира в молоке наблюдалась у животных, позитивных по аллелю К по DGAT1, а величина удоя возрастала у особей при наличии аллеля V по BGH.

В последние годы селекционеры и генетики получили возможность получать результаты генотипирования животных по сотням тысяч и даже нескольким миллионам нуклеотидов, полученных с помощью полногеномного секвенирования и ДНК-чипов. Это в совокупности позволяет достичь высокой степени точности биометрических моделей, но и приводит к увеличению сложности обработки огромных входных объёмов данных с применением современных инструментов биоинформатики. Высокие затраты на генотипирование также ограничивают крупномасштабное применение таких подходов. Особую озабоченность в этом аспекте вызывают селекционные программы популяций местных пород. Если требуется лишь подтверждение происхождения или изучение различий между популяциями с помощью дистанций, то для этих целей может быть достаточно микросателлитного анализа.

1.8 Микросателлитный анализ

Развитие биотехнологии и генетики привели к возможности исследований посредством микросателлитов. В молочном скотоводстве микросателлитные исследования позволяют подтвердить с высокой точностью достоверность происхождения, выявить генетические различия между популяциями и оценить её структуру, тем самым помогая в селекции и планировании направления племенной работы [43, 44]. Так, благодаря микросателлитному анализу, Л.К. Эрнст с соавторами [134] изучали процесс по голштинизации местных аборигенных пород. В сво-

ей работе авторы пришли к выводу, что голштинизация приводит к межпородному нивелированию генетических различий и сближению генотипов в породах.

Согласно утверждениям Д.Б. Голдштейна [172] и П. Жарне [185] в основе микросателлитов лежат, распространённые по всему эукариотическому геному, участки ДНК длиной от двух до шести пар оснований. При относительно малых размерах, микросателлиты могут определяться при использовании ПЦР на ДНК, выделяемой из крови, кожи, корней волос и т.д. Полиморфизмы могут быть провизуализированы с использованием секвенирующего геля в промышленных масштабах. Высокая скорость выделения микросателлитов позволяют оценивать меж- и внутривидовое генетическое разнообразие и генетическое смешивание даже близкородственных пород. Это привело к высокой популярности микросателлитного анализа при изучении популяций сельскохозяйственных животных [86, 108, 245] во всем мире.

Так, Д.Н. Кольцов с соавторами используя метод микросателлиты смогли определить генетическое различие между популяциями крупного рогатого скота сычевской породы и её внутривидовым типом «Вазузский» [56].

В 2011 Е.А. Гладырь, Г.М. Гончаренко, П.В. Горелов и другие соавторы [25] опубликовали результаты исследования по изучению изменчивости микросателлитов, наблюдаемую при создании нового типа мясного скота в Сибири. Авторами было установлено наличие 13 микросателлитов у крупного рогатого скота симментальской, герфордской пород и их помесей. Был отмечен высокий уровень генетического разнообразия при варьировании среднего числа аллелей на локус от 6,92 до 8,83. 95,3% всего разнообразия, как сообщают авторы, может быть объяснено наблюдаемыми внутривидовыми различиями у помесей F1. В работе подчёркивается, что основанная на микросателлитном анализе система является мощным инструментом для оценки генофонда интродуцированных пород.

Годом позднее вышла ещё одна статья под авторством Е.А. Гладырь, Н.А. Зиновьевой, В.А. Багирова с соавторами [26], в которой представлены данные по изучению молекулярно-генетической характеристики аллелофонда крупного рогатого скота в республике Таджикистан. В работе описаны результаты по иссле-

дованию популяций чёрно-пёстрой и швицкой пород таджикского типа, местного зебувидного скота и 2 локальных популяций швицезебувидного скота. Микросателлиты (11 локусов) использовались в качестве ДНК-маркеров. В отдельных локусах микросателлитов число аллелей изменялось от 2 (локус TGLA122) до 14 (локус TGLA227), в сумме всего 108 аллелей. При исследовании выбранного массива скота было обнаружено 15 частных аллелей. Ценность работы состоит в доказательстве своеобразности аборигенных пород в Таджикистане, животные которых являются носителями редкого биологически ценного аллелофонда и потенциально могут стать источником уникальных ресурсов.

В 2013 году коллективом авторов Е.А. Гладырь, Н.А. Зиновьева и др. [27] опубликована работа по изучению аллелофонда казахской белоголовой породы на Южном Урале, а также 2-х различных популяций галловейской породы и породы геррефорд в Западной Сибири. Исследование показало генетическую общность геррефордского скота и казахской белоголовой породы.

Микросателлитный анализ продолжает использоваться во множестве исследований [33, 50, 82] и зарекомендовал себя как один из наиболее доступных методов изучения меж- и внутривидовых взаимосвязей в популяционной генетике. При изучении генетического разнообразия животных, их принадлежности к породной группе или популяции, микросателлиты могут использоваться в качестве молекулярных маркеров [175, 185].

1.9 Молекулярно-генетические маркеры SNP

Молекулы ДНК, состоящие из нуклеотидов, несут наследственную информацию организма. Изменения на определённом участке ДНК последовательности нуклеотидов называют мутациями. Когда заменяется одна пара нуклеотидов на другую, возникает однонуклеотидный полиморфизм SNP (Single Nucleotide Polymorphism). Исследователи используют полиморфные нуклеотидные последовательности ДНК в качестве маркерных систем, чтобы изучать генетический полиморфизм непосредственно на уровне геномов [73, 119, 133, 231]. К середине 2000 годов технологии с использованием SNP были признаны в качестве перспек-

тивных методов для исследований в области популяционной генетики. Было предложено разработать статистические модели, которые бы позволили оценивать генетический потенциал животных и растений [221]. К концу 2000-х появились первые результаты, которые подтверждали связь последовательностей ДНК-маркеров с рядом количественных признаков. И уже с января 2009 года в США на официальном уровне введена геномная оценка молочного скота, как одного из способов повышения точности селекции. В последующие несколько лет программа по геномной оценке была введена ещё в 12 странах мира. В 2014 году опубликована статья под авторством А.А. Сермягина с соавторами [111], в которой рассматриваются сравнительные данные геномной оценки голштинского скота, представлены пути развития и текущее состояние методов, основанных на геномной оценке в молочном скотоводстве России и других странах мира. Также в работе рассматриваются вопросы о применении геномной оценки в селекционном процессе скота молочного направления и её точность. Позднее, в ряде исследований доказана высокая информативность SNP в геноме, т.е. молекулярно-генетических маркеров однонуклеотидных полиморфизмов, отвечающих за сочетания хозяйственно полезных показателей [128].

В настоящее время разработанные на уровне ДНК молекулярно-генетические методы позволяют проводить отбор животных непосредственно по генотипу.

Тем не менее, метод ограничивается недостаточным уровнем понимания основных принципов функционирования генома и реальных механизмов проявления фенотипических признаков. Возникают фундаментальные вопросы о том, как генетическая изменчивость оказывает воздействие на генетические системы, формируя конечные проявления в виде фенотипов. Несмотря на значительные достижения в разработке методов для прогнозирования количественных признаков на уровне генома, по-прежнему существуют значительные трудности в полном понимании эффектов очень малых воздействий на системы организма. Кроме того, существует ограниченное понимание функционирования и взаимодействия клеточных сетей. Сложные фенотипические признаки обусловлены огромным числом вариантов малого эффекта, которые потенциально могут влиять на большое коли-

чество регуляторных вариантов, активных в тканях, связанных с данными признаками.

1.10 Прогноз продуктивности животных на основании геномной оценки

Методология аналитической оценки племенной ценности на основе ДНК-маркеров, охватывающих весь геном животного, впервые была предложена Т.Н.Е. Meuwissen с соавторами [210, 211]. Они в начале 2000-х годов опубликовали подходы, которые могли бы обеспечить внедрение в селекцию методов геномной оценки. Так, согласно им селекция с применением нововведений должна включать следующие шаги:

- генотипирование животных с достоверными EBV, полученными путём классических подходов, оценкой по потомству и родословной;
- разработка и внедрение уравнений прогноза для референтной популяции;
- создание регрессионных моделей, основанных на SNP для оценки молодых быков;
- отбор оценённых быков, которые удовлетворяли бы требованиям программы разведения.

Известно, что Т.Н.Е. Meuwissen [211] с соавторами в своей работе использовали смоделированные данные для проведения анализа с большим количеством равномерно расположенных маркеров, пытаясь идентифицировать не QTL, а лишь некоторые маркеры, которые случайно оказались тесно связанными с QTL. В расчётах были учтены гаплотипы. Анализируя данные методами BayesA и BayesB предполагалось различное распределение эффектов гаплотипов. Прогенотипировав 2200 животных, авторы смогли получить точность прогнозирования 0,85. Учитывая такой высокий показатель точности, при достаточно небольших данных, работа подавала большие надежды во всем сообществе животноводов.

Позднее, с развитием технологий появилась возможность усовершенствовать процедуры оценки племенной ценности животных, используя информацию по все большему числу SNP. Сегодня прочно укрепился термин геномный отбор, который подразумевает форму отбора, при которой используются генетические

маркеры, покрывающие весь геном, при этом все количественные факторы признаков находятся в неравновесной связи хотя бы с одним [254].

1.11 Оценка генетического потенциала животных с использованием SNP

За последние десять лет, область молекулярной генетики и секвенирование ДНК животных претерпели значительные изменения в лучшую сторону. Прорывные технологии позволили создать SNP-чипы, с возможностью генотипирования животных по очень большому количеству SNP-маркеров (50-700 тыс.) [113]. Эти инновации нашли широкое применение и в молочном скотоводстве. Повсеместно наблюдается переход от традиционной племенной оценки быков по дочерям к технологиям геномной племенной оценки. Оценённые по геному быки на международном рынке все чаще заменяют быков, которые оценивались традиционным способом. Базы данных по геномной оценке быков из разных стран объединены в рамках международной организации INTERBULL. Это объединение способствует повышению точности оценки племенной ценности быков и консолидации применяемых подходов [114, 115].

Доступная и точная технология генотипирования стала возможной благодаря внедрению чипа BovineSNP50 [207]. С введением в практику чипов подобного типа уже заговорили о крупномасштабном генотипировании. В последующих работах, посвящённых методологии геномного отбора, исследователи сосредоточились на эффектах SNP и на расчёте матриц геномного родства [250]. Была разработана концепция, позволяющая проводить сравнения между предсказаниями с использованием генеалогических данных и генома.

В период 2004 по 2008 год, корпорация Illumina Corporation (San Diego, CA), при участии Министерства сельского хозяйства США (USDAARS), разработали платформу и SNP-чипы, позволяющие проводить доступное полногеномное генотипирование животных [249]. При создании SNP-чипа учитывался широкий ресиквенс генома 392-х животных, среди которых было представлено 14 пород крупного рогатого скота молочного и мясного направления, а также 166 геномов животных африканских пород [207]. Работа позволила обнаружить 444 792 SNPs

и выделить 54 000 SNPs, которые имели высокую степень детектирования и частоту минорного аллеля (Minor Allele Frequency (MAF) более 5%. Полученные данные легли в основу разработки SNP-чипа BovineSNP50 BeadChip. Популярность и точность данного чипа позволила пережить три модификации и дойти до наших дней [64].

Начиная с 2007 года, благодаря относительно недорогим SNP-чипам, полногеномное секвенирование геномов животных не прерывается. Идентифицировано несколько миллионов SNP. Компания Illumina за эти годы внедрила в практическое использование ещё несколько видов чипов. Так, для малой плотности используются Bovine3K и 6K (2900 и 6909 SNP), а для высокой плотности BovineHD (777 962 SNP). Существуют индивидуальные версии чипов GGP (GeneSeek Inc.) и IDB (ICBF), предназначенные для выявления мажорных генов, мутаций и рецессивных аллелей.

В последние годы на рынок вышли компании, которые предлагают альтернативные SNP-чипы и оборудование для генотипирования. К примеру, Affymetrix с чипами Axiom Genome-Wide Bos 1 Array Plate (648 855 SNP) и др.

SNP-чипы позволяют получить данные, которые используются для раннего прогнозирования мясной и молочной продуктивности у крупного рогатого скота. Так, набор Illumina BovineSNP50K позволил обеспечить прогнозирование племенной ценности скота по 15 хозяйственно-полезным признакам [139, 178, 215]. Также следует отметить работы, в которых опубликованы исследования с применением геномного анализа на других видах животных, к примеру лошадях [160, 180, 205, 208].

Технология полногеномной оценки позволила с высокой точностью рассчитывать племенную ценность животного через индекс племенной ценности. Многократно сократить время, а также затраты на отбор животных с высоким генетическим потенциалом, чтобы использовать их в селекционных программах [76, 103, 225].

Перспективные технологии генотипирования развивались параллельно совершенствованию прогнозных моделей, в которых использовались данные, полу-

ченые в процессе генотипирования. Страны с развитым молочным животноводством собрали большой банк данных генотипов (только в США зарегистрировано более 2 миллионов генотипов). Такой объем накопленных данных привёл к потребности в разработке более совершенных моделей прогнозирования. Так, была разработана модель, позволяющая строить матрицу генотипов в соответствии с количеством маркеров и не зависящая от количества животных - SNP BLUP [213]. Объединить геномную и родословную структуру в единое расчётное пространство позволили перспективные модели, основанные на ssGBLUP [199].

В разработке моделей для геномного отбора ключевым фактором стало тестирование и оптимизация моделей с целью повышения их точности. Особое внимание было уделено методам, направленным на улучшение точности прогнозирования, включая отбор отдельных SNP или использование методов дифференциального взвешивания. Предполагается, что большинство SNP, связанных с ассоциированными с фенотипом локусами (QTL), могут быть выявлены по данным SNP. Такой подход позволяет улучшить прогнозирование фенотипических показателей путём более точного определения значимых генетических маркеров и их вклада в формирование желаемых признаков.

При исследовании ассоциаций на уровне всего генома (полногеномные ассоциативные исследования - GWAS) используется модель, в которой в качестве фиксированного эффекта принимается однократно анализируемый маркер [195], в качестве примера может служить EMMAX - смешанная модель ускоренной ассоциации [192]. Для исключения помех, возникающих из-за структурной составляющей популяции, модель дополняется информацией о родословной животного или матрицей геномного родства (G). [195]. Альтернативой методу, использующему данную модель, являются байесовские методы (BayesB / BayesR), в которых все SNP принято рассматривать вместе и интерпретировать большие сигналы в качестве маркеров для близлежащих QTL (Quantitative Trait Loci).

При геномной оценке крупного рогатого скота принято включать в модель ssGBLUP как можно большее количество животных, тем самым повышая точность геномной оценки. При использовании таких моделей удаётся достичь 75

процентной достоверности геномной оценки по признакам молочной продуктивности, в то время как результаты традиционной племенной оценки находятся в диапазоне между 95% и 99%. Геномная оценка позволяет значительно сократить временные рамки селекционного процесса с 5-6 лет до 2 лет и сэкономить до 90% средств, которые обычно направляются на внедрение традиционных методов и представляется наиболее перспективной и востребованной.

Данный метод представляет собой довольно сложную технологию, которая в своём изначальном виде опирается на характеристику животных (быков), с высокой точностью оценки и использованием генеалогической информации. Этот метод основан на применении линейных смешанных моделей, в которых учитывается матрица родства. Одной из распространённых вариаций этого метода в настоящее время является ssGBLUP [212, 250]. В 2009 году Legarra и его коллеги [199] впервые представили комбинированную матрицу, которая объединяла эти разные типы отношений. Полный анализ, включая использование одношаговой геномной BLUP (ssGBLUP), был представлен в 2010 году Агиларом и соавторами [138, 151]. Этот подход позволил сделать возможным выполнение анализа с использованием модели BLUP, путём замены матрицы родства.

Использование методов GBLUP или байесовских методов предполагает постобработку, чтобы включить в модель информацию о родословной, поэтому традиционная BLUP всё ещё остаётся актуальной. Определение EBV (GEBV) проходит в несколько итераций. К преимуществам данного подхода можно отнести более низкую себестоимость проведения исследования (только кандидаты на отбор подвергаются генотипированию), сохраняется традиционная оценка BLUP, геномный отбор при таком подходе проходит с использованием вспомогательных анализов. К недостаткам можно отнести то, что генерирование дерегрессированных геномных значений (Deregressed Genomic Values, сокр.: DVG) возможно только для простых моделей (например, модели с одним признаком) и только генотипированные животные могут быть включены модель [200].

В современной постгеномной информатике все чаще употребляется понятие «генные сети». Согласно Н.А. Колчанову, генная сеть представляет собой группу

связанных генов, которые обеспечивают формирование в организме фенотипических признаков (биохимических, молекулярных, физиологических) [55].

Процессы, протекающие в организме, можно рассматривать как результат взаимодействия его генных сетей. Каждая из локальных генных сетей может быть представлена как функционально автономное, дискретное сообщество генов и продуктов экспрессии этих генов. Вся совокупность таких сетей организма интегрируется в одну глобальную. Согласно концепции биоинформатики, каждый живой организм может рассматриваться как «сеть сетей», т.е. являться глобальной генной сетью, состоящей из множества локальных сетей [8]. Изучение генных сетей, их взаимосвязей и влияние на формирование того или иного фенотипического признака могут вывести геномную селекцию на новый уровень.

В последние несколько десятилетий активно ведутся исследования по поиску генетических маркеров, которые отвечают за желаемые хозяйственно-полезные показатели. Одним из таких направлений является поиск генов количественных признаков (QTL) - картирование основных. При таком подходе выявляются ассоциации между полиморфными вариантами генов и уровнем изменчивости хозяйственно-полезных признаков в линиях и семействах разводимых животных. Качество прогноза уровней изучаемых признаков можно повысить, учитывая полученные комплексные генотипы, включены в оценку показателей племенных животных, тем самым повышая их племенную ценность [28].

Использование GWAS предоставляет большой спектр данных об SNP, ассоциированных с различными показателями продуктивности. Множество исследовательских коллективов пополняют базы данных, позволяющие проводить поиск информации по накопленным SNP и вести исследования в области геномной селекции. Однонуклеотидные локусы, ассоциированные с признаками молочной продуктивности, описаны во множестве работ. Приведём лишь некоторые из таких исследований.

В работе Cole J.B. с соавторами [154] описан полиморфизм rs110527224, как имеющий ассоциации с большим количеством признаков молочной продуктивности и здоровья голштинского скота. Для генотипирования животных использовал-

ся чип Illumina BovineSNP50 BeadChip. Исследователи проанализировали 31 признак продуктивности, размножения, здоровья и экстерьера у коров голштинской породы в США. Было выявлено, что к высокозначимым генам и участкам хромосом относятся: область GNAS BTA13, отвечающая за удои, жир и белок; область INSR BTA7 и LOC520057 BTAХ и GRIA3, отвечающие за стельность, количество соматических клеток и период хозяйственного использования; область LRP1B BTA2 за соматические клетки; область DGAT1-NIBP BTA14 за процентное содержание жира; FKBP2 от BTA1 за белок в кг. и его процентное содержание, MGMT от BTA26 и PDGFRA от BTA6 за процентное содержание белка; BTA18 за лёгкость отёла и мертворождение; CD82 от BTA15, DST от BTA23 и область MOCS1-LRFN2 за мертворожденность дочерей; LOC520057 и GRIA3 от BTAХ за частоту беременности дочерей. Что касается особенностей строения тела, то BTA11, BTAХ, BTA10, BTA5 и BTA26 имели наибольшую концентрацию SNP-эффектов, а RHKA2 из BTAХ и REN из BTA16 имели наиболее значимые эффекты, отвечающие за размер животного. BTAХ, BTA19 и BTA3 оказывали большее влияние на формирование формы тела. На признаки вымени влияли BTA16, BTA22, BTAХ, BTA2, BTA10, BTA11, BTA20, BTA22 и BTA25, на признаки сосков BTA6, BTA7, BTA9, BTA16, BTA11, BTA26 и BTA17, а на признаки конечностей влияли BTA11, BTA13, BTA18, BTA20 и BTA26. В анализе использовалось 45 878 SNP. При расчётах применяли регрессию методом наименьших квадратов.

Iung L.H.S. с соавторами [183], благодаря чипам Illumina BovineLD BeadChip и Neogen GeneSeek Genomic Profiler HDv3 смогли установить ассоциацию rs43734861 с показателями молочной продуктивности. Исследования проводились на голштинском скоте бразильской селекции. Для этого были использованы 75 228 фенотипических записей от 5 981 коров и генотипические данные 56 256 однонуклеотидных полиморфизмов (SNP) от 1 067 коров в рамках взвешенного одноступенчатого GWAS. В исследовании были обнаружены потенциальные новые гены, участвующие в повреждении тканей и восстановлении молочной железы (COL18A1), иммунном ответе (LTTC19), гомеостазе глюкозы (SLC37A1), синтезе

ненасыщенных жирных кислот (LTBP1) и транспорте сахара (SLC37A1 и MFS-D4A), влияющие на удой молока, количество соматических клеток, процентное содержание жира и состав жирных кислот. Анализ ассоциаций в работе проводился с использованием взвешенного одноэтапного геномного метода BLUP.

В работе Nayeri S. с соавторами [219] была обнаружена ассоциация локализованного в внегенном пространстве нуклеотидного полиморфизма rs110560119 с процентным содержанием белка у коров голштинской породы в Канаде. При генотипировании исследователи использовали 602 095 SNP и чипы Illumina BovineSNP50 BeadChip (с присвоением HD) или BovineHD BeadChip. Анализ проводился на молочную продуктивность и признаки фертильности. Идентифицированные области QTL для хозяйственно-полезных признаков (включая удой) частично совпадают с генами с известными соответствующими биологическими функциями, выявленными в более ранних исследованиях, такими как DGAT1 и CPSF1. В пределах данного исследования описан ещё один однонуклеотидный полиморфизм rs137202311, который локализован в пределах гена SPAG17 ($p < 0,01$), играющего важную роль в функции и структуре подвижных ресничек эпителия респираторной системы и описанный ранее в исследованиях на людях.

Локализованный в пределах гена BAALC, однонуклеотидный полиморфизм rs109719021 имеет ассоциацию с процентным содержанием молочного белка. Об этом написано в публикации Jiang J. с соавторами [186]. Для исследования был взят американский голштинский скот и проведена идентификация генотипов по 60 671 SNP. В анализе использовались байесовская линейная смешанная модель и метод наименьших квадратов.

В работе Schopen G.C. с соавторами [241] отмечается, что однонуклеотидный полиморфизм rs43629218, который локализован в пределах гена MYO1E, связан с молочным белком, а точнее его процентным содержанием, у коров голштино-фризского скота. После генотипирования 1713 животных по 50 228 SNP, данные анализировали на белковый состав молока. Ассоциации с содержанием и составом молочного белка были обнаружены на 20 бычьих аутосомах. Основные области генома, связанные с составом молочного белка или процентным содержа-

нием белка в молоке, были обнаружены на хромосомах 5, 6, 11 и 14. В работе использовалась общая линейная модель и модель на животных.

Полиморфизмы rs110255620, rs110456580, rs110323597 ассоциированы с жирностью молока. Так, в работе Pedrosa V.V. с соавторами [224] было описано, что rs110255620, который локализован во внегенном пространстве, ассоциирован с процентом белка и молочного жира, а также уровнем удоя. Исследования проводились на североамериканской популяции голштинского скота с использованием диапазона SNP от 5 108 861 до 6 101 357. Для генотипирования применялись SNP-панели Illumina средней плотности. В работе были подтверждены ранее известные ассоциации, а также выявлено несколько новых генов-кандидатов, включая ARHGAP35, NPAS1, TMEM160, ZC3H4, SAE1, ZMIZ1, PPIF, LDB2, ABI3, SERPINB6 и SERPINB9 для признаков молочной продуктивности (LP); NIM1K, ZNF131, GABRG1, GABRA2, DCHS1 и SPIDR для удоя; NR6A1, OLFML2A, EXT2, POLD1, GOT1 и ETV6 для жира; DPP6, LRRC26 и семейство генов KCN для жира в %; CDC14A, RTCA, HSTN и ODAM для белка; и HERC3, HERC5, LALBA, CCL28 и NEURL1 для процентного содержания белка. GWAS был выполнен с применением смешанных линейных моделей. Используя байесовский метод и регрессионную модель с одним SNP, Meredith B.K. и соавторами [209], также удалось обнаружить ассоциацию rs110255620 с процентным содержанием жира и белка и удоём в ирландской популяции голштино-фризского скота. Также авторами описана ассоциация rs110456580, локализованного в пределах белок-кодирующего гена TG (Thyroglobulin), с уровнем молочного жира у животных той же популяции.

Полиморфизм rs110323597 зарегистрирован M.M.I. Salem с соавторами [239] в базе QTL, которые в своём исследовании показали его ассоциирование с молочным жиром у скота голштинской породы в Португалии. В своей работе исследователи использовали 37 031 SNP, охватывающих 29 аутосом. Было обнаружено двенадцать значимых SNP расположенных на BTA3, и 10 других SNP, расположенных на BETA3, BTA8, BTA14, BTA17, BTA18 и BTA21. SNP ARS-BFGL-

NGS-57820 и ARS-BFGL-NGS-4939 на BTA14 были обнаружены внутри одного и того же блока гаплотипов с геном DGAT1.

Опубликовано множество работ, в которых описаны ассоциации полиморфизмов с содержанием органических и минеральных компонентов в молоке [146, 147, 169], длительностью лактации и продуктивным долголетием [154; 197], скоростью молокоотдачи [196, 206], числом и показателями соматических клеток [181, 209, 219, 222]. Представляют большой интерес работы посвященные идентификации полиморфных вариантов генов липидного и белкового обмена с применением ДНК-технологий и изучение их влияния на показатели продуктивности коров [38, 39].

1.12 Значение информационных технологий в селекционном процессе молочного скотоводства

Последние десятилетия ознаменовались бурным развитием информационных технологий, которые повсеместно заняли ниши в управлении организационной структурой предприятий, бухгалтерским учётом, системой документооборота, фиксированием, передачей и анализом потоков поступающей информации. Все это значительно упростило процесс управления производством любого типа [35]. Кроме этого, информационные технологии позволяют оперативно создавать и использовать огромные объёмы данных, применять элементы искусственного интеллекта для анализа и исследований в производственной и научной деятельности, накапливать информацию в банках и базах данных. Несмотря на некоторые сдерживающие факторы [4], цифровые технологии рассматриваются как главный движущий ресурс, определяющий рост благосостояния и благополучия современного общества.

Внедрение новых методов хранения и обработки зоотехнических данных, автоматизация многих элементов управления селекционным процессом и т.д., позволяют совершенствовать отрасль молочного скотоводства.

С внедрением цифровизации в производственные цепочки сельскохозяйственных предприятий животноводческой направленности, работа селекционеров

может стать более эффективной и позволит обрабатывать огромные массивы данных. Все это даст возможность делать своевременный анализ результатов разведения, методов подбора и выбор оценённых быков по качеству их потомства, тем самым корректируя принятые программы разведения [74, 75, 79, 122].

Ещё В.Ф. Красота с соавторами при определении элементов крупномасштабной селекции писал, что нужно применять новейшие технические средства для моделирования селекционных программ [62].

Х.А. Амерханов в своей работе [2] отмечал, что отечественное животноводство, может стать конкурентоспособным при существенной корректировке управления племенным делом, а также внедрением эффективных методов селекции на базе использования информационного обеспечения, систем идентификации и мечения животных.

Но, к сожалению, если перейти к государственной составляющей или даже отдельно взятому региону, то есть ряд вопросов, которые до сих пор не решены в отрасли молочного скотоводства. Активная голштинизация в нашей стране привела к уязвимости отрасли перед западными компаниями. Ведь программное обеспечение, позволяющее осуществлять отбор и подбор, разработано зарубежными компаниями и апробировано для стад зарубежной селекции, как следствие такой подход пренебрегает учётом хозяйственных условий местных предприятий. Все это ведёт к существенным изменениям в генеалогической и генетической структуре маточного поголовья и создаёт серьёзные риски для отечественной племенной работы. Это косвенно подтверждается тем, что в Российской Федерации отсутствует второе и последующее поколение быков отечественной селекции, которые были бы сопоставимы по качеству потомства с зарубежными.

Применяемые компаниями-поставщиками племенного материала схемы часто не предусматривают оценку качества потомства быков и включают только иностранных производителей, оценённых за рубежом с использованием геномного подхода. Это приводит к искажению оценки племенной ценности зарубежных быков в различных средовых условиях регионов нашей страны.

Усиливающаяся потребность в сохранении и совершенствовании имеющегося генетического потенциала с учётом многочисленных природно-климатических и хозяйственных условий регионов России, а также стремление к обеспечению независимости в области племенного молочного скотоводства, делают неотложной необходимостью изучения генетической и генеалогической структур молочного скота на уровне страны и регионов в целом, а не только на уровне отдельных хозяйств. Все это требует разработки и внедрения собственных программных решений, единых баз данных, ПО направленного на автоматизацию процессов подбора родительских пар с учётом средовых условий, программ, которые позволят централизованно управлять селекционными процессами. Применение отечественных современных и перспективных подходов в области селекции, биоинформатики, математического и компьютерного моделирования в животноводстве [87, 190, 191] представляется крайне важным.

Для отрасли молочного скотоводства на рынке России существуют внедрённые программные решения, к примеру, СЭЛЕКС-Молочный скот [79, 110], но глубина проработанности данного программного решения даёт основание беспокоиться о проблеме достоверности данных первичного зоотехнического учёта в связи с отсутствием централизованного подхода к их проверке [47]. Подобная озабоченность относится и к возможности эффективно вести крупномасштабную селекцию на региональном и федеральном уровнях. Децентрализованность баз данных и низкий уровень достоверности данных приводят к потребности поиска и разработки альтернативных инструментов. Согласно Красоте [62], ведение крупномасштабной селекции на уровне страны немыслимо без централизованного руководства ведущими звеньями селекционного процесса при соответствующей организации племенной базы и генеалогической структуры породы, обеспечения систематического анализа фактического эффекта селекции и корректировки действующих селекционных программ для получения максимального селекционно-экономического эффекта с использованием для этих целей самых современных из доступных технических средств. Именно «доступные технические средства» являются ключевыми в сложившейся ситуации. Для выполнения основных постула-

тов, предложенных ещё классиками нужна централизованная система управления на уровне не отдельно взятого предприятия, а на уровне всей страны. Такая система должна иметь доступную для идентификации генеалогию по каждому племенному животному, единую выверенную базу данных фенотипов и генотипов, доступные средства для осуществления процессов идентификации животных, подбора родительских пар и прогнозно-аналитических инструментов для осуществления целенаправленного отбора. Таким решением может стать федеральная государственная информационно-аналитическая система племенных ресурсов.

Если рассматривать опыт стран с развитым молочным животноводством, то путь создания ИАС на уровне страны ими уже пройден. Так, в Германии реализована единая информационная система, которая обеспечивает запись и хранение всей совокупной информации по регистрации племенных животных и данных по их продуктивности. Разработчиками выступили Федеральное министерство продовольствия и сельского хозяйства со Специализированной породной ассоциацией Германского международного центра генетики – CGI. Для реализации исполнения проекта ими была создана дочерняя IT-компания VIT. Об успешности проекта говорит то, что ежегодно базы данных информационной системы пополняются показателями по продуктивности для более чем 1,8 млн. животных. Число постоянных пользователей данной системы уже превысило 19 тыс. хозяйств и несколько племенных центров. Кроме информации о хозяйственно-полезных признаках в рамках системы доступны данные об единой идентификационном обозначении животных, их родословная, данные об искусственном осеменении, экстерьерные показатели, сведения о хозяйственном использовании, расчетно-аналитические данные для отчёной документации и тд. Перечень доступных составляющих позволяет в рамках системы формировать сертификаты на племенных животных, вести племенные книги, распечатывать свидетельства оценки племенной ценности по молодняку и быкам-производителям, оценённых по качеству потомства.

Централизованное управление селекционными процессами в рамках единой платформы проработано и во Франции при участии Национального сельскохозяй-

ственного научно-исследовательского центра INRA. По всей территории страны множество служб осуществляет организационные вопросы по идентификации, регистрации животных, собирают показатели по продуктивности, осуществляют генетическую оценку скота. Но, вся эта работа проводится под жёстким «вертикальным» контролем служб Министерства сельского хозяйства Франции, породными ассоциациями и Институтом животноводства (Institut de l'Elevage). Результаты стекаются в единый банк данных, который доступен для всех профильных организаций. Благодаря высококвалифицированным специалистам и их поисково-исследовательским работам, в последние 2 десятилетия INRA превратилась в «центр верификации» данных, которые поступают с различных ассистентских служб, оценок бонитёров и данных первичного зоотехнического учёта предприятий, а затем добавляются единый банк данных.

На территории постсоветского пространства также известны проекты по созданию информационно-аналитических систем. Так на территории Белоруссии при координации республиканского унитарного предприятия по племенному делу в животноводстве «Белплемяживобъединение» создаются решения для централизованного ведения селекционно-племенной работы с использованием передовых информационных технологий. Проводимая работа сопряжена с созданием в рамках одной цифровой платформы единой верифицированной базы данных и систем прогнозного моделирования различных элементов селекционного процесса [11].

В целях автоматизации процессов селекционной и племенной работы в Республике Казахстан с 2010 года функционирует информационно-аналитическая система «Республиканская система животноводства». Основной функцией которой является накопление и анализ информации по племенным и товарным животным [97].

В США, Канаде, странах Западной Европы и ряде других государств, животное принято считать племенным, если оно зарегистрировано в национальной информационной системе и ему присвоен международный номер согласно правилам Международного комитета информационного обеспечения в племенном животноводстве (ICAR – International Committee on Animal Recording), который на-

ходится в ведении Всемирной и Европейской ассоциациях по животноводству (WAAP и EAAP). Данные организации выработали определённый свод требований, рекомендаций и нормативной документации, для обеспечения единого информационного пространства мирового племенного животноводства [163, 252].

С развитием геномной оценки, в мировой практике проведена огромная работа по созданию цифровых банков данных генотипов в странах с развитым молочным скотоводством. Там можно отметить США > 2 млн генотипов [150], Ирландия > 1 млн генотипов [182], EuroGenomics (Германия, Франция, Голландия, Бельгия, Дания, Швеция, Финляндия, Польша, Испания) > 1,6 млн [162].

В Российской Федерации в 2023 году было официально объявлено о создании Федеральной государственной информационно-аналитической системы племенных ресурсов (ФГИАС ПР) [125].

1.13 Особенности работы с большими массивами данных при проведении исследования

Современное животноводство неразрывно связано с применением цифровых технологий и реализованных на их основе методов и подходов для анализа и обработки данных. Зачастую объем данных может составлять большие выборки и интерпретировать их становится все сложнее без специализированных инструментов. К таким инструментам можно отнести элементы прикладного искусственного интеллекта, многомерную статистику, моделирование, базы данных и решения, которые приходится программировать под конкретные задачи. Такой подход обеспечивает получение более точных результатов анализа при раскрытии определённых подзадач и позволяет выявить некоторые скрытые закономерности в данных. Важность применения программных инструментов также обусловлена рядом уникальных особенностей племенного учёта в России среди которых можно выделить выраженную фрагментированность и децентрализованность в исходных информационных массивах (базах данных). При этом, сложившиеся за последние годы условия, при ограничении импорта маточного поголовья, увеличении срока поставок племенного материала и прочее, предъявляют новые вызовы

для сохранения генетического потенциала популяций чёрно-пёстрого скота, имеющегося в Сибири и в целом по стране. При сохранении классических методов селекции, в условиях многолетнего межлинейного кроссирования за рубежом, решение поставленных задач становится малоэффективным и также требует новых подходов с применением цифровых технологий [13, 227]. В странах с развитым скотоводством имеется практика применения линейных, линейных и нелинейных смешанных моделей, единые информационные базы данных с фенотипами и генотипами скота, но все это требует высокого качества исходных данных первичного зоотехнического учёта. Исследование генеалогической и генетической структуры популяций при таком подходе значительно упрощается. В России фрагментированность данных, их обособленная локализованность, доступность и «человеческий фактор» значительно нивелируют качество исследований по изучению больших популяций скота на уровне региона или страны, что, как следствие, сказывается на прогнозировании и сохранении генетического потенциала изучаемого скота [47]. В настоящем исследовании представлен ряд инструментов и подходов, которые позволяют централизовать исходные данные [107, 109], значительно повысить уровень контроля используемой информации и обеспечить изучение популяции чёрно-пёстрого на уровне всего региона.

2. МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЙ

2.1 Создание единой региональной базы верифицированных данных

Общая схема проведённого исследования генеалогической и генетической структуры чёрно-пёстрого скота Западной Сибири представлена на рисунке 1.



Рисунок 1 - Схема исследования генеалогической и генетической структуры чёрно-пёстрого скота Западной Сибири

Исследования выполнены на базе лаборатории прикладной биоинформатики ФГБОУ ВО Новосибирский государственный аграрный университет. При выполнении диссертационной работы были использованы данные о фенотипах, генотипах и генеалогии чёрно-пёстрого скота из баз данных племенных предприятий Сибири и открытых источников. Данные первичного зоотехнического учёта были извлечены из локальных баз данных СУБД Firebird [167] предприятий и перемещены в единую базу данных СУБД MySQL [217] лаборатории, архитектура которой специально была разработана под выполнение поставленных задач. Работа с базами данных велась совместно с региональным информационно-селекционным центром (РИСЦ) АО «Новосибирскагроплем» по договору о совместном

сотрудничестве. При слиянии исходных массивов в единое цифровое пространство, данные были дополнены показателями продуктивности и родословными крупного рогатого скота молочного направления продуктивности из открытого источника, который находится в свободном доступе на сайте ВНИИплем [80]. Всего для работы были взяты исходные материалы первичного зоотехнического учёта по 22 товарным и племенным предприятиям Сибирского федерального округа.

При объединении баз данных предприятий анализировалась информация о происхождении животных и выявлялись возможные ошибки в генеалогии. Для анализа заполненности родословных были извлечены инвентарный номер, кличка и дата рождения каждого животного и его предков (3 ряда) как со стороны отца, так и со стороны матери. На их основе построен «уникальный ключ», который позволил идентифицировать запись о животном и не допустить дублирование строк. В процессе работы были получены 463 107 записей о животных из баз данных 22 предприятий (учтено хозяйство), после идентификации 299 102 жив. С учётом того, что животное может несколько раз встречаться в разных рядах родословных, общее количество записей в 4 рядах родословной составило 6 018 698 на 18 хозяйств (4 фермы были добавлены в анализ позднее на этапе исследования количества потомков быков). Для анализа родословной брались животные всех рядов и многие из них имеют зарубежное происхождение. По каждому из предприятий проанализирована заполненность генеалогии по группам: быки-осеменители, все быки, коровы и тёлки (прил. А,Б). На языке R были написаны программные инструменты для визуализации родословных в виде нисходящих диаграмм от предка к потомкам, с учётом боковых родственников. Построение и анализ таких генеалогических деревьев позволяют обнаружить искажения в родословных, не выявленные на предыдущих этапах верификации.

На этапе электронной идентификации животного при внесении записи о нем в базу данных, был проведён анализ на возможные ошибки при внесении даты рождения животного. Среди частых ошибок в локальных базах данных предприятий можно выделить замену местами даты и месяца рождения. Такие ошибки

корректировались с учётом контроля записей об отце и матери. Наряду с этим в исследовании были учтены генеалогические циклы и несоответствие дат рождения животных по предкам, когда предки имеют дату рождения позже, чем у пробанда. Следующим этапом было дополнение родословных записями из открытых зарубежных источников.

К сожалению, до настоящего времени полные и верифицированные родословные, включающие актуальные данные первичного цифрового учёта так и не были построены, что затрудняет объективно оценивать как генеалогическую, так и генетическую структуру маточного поголовья. Для нивелирования этой проблемы была разработана архитектура базы данных, которая включала животных, как отечественного происхождения, так и представителей зарубежных популяций. В качестве отправной точки для создания единой базы данных была выбрана информация, размещённая в свободном доступе на электронном ресурсе CDN [149]. Представленные здесь данные регламентированы Interbull - постоянным подкомитетом международного комитета по регистрации животных (ICAR) в скотоводстве. Разработка велась с учётом того, что база будет пополняться на постоянной основе, а также включать в себя всех племенных животных, независимо от страны происхождения.

С сайта CDN были взяты данные в формате .csv и разложены в массивы на языке PHP [228], верифицированы и записаны в ранее подготовленную структуру таблиц разработанной базы данных, согласно классическим правилам реляционной алгебры. Каждое животное было уникально идентифицировано и сопоставлено со страной происхождения. Информация о показателях продуктивности, общие данные о животных (номер, кличка, дата рождения), страна происхождения и данные о родословной, выделены в отдельные таблицы в архитектуре базы данных и связаны между собой в отношении посредством внешних ключей. При записи исходного массива в базу данных, животные с неполной или недостающей информацией по происхождению или по показателям продуктивности получали дополнительный флаг, который явно указывал о том, что информацию по таким животным использовать в исследовании нельзя. Все достоверно подтверждённые записи

о животных CDN, т.е. с полным набором данных, связывались в единое целое с отечественным массивом записей о животных, полученным из баз данных племенных предприятий (региональная база данных), чтобы сформировать целостную «картину» генеалогических представлений популяции чёрно-пёстрого скота СФО.

В общем виде структура базы данных представлена на рисунке 2.

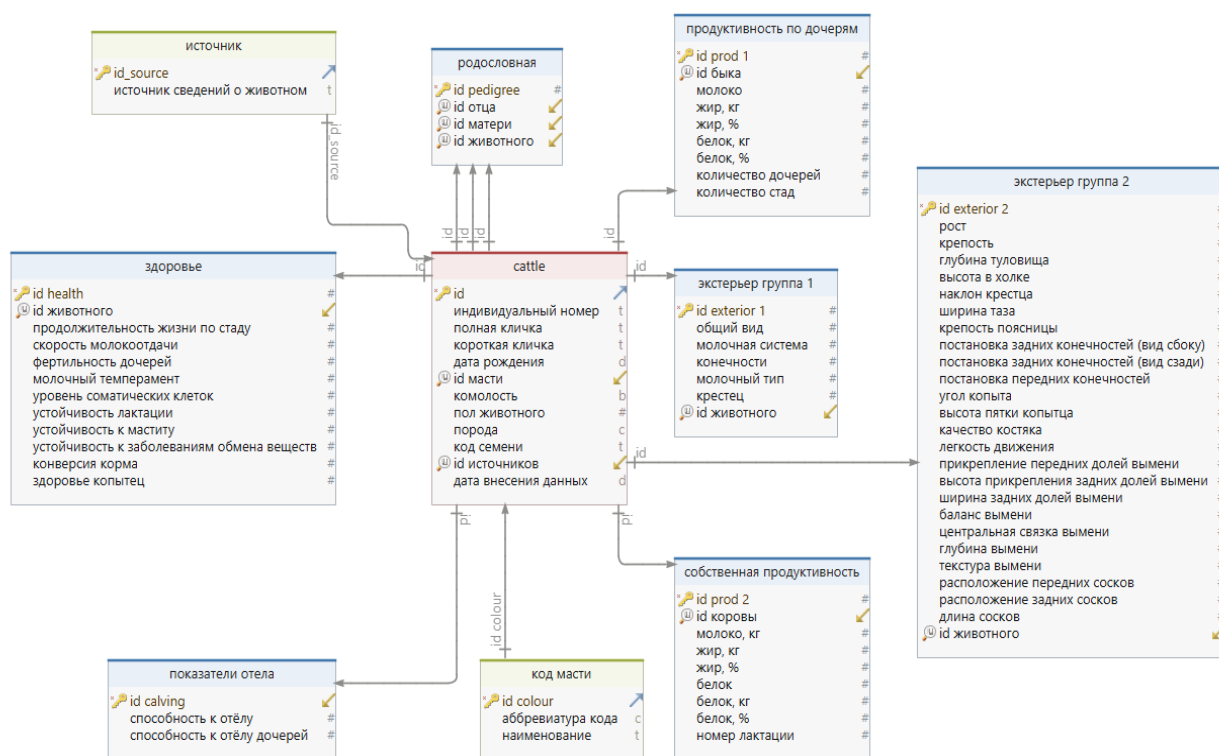


Рисунок 2 - Схематическое представление базы данных

Следует отметить, что первоначально в базу данных были добавлены быки, по которым в CDN имеется необходимая для применения в РФ информация, так как подавляющая доля чёрно-пёстрого скота имеет зарубежное происхождение. После этого по каждому животному определялись предки по отцовской и материнской линии и также добавлялись в основную таблицу о животном «cattle» (см. рис. 3).

id_cattle	ind_num	name_full	name_short	date_bd	colour_code_id	polled	sex	breed	ic	naab_code	naab_m_code	first_official_prod_prof	source
1	HOUSAM1563453	WILLOW-FARM ROCKMAN IVANHOE	WILLOW	1969-08-04		1		1 HO	0.00	008HO09527	0008		9601 1
2	HOUSAM1392858	NO-NA-ME FOND MATT	FOND MATT	1960-12-20		1		1 HO	0.00	007HO00091	0007		9601 1
3	HOUSAM1650414	S-W-D VALIANT	VALIANT	1973-06-28		1		1 HO	0.23	029HO02851	0029		9601 1
4	HOUSAM1626813	MARSHFIELD ELEVATION TONY	MARS	1972-08-06		1		1 HO	0.42	007HO00401	0007		9601 1
5	HOUSAM1667366	CARLIN-M IVANHOE BELL	BELL	1974-05-16		1		1 HO	0.46	007HO00543	0007		9601 1
6	HOCANM343514	GLENAFTON ENHANCER	ENHANCER	1977-03-31		1		1 HO	5.29	073HO00332	0200		9601 1
7	HOUSAM1964484	SOUTHWIND BELL OF BAR-LEE	SOUTHWIND	1984-09-14		1		1 HO	1.78	029HO05296	0029		9601 1
8	HOUSAM1629391	HANOVER-HILL TRIPLE THREAT *BR	TRIPLE THREAT	1972-04-24		2		1 HO	0.43	029HO02775	0029		9601 1
9	HOUSAM1672151	OCEAN-VIEW SEXATION	SEXATION	1973-09-11		1		1 HO	1.17	029HO02960	0029		9601 1
10	HOCANM391624	KENTVILLE STARDUST	STARDUST	1987-02-20		1		1 HO	0.37	073HO01237	0200		9601 1
11	HOUSAM1430145	CITATION R MAPLE	R MAPLE	1962-08-03		1		1 HO	0.00	029HO01883	0029		9601 1
12	HOUSAM1491007	ROUND OAK RAG APPLE ELEVATION	ELEVATION	1965-08-30		1		1 HO	0.00	007HO00058	0007		9601 1
13	HOUSAM1879149	ARLINDA MELWOOD-ET	MELWOOD	1982-03-06		1		1 HO	2.70	029HO05730	0029		9601 1
14	HOCANM311497	A NORTHCROFT ADMIRAL CITATION	CITATION	1968-11-22		1		1 HO	4.01		0200		9601 1
15	HOCANM340909	CAL-CLARK CUTLASS	CUTLASS	1976-08-22		1		1 HO	0.73	073HO00281	0200		9601 1
16	HOCANM357278	SESSINK JETSTAR LARNO		1980-03-21		3		1 HO	0.00				0 1
17	HOCANM341536	A BRIGEEEN FOND CITATION	CITATION	1975-07-09		1		1 HO	2.54	072HO00239	0200		9601 1
18	HOCANM364890	RIVERDOWN SENATOR BONANZA		1981-10-30		3		1 HO	0.19				0 1
19	HOCANM347045	A MARIES THUNDER	THUNDER	1977-01-29		1		1 HO	0.82	072HO00269	0200		9601 1
20	HOCANM344610	STOLTZFUS ASTRONAUT SYN	SYN	1977-03-12		1		1 HO	3.56	069HO00020	0200		9601 1
21	HOCANM337778	MONOAK MARK SNOW	SNOW	1975-03-05		1		1 HO	3.79	069HO00014	0200		9601 1
22	HOCANM345653	CHERRY LANE SUPERSTAR	SUPERSTAR	1977-08-27		1		1 HO	3.47	072HO00274	0200		9601 1
23	HOCANM322172	A NELACRES JOHANNA SENATOR	SENATOR	1966-10-25		1		1 HO	1.45	070HO00084	0200		9601 1
24	HOCANM344905	LIME-HOLLOW BREEZE	BREEZE	1976-11-23		1		1 HO	1.87	073HO00292	0200		9601 1
25	HOCANM359136	MAPLELAWN CELTIC ET-RED	MAPLELAWN CELTI	1980-06-26		4		1 HO	0.00	070HO00004	0200		9601 1

Рисунок 3 - Графическое представление данных в основной таблице записей
О ЖИВОТНОМ

Параллельно создавались записи в таблице «родословная», где находится информация об идентификаторе животного, его отца и его матери (рисунок 4).

id_ped	id_cattle_ped	1	o_id	m_id
876805			876805	349 1277702
876804			876804	11874 936840
876803			876803	5 871379
876802			876802	2100 863611
876801			876801	15763 866111
876800			876800	2950 936081
876799			876799	37 1277701
876798			876798	8909 1277700
876797			876797	202 1277699
876796			876796	7565 1277698
876795			876795	9427 1277697

Рисунок 4 - Представление данных в таблице «родословная»

Таким образом, удалось сформировать связи между животным и его предками при условии, что все ряды предков находятся в одной таблице. Для тех животных, отца и/или мать которых не удалось обнаружить, в качестве идентификатора отца и идентификатора матери в таблице «родословная» используется нулевое значение. К моменту окончания работы в базу данных уже добавлена информация о 2 133 863 уникальных животных.

Разработка подобного инструмента была продиктована необходимостью добавления рядов предков в отечественных базах данных племенных предприятий и исправления ошибок в родословной. Это позволило сформировать наиболее полное несмещённое представление о генеалогии чёрно-пёстрого скота Западной Сибири по генеалогическим линиям и семействам. Данное исследование частично согласуется с пунктом 9 паспорта специальности *«Совершенствование и разработка системы организации племенного дела и нормативно-законодательной базы племенного животноводства»*, так как единая региональная база с верифицированными данными может помочь в совершенствовании системы организации племенного дела на предприятиях Западной Сибири.

По материалам данного раздела были получены два свидетельства о регистрации баз данных:

- свидетельство о государственной регистрации базы данных №2024622398 Российская Федерация. «Комплексная линейная оценка вымени популяции скота канадской селекции»: №2024622072 : заявл. 16.05.2024 : опубл. 31.05.2024 / А.Ф. Петров, Е.В. Камалдинов, К.Н. Нарожных и др.; заявитель и правообладатель ФГБУ ВО Новосибирский ГАУ;

- свидетельство о государственной регистрации базы данных №2024625179 Российская Федерация. «База данных генеалогической структуры быков голштинской породы канадского происхождения»: №2024624918 : заявл. 01.11.2024 : опубл. 14.11.2024 / А.Ф. Петров, Е.В. Камалдинов, К.Н. Нарожных; заявитель и правообладатель ФГБУ ВО Новосибирский ГАУ.

2.2 Статистический анализ

На этапе анализа данных первичного электронного зоотехнического учёта для оценки достоверности показателей молочной продуктивности применялся критерий Андерсона-Дарлинга, который позволяет дать оценку соответствия эмпирического распределения нормальному гауссовскому. Этот критерий позволил судить о соответствиях в данных нормальному распределению, и как следствие, эмпирически оценить их надёжность и точность. Те показатели молочной продуктивности, которые не прошли проверку, были исключены из исследования. По всем оставшимся показателям рассчитаны показатели описательной статистики.

Базы данных предприятий региона были проверены на соответствие приемлемым границам. Рассчитаны медианные значения лимитов изменчивости. Проверку проводили по ряду хозяйственно-полезных признаков с непрерывным характером изменчивости. Весь анализ проводился на языке R [233].

Подробно данная методика описана автором в статье:

- Камалдинов, Е.В. Достоверность данных первичного зоотехнического учёта в молочном скотоводстве / Е.В. Камалдинов, А.Ф. Петров, К.С. Шатохин и др. // Вестник НГАУ. – 2022. – Т.63, №2. – С. 76-83.

2.3 Оценка генеалогической структуры

Оценка генеалогической информации проводилась по базам данных племенных предприятий, занимающихся разведением крупного рогатого скота молочного направления продуктивности, Новосибирской, Кемеровской и Томской областей. После встройки исходных массивов в единое цифровое пространство общий объем записей для последующей обработки превысил 6 млн. записей. Для сравнительного анализа и сопоставления рядов родословных отдельных групп с неполными данными о происхождении по всем генеалогическим данным предприятий региона, была осуществлена процедура ранжирования сумм частот, где минимальному ранговому значению соответствует минимальное значение частот.

ты. Дополнительно племенные предприятия были проранжированы по степени полноты (заполненности) генеалогических массивов доступных в их базах данных (приложение А,Б). Для этой цели был применён интегральный показатель, описывающий количество отсутствующих записей о происхождении животных, исключая группу «быки-производители». По каждому хозяйству были построены генеалогические деревья родственных групп по всем племенным животным представленным в их популяциях. Для сортировки массивов с данными и их предварительной подготовки использовалась библиотека `dplyr` языка R, доступная в открытых репозиториях. Оценка степени сходства предприятий региона по качеству и полноте заполнения рядов родословных, представленных в их базах данных, применялся кластерный анализ. При построении матрицы сходства использовались расстояния Чебышева [152], а характеристика расстояния между кластерами осуществлялось с помощью метода полной связи (дальнего соседа) [37]. Для работы с рангами абсолютных частот была выбрана дистанция Евклида [90] и метод кластеризации Уорда [216]. Визуализация результатов и кластерный анализ осуществлялись в пакетах `factoextra` и `cluster`. Входные данные для кластерного анализа представляли собой матрицы. В первой были описаны проценты пропущенных записей в родословной по группам животных (быки, коровы, телки) каждого из 18-ти предприятий.

Во второй ранговые значения абсолютных частот отсутствующих данных о происхождении в группах животных до третьего поколения предков. Для визуализации процесса объединения хозяйств в кластеры на дендрограмме, при проведении кластерного анализа, использовалась иерархическая агломеративная кластеризация [49], которая позволила вывести результат в наглядном виде.

На основе выверенных данных, было рассчитано какие насыщенные генеалогические ветви наблюдаются в популяциях хозяйств Западной Сибири. Предприятия были сравнены по количеству общих предков по быкам и дана сравнительная характеристика генеалогии всей популяции.

С целью оценки качества входных количественных признаков (удой (кг), молочный жир (кг) и массовая доля молочного жира (%) за 305 дней лактации)

были рассчитаны показатели описательной статистики, включая оценку степени соответствия эмпирических распределений распределению Гаусса с помощью специализированных статистических критериев Андерсона-Дарлинга (AD) и Шапиро-Франсиса (SF) (приложение Г). Предварительно все данные (за 305 дней лактации) были обрезаны до следующих границ (таблица 1):

Таблица 1 - Минимальные границы при фильтрации входящих данных по всем лактациям

Признак	Min	Max
Удой, кг	3500	30000
Белок за 305, %	3,0	4,8
Жир за 305, %	3,6	5,0
Сервис-период, дней	21	240

2.4 Оценка генетической структуры

2.4.1 Генетическая структура по микросателлитным локусам

Генетическая структура чёрно-пёстрого скота Западно Сибири изучалась по STR-локусам микросателлитов и данным полногеномного SNP-типирования. Для анализа микросателлитов из баз данных отобраны животные с известными профилями за последние пять лет. Из выборки исключены некорректные записи (например, 2_103/22), дубликаты (по кличке, номеру и дате рождения) и профили с нулевой длиной аллелей. Окончательно, анализ проведён по 10 228 уникальным микросателлитным профилям.

Поскольку информация по локусам и референтным значениям аллелей была получена из разных источников в результате работы нескольких лабораторий генетической экспертизы достоверности происхождения, то и перечень генотипов в общей выборке был неоднороден. Поэтому, было принято решение ограничиться только теми STR - маркерами (микросателлитами), которые обычно принято

брать для исследования. Все исследуемые локусы рекомендованы решением Коллегии Евразийской экономической комиссии от 2 июня 2020 г. №74 «Об утверждении Положения о проведении молекулярной генетической экспертизы племенной продукции государств – членов Евразийского экономического союза» для крупного рогатого скота [98]. В этот перечень вошли следующие локусы: BM1818, BM1824, BM2113, ETH3, ETH10, ETH225, INRA023, SPS115, TGLA53, TGLA122, TGLA126, TGLA227. В исходных данных было выявлено значительное варьирование размеров амплифицируемых фрагментов STR-локусов, поэтому их набор был также ограничен в соответствии с рекомендациями Международного общества генетики животных (ISAG) и согласно встречаемому списку локусов в научной литературе (таблица 2).

Таблица 2 - Характеристика микросателлитных локусов

Локус	Хромосома	Структура последовательности	Источник	Референтный интервал
BM1818	D23S21	(TG) _n	Bishop et al. (1994) [143]	253-277
BM1824	D1S34	(GT) _n	Barendse et al. (1994) [141]	176-188
BM2113	D2S26	(CA) _n	Sunden et al. (1993) [244]	124-146
ETH3	D19S2	(GT) _n AC(GT) ₆	Solinas-Toldo et al. (1993) [247]	100-128
ETH10	D10S5	(AC) _n	Solinas-Toldo et al. (1993) [247]	206-222
ETH225	D9S2	(TG) ₄ CG(TG)(CA) _n	Steffen et al. (1993) [243]	139-157
INRA023	D3S10	(AC) _n	Vaiman et al. (1994) [248]	201-225
SPS115	D15	(CA) _n TA(CA) ₆	Baylor College of Medicine Human Genome Sequencing Center (2006) [145]	247-261
TGLA53	D16S3	(TG) ₆ CG(TG) ₄ (TA) _n	Georges & Massey (1992) [170]	151-187
TGLA122	D21S6	(AC) _n (AT) _n	Georges & Massey (1992) [170]	136-182
TGLA126	D20S1	(TG) _n	Georges & Massey (1992) [170]	111-127
TGLA227	D18S1	(TG) _n	Georges & Massey (1992) [170]	76-104

Анализ генетической структуры осуществлялся согласно общепринятым методикам [71, 176, 189, 193, 257, 259] и состоял из расчёта следующих показателей:

- число аллелей на локус (N_a);
- частота встречаемости аллеля для диаллельного представления генотипа

$$p = \frac{n}{2N} \quad (1)$$

где p - частота аллеля, n - число животных носителей аллеля; N - общее количество исследованных животных. Частота рассчитывается по каждому локусу;
 - число эффективных аллелей (уровень полиморфности)

$$Ne = \frac{1}{\sum p_{ij}^2} \quad (2)$$

где Ne - число эффективных аллелей (уровень полиморфности); p - частота встречаемости j -го аллеля для i -го локуса;

- наблюдаемая гетерозиготность

$$Ho = \frac{h_j}{n} \quad (3)$$

где Ho - наблюдаемая гетерозиготность; h_j - количество гетерозиготных генотипов в локусе; n - общее количество генотипов в локусе;

- ожидаемая гетерозиготность

$$He = 1 - \sum p_i^2 \quad (4)$$

где He - ожидаемая гетерозиготность; p_i - частота встречаемости i -го аллеля;
 - индекс фиксации Райта

$$F_{is} = \frac{He - Ho}{He} \quad (5)$$

где F_{is} - индекс фиксации Райта; He - ожидаемая гетерозиготность; Ho - наблюдаемая гетерозиготность;

- индекс Шеннона (индекс разнообразия Шеннона)

$$I = - \sum (p_i * \ln(p_i)) \quad (6)$$

где I - индекс Шеннона; p_i - частота встречаемости i -го аллеля.

- индекс Симпсона (филогенетический индекс или индекс разнообразия Симпсона) [131]:

$$D_{sm} = \sum p_i^2 \quad (7)$$

где D_{sm} - индекс Симпсона; p_i - частота встречаемости i -го аллеля.

В отличие от работы, опубликованной автором ранее [86] информативными были выбраны аллели с частотой более 5%. Стандартная ошибка частоты аллеля рассчитывалась. Соответствие ожидаемой и наблюдаемой гетерозиготности про-

верялось методом хи-квадратом, сравнение средних частот аллелей проводилось – точным тестом Фишера (100 000 симуляций). Для расчета была написана программа, верифицирующая данные и рассчитывающая показатели генетической изменчивости. Программа протестирована путем сравнения результатов (число аллелей, индекс фиксации Райта, частоты аллелей, индекс Шеннона, число эффективных аллелей) с результатами, полученными с помощью пакетов *adegenet* [188], *DiveRsity* [194] в R и *GenAlEx* [223, 242] (*MSEXcel*) на меньших выборках. Разработанное ПО позволило ускорить обработку больших массивов данных и выполнить все расчеты в единой программной среде.

2.4.2 Кластеризации предприятий по общим быкам, генетическому сходству по микросателлитам и вклад быков производителей в генеалогическую структуру популяции чёрно-пёстрого скота Западной Сибири

Предприятия сравнивались методами кластерного анализа. Для проведения кластерного анализа по встречаемости общих быков была создана частотная таблица отражающая сколько быков являются общими предками животных на предприятиях Западной Сибири. В выборку были взяты 9 предприятий, так как не в каждом хозяйстве представлена полная и достоверная информация о родословной и внесены микросателлитные профили у животных. Методы, которые использовались для построения кластеров были основаны на расчёте дистанций Евклида (*Euclidean Distance*), максимального расстояния/расстояние Чебышева (*Maximum Distance/Chebyshev Distance*), расстояния Манхэттена/расстояние «городских кварталов» (*Manhattan Distance/City Block Distance*), дистанции Канберры (*Canberra Distance*), бинарного расстояния/расстояние Жаккара (*Binary Distance/Jaccard Distance*) и расстояния Минковского (*Minkowski Distance*). Расположение кластеров основывалось на таких методах агломеративной иерархической кластеризации как метод Варда (*ward.D*), квадратов Варда (*ward.D2*), «ближайшего соседа» (*single*), «дальнего соседа» (*complete*), метод средней связи (*average*) и метод

медианной связи (median). Всего было построено 36 возможных вариантов отображения кластеров.

В основе кластеризации по микросателлитам использовались методы для расчёта генетических дистанций: Эдвардса (Edwards' Distance), Рэйнолдса (Reynolds' Distance), классические евклидовы дистанции Роджерса (Rogers' Distance), Провести (Provesti 's Distance) и Нея (Nei's Distance). К набору возможных вариантов иерархической кластеризации, описанному выше, были добавлены еще два метода: метод кластеризации МакКитти (mcquitty) и центроидный метод (centroid). Итого 40 возможных вариантов отображения кластеров по генотипам.

При расчёте генетических дистанций и проведении последующего кластерного анализа были взяты данные по тем же 9 предприятиям, так как именно такое число хозяйств содержали в базах данных информацию о микросателлитном профиле животных. Такой подход позволил сопоставить предприятия друг с другом, используя различные подходы.

Вклад быков производителей оценивался как общее число потомков (коров и нетелей в стаде) у быков, которые являются общими в независимости от хозяйства.

2.4.3 Генетическая структура по SNP

Материалом для исследования генетической структуры SNP служили образцы биоматериала - ушные выщипы у крупного рогатого скота голштинской породы ($n = 100$). Геномную ДНК выделяли с использованием колонок Nexttec Biotechnology GmbH, согласно инструкциям и методикам производителя.

Концентрацию ДНК определили путём её поглощения при длине волны $\lambda = 260$ нм. Электрофорез в агарозном геле позволил проверить состояние ДНК и оценить её качество, определив степень деградации и наличие примесей. С помощью чипов BovineSNP50 BeadChip провели скрининг SNP.

Для исследования были взяты только SNP-локусы, которые являлись общими для всех животных в исследуемой популяции. Проверка качества генотипирования, включала подготовку, контроль качества наборов данных, оценку основ-

ных популяционно-генетических показателей с применением библиотеки snpReady [174] (R) и программного обеспечения PLINK 1.9 [229] (усечение проводилось согласно GenCall (GC) – 0,5 и GenTrain (GT) – 0,3 [164, 235]). На втором этапе с помощью пакета snpReady к данным были применены фильтры: пороговое значение для удаления выборок из данных по частоте пропуска (sweep.sample = 0,8), пороговое значение для удаления маркера по частоте отсутствия генотипа (call.rate = 0,95), частота минорных аллелей менее 5 % (maf = 0,05). SNP с неизвестной позицией были исключены из исследования.

После предварительной подготовки данных исходных результатов геномной оценки рассчитывались показатели генетической изменчивости по тем SNP, которые согласно источникам литературы ассоциированы с признаками молочной продуктивности (приложение В).

Так, по всем отобраным локусам j с аллелями A_1 и A_2 были рассчитаны:

- частоты аллелей:

$$\begin{aligned} f(A_1) &= p_j = \frac{n A_1}{2N} \\ f(A_2) &= q_j = 1 - p_j \end{aligned} \quad (8)$$

- наблюдаемая гетерозиготность:

$$H_o = \frac{n H_j}{N} \quad (9)$$

- ожидаемая гетерозиготность:

$$H_e = 2 p_j q_j \quad (10)$$

- частота минорных аллелей:

$$maf = \min(p_j, q_j) \quad (11)$$

- генетическое разнообразие Нэя:

$$DG = 1 - p_j^2 q_j^2 \quad (12)$$

- уровень полиморфизма:

$$PIC = 1 - (p_j^2 + q_j^2) - 2 p_j^2 q_j^2 \quad (13)$$

- статистика равновесия Харди-Вайнберга:

$$\chi^2 = \frac{1}{d} \sum_{k=1}^3 \frac{(O_k - E_k)^2}{E_k} \quad (14)$$

где nA_1 количество повторений аллеля A_1 в популяции, N количество особей, nH_j количество гетерозиготных генотипов (типа A_1A_2 и A_2A_1) в локусе j . O_k это наблюдаемые значения для генотипов 0, 1 и 2, E_k это ожидаемые значения для $0 = N * (1 - p_j)^2$, $1 = N * 2p_j(1 - p_j)$ и $2N * 2p_j^2$. Статистическая проверка данных осуществлялась методом хи-квадрат и расчётом стандартной ошибки по частотам аллелей.

На основе полученных результатов были рассчитаны популяционные параметры (средние значения и границы DG, PIC, MAF, Ho, He).

2.5 Валидация баз данных на примере кластеризации стран по признакам молочной продуктивности

2.5.1 Кластеризация стран по признакам молочной продуктивности

Для проведения валидации зарубежных баз данных и сравнительного анализа популяций чёрно-пёстрого скота в различных странах была применена методология кластеризации, основанная на комплексе признаков молочной продуктивности, с использованием данных из международной базы CDN. Исходные данные для кластеризации были извлечены из сводных таблиц базы данных CDN. Для обработки и верификации данных использовался язык программирования PHP, что позволило создать структурированные массивы данных, которые в дальнейшем были импортированы в реляционную базу данных в СУБД MySQL, разработанную в рамках диссертационного исследования. Выборка данных для кластеризации была сформирована на основе критерия наличия информации о молочной продуктивности не менее чем для 50 оцененных животных из каждой страны, что обеспечило репрезентативность анализа. В итоговую выборку вошли данные по 819 358 животным голштинской породы из 28 стран мира, рожденных не ранее 1 января 2018 года. Анализ проводился по комплексу признаков молочной продуктивности, включающему оценочные племенные значения (EBV) по удою молока, выходу молочного жира (в килограммах и процентах), а также выходу молочного белка (в килограммах и процентах). Для оценки степени сходства между странами

по комплексу признаков молочной продуктивности был использован метод кластерного анализа, основанный на евклидовой метрике. В качестве алгоритма кластеризации был выбран метод Уорда, минимизирующий внутрикластерную дисперсию и позволяющий эффективно выявлять группы стран со схожими характеристиками молочной продуктивности. С целью снижения размерности исходных данных и устранения эффекта мультиколлинеарности был применен метод главных компонент (РСА). Отбор значимых главных компонент осуществлялся на основании критерия Кайзера-Гарриса, согласно которому учитывались компоненты с собственными значениями более 1, а также с использованием графического метода «каменистой осыпи», позволяющего визуально определить перелом на графике собственных значений и выделить наиболее значимые компоненты. Перед применением РСА была проведена стандартизация исходных данных для приведения признаков к единой шкале и устранения влияния различий в их масштабах. Статистическая обработка данных и визуализация результатов кластеризации и РСА осуществлялись с использованием специализированных пакетов, включая `factoextra`, `cluster` и пакет `stats` (R). Оценка качества кластеризации проводилась на основе визуального анализа дендрограмм, позволяющего интерпретировать структуру кластеров и выявить группы стран со схожими профилями молочной продуктивности.

2.5.2 Моделирование удоя по исходным данным региональных предприятий

Оценка вклада генетических и паратипических факторов в фенотипическую изменчивость исследуемых признаков осуществлялась с использованием линейных моделей и линейных смешанных моделей (`linear models and linear mixed models`) [214], реализованных в статистической среде R с применением пакетов `lme4` [168, 185], `lm` и `merTools`.

В качестве фиксированных эффектов в моделях рассматривались следующие категориальные предикторы: «Сезон отёла» (СО), «Год отёла» (ГО), «Сезон запуска» (СЗ), «Год запуска» (ГЗ), а также их взаимодействия второго порядка: «Сезон отёла : Год отёла» (СО:ГО), «Сезон запуска : Год запуска» (СЗ:ГЗ), «Сезон

отёла : Сезон запуска» (СО:СЗ), «Год отела : Год запуска» (ГО:ГЗ).

В качестве случайных эффектов (random effects) были включены: «Номер лактации», «Возраст первого плодотворного осеменения» (В1ПО), «Возраст первого отёла» (В1ОТ), «Идентификатор отца» (О) и «Индивидуальный идентификатор животного» (ID).

Оценка вклада фиксированных эффектов в общую дисперсию зависимых переменных производилась посредством дисперсионного анализа (ANOVA). Общая структура линейной смешанной модели представлена формулой 15.

$$y = Xb + Za + e \quad (15)$$

где:

$y = n \times 1$ вектор наблюдений; n = количество записей;

$b = p \times 1$ вектор фиксированных наблюдений; p = количество уровней для фиксированных эффектов;

$a = q \times 1$ вектор случайных эффектов; q = количество уровней для случайных эффектов;

$e = n \times 1$ вектор остаточных случайных эффектов;

X = проектная матрица порядка $n \times p$, которая связывает записи с фиксированными эффектами;

Z = проектная матрица $n \times q$, которая связывает записи со случайными эффектами.

В рамках построения смешанной модели случайные эффекты были представлены матрицей Z_i , отражающей структуру группирующих переменных. Каждый i -й случайный фактор характеризовался количеством уровней l_i . Для каждого такого фактора формировался соответствующий набор столбцов матрицы Z_i , общее количество которых (q_i) определялось произведением числа уровней l_i на число оцениваемых параметров p_i , относящихся к данному случайному фактору. Размерность матрицы Z_i напрямую зависела от количества этих столбцов (формула 16).

$$q = \sum_{i=1}^k q_i = \sum_{i=1}^k l_i p_i \quad (16)$$

Матрица Z_i , соответствующая i -му случайному эффекту, строилась посредством тензорного произведения Кронекера (Kronecker product) блоков J и X (17).

$$Z_i = (J_i^T * X_i^T)^T = \begin{bmatrix} J_{i1}^T & \otimes & X_{i1}^T \\ J_{i2}^T & \otimes & X_{i2}^T \\ \dots & \dots & \dots \\ J^T & \otimes & X^T \end{bmatrix} \quad (17)$$

где $*$ и \otimes - произведения Хатри-Рао и Кронекера [202], J_{ij}^T и X_{ij}^T - векторы строк j -й строки J_i и X_i . Эти строки сопоставлены с j -м элементом зависимого признака.

Матрица модели случайных эффектов Z , строилась из $k \geq 1$ блоков (18),

$$Z = [Z_1 Z_2 \dots Z_k] \quad (18)$$

Для оценки эффективности и предотвращения переобучения модели исходная выборка была разделена на две части: обучающую ($n = 166\ 766$) и тестовую ($n = 35\ 762$), реализуя, таким образом, процедуру однократной кросс-валидации (hold-out cross-validation). Модель обучалась на данных обучающей выборки, а ее прогностическая способность оценивалась на независимой тестовой выборке. Использовались только записи о дочерях быков, имеющих не менее 7 потомков в стаде, вследствие ограничения ряда критериев по объему совокупностей. Сопоставление моделей осуществлялось с использованием критериев Акаике и Шварца [203].

3. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

3.1 Оценка генеалогической структуры

3.1.1 Оценка качества входных генеалогических данных региональных предприятий

Для эффективной оценки качества первичных данных электронного зоотехнического учёта были предложены методы по поиску ошибок и их исправлений в генеалогических данных, а также по признакам молочной продуктивности, на основе материалов описанных в статье «Достоверность данных первичного зоотехнического учёта в молочном скотоводстве» [47].

Исследование родословных по племенным предприятиям Сибири выявило значительную вариабельность качества генеалогических данных в зависимости от исследуемых хозяйств. Широкое распространение искусственного осеменения и доступная информация о происхождении животных объясняют преобладающее число записей по отцовским рядам предков (таблица 3).

Таблица 3 - Ранги абсолютных частот пропущенных записей о происхождении в группах животных региональных предприятий

Группа	О	О	О	О	О	О	О	М	М	М	М	М	М	М	$r_{\text{ряд}}$ родословной / группа
	О	О	М	О	О	М	М	М	М	М	М	М	М		
	1	2	2	3	3	3	3	1	2	2	3	3	3		
Всего	1	3	7	5	9	11	13	2	8	4	14	12	10	6	0,778
Всего по быкам-осеменителям	2	3	5	6	11	9	13	1	8	4	14	10	12	7	0,827
Всего по быкам	1	3	8	5	11	7	13	2	10	4	14	9	12	6	0,705
Всего по коровам	2	4	3	7	9	12	13	1	5	6	14	11	10	8	0,876
Всего по телкам	2	6	3	9	7	13	11	1	4	5	12	14	8	10	0,876

Был рассчитан процент неполных записей о происхождении животных по трём рядам предков (приложение А,Б). Выявлено, что средний показатель такой

доли составляет в среднем около 7% по региону. Замечена неоднородность в пропущенных данных. Так, наибольший показатель по числу пропусков наблюдается по материнской линии (МММ) в третьем ряду родословной. При этом по отцовской линии (О) число записей характеризуется наибольшей полнотой.

Существует немало предприятий, где отбор производителей не носит системный, но часто стохастический характер, о чем может говорить 85,5% пропусков в записях разных уровней родословных. Это затрудняет эффективность оценки племенной ценности животных, а также снижает возможности для проведения качественного подбора родительских пар, который направлен на увеличение генетического потенциала животных. Создаётся угроза повышения степени гомозиготности маточного поголовья, уровень которой может быть скрыт от профильных специалистов.

Кластерный анализ, выполненный на основе ранжирования частоты отсутствующих генеалогических данных до третьего ряда родословной (рисунок 5), позволил выявить наличие двух основных кластеров. Это свидетельствует о значительных расхождениях в генеалогической структуре в связи с недостающими сведениями в родословной в представленных группах.

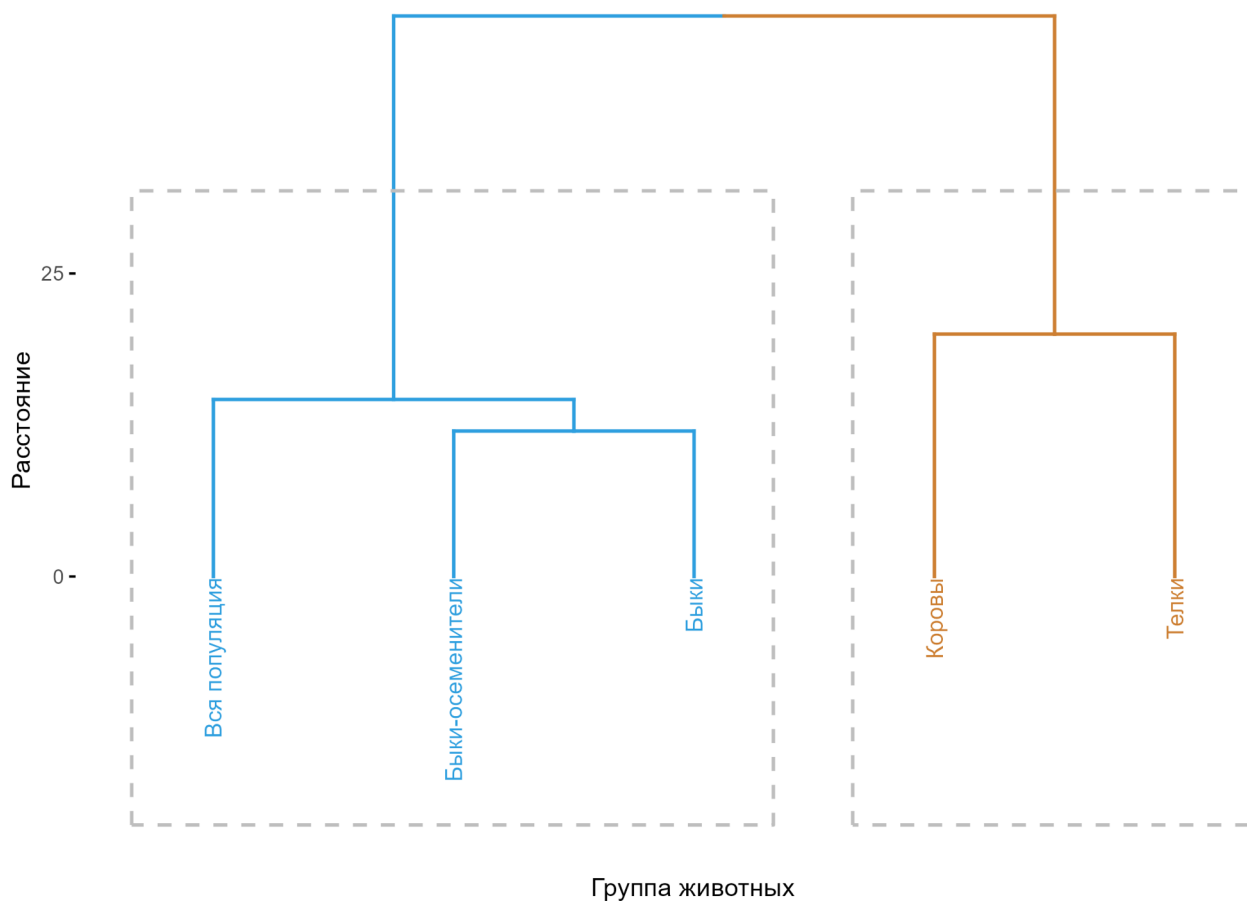


Рисунок 5 - Дендрограмма групп животных по рангам пропущенных генеалогических записей

Первый кластер объединяет всю популяцию, быков-производителей и племенных быков, демонстрируя высокую степень сходства в распределении рангов пропущенных записей. Это свидетельствует о практически идентичном характере распределения пропусков по предкам для быков и быков-производителей по сравнению с общей совокупностью животных. Подобный результат обусловлен тем, что быки составляют значительную часть популяции и их происхождение играет определяющую роль в селекционном процессе.

Второй кластер включает только коров и телок, что указывает на отличия в структуре рангов пропущенных генеалогических данных для этих групп относительно быков и всей популяции. Данное расхождение связано с особенностями ведения племенного учёта для различных половозрастных групп. Информация о происхождении коров и телок зачастую регистрируется менее тщательно по срав-

нению с данными о происхождении быков.

Высокое значение расстояния между двумя кластерами на дендрограмме подтверждает существенные различия в характере пропусков генеалогических данных. Использование рангов вместо абсолютных частот позволило нивелировать влияние аномальных значений и различий в общем количестве пропусков для разных групп, сфокусировавшись на сравнительном анализе структуры пропусков.

Кластеризация выполненная на основе ранжирования рангов частот недостающих генеалогических данных до третьего ряда родословной (рисунок 6), продемонстрировал наличие уже пяти кластеров, характеризующих структуру пропусков информации о происхождении животных. Кластеры формируются группами предков со схожими моделями отсутствующих сведений, что позволяет выявить закономерности в ведении генеалогического учёта.

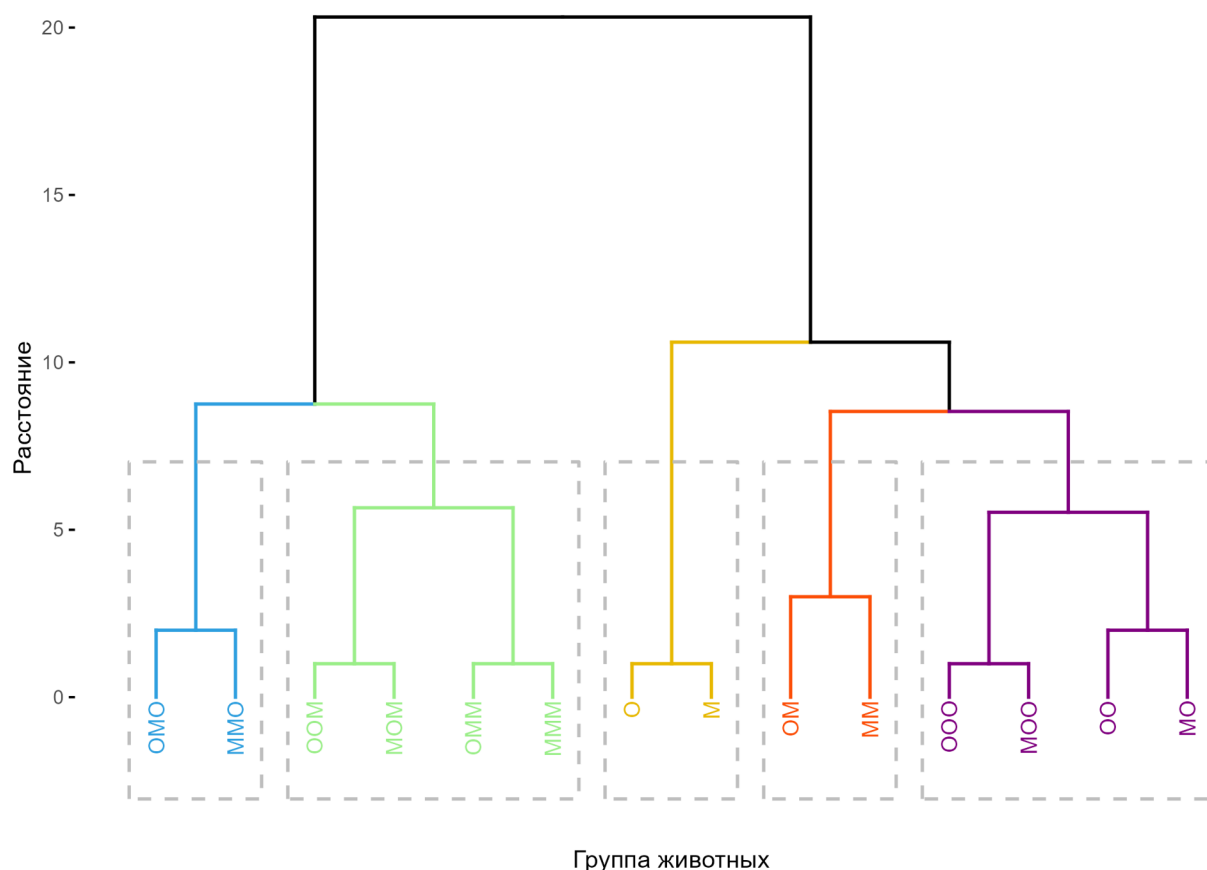


Рисунок 6 - Кластеризация комбинаций предков до третьего ряда родословной по рангам частот пропусков

Кластер 1 объединяет группы ОМО и ММО, представляющих собой предков третьего ряда родословной. Попадание этих групп на дендрограмме в один сегмент неслучайно и указывает на наличие общих факторов, обуславливающих пропуски информации о происхождении именно для этих комбинаций предков. Схожесть в структуре пропусков генеалогической информации может быть объяснена тем, что получение данных о родственниках третьего и более отдалённых поколений по материнской линии, представляет трудности исторического характера, когда записи велись преимущественно на бумажных носителях и часто оказываются несогласованными при оцифровке.

Во второй кластер вошли четыре группы включающие предков третьего ряда родословной (ООМ, МОМ, ОММ и МММ). Представлены варианты с двумя материнскими и одним отцовским предком, а также с одним материнским и двумя отцовскими.

Кластер 3 характеризуется наличием в нем предков первого ряда (отца и матери), О и М соответственно. Данный факт указывает на закономерность пропусков данных для непосредственных предков, что может быть связано с принятой практикой одновременной регистрации информации об отце и матери; следовательно, при отсутствии сведений об одном родителе высока вероятность отсутствия сведений и о другом.

Предки второго ряда, относящиеся к группе ОМ и ММ обособлены в 4 кластере. Наблюдаемая близость кластеров 4 и 3, вероятно, обусловлена закономерностью, при которой информация о родителях (предках первого порядка) оказывается более полной, чем информация о предках второго и более высоких порядков.

Пятый кластер объединяет предков ООО, МОО, ОО и МО, представляющих собой варианты групп предков как второго, так и третьего поколений. Наличие в одном кластере предков разных поколений может свидетельствовать о неоднородности процесса заполнения генеалогической информации. Вероятно, в практике племенного учёта заполнение данных до второго или третьего поколения часто обусловлено не столько стремлением к полноте и детализации родословных,

сколько наличием или отсутствием необходимой племенной документации. Решение о заполнении информации о предках определённого поколения принимается не систематически, а исходя из практических нужд в конкретный момент времени. Это объясняет, почему предки второго и третьего поколений оказались объединены в одном кластере, демонстрируя схожие модели пропусков генеалогической информации.

Полученные результаты кластерного анализа открывают возможности для дальнейших, более детальных исследований. Анализ конкретных рангов внутри каждого кластера может способствовать более глубокому пониманию причин возникновения пропусков в генеалогической информации и выявлению тех факторов и причин, которые влияют на полноту и точность ведения племенного учёта.

Анализ полноты заполнения генеалогической информации в базах данных (приложение А, Б) хозяйств Западной Сибири позволил выявить значительную вариабельность между предприятиями по качеству и заполненности представленных родословных, что оказывает прямое влияние на эффективность программ селекции и сохранения генетического разнообразия в популяциях животных.

Ранжирование хозяйств по количеству пропусков в родословных демонстрирует, что лишь небольшая часть предприятий уделяет достаточное внимание ведению и актуализации племенного учёта, что приводит к снижению точности оценки племенной ценности животных, ограничивает возможности для эффективного подбора родительских пар и целенаправленного отбора быков.

Выявлено, что хозяйства 16, 18 и 14 занимают высокие позиции по рангам частот недостающих данных в родословных, демонстрируя высокий уровень качества генеалогической информации. Такие результаты свидетельствуют о целенаправленной и систематической работе по корректировке и актуализации данных о происхождении животных в этих хозяйствах, что создаёт благоприятные условия для проведения эффективной селекции, направленной на улучшение продуктивных качеств и сохранение генетического потенциала в популяциях их животных.

Дендрограмма, построенная с использованием кластерного анализа, осно-

ванного на метрике Чебышева и методе полной связи, позволила отобразить максимальное расхождение между хозяйствами по проценту пропущенных записей для каждой из трёх категорий животных (быки, коровы, телки) (рисунок 7). Выбор метрики Чебышева был обусловлен тем, что она чувствительна к выбросам и позволяет выявить хозяйства, которые значительно отличаются от остальных по качеству генеалогической информации хотя бы для одной категории животных. Метод полной связи обеспечивает формирование компактных и изолированных кластеров, что облегчает интерпретацию результатов.

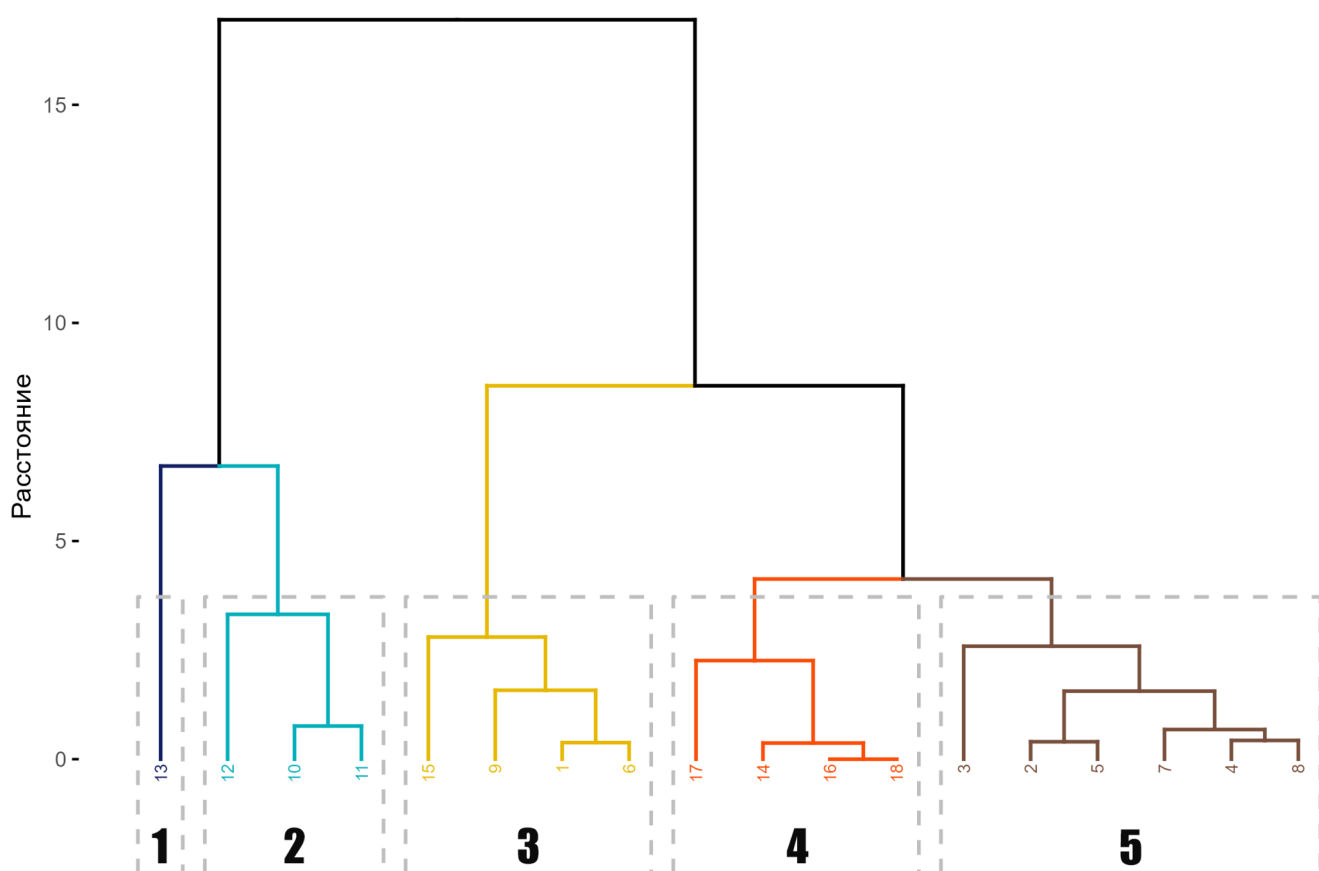


Рисунок 7 - Дендрограмма сходства хозяйств по качеству генеалогической информации

Анализ кластеров, сформированных на основе процента пропущенных генеалогических записей, позволил выявить группы хозяйств со схожим уровнем точности и полноты генеалогической информации, что может иметь важное значение для оценки эффективности селекционно-племенной работы в изучаемых

хозяйствах на региональном уровне.

Деление хозяйств на 5 кластеров, представленных на дендрограмме, позволяет провести детальный сравнительный анализ качества генеалогической информации в разных предприятиях. В обособленный кластер вошло хозяйство 13, которое характеризуется экстремально высоким уровнем пропусков генеалогических данных для быков (почти 17%), что значительно превосходит аналогичные показатели во всех остальных предприятиях и может свидетельствовать о серьезных недостатках в системе племенного учёта. Высокий процент пропущенных записей о происхождении быков (более 10% в хозяйствах 11 и 12) наблюдается в группе, которая входит во второй кластер, объединяющий хозяйства 12, 10 и 11. Примечательно, что в хозяйстве 12 наблюдается повышенный уровень пропусков генеалогической информации по телкам (более 1%). В то же время, хозяйство 10 характеризуется средними значениями пропусков в данных для коров и телок, не превышающими 1%.

Кластер 3 демонстрирует разнородность в отношении качества генеалогической информации, объединяя хозяйства 15, 9, 1 и 6, для которых характерен более высокий процент пропусков генеалогических записей в сравнении со средним значением по выборке. Внутри кластера наблюдается некоторая вариабельность в структуре пропусков. Так, хозяйства 9 и 15 демонстрируют повышенный процент пропусков по материнской линии - около 2% для коров в хозяйстве 9 и около 3% для телок в хозяйстве 15, что может свидетельствовать о недостаточном внимании к заполнению данных по этим группам животных в представленных хозяйствах. При этом хозяйства 1 и 6 демонстрируют более равномерное распределение пропусков генеалогической информации, составляющих 3-6% для быков и не более 0,7% для коров и телок. Такая структура пропусков, хотя и немного превышает средние значения по выборке, но может быть обусловлена случайными факторами и не оказывает существенного влияния на точность оценки племенной ценности.

Относительно более высоким качеством генеалогической информации выделяются предприятия 17, 14, 16 и 18, которые объединены в пределах кластера 4.

Минимальное количество пропусков (близкое к нулю), с учётом всех половозрастных групп животных, наблюдается по хозяйствам 16, 14 и 18. Но хозяйство 17 на фоне остальных предприятий представленных в данном кластере, выделяется большим процентом пропусков по коровам (более 2%), при этом сохраняя высокую степень заполненности родословной по быкам.

Кластер 5, включающий хозяйства 3, 2, 5, 7, 4 и 8, характеризуется высокой степенью полноты и точности генеалогической информации. Низкий процент пропущенных записей по всем трём анализируемым группам животных (быки, коровы, телки) свидетельствует об эффективной системе племенного учёта в этих хозяйствах, что создаёт благоприятные условия для реализации селекционных программ. Некоторое исключение составляет хозяйство 3, где отмечается несколько более высокий процент пропусков записей в родословных телок, что может указывать на необходимость уделять большее внимание ведению родословных для молодняка.

Проведённый кластерный анализ наглядно демонстрирует возможности данного метода для сравнительного ранжирования хозяйств по качеству генеалогической информации. Выявление групп хозяйств со схожими характеристиками может позволить разрабатывать дифференцированные рекомендации по улучшению ведения генеалогического учёта, учитывающие специфику каждой группы хозяйств. А это, в свою очередь, может способствовать повышению эффективности племенной работы как на региональном, так и федеральном уровнях.

В целом по региону доля отсутствующих записей до третьего ряда по всем группам животных достаточно высока и приближается к 7%. В структуре этого показателя на долю быков-производителей, происхождение которых неизвестно, приходится 85,5% от общего числа пропусков. Учитывая дублирование отдельных ветвей генеалогического древа стоит отметить, что этот показатель на практике может быть значительно выше и представляет серьёзное препятствие на пути к крупномасштабной селекции. Стоит также подчеркнуть, что доля животных, у которых лабораторно оценена достоверность происхождения, остаётся невысокой. Это приведёт к пересмотру генеалогической структуры в процессе массового

генотипирования маточного поголовья в регионе.

Созданные во время выполнения исследования инструменты (единая база данных голштинизированного скота, алгоритмы и методы визуализации генеалогических массивов данных и т.д.) позволили построить генеалогические деревья от пробанда до родоначальников линий и семейств по всем животным в исследуемых предприятиях. При этом выявить ещё ряд недостатков и искажений в родословных животных, представленных на предприятиях Западной Сибири. Примеры визуализации корректной генеалогической структуры одного из племенных предприятий представлены на рисунках 8-9.

Следует отметить, что детальный анализ генеалогических деревьев по данным региональных предприятий позволил выявить ещё ряд ошибок и манипуляций с данными. Одной из наиболее часто встречаемых ошибок было зацикливание и дублирование, которые могут возникать в процессе составления генеалогических деревьев. Эти ошибки искажают информацию о родственных связях и неизменно приводят к смещённой оценке племенной ценности.

В процессе обработки данных обнаружены случаи преднамеренного искажения данных. Так, наблюдалось присвоение одной матери непропорционально большого количества потомков (рисунок 10). Автоматизация идентификации таких ошибок достаточно простой процесс и предполагает создание глобальных генеалогических баз данных.

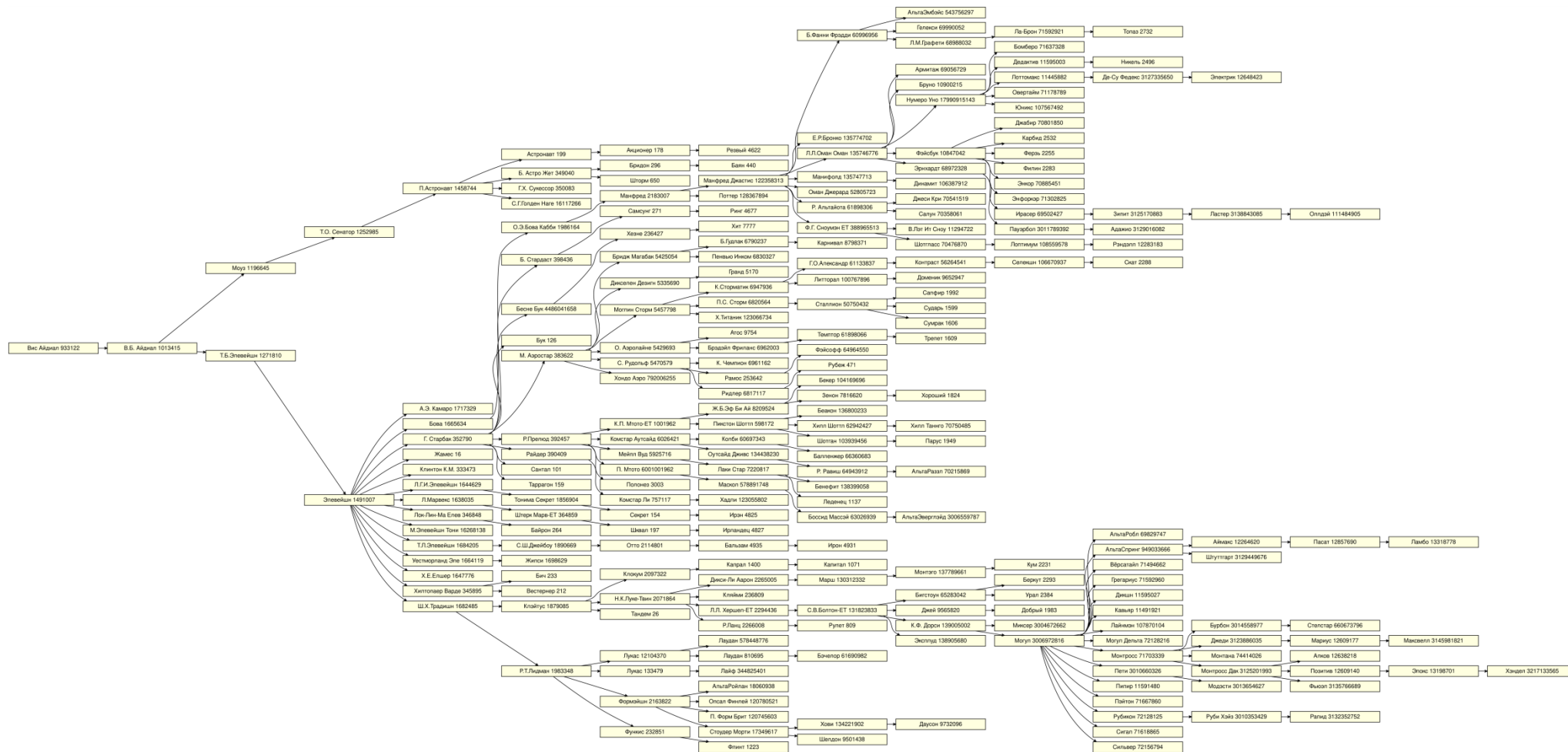


Рисунок 8 - Генеалогическое древо линии Вис Бэк Айдяла 1013415

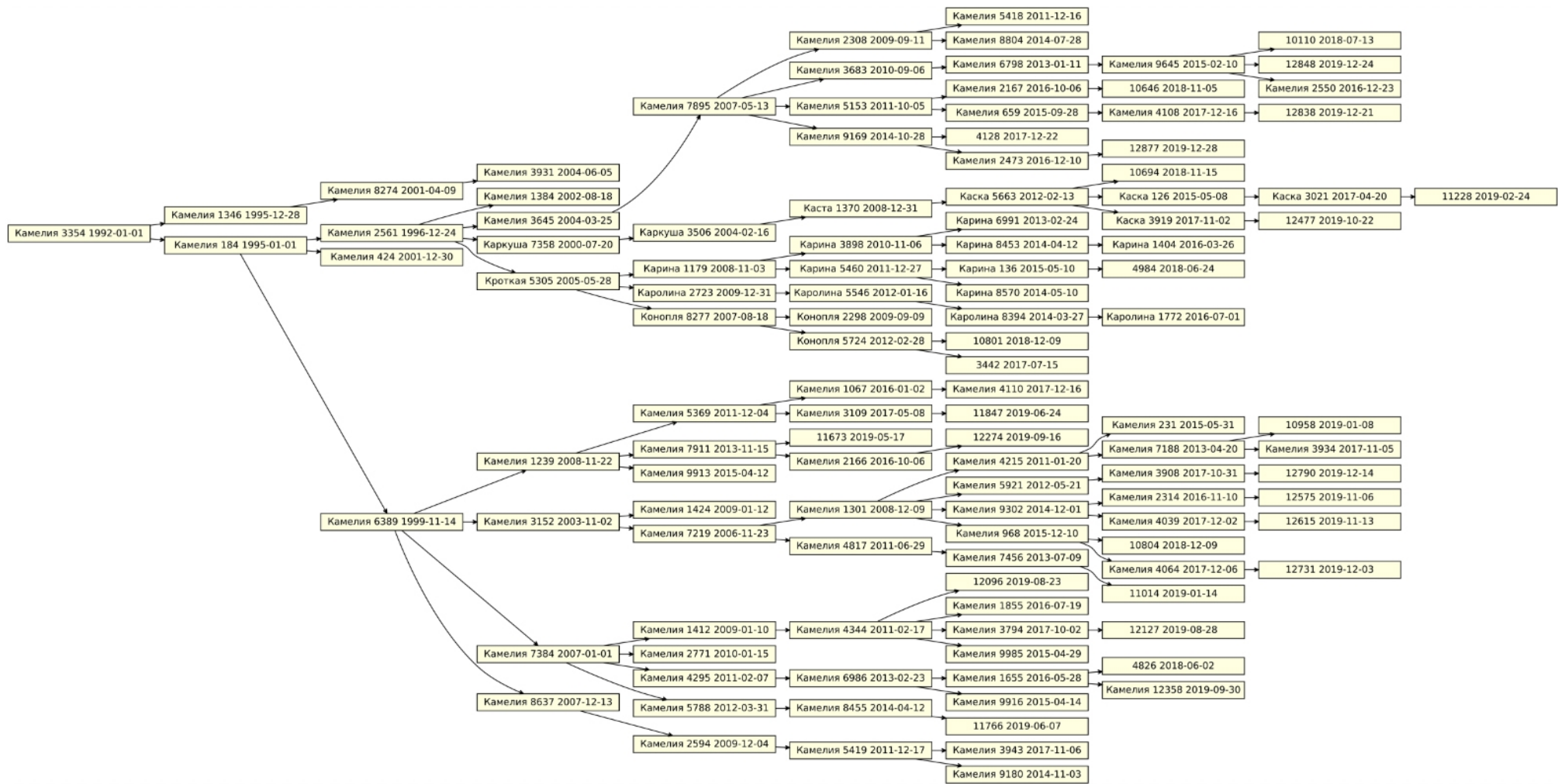


Рисунок 9 - Генеалогическое древо семейства Камелии 3354

ние искажений не представляется возможным, данные исключались из общей исследуемой выборки. При проведении анализа исходных родословных по базам данных животноводческих предприятий и описании подходов к выявлению искажений, пропусков, заикливания и дублирования, описанных в данном разделе, были получены материалы, опубликованные в следующих статьях рецензируемых журналов:

- Камалдинов, Е.В. Оценка качества генеалогических данных в племенных предприятиях Западной Сибири / Е.В. Камалдинов, А.Ф. Петров, К.Н. Нарожных и др. // Животноводство и кормопроизводство. – 2024. – Т.107, №4. – С. 53-67.

- Камалдинов, Е.В. Достоверность данных первичного зоотехнического учёта в молочном скотоводстве / Е.В. Камалдинов, А.Ф. Петров, К.С. Шатохин и др. // Вестник НГАУ. – 2022. – №2(63). – С. 76-83.

3.1.2 Генеалогическая структура чёрно-пёстрого скота Западной Сибири

Проанализировав число потомков (коров и нетелей) от различных быков-производителей в популяции черно-пестрого скота Западной Сибири (таблица 4) можно выделить ключевую особенность генеалогической структуры – интенсивное использование ограниченного числа производителей, преимущественно голштинской породы североамериканской селекции. Непропорционально большое количество дочерей у небольшой группы быков говорит о несбалансированном использовании генетических ресурсов с преобладанием импортного материала.

Генеалогические линии Рефлекшн Соверинг и Вис Бэк Айдиал доминируют в региональной популяции, что неудивительно, учитывая их широкое распространение и влияние голштинской породы по всему миру. Известно, что линия Рефлекшн Соверинг исторически закрепила за собой репутацию «молочной» линии, делающей акцент на высокой молочной продуктивности. Внутри линии Рефлекшн Соверинг наибольшее количество потомков имеют быки SOCRATES (2514), BOLIVER (1803), PLANET (1606) и SUPER (1554). Эти быки, как и многие другие в этой линии (MCCUTCHEN, GOLDWYN, SUPERSIRE), известны своей способ-

ностью передавать потомству высокую молочную продуктивность и хороший экстерьер. В линии Вис Бэк Айдиал выделяются DORCY (6042), O MAN (2891) и MAN-O-MAN (2706), также широко известные и ценимые за продуктивность и другие хозяйственно-полезные качества дочерей. Присутствие быка EARNHARDT P (298 потомков) интересно тем, что он является носителем гена безрогости (polled), что отражает современные тенденции в селекции. MOGUL, представитель более современной селекции, отличается сбалансированным набором качеств, улучшая как продуктивность, так и функциональные признаки. MANFRED, представитель более раннего поколения, заложил основы для многих последующих достижений линии, будучи надежным улучшателем продуктивности и типа телосложения. DORCY, лидер по количеству потомков из этой линии, сочетал в себе высокую продуктивность с улучшением экстерьера и, что немало важно, повышением резистентности к заболеваниям. O MAN и MAN-O-MAN известны продолжением тенденции к универсальности. Они передают потомству высокую и стабильную продуктивность, хороший экстерьер и отличную фертильность, что особенно ценится в условиях интенсивного воспроизводства.

Обращает на себя внимание явление, когда в качестве ООО встречаются представители разных поколений. Это служит свидетельством ретроспективной дискретизации генеалогической структуры, указывающей на определённую стратегию в селекционной работе. С одной стороны, имеются предприятия, использующие преимущественно молодых (геномно оцененных) быков, тогда как другая группа предприятий отбирает производителей с относительно большим количеством оцененных дочерей. В результате может наблюдаться некоторый смешанный эффект разных поколений одной родственной группы, вклад которой может значительно консолидировать маточное поголовье по тем или иным признакам отбора по сравнению с таковыми без соответствующего эффекта. В дальнейшем будет сделана попытка оценки дискретизации генеалогических данных, источником которой являются региональные предприятия в рамках стратегий реализуемых там стратегий селекционных программ.

Таблица 4 - Наиболее наполненные ветви чёрно-пёстрого скота Западной Сибири

Кличка	Потом-ков, гол.	Линия	EFI, %	Потом-ков, %	0000	00000	000000
DORCY 139005002	6042	ВБА	4,98	13,82	BOLTON 131823833		
O MAN 122358313	2891	ВБА	4,96	6,61	MANFRED 2183007		
MAN-O-MAN 135746776	2706	ВБА	6,22	6,19	O MAN 122358313	MANFRED 2183007	
MOGUL 3006972816	1802	ВБА	7,58	4,12	DORCY 139005002	BOLTON 131823833	
MTOTO 1962A	1140	ВБА	4,96	2,61			
MANFRED 2183007	1109	ВБА	5,75	2,54			
SNOWMAN 388965513	763	ВБА	9,03	1,74	O MAN 122358313	MANFRED 2183007	
MASCOL 578891748	725	ВБА	4,53	1,66	MTOTO MVT1962A		
SHOTTLE 598172	533	ВБА	6,01	1,22	MTOTO MVT1962A		
FREDDIE 60996956	408	ВБА	7,05	0,93	O MAN 122358313	MANFRED 2183007	
NUMERO UNO 17990915143	385	ВБА	4,37	0,88	MAN-O-MAN 135746776	O MAN 122358313	MANFRED 2183007
MONTROSS 71703339	371	ВБА	12,38	0,85	MOGUL 3006972816	DORCY 139005002	BOLTON 131823833
RUDOLPH 5470579	356	ВБА	2,1	0,81			
EARNHARDT P 68972328	298	ВБА	5,57	0,68	MAN-O-MAN 135746776	O MAN 122358313	MANFRED 2183007
BACCULUM-RED 9104500	284	ВБА	6,41	0,65			
JOCKO BESN 5694028588	281	ВБА	5,43	0,64			
JET STREAM 130558361	271	ВБА	4,36	0,62	JOCKO BESN 5694028588		
HILL 62942427	260	ВБА	8,7	0,59	SHOTTLE 598172	MTOTO 1962A	
BOLTON 131823833	252	ВБА	6,12	0,58			
RUBICON 72128125	247	ВБА	8,26	0,56	MOGUL 3006972816	DORCY 139005002	BOLTON 131823833
OUTSIDE 6026421	221	ВБА	3,41	0,51			
STORMATIC 6947936	214	ВБА	3,07	0,49	STORM 5457798		
ALMERE PABLO 198991265	213	ВБА	1,86	0,49			
LAWN BOY P-RED 133080890	200	ВБА	6,67	0,46	BACCULUM-RED 9104500		
STORM 5457798	188	ВБА	1,11	0,43			
FACEBOOK 10847042	185	ВБА	8,62	0,42	MAN-O-MAN 135746776	O MAN 122358313	MANFRED 2183007
LOTTOMAX 11445882	167	ВБА	9,02	0,38	NUMERO UNO 17990915143	MAN-O-MAN 135746776	O MAN 122358313
WINCHESTER 2205082	164	ВБА	4,74	0,38			
DELTA ELDORADO 190571881	158	ВБА	4,08	0,36			
JEEVES 134438230	156	ВБА	7,27	0,36	OUTSIDE 6026421		
LUKAS 133479	124	ВБА	5,2	0,28			
DURHAM 2250783	509	МЧ	6,6	1,16			
GENE 1887096	235	МЧ	3,09	0,54			

Кличка	Потом- ков, гол.	Линия	EFI, %	Потом- ков, %	0000	00000	000000
SOCRATES 133126053	2514	PC	5,55	5,75	BLITZ 17013604		
BOLIVER 123586443	1803	PC	6,37	4,12	AMEL 2231596		
PLANET 60597003	1606	PC	7,17	3,67	TABOO 17121203		
SUPER 62065919	1554	PC	5,53	3,55	BOLIVER 123586443	AMEL 2231596	
TABOO 17121203	1140	PC	4,97	2,61			
MCCUTCHEN 69990138	982	PC	5,95	2,25	BOOKEM 66636657	PLANET 60597003	TABOO 17121203
BOOKEM 66636657	963	PC	5,73	2,2	PLANET 60597003	TABOO 17121203	
EPIC 11104016	910	PC	7,58	2,08	SUPER 62065919	BOLIVER 123586443	AMEL 2231596
BW MARSHALL 2290977	879	PC	6,52	2,01			
GOLDWYN 10705608	777	PC	15,45	1,78	JAMES 5902195		
ALTAOAK 69169951	686	PC	6,0	1,57	NIAGRA 62067753	BOLIVER 123586443	AMEL 2231596
SUPERSIRE 69981349	629	PC	7,91	1,44	ROBUST 64966739	SOCRATES 133126053	BLITZ 17013604
ROBUST 64966739	627	PC	5,78	1,43	SOCRATES 133126053	BLITZ 17013604	
JUROR M2124357	585	PC	5,18	1,34			
BLITZ 17013604	513	PC	10,76	1,17			
MONTEREY 69087180	505	PC	8,9	1,15	MCCUTCHEN 69990138	BOOKEM 66636657	PLANET 60597003
OCTOBERFEST 3013129308	438	PC	6,34	1,0	ALTAOAK 69169951	NIAGRA 62067753	BOLIVER 123586443
DOMAIN 137974489	433	PC	4,59	0,99	TOYSTORY 60372887	BW MARSHALL 2290977	
NIAGRA 62067753	364	PC	5,12	0,83	BOLIVER 123586443	AMEL 2231596	
JAMES 5902195	285	PC	3,41	0,65			
AMEL 2231596	269	PC	6,1	0,62			
DOORSOPEN 3013129323	243	PC	7,09	0,56			
ADDISON 839380546	233	PC	2,42	0,53			
FACTOR 2129913	182	PC	1,42	0,42			
ABBOTT 388380500583	158	PC	8,46	0,36			
JESTHER 5994022699	135	PC	3,49	0,31			
OBSERVER 65917481	127	PC	6,22	0,29	PLANET 60597003	TABOO 17121203	
TUFFENUFF 71451770	119	PC	8,27	0,27			
JEROD 69990187	111	PC	6,82	0,25	OBSERVER 65917481	PLANET 60597003	TABOO 17121203
TOYSTORY 60372887	104	PC	4,68	0,24	BW MARSHALL 2290977		

* ВБА - Вис Бэж Айдиал; МЧ - Монтвик Чифтейн; РС - Рефлекшн Соверинг

3.1.3 Роль отцовских предков в формировании генеалогической структуры маточного поголовья Западной Сибири

Для оценки роли отдельных быков-производителей в формировании генеалогической структуры чёрно-пёстрого скота Западной Сибири был проведен анализ данных о потомстве, полученных из баз данных племенных предприятий региона. В частности, были выделены топ-20 быков-производителей по количеству дочерей в популяции, представленных как среди коров основного стада (таблица 5), так и среди нетелей (таблица 6). Анализ данных таблиц позволяет выявить наиболее частотных представителей отцовских предков, оказавших наибольшее влияние на генеалогическую структуру популяции.

Таблица 5 – Топ-20 быков-производителей по количеству коров (дочерей) в популяции чёрно-пёстрого скота Западной Сибири

Индивидуальный номер	Кличка	Дата рождения	Количество потомков
HOUSAM68999396	BLUE-HORIZON ALTASUPLEX	2011-02-12	1114
HOUSAM66011447	SULLY ALTAMETEOR-ET	2009-02-20	680
HOUSAM71864244	SYNERGY ALTAPARQUET-ET	2012-10-27	642
HOUSAM70715474	DYCREST ALTASANFORD-TW	2011-10-02	630
HOCANM12302930	PROGENESIS FORTUNE	2015-08-12	589
HOUSAM71588596	BOMAZ ALTABRANSON-ET	2013-05-11	479
HOUSAM69916619	MILLER-FF ALTAELLIOT-ET	2011-06-23	458
HOUSAM71088584	FARNEAR PAR ALTAPRESET-ET	2011-12-01	437
HODEUM537222111	ALTA ECHELON	2012-02-10	420
HOUSAM61898306	REGANCREST ALTAIOTA-ET	2005-06-12	404
HOUSAM67681058	DIRT-ROAD ALTACHATTER-ET	2012-09-17	399
HOCANM107567492	CROTEAU LESPERRON UNIX	2012-06-17	386
831337	Интендант-М	2002-03-04	317
HOUSAM69829746	LONE-OAK-ACRES ALTARABO-ET	2012-10-14	275
HOUSAM69829747	LONE-OAK-ACRES ALTAROBLE-ET	2012-10-12	266
HONLDM426436885	WALRICH DELTA OAKLAND	2006-07-30	240
HOCANM106627797	DULET KICKBALL	2011-01-30	237
HO840M3006559787	NO-FLA ALTAEVERGLADE-ET	2010-10-05	235
HOCANM12192429	SILVERRIDGE V EUGENIO	2015-03-10	230
HOUSAM71631016	SPRUCE-HAVEN ALTATD-RED-ET	2013-02-13	203

Таблица 6 – Топ-20 быков-производителей по количеству нетелей (дочерей) в популяции чёрно-пёстрого скота Западной Сибири

Индивидуальный номер	Кличка	Дата рождения	Количество потомков
HOUSAM69829747	LONE-OAK-ACRES ALTAROBLE-ET	2012-10-12	1417
HOUSAM69829746	LONE-OAK-ACRES ALTARABO-ET	2012-10-14	1074
HODEUM537222111	ALTA ECHELON	2012-02-10	698
HOUSAM142289925	LEADERWIN MUSCADET-ET	2012-12-23	549
HOUSAM71286351	HAR-DALE-ACRES-JP ALTACR-ET	2012-11-09	387
HO840M3129038326	IHG ALTAMAXIMUS-ET	2015-09-03	377
HO840M3125475825	PEAK ALTAZ28-ET	2014-08-09	367
HO840M3006559787	NO-FLA ALTAEVERGLADE-ET	2010-10-05	363
HO840M3129128746	PEAK ALTATURNKEY-ET	2015-03-27	358
HOUSAM71631016	SPRUCE-HAVEN ALTATD-RED-ET	2013-02-13	341
HO840M3130854061	VIEW-HOME CARDINALS-ET	2015-05-19	335
HOUSAM70715474	DYCREST ALTASANFORD-TW	2011-10-02	333
HOFRAM7261003981	EBA ALTAJETHRO	2014-06-27	258
HOCANM11446030	ROCKYMOUNTAIN ALTAUNDADO	2015-01-27	242
HOUSAM69990160	DE-SU ALTALAKER-ET	2011-02-02	240
HOITAM17991056452	AQUILA ALTASEDARA	2011-07-25	231
HOCANM12719190	PROGENESIS MISSION	2017-03-27	218
HOUSAM70694639	ROSYLANE-LLC ALTASTALWART	2012-10-14	204
HO840M3129449398	NO-FLA ALTAEXEMPLAR-ET	2015-05-13	203
HOUSAM66726504	SANDY-VALLEY ALTAXXX-RED-ET	2009-07-19	197

Полученные результаты указывают на ряд важных тенденций и особенностей генеалогической структуры популяции чёрно-пёстрого скота Западной Сибири. Так, отмечается высокая интенсивность использования отдельных производителей, в частности, быка BLUE-HORIZON ALTASUPLEX с 1114 дочерьми (коровы), что может свидетельствовать о его выдающихся племенных качествах. В списке представлены быки, относящиеся к различным линиям и ветвям, что положительно сказалось на поддержании генетического разнообразия. Тенденция использования быков североамериканской селекции является характерной чертой современных программ селекции молочного скота, ориентированных на получение максимального генетического прогресса. Выбранная стратегия, с одной стороны, позволяет быстро улучшить продуктивные качества стада, но, с другой стороны, может потенциально привести к снижению генетического разнообразия и увеличению инбридинга в популяции. Разброс по годам рождения быков (2002-2015 гг.) указывает на использование как проверенных временем, так и более мо-

лодых производителей преимущественно геномно-оцененных.

Быки, лидирующие по количеству дочерей-коров, могут представлять более раннее поколение производителей, активно использовавшихся в прошлые годы, в то время как быки, лидирующие по количеству дочерей-тёлок, могут отражать современные селекционные тренды и предпочтения, ориентированные на использование более молодых и перспективных производителей.

Анализ таблицы с топом 20 быков-производителей по количеству дочерей-тёлок (нетелей) показал еще более высокую интенсивность использования отдельных быков в популяции телок, с максимальным количеством потомков у LONE-OAK-ACRES ALTAROBLE-ET (1417 дочерей), что может быть связано с уменьшением интервала между поколениями, широким использованием сексированного семени и общей тенденцией к увеличению использования наиболее востребованных производителей. Отмечено пересечение некоторых быков в обоих топовых списках, что указывает на их стабильно высокую оценку.

В целом, по обоим спискам, наблюдается высокая степень зависимости исследуемой популяции чёрно-пёстрого скота Западной Сибири от ограниченного числа быков-производителей, преимущественно зарубежной селекции. Это, способствуя быстрому генетическому улучшению стада по молочной продуктивности, создает риски повышения инбридинга и снижения генетического разнообразия. В связи с этим, для оптимизации селекционно-племенной работы, необходима диверсификация быков с включением производителей отечественной селекции, внедрение методов геномной селекции, разработка программы сохранения генетического разнообразия, включающей криоконсервацию генетического материала, а также использование более широкого спектра молекулярно-генетических маркеров, ассоциированных с хозяйственно полезными признаками.

Анализ кличек отцов и их индивидуальных номеров, представленных в таблицах, подтверждает тенденцию к голштинизации, доминирующую в отечественном молочном скотоводстве на протяжении последних десятилетий. Наряду с иностранными производителями, в топ-20 быков-производителей по количеству дочерей-коров в популяции чёрно-пёстрого скота Западной Сибири представлен и

бык зарубежной селекции, но работающий на территории РФ, такой как Интендант-М. Однако, его вклад значительно ниже вклада иностранных производителей, что объясняется отсутствием в России эффективно работающей системы оценки генетического потенциала чёрно-пёстрого скота. Необходимо отметить, что в топ-20 быков-производителей по количеству дочерей-нетелей в исследуемой популяции не вошел ни один бык отечественной селекции. Отсутствие в топ-20 быков отечественной селекции второго и последующих поколений, сопоставимых по качеству потомства с зарубежными аналогами, как отмечалось ранее, является серьезным вызовом для отечественной племенной работы и требует принятия мер по развитию собственной селекционной базы и повышению конкурентоспособности отечественных быков-производителей.

3.2 Валидация данных первичного цифрового учёта на примере признаков молочной продуктивности предприятия

В ходе анализа были рассчитаны медианные значения максимальных и минимальных признаков, которые могут влиять на продолжительность лактации. Полученные результаты дают общее представление о доступном цифровом материале без предварительной фильтрации и верификации. Эти результаты являются важным показателем качества доступного цифрового материала и дают представление о его полноте и достоверности в контексте исследуемых регионов.

Валидация данных на примере показателей молочной продуктивности предполагала фильтрацию животных по ряду репродуктивных признаков, которые позволили выделить отобрать животных с приемлемыми границами варьирования. Установлено значительное превышение средних уровней по лимитам значений таких признаков относительно приемлемых границ (таблица 7).

Таблица 7 - Приемлемые границы некоторых признаков, связанных с лактацией

Название параметра	Значения по базам*	Приемлемые границы
возраст первого плодотворного осеменения (мес.)	от $6,9 \pm 1,04$ до $116,4 \pm 8,7$	13-18
сухостойный период	от $-1,2 \pm 0,2$ до $373,2 \pm 72,4$	30-120
сервис-период	от $13,2 \pm 1,8$ до $1012,6 \pm 106,7$	30-120
межотельный период	от $236,9 \pm 14,9$ до $1226,5 \pm 98,4$	350-470

* - медианные значения лимитов изменчивости

Применение фильтрации (таблица 1) по сервис-периоду позволило исключить из совокупности животных с возможными проблемами, связанными с репродуктивными качествами. Усечение данных привело к снижению среднего значения критерия Андерсона-Дарлинга с 2,905 до 2,89, уменьшив объём совокупности и количество групп (с 301 до 268). Без фильтрации на долю совокупностей с негауссовским распределением приходилось 56,1%, в то время как с фильтрацией эта доля была несколько выше (58,2%).

Согласно результатам тестирования, представленным в приложении Г, были визуализированы варианты распределения признаков с самыми высокими значениями критериев, среди которых преимущественно встречались такой признак, как массовая доля молочного жира. Также встречался в списке признаков удой за 305 дней лактации который характеризовался двухвершинной кривой распределения (рисунок 11-13).

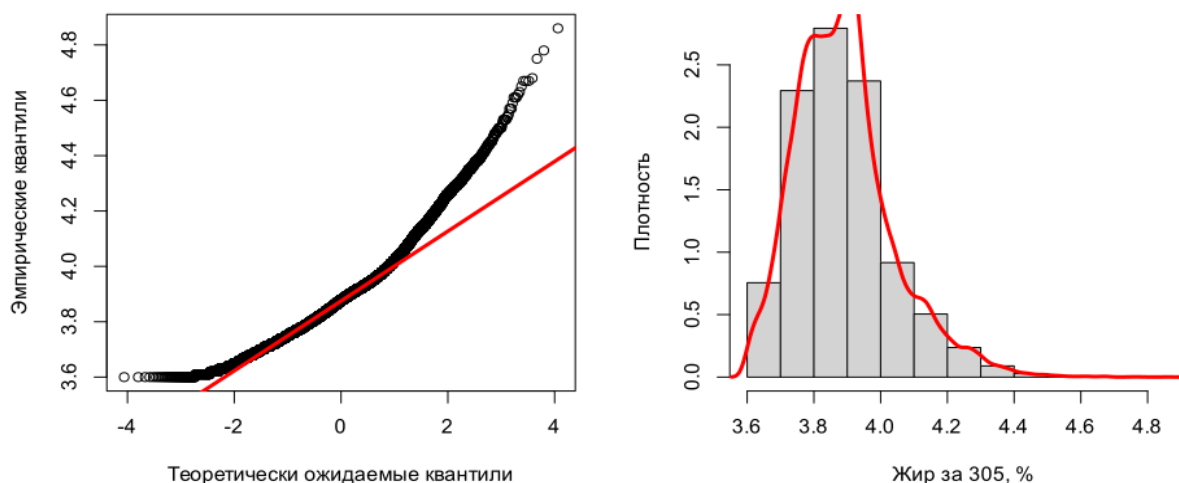


Рисунок 11 - Распределение массовой доли молочного жира за 305 дней первой лактации (Хозяйство 4, AD=47,0***)

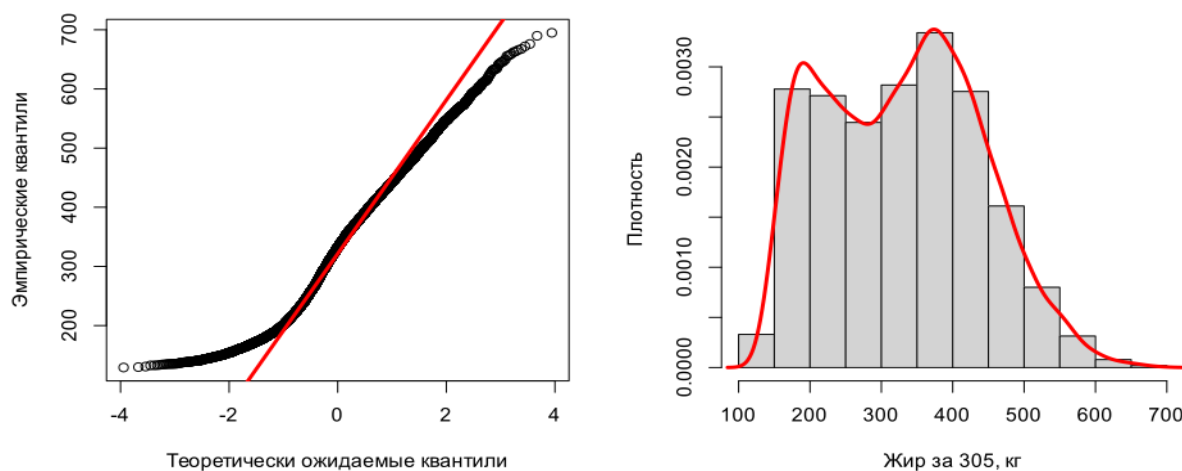


Рисунок 12 - Распределение содержания молочного жира (кг) за 305 дней второй лактации (Хозяйство 3, AD=27,0***)

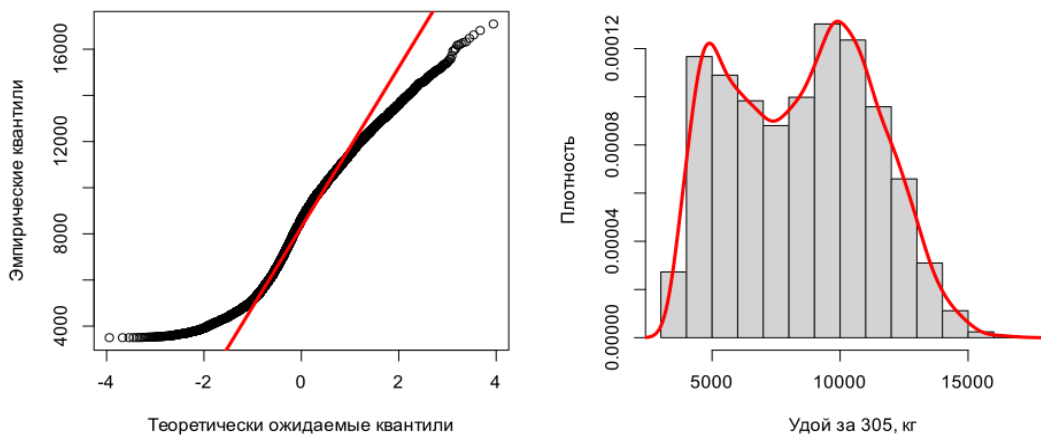


Рисунок 13 - Распределение удоя (кг) за 305 дней второй лактации
(Хозяйство 3, AD=23,2***)

Характерной особенностью всех тестируемых распределений по массовой доле жира является наличие длинного «хвоста» справа от средней арифметической. По удою за 305 дней лактации, как и по концентрации молочного жира (кг) отмечается присутствие двух вершин, что возможно говорит о трансгрессии, когда распределения двух совокупностей накладываются друг на друга. Скорее всего на предприятиях не практикуется выбраковка низкопродуктивных животных, которых оставляют в стаде для получения молока. Генетический потенциал таких животных в условиях племенных предприятий не только не реализуется, но и создаются предпосылки для формирования соответствующего генофонда маточного стада. Данный процесс становится ещё более выраженным в связи с деятельностью иностранных компаний, импортирующих племенной материал из-за рубежа не всегда высокого качества. Отсутствие национальной системы оценки племенной ценности в РФ также способствует снижению генетического потенциала особенно в случае возможного эмбарго на поставку семени, эмбрионов или племенных животных. В целом картина распределения содержания молочного жира (кг) повторяла таковое по удою (кг) за 305 дней лактации вне зависимости от лактации.

Предложенная выше методика может способствовать совершенствованию и разработке новых методов оценки племенных и продуктивных качеств сельскохозяйственных животных, выявляя недостатки в ведении племенного учета, которые могут исказить будущую оценку. Данный раздел опирался на методику опубликованную в статье:

- Камалдинов, Е.В. Достоверность данных первичного зоотехнического учёта в молочном скотоводстве / Е.В. Камалдинов, А.Ф. Петров, К.С. Шатохин и др. // Вестник НГАУ. – 2022. – №2(63). – С. 76-83.

3.3 Валидация входных данных первичного цифрового учёта на примере признаков молочной продуктивности зарубежного происхождения

В результате кластерного анализа было выделено три основных кластера стран, различающихся по уровню молочной продуктивности (рисунок 14).

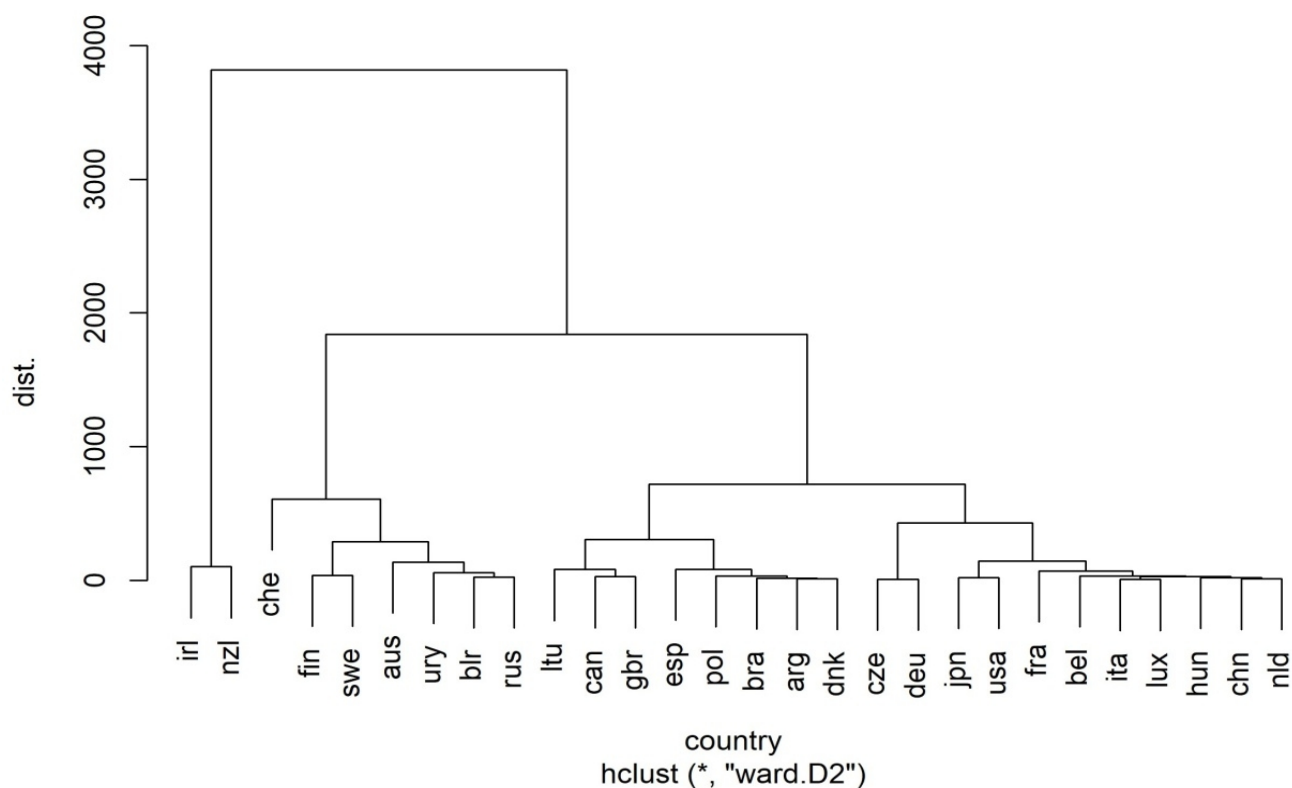


Рисунок 14 - Кластеризация стран по признакам молочной продуктивности

Первый кластер (изолированный) включил Ирландию (IRL) и Новую Зеландию (NZL). Данные страны характеризуются наиболее низкими показателями абсолютной молочной продуктивности (удой, выход жира и белка в кг), но при этом демонстрируют высокие значения относительных показателей (содержание жира и белка в процентах) (см. таблицу 8).

Второй кластер объединил Швейцарию (CHE), Финляндию (FIN), Швецию (SWE), Уругвай (URY), Беларусь (BLR) и Россию (RUS). Для этих стран характерна низкая вариабельность рангов, что может свидетельствовать о более сбалансированном подходе к селекции, ориентированном не только на удой, но и на другие признаки, такие как адаптивность и здоровье животных.

Таблица 8 - Ранги стран по признакам молочной продуктивности

SS*	Средний ранг	Жир (%)	Жир (кг)	Белок (%)	Белок (кг)	Удой (кг)	Страна
2,8	3,7	6,5	4	3	1	4	США
2,2	5,8	5	8	4	6	6	Нидерланды
10,2	6,8	3	2	10	2	17	Великобритания
4,0	7,0	11	5	7	3	9	Китай
3,8	7,2	11	9	7	4	5	Бельгия
8,5	7,5	16	13,5	2	5	1	Чехия
12,8	8,2	8	21	1	9	2	Германия
6,4	9,6	16	15	7	7	3	Япония
7,3	10,2	6,5	17,5	7	10	10	Венгрия
3,7	10,3	11	6,5	14	8	12	Польша
7,6	12,4	4	20	7	17	14	Дания
4,5	13,0	11	17,5	12,5	13	11	Франция
6,7	14,8	21,5	11	15,5	11	15	Аргентина
5,1	14,9	16	6,5	20	13	19	Литва
5,0	15,0	11	10	19	15	20	Австралия
7,3	15,2	16	22,5	12,5	17	8	Италия
8,6	16,4	20	25	11	19	7	Люксембург
10,4	16,6	1	1	27	27	27	Новая Зеландия
6,3	16,7	23	19	15,5	13	13	Бразилия
4,0	17,5	21,5	13,5	17,5	17	18	Канада
10,2	17,8	2	3	28	28	28	Ирландия
6,0	21,0	24,5	27	17,5	20	16	Испания
3,4	21,1	24,5	12	24	21	24	Швеция
2,6	22,4	16	24	22	25	25	Финляндия
3,4	22,6	19	16	26	26	26	Швейцария
2,9	23,6	26,5	22,5	24	22	23	Уругвай
3,9	24,1	26,5	28	21	23	22	Беларусь
3,4	24,6	28	26	24	24	21	Россия

*SS - средняя сумма квадратов отклонений рангов.

Третий, наиболее многочисленный кластер, включал остальные 19 стран, во

главе с США и Нидерландами, которые продемонстрировали наивысшие показатели молочной продуктивности. Данный кластер характеризуется выраженной дифференциацией стран по признакам молочной продуктивности, что может быть обусловлено различиями в интенсивности селекции и применяемых технологиях содержания и кормления скота.

Для визуализации дифференциации стран в пространстве признаков молочной продуктивности был использован PCA (рисунок 15). Первая главная компонента (Dim1) объясняет 64% дисперсии и отражает общую молочную продуктивность (абсолютные значения удоя, выхода жира и белка). Вторая главная компонента (Dim2) объясняет 30.8% дисперсии и характеризует относительные показатели – содержание жира и белка в молоке (%).

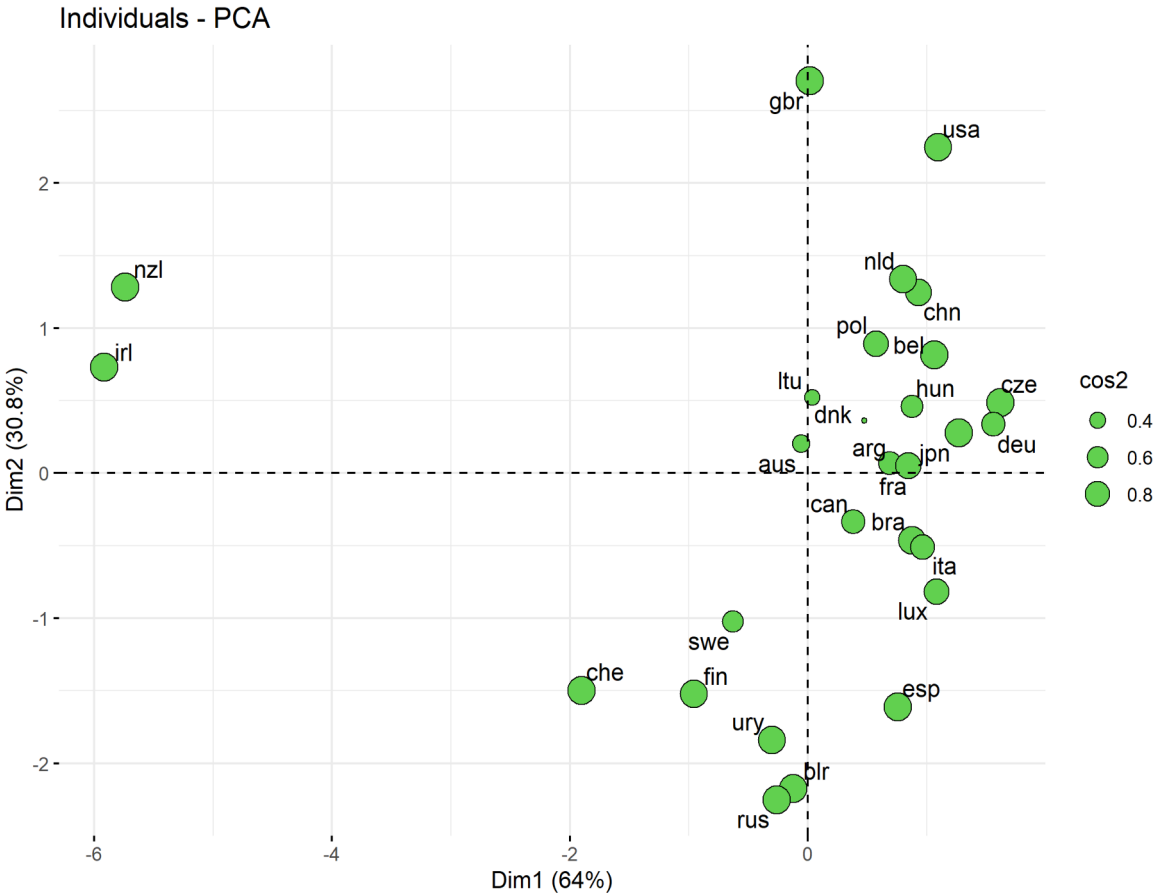


Рисунок 15 - Распределение стран по признакам молочной продуктивности в пространстве главных компонент

Размещение стран на диаграмме PCA подтверждает результаты кластерного

анализа. США и Нидерланды, лидирующие по общей молочной продуктивности, располагаются в правой части диаграммы. Ирландия и Новая Зеландия, с акцентом на пастбищное содержание, выделяются в нижней части диаграммы.

Кластеризация стран по признакам молочной продуктивности выявила существенные различия в селекционных стратегиях и подходах к ведению молочного скотоводства в разных регионах мира. Лидирующие позиции США и Нидерландов, объединённых в третий кластер, подтверждают эффективность интенсивной селекции, направленной на максимальное увеличение молочной продуктивности. В то же время, выделение в отдельный кластер Новой Зеландии и Ирландии, ориентированных на пастбищное содержание и относительные показатели качества молока, свидетельствует о существовании альтернативных, экономически оправданных моделей молочного скотоводства.

Второй кластер, объединивший ряд европейских и постсоветских стран, характеризуется более сбалансированным подходом к селекции, что может быть обусловлено стремлением к устойчивости производства и учёта не только продуктивных, но и адаптивных качеств животных. В частности, низкие ранги России и Беларуси могут указывать на необходимость интенсификации селекционной работы и повышения генетического потенциала молочного скота в данных странах.

Проведённая валидация зарубежных баз данных на примере CDN с использованием методов кластерного и многомерного анализа подтвердила наличие значительной дифференциации стран по признакам молочной продуктивности голштинского скота. Результаты исследования позволяют:

- Объективно оценить позиции различных стран в мировом молочном скотоводстве по комплексу признаков молочной продуктивности.
- Выявить группы стран со схожими селекционными стратегиями и подходами к ведению молочного скотоводства.
- Подтвердить валидность данных CDN для проведения международных сопоставлений и бенчмаркинга в молочном скотоводстве.
- Использовать полученные результаты для разработки селекционных программ, адаптированных к специфическим условиям различных регионов, в

том числе и Сибири.

Данный раздел демонстрирует возможность применения методов многомерного статистического анализа для валидации больших массивов данных из зарубежных источников и их использования для сравнительной оценки генетического потенциала и селекционных стратегий в мировом молочном скотоводстве. Полученные результаты могут быть использованы в дальнейших исследованиях, направленных на совершенствование селекционно-племенной работы с чёрно-пёстрым скотом Западной Сибири с учётом мировых тенденций и региональных особенностей.

Материалы представленные в данном разделе отражены в статье:

- Petrov, A.F. Clustering of countries based on dairy productivity characteristics of Holstein cattle for breeding material selection / A.F. Petrov, O.V. Bogdanova, K.N. Narozhnykh and etc. // *Veterinary World*. – 2024. – Vol. 17. – С. 1108-1118.

3.4 Роль паратипических факторов в изменчивости удоя региональной субпопуляции черно-пёстрого скота

Возможность использования верифицированных данных с целью прогноза уровней селекционируемых признаков является главным критерием качества входных данных. С этой точки зрения, построение и интерпретация математических моделей является важным звеном валидации данных в ретроспективном аспекте. Имплементация такого процесса необходима не только для оценки качества иностранных источников данных, но и для характеристики пригодности уточнённых региональных баз данных животноводческих предприятий. Это позволит не только охарактеризовать применимость уже разработанных цифровых решений для цифровой крупномасштабной селекции, но и перспективных. В нашем исследовании для решения такой задачи проведено изучение факторов, влияющих на изменчивость удоя черно-пёстрого скота в условиях промышленных комплексов. На примере одного из ведущих племенных заводов Новосибирской области. Массив данных охватывал период с 2000 по 2020 годы и включал информацию о 319 210 головах скота для оценки значимости паратипических факторов

в вариабельности удоя, для чего был применен метод математического моделирования с использованием смешанных регрессионных моделей.

В качестве фиксированных эффектов в модель были включены «Год отёла», «Сезон отёла», «Год запуска», «Сезон запуска», а также их взаимодействия. Среди случайных эффектов рассматривались «Отец», «Лактация», «Возраст первого плодотворного осеменения», «Возраст первого отёла» и «Идентификационные данные животного». Анализ проводился с применением линейных смешанных регрессионных моделей и соответствующих статистических критериев, используя библиотеки lme4, lm и merTools в среде статистического программирования R. Для повышения точности оценок и нивелирования влияния выбросов, исходные данные были подвергнуты усечению, с установлением пороговых значений для удоя, содержания молочного жира и длительности сервис-периода, как описано в разделе 2.3.

Результаты моделирования показали значимую роль фиксированных факторов в изменчивости удоя. Значимость отдельных факторов оценивалась на основе дисперсионного анализа, а величина их вклада – по значениям F-критерия (таблица 9).

Таблица 9 - Роль фиксированных эффектов в изменчивости удоя

Фактор	df	SS Effect	MS Effect	F
Год отёла (ГО)	20	3867006343	193350317	186,6
Год запуска (ГЗ)	20	846344821	42317241	40,8
Сезон отёла (СО)	3	523622364	174540788	168,5
Сезон запуска (СЗ)	3	136825103	45608368	44,0
ГО:СО	58	645445340	11128368	10,7
ГЗ:СЗ	58	139749278	2409470	2,33
ГО:ГЗ	44	90102808	2047791	1,98
СО:СЗ	9	498777656	55419740	53,5

Наибольшее влияние на изменчивость удоя оказывал фактор «Год отёла», что отражает динамику условий кормления и содержания скота на протяжении

20-летнего периода наблюдений. Значимым также оказалось влияние «Сезона отёла», что подчеркивает сезонные колебания продуктивности. Взаимодействия факторов «Год отёла:Сезон отёла» и «Сезон запуска:Год запуска» также внесли существенный вклад в общую дисперсию удоя.

Оценка роли изучаемых факторов (таблица 10) выявила, что наибольшая доля дисперсии приходится на «Идентификационные данные животного» (ID), что подчеркивает индивидуальную вариабельность молочной продуктивности.

Таблица 10 - Роль исследуемых факторов в изменчивости удоя

Источник изменчивости	n	σ^2	σ	$\sigma^2, \%$	$\sigma, \%$
ID	16011	756856	869,97	30,3	26,6
Возраст первого отёла (В1ОТ)	2619	33592	183,28	1,40	5,6
Отец	165	264306	514,11	10,6	15,7
Возраст первого плодотворного осеменения (В1ПО)	23	3008	54,84	0,10	1,7
Лактация	10	403068	634,88	16,1	19,0
Остаточный	-	1036127	1017,90	41,5	31,0
Общая	34387	2496957	3274,98	100	100

Визуализация вклада случайных предикторов (рисунок 16) позволила представить не только величину эффекта, но и плотность распределения остатков в пределах градаций факторов.

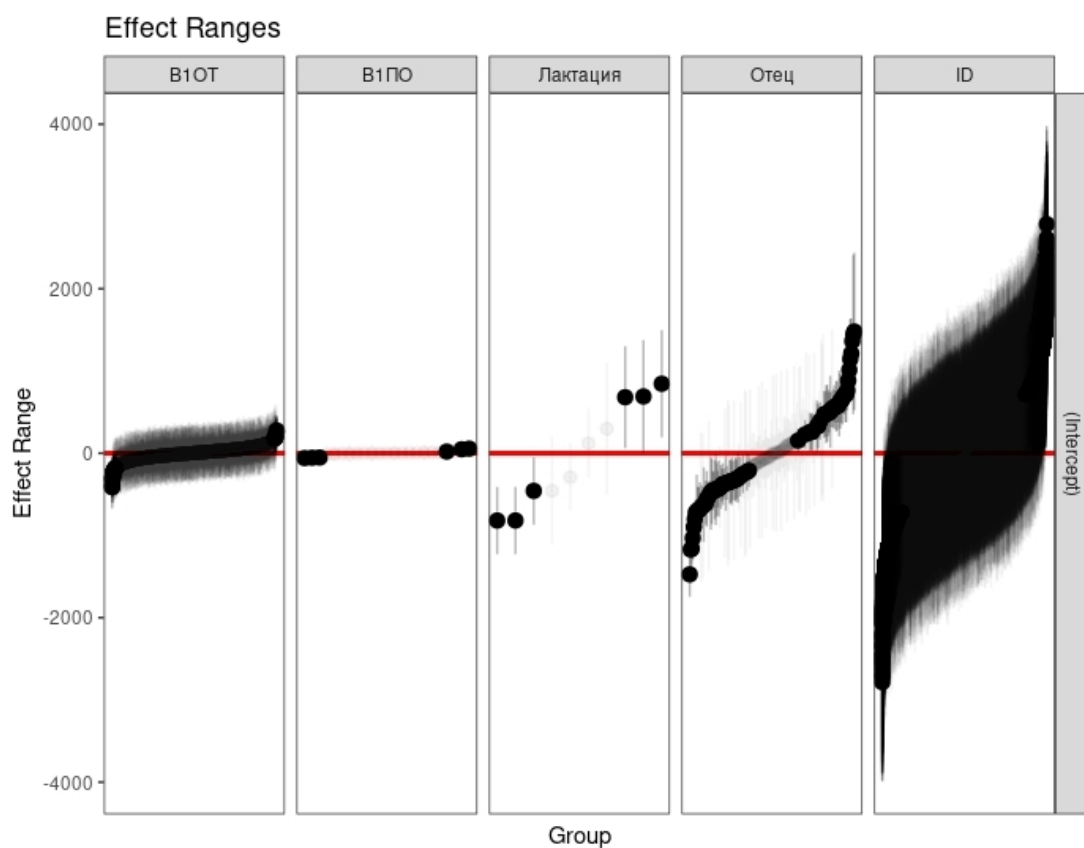


Рисунок 16 - Визуализация эффектов модели

Построенная модель продемонстрировала достаточно высокую точность прогноза, подтвержденную коэффициентами корреляции Пирсона, Спирмена и Кендалла ($r \approx 0,90$). Коэффициент детерминации R^2 для тестовой выборки составил 0,819, что свидетельствует о высокой доле объясненной дисперсии удоя при использовании фенотипических данных (рисунок 17).

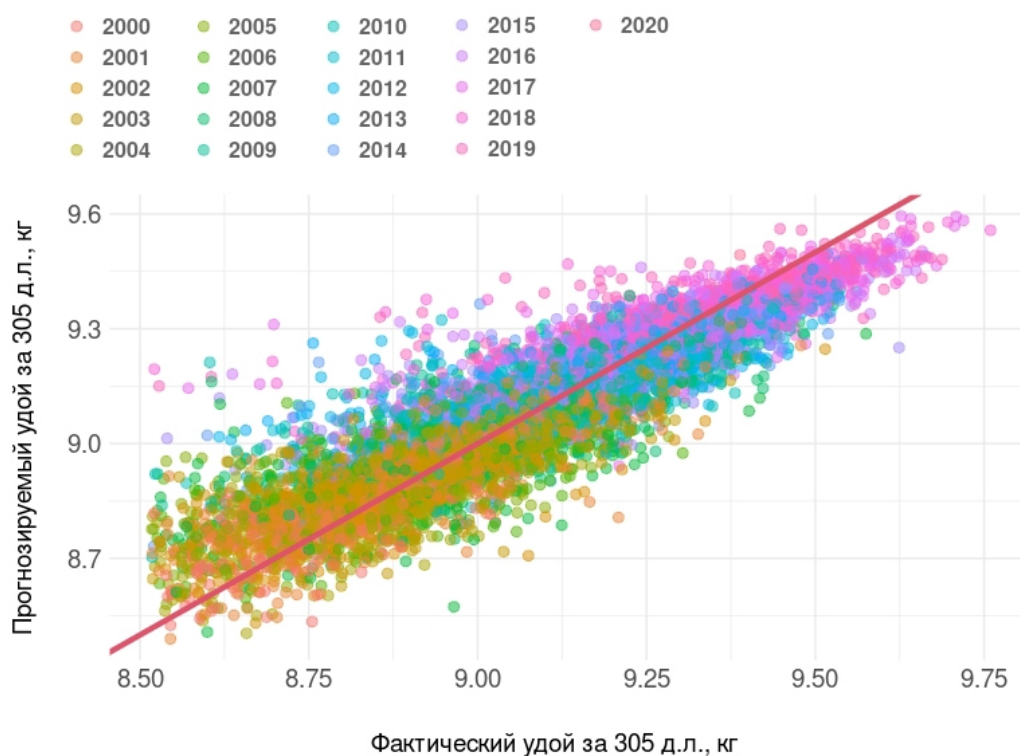


Рисунок 17 - Регрессионная зависимость эмпирических и прогнозируемых значений

Результаты моделирования подтвердили значительное влияние фиксированных факторов на изменчивость удоя черно-пёстрого скота в условиях промышленного комплекса. Факторы «Год отёла» и «Сезон отёла», а также их взаимодействия, оказывают наибольшее влияние на вариабельность удоя. Включение случайных эффектов, особенно «Идентификационные данные животного», существенно повышает точность прогноза. Полученная модель может быть использована для прогнозирования удоя и оценки племенной ценности животных, однако ее применение на других субпопуляциях может потребовать дополнительной калибровки и учета региональных особенностей. В целом, исследование подчеркивает важность комплексного учета генетических и паратипических факторов для повышения эффективности селекционно-племенной работы в молочном скотоводстве. Проведённая работа с данными первичного зоотехнического учёта на примере смешанной модели показала не только возможность интерпретации результатов и их согласованности, но и достаточно высокую прогностическую способность. Это является свидетельством необходимости проведения работ по

повышению качества цифрового учёта в рамках крупных региональных цифровых платформ для принятия решений. Материалы, опубликованные в данном разделе, ранее были представлены в статье:

- Петров, А.Ф. Роль фиксированных факторов в изменчивости удоя скота ирменского типа в условиях промышленного комплекса / А.Ф. Петров, Е.В. Камалдинов, О.В. Богданова и др. // Вестник НГАУ. – 2022. – №4(61). – С. 137-149.

3.5 Генетическая структура популяции по STR-локусам

Выявлен 31 локус по микросателлитным профилям (таблица 11).

Часть выявленных STR-локусов (BM2113, ETH10, TGLA122, ETH225, SPS115, BM1824, TGLA227, INRA023, BM1818, TGLA126, ETH3 и TGLA53) составляет стандартную панель маркеров при исследовании популяций крупного рогатого скота рекомендованную ISAG и согласуется с решением №74 коллегии ЕАЭС от 2 июня 2020. Данная группа характеризуется наибольшими частотами встречаемости от 1962 до 10228 по количеству животных. При этом выявлено 1220 уникальных генотипов по локусам.

Таблица 11 - Микросателлитный профиль молочного скота Западной Сибири

Локус	Количество, %	Локус	Количество, %	Локус	Количество, %
BM2113*	100	TGLA53	19,18	ILST005	0,44
ETH10*	99,92	ILST006	3,56	RM067	0,32
TGLA122*	99,83	CSSM66	3,23	BL42	0,06
ETH225*	99,76	CSRM60	3,22	BM203	0,06
SPS115*	99,65	HAUT27	2,24	BM4107	0,06
BM1818*	99,44	HEL1	0,69	BM4208	0,06
BM1824*	99,41	CSSM36	0,68	BM5004	0,06
TGLA227*	99,01	SPS113	0,60	ETH152	0,06
INRA023*	98,83	MGTG4B	0,59	INRA032	0,01
TGLA126*	98,80	CYP21	0,47		
ETH3*	82,37	ETH185	0,44		

* выделены локусы, рекомендованные ISAG

После выбора только 12 локусов с референтными интервалами аллелей согласно данным, представленным в таблице 2, количество уникальных генотипов по локусам составило 502 (рисунок 18), где их преобладающее количество при-

шло на локусы TGLA122 (87), TGLA53 (81) и TGLA227 (78), а наименьшее разнообразие было выявлено в локусе BM1824 (20). По всем остальным локусам разнообразие варьировало от 20 до 38. Локусы BM2113, ETH225, INRA023 и SPS115 имели схожее количество уникальных генотипов по локусам, значения которых располагались в диапазоне от 31 до 38. Полученные данные свидетельствуют о значительной вариабельности их числа в зависимости от исследуемого микросателлитного локуса у чёрно-пёстрого скота популяции Западной Сибири. Локусы TGLA122, TGLA53 и TGLA227 демонстрируют высокий уровень полиморфизма, что может указывать на их высокую информативность для генетических исследований данной породы в представленных регионах.

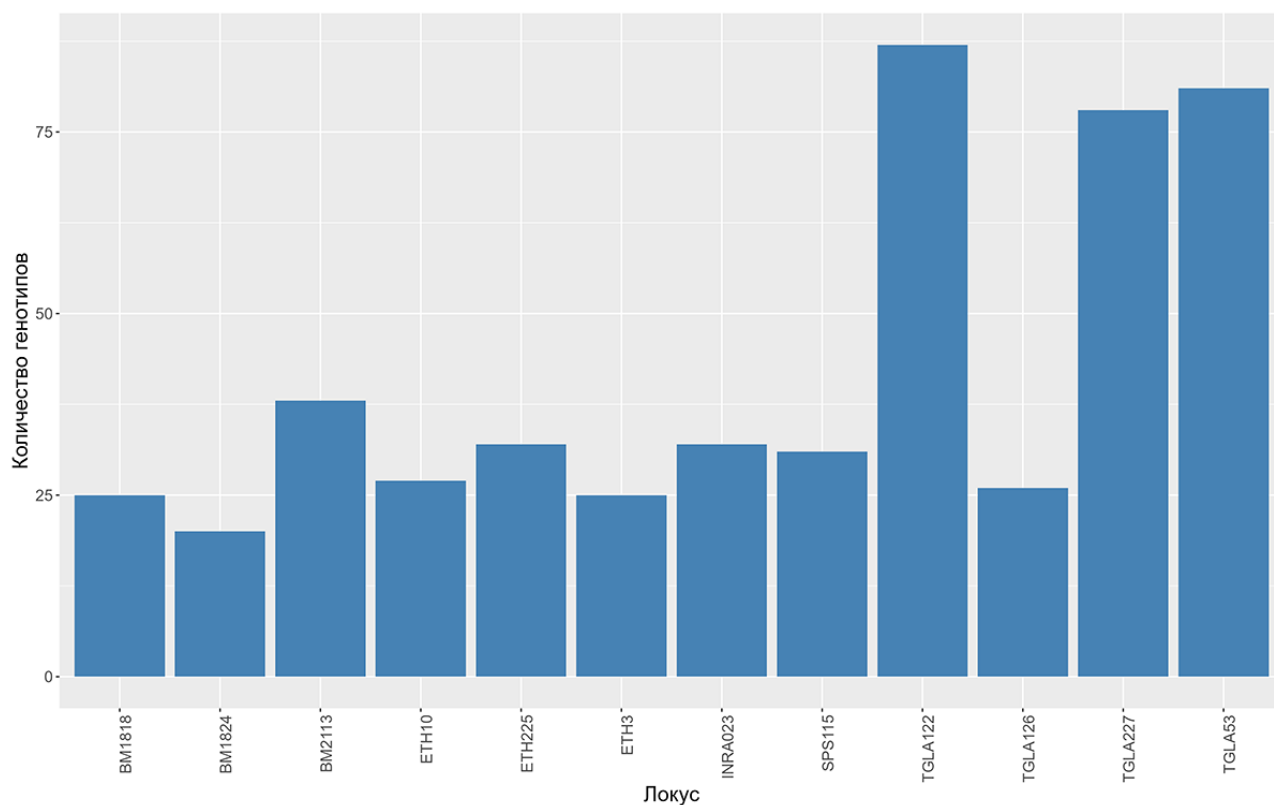


Рисунок 18 – Распределение уникальных генотипов по локусам без поправки на число информативных аллелей

Количество уникальных генотипов, обнаруженных по микросателлитным локусам, является одним из важных показателей генетического разнообразия. Наличие редких генотипов, встречающихся с низкой частотой, могут быть результа-

том случайных ошибок при исследовании или представлять собой недавние мутации, не отражают общую картину генетической изменчивости популяции.

Поправка в 5 % на число информативных аллелей ограничила число уникальных генотипов до 130 по всем локусам (рисунок 19).

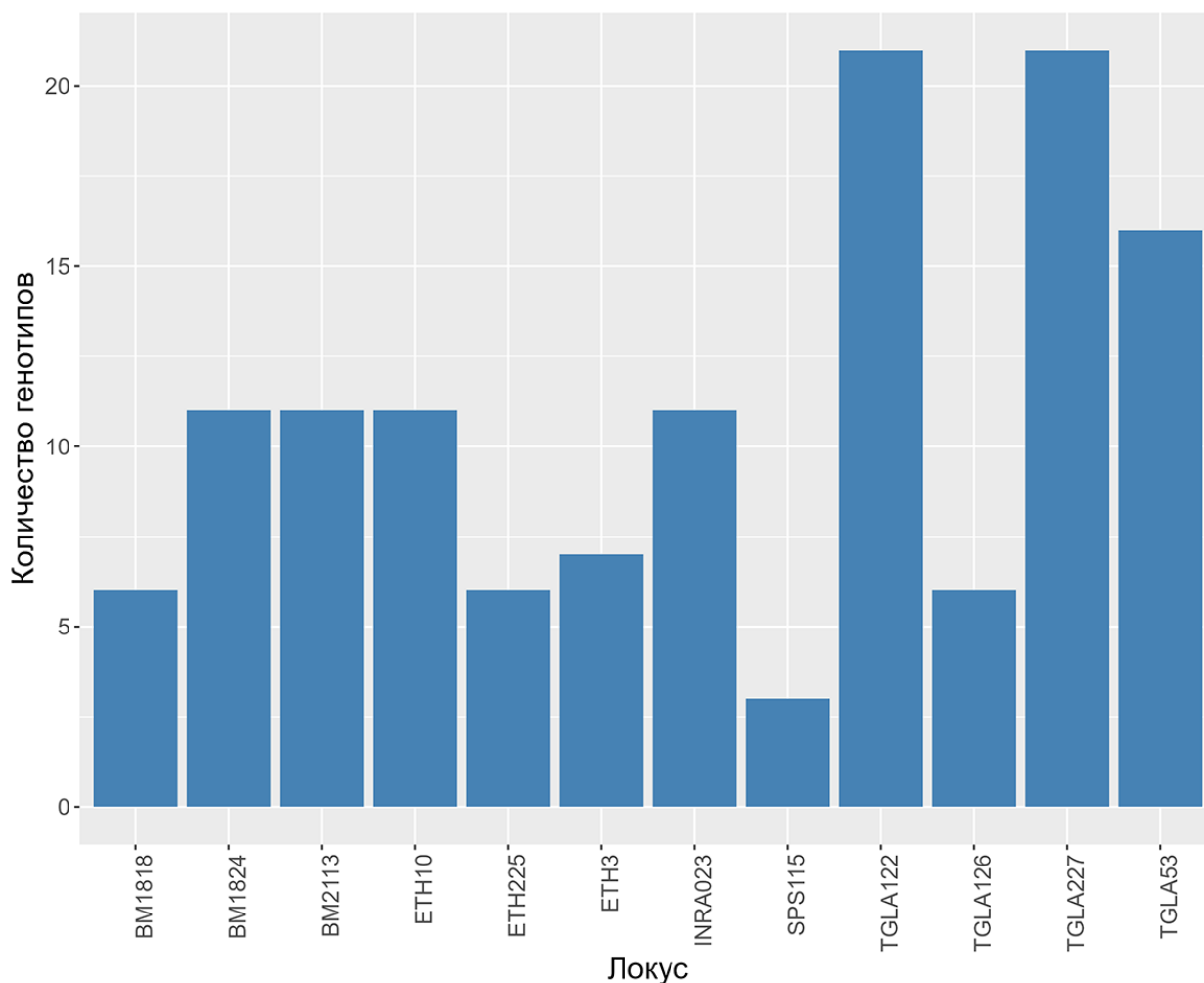


Рисунок 19 – Распределение уникальных генотипов по локусам с поправкой 5% на число информативных аллелей

Установлено, что количество уникальных генотипов, после поправки на число информативных аллелей, варьирует от 3 (локус SPS115) до 21 (локусы TGLA122 и TGLA227). Наибольшее их количество было обнаружено для локусов TGLA122 и TGLA227 (по 21 генотипу), за ними следовал локус TGLA53 (16 генотипов). Примечательно, что концентрация BM1824, BM2113, ETH10 и INRA023 в популяции была одинаковой (11). Локус SPS115, обладал низкой информативно-

стью и может быть менее полезен для анализа генетической структуры изучаемой популяции.

Установлены частоты уникальных генотипов (рисунок 20). Выявлены наиболее распространённые генотипы SPS115 248/248 (4769 особей), TGLA126 115/117 (3900 особей) и BM1818 262/266 (3846 особей) (приложение Д). К наименее распространённым, которые встречаются всего 1 раз, можно отнести: BM2113 127/125, ETH3 127/125, TGLA53 162/160, BM1824 180/178, INRA023 214/210, ETH10 219/213.

Выявленное распределение частот генотипов свидетельствует о значительной гетерогенности исследуемой популяции голштинского скота. Наличие высокочастотных генотипов может указывать на действие отбора по данным локусам. Редкие генотипы, встречающиеся только у одной особи, могут быть результатом спонтанных мутаций или интрогрессии генетического материала из других пород.

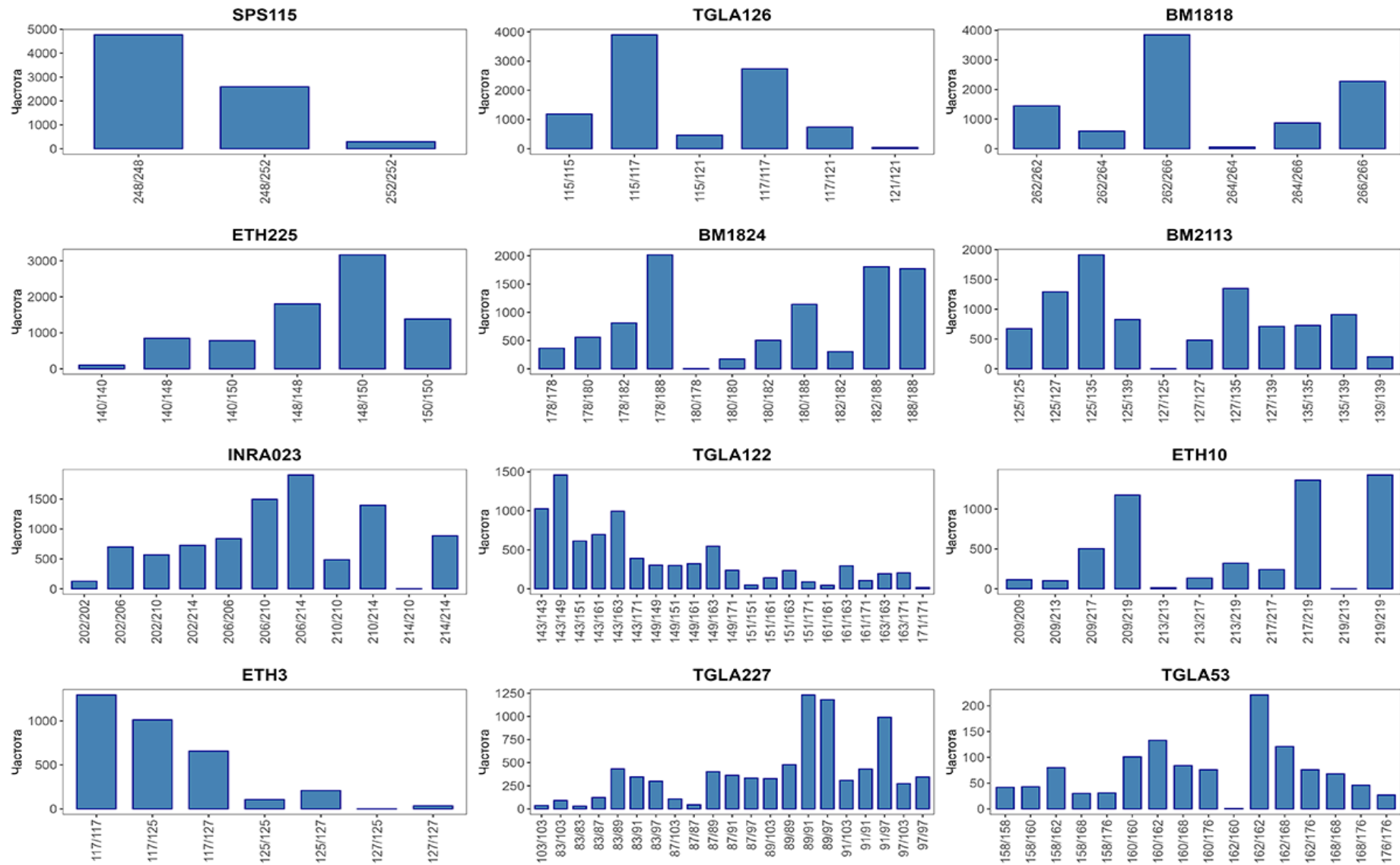


Рисунок 20 – Количество генотипов по локусам

Анализ 12 выбранных микросателлитных локусов по заданным референтным интервалам аллелей исследуемой популяции скота позволил выявить 47 вариантов аллелей (см. рисунок 21). Наиболее полиморфными оказались локусы TGLA122 ($N_a = 6$), TGLA227 ($N_a = 6$) и TGLA53 ($N_a = 5$), а наименьшее количество аллелей обнаружено в локусе SPS115 ($N_a = 2$).

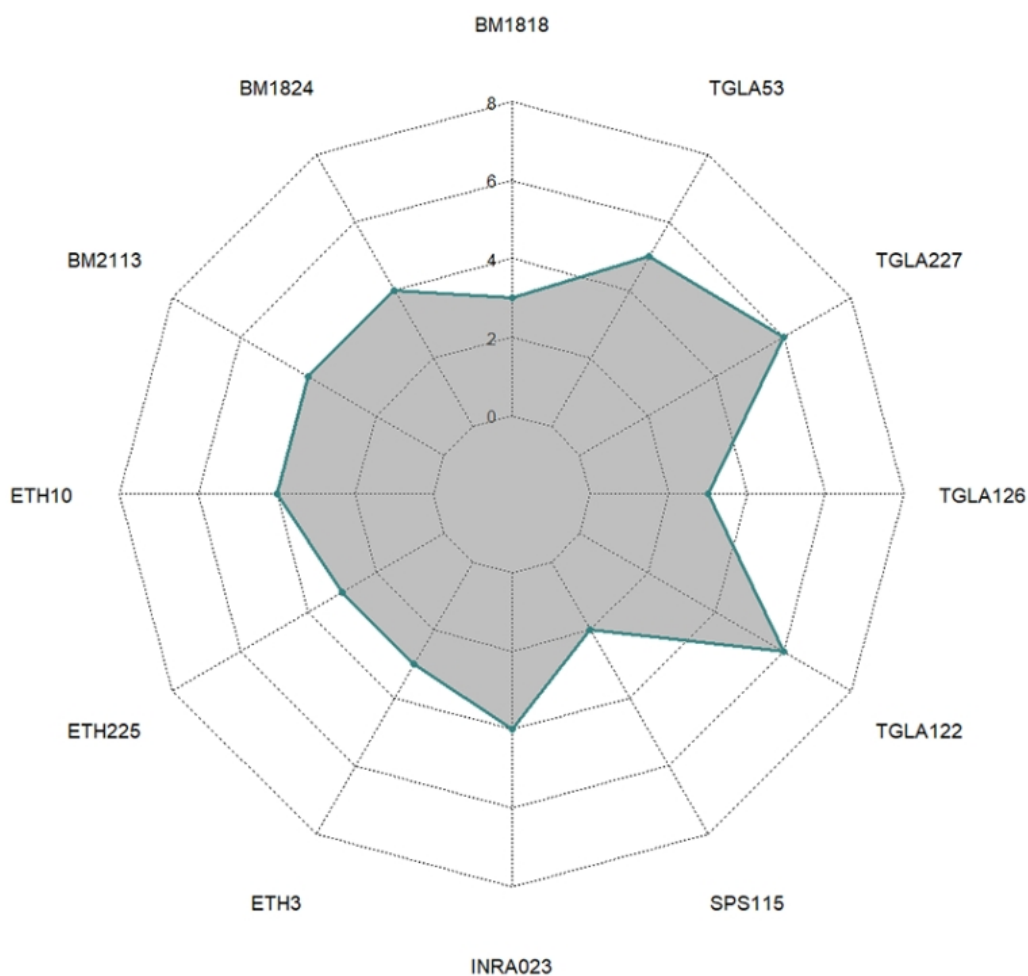


Рисунок 21 – Количество аллелей по группе локусов

Частота встречаемости аллелей (таблица 12) варьировала от 0,054 в локусе ETH10 (аллель 213) до 0,791 в локусе SPS115 (аллель 248). Наблюдались локусы с выраженным преобладанием одного аллеля (например, SPS115 с аллелем 248), с более равномерным распределением частот аллелей (например, TGLA227 или TGLA53, рисунок 22). Наблюдаемая высокая вариабельность частот аллелей яв-

ляется индикатором существенного генетического разнообразия в пределах исследуемой популяции. Увеличение генетической изменчивости популяции определяется как ростом числа аллелей, так и повышением равномерности их распределения, то есть приближением частот аллелей друг к другу. Выявлена высокая вариабельность частот, что указывает на генетическое разнообразие исследуемой популяции.

Следует отметить, что чем большим числом аллелей представлены STR локусы, тем более информативными они являются при оценке популяции.

Таблица 12 - Встречаемость аллелей по изучаемым микросателлитным локусам

Локус	Аллель	Число аллелей	Частота	Локус	Аллель	Число аллелей	Частота
BM1824	178	4119	0,217	ETH10	209	2010	0,186
	180	2554	0,134		213	593	0,054
	182	3735	0,197		217	2483	0,229
	188	8512	0,449		219	5712	0,528
BM2113	125	5382	0,295	ETH3	117	4259	0,641
	127	4317	0,237		125	1437	0,216
	135	5631	0,309		127	940	0,141
	139	2854	0,156	INRA023	202	2248	0,123
SPS115	248	12133	0,791		206	5776	0,316
	252	3191	0,208		210	4434	0,242
TGLA122	143	6208	0,375	TGLA53	214	5796	0,317
	149	3474	0,209		158	268	0,113
	151	1479	0,089		160	539	0,228
	161	1653	0,099		162	853	0,361
	163	2663	0,160		168	417	0,176
	171	1069	0,064		176	283	0,119
TGLA227	83	1350	0,082	TGLA126	115	6745	0,371
	87	1423	0,086		117	10114	0,557
	89	4531	0,276		121	1295	0,071
	91	4102	0,250	BM1818	262	7339	0,403
	97	3770	0,230		264	1582	0,086
	103	1182	0,072		266	9271	0,509
ETH225	140	1834	0,113				
	148	7626	0,471				
	150	6720	0,415				

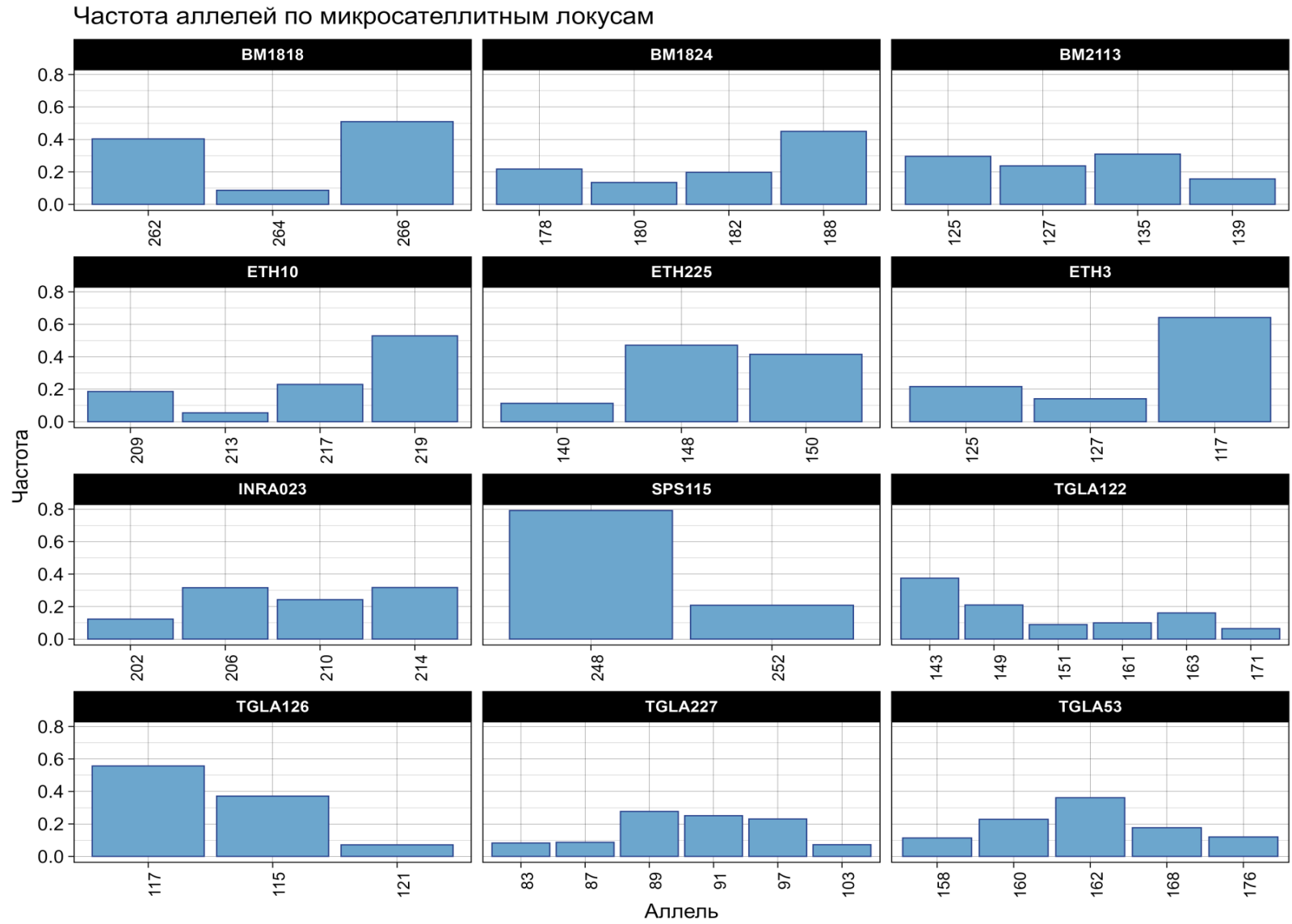


Рисунок 22 - Частота аллелей по микросателлитным локусам

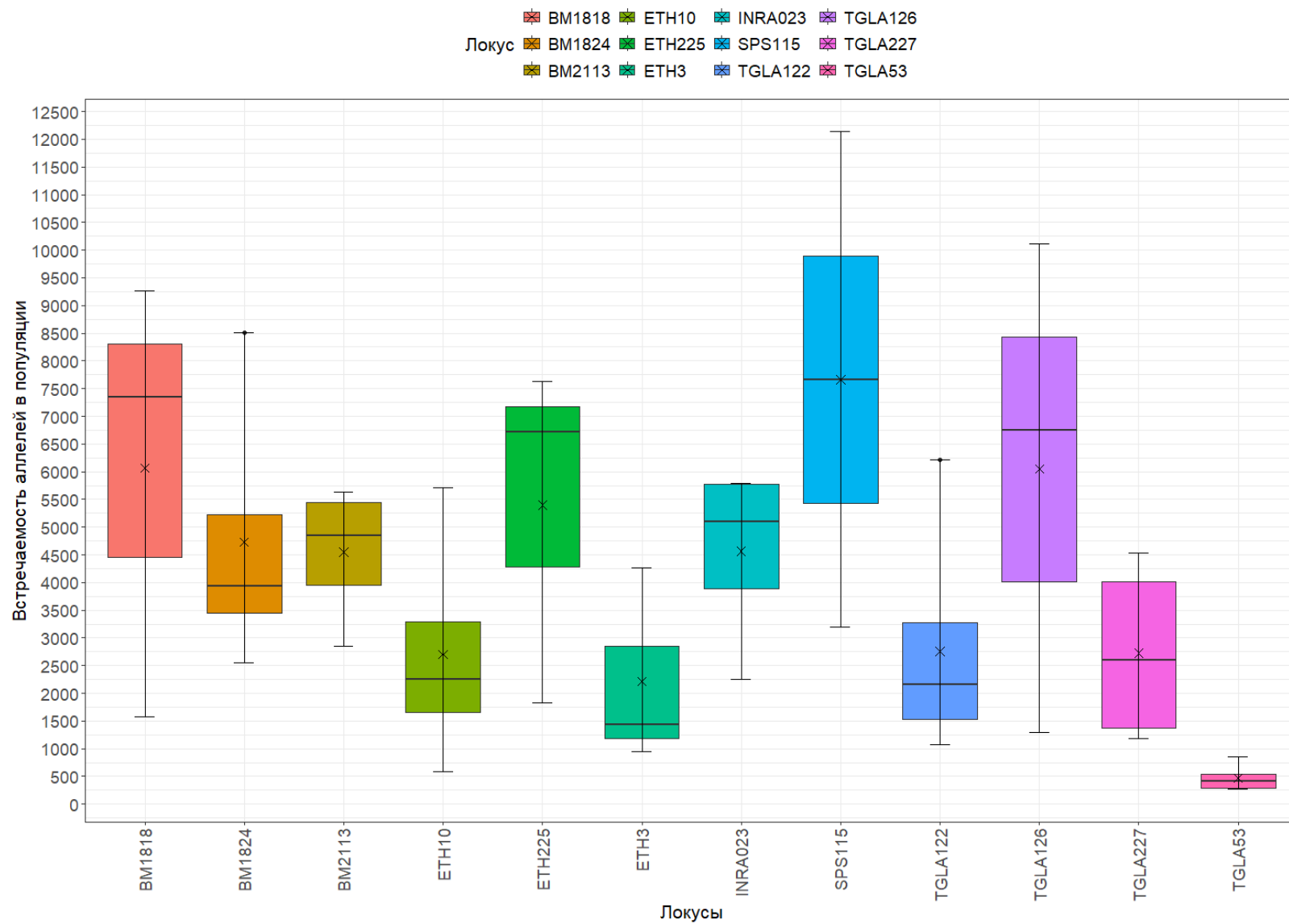


Рисунок 23 – Встречаемость аллелей по группам локусов в изучаемой популяции

Встречаемость аллелей существенно варьируется в зависимости от локуса (рисунок 23). Так, локус SPS115 демонстрирует наибольшую медианную встречаемость аллелей (аллели имеют тенденцию встречаться с более высокими и, что более важно, более равномерными частотами по сравнению с другими локусами) ($n=7762$), где на аллель 248 ($n_{248}=12133$) приходится 79.17% от всех аллелей в этом локусе. В тоже время в локусе TGLA53 обнаружена самая низкая медианная встречаемость равная всего 417, причём ни один аллель не имеет частоту выше 36.14% (аллель 162). Локусы BM1818 и BM1824 демонстрируют значительный разброс во встречаемости аллелей. В BM1818 аллель 266 присутствует с частотой 50.96%, в то время как аллель 264 - только 8.69%. В BM1824 аллель 188 встречается с частотой 44.98%, а аллель 178 - 21.77%. В других локусах, например, ETH3 и INRA023, разброс меньше, что указывает на более равномерное распределение встречаемости аллелей. Различия во встречаемости аллелей между локусами могут быть обусловлены различными факторами. Это может быть селективное давление, когда некоторые аллели могут давать организмам преимущества в определённых условиях среды, что и приводит к их более высокой встречаемости в популяции. С другой стороны, высокая встречаемость аллеля 248 в локусе SPS115 является характерной для голштиinizированного чёрно-пёстрого скота. Полученные результаты подтверждаются результатами исследований других авторов.

Выявлено, что в локусе BM1818 аллели 262 и 266 составляют более 90% всех аллелей, в то время как в TGLA122 ни один аллель не имеет частоту выше 37.52%. Установлено, что увеличение числа генотипов в локусе ведёт к росту доли редких аллелей, что связано с повышением уровня генетического разнообразия в таких локусах.

На рисунке 24 приведена схема, отражающая количество наблюдаемых и ожидаемых гомозигот по каждому аллелю в группах локусов изучаемой популяции чёрно-пёстрого скота Западной Сибири.



Рисунок 24 – Количество наблюдаемых и ожидаемых гомозигот

Локусы VM1818, VM1824, VM2113, ETH10, ETH225, INRA023 демонстрируют незначительные отклонения наблюдаемого количества гомозигот от ожидаемого, что согласуется с предположением о равновесии Харди-Вайнберга (HWE). Вероятно, эти локусы либо нейтральны по отношению к отбору, либо находятся под действием стабилизирующего отбора, который поддерживает существующее распределение аллелей. Незначительные отклонения могут быть обусловлены случайными флуктуациями частот аллелей в результате генетического дрейфа, особенно если популяция ограничена по размеру. Для локуса ETH3 наблюдается заметное отклонение от HWE по аллелю 127, где наблюдаемое количество гомозигот значительно ниже ожидаемого (72 против 133). Такое отклонение может указывать на селективное преимущество гетерозигот, известное как сверхдоминирование. В этом случае гетерозиготы обладают более высокой приспособленностью, чем гомозиготы, что приводит к увеличению их частоты в популяции. В локусе TGLA122 наблюдается последовательное снижение наблюдаемого количества гомозигот по всем аллелям по сравнению с ожидаемым. Это может быть результатом отбора в пользу гетерозигот, либо действия других эволюционных сил, таких как поток генов из популяции с иными частотами аллелей или генетический дрейф, приводящий к случайному изменению частот аллелей. Локус TGLA227 демонстрирует отклонения от HWE с несколькими аллелями, проявляющими значимые различия между ожидаемым и наблюдаемым количеством гомозигот. В частности, аллель 83 демонстрирует значительно меньшее наблюдаемое количество гомозигот (58 против 111), что может указывать на селекцию против гомозигот по данному аллелю. Наиболее заметные отклонения от HWE проявляются в локусе TGLA53, причём наблюдаемое количество гомозигот по аллелям 158, 160 и 162 значительно превышает ожидаемое. Такое распределение может указывать на локальное повышение частоты отдельных аллелей, возможно, вследствие отбора или дрейфа в небольшой популяции.

Минимальные различия между наблюдаемым и ожидаемым количеством гомозигот демонстрирует локус SPS115 по обоим аллелям (248 и 252). Это гово-

рит о том, что данный локус, вероятно, находится в состоянии, очень близком к равновесию Харди-Вайнберга. Отсутствие существенных отклонений указывает на то, что этот локус вероятно относится к селективно нейтральным.

Анализ данных по 12 микросателлитным локусам позволяет в целом сделать вывод о том, что большинство локусов находятся в состоянии, близком к равновесию Харди-Вайнберга.

Эффективным инструментом для оценки генетического разнообразия популяции может служить число эффективных аллелей. Данный показатель отражает пропорции полиморфных локусов в геноме, количество аллелей на локус и степень их выравнивания. Величина эффективных аллелей выступает в качестве комплексной метрики полиморфизма популяции, являясь значимым дескриптором её генетической структуры и адаптивного потенциала. Поэтому расчёт числа эффективных аллелей является важным для оценки генетического разнообразия любой популяции.

В проведённом исследовании было установлено, что среднее значение числа эффективных аллелей (N_e) по группе локусов (таблица 13) составило $3,09 \pm 0,29$. Наименьшее значение N_e было выявлено в локусе SPS115 (1,49), а наибольшее в TGLA227, TGLA122 и TGLA53 (более 4). Локусы были разделены на две группы по среднему числу эффективных аллелей. В первую группу вошли те локусы, среднее значение N_e которых меньше чем 3,09 (BM1818, ETH10, ETH225, ETH3, SPS115, TGLA126), во вторую превышающие указанную отметку (BM1824, BM2113, INRA023, TGLA122, TGLA227, TGLA53). Чем меньше число эффективных аллелей, тем меньше вариантов генов присутствует в популяции, и тем ниже её генетическое разнообразие. Это, в свою очередь, может негативно сказываться на жизнеспособности и эволюционном потенциале популяции. Анализ выявил значимую вариабельность числа эффективных аллелей (N_e) по исследованным локусам в изучаемой популяции. Вторая группа локусов имеет высокую степень полиморфизма, в то время как по первой группе наблюдается снижение генетического разнообразия, особенно это проявляется в локусе SPS115.

Наблюдаемая гетерозиготность (H_o) является показателем генетического разнообразия популяции, отражающим долю гетерозиготных особей по определённому локусу или набору локусов. Частота гетерозигот свидетельствует о присутствии различных аллелей в генофонде и служит индикатором аллельного разнообразия. Для более точной оценки генетической variability популяции используют показатель ожидаемой гетерозиготности (H_e), который рассчитывается на основе частот аллелей и позволяет оценить потенциальный уровень гетерозиготности в условиях случайного скрещивания.

В ходе исследования генетической структуры популяции, проведённого с использованием 12 микросателлитных локусов (рисунок 25), были определены уровни наблюдаемой (H_o) и ожидаемой (H_e) гетерозиготности. Средние значения H_o и H_e составили, соответственно, $0,64 \pm 0,03$ и $0,65 \pm 0,03$. Близость этих значений указывает на то, что популяция в целом находится в состоянии, близком к равновесию Харди-Вайнберга.

Анализ отдельных локусов выявил, что у восьми из них (BM1818, BM2113, ETH10, ETH225, ETH3, INRA023, SPS115, TGLA126), наблюдаемая гетерозиготность соответствовала ожидаемой, что может свидетельствовать о нейтральном поведении этих локусов, т.е. они являются селективно-нейтральными. Для некоторых локусов наблюдалось значительное отклонение от ожидаемых значений. Наиболее выраженное различие между H_o и H_e (0,148) было выявлено для локуса TGLA53. Такое значительное отклонение может быть обусловлено различными факторами, включая влияние отбора, или действия генетического дрейфа.

Локус SPS115 продемонстрировал наименьший уровень как наблюдаемой, так и ожидаемой гетерозиготности, что может быть связано с действием проводимого отбора или эффектом генетического дрейфа. Локус TGLA227, напротив, отличался наибольшим уровнем как H_o , так и H_e , что может указывать на повышенную генетическую изменчивость в этом локусе.

Полученные результаты позволяют предположить наличие процессов, влияющих на генетическое разнообразие в исследуемой популяции. Дальнейшее изу-

чение этих процессов, включая оценку влияния отбора, генетического дрейфа и миграции, может помочь найти причины отклонения от ожидаемых уровней гетерозиготности в отдельных локусах.

Таблица 13 - Генетическая изменчивость молочного скота (n = 10228 гол.) чёрно-пёстрого скота Западной Сибири

Локус	Na	Ho	He	Fis	Ne	I	D	χ^2	p	p (F)
BM1818	3	0,584	0,570	-0,025	2,325	0,922	0,570	17,295	0,003	0,041
BM1824	4	0,723	0,693	-0,044	3,257	1,282	0,693	61,581	< 0,05	< 0,05
BM2113	4	0,771	0,736	-0,048	3,781	1,355	0,736	85,916	< 0,05	< 0,05
ETH10	4	0,667	0,630	-0,059	2,700	1,147	0,630	56,682	< 0,05	< 0,05
ETH225	3	0,593	0,593	-0,001	2,454	0,966	0,593	1,066	0,957	0,982
ETH3	3	0,567	0,521	-0,087	2,088	0,893	0,521	46,763	< 0,05	< 0,05
INRA023	4	0,744	0,725	-0,026	3,635	1,330	0,725	23,513	< 0,05	< 0,05
SPS115	2	0,339	0,330	-0,027	1,492	0,512	0,330	5,629	0,059	0,148
TGLA122	6	0,802	0,767	-0,046	4,294	1,612	0,767	104,303	< 0,005	< 0,05
TGLA126	3	0,563	0,546	-0,030	2,205	0,882	0,546	12,132	0,03	0,154
TGLA227	6	0,833	0,788	-0,058	4,710	1,649	0,788	124,900	< 0,05	< 0,05
TGLA53	5	0,611	0,759	0,195	4,144	1,513	0,759	183,695	< 0,05	< 0,05
$\bar{x} \pm s_{\bar{x}}$	-	0,65±0,04	0,64±0,04	-0,02±0,02	3,09±0,29	1,17±0,1	0,64±0,04	-	-	-

* \bar{x} - среднее арифметическое; $s_{\bar{x}}$ - стандартная ошибка; Na - число аллелей на локус; Ho - наблюдаемая гетерозиготность; He - ожидаемая гетерозиготность; Fis - индекс фиксации Райта; Ne - число эффективных аллелей на локус; I - индекс Шеннона; D - индекс Симпсона; χ^2 – значение хи-квадрат.

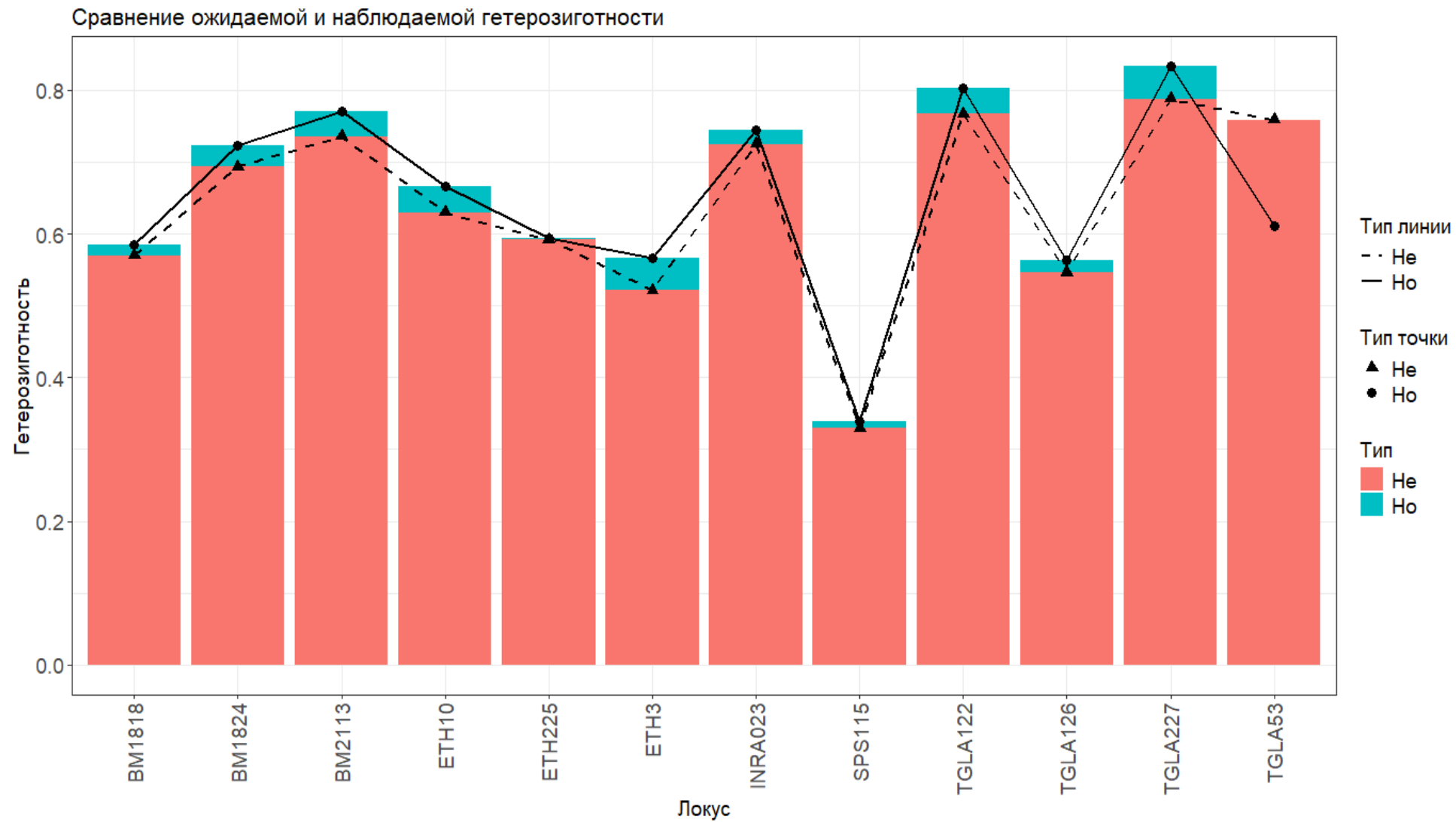


Рисунок 25 - Сравнение ожидаемой и наблюдаемой гетерозиготности

Полученные результаты исследования показали, что в целом F_{is} в исследуемой популяции был близок к нулю, несмотря на небольшие отклонения в отрицательную сторону, что свидетельствует о грамотной работе зоотехников и отсутствии выраженной гомозиготности. Наблюдался некоторый дефицит гетерозигот по локусу TGLA53. Это может быть объяснено использованием быков-производителей голштинской породы с высокой степенью инбридинга, что также играет роль в увеличении генетического родства в популяции.

Прогнозируется, что продолжение этой тенденции может привести к повышению F_{is} в положительном диапазоне и усилению эффектов инбредной депрессии. Инбредная депрессия является результатом накопления рецессивных аллелей в гомозиготном состоянии, что может привести к снижению продуктивных качеств животных и периоду их хозяйственного использования.

Зная, что индекс Шеннона указывает на генетическое разнообразие по локусу, учитывая как число аллелей, так и их частоту. Чем выше его значение, тем больше генетическое разнообразие в популяции. Такая картина наблюдалась по локусам TGLA227 (1,649), TGLA122 (1,612) и TGLA53 (1,513) при среднем значении по всем локусам $1,17 \pm 0,1$. Индекс Симпсона (D) установлен на уровне 0.64 ± 0.04 , что указывает на отсутствие сильного доминирования одного или нескольких аллелей.

По большинству локусов наблюдалось нарушение равновесия Харди-Вайнберга ($p < 0.05$), что может быть связано с недостаточной выборкой или с действием других факторов. Имеет место сложная структура популяции, включающая большое количество животных в разных предприятиях, в которых вектор отбора зависит от видения зоотехников-селекционеров. Следует также подчеркнуть, что данные по микросателлитным профилям были получены из различных лабораторий.

Точный тест Фишера для локусов SPS115 и TGLA126 показал высокие значения p , что не дает оснований отвергнуть нулевую гипотезу о соответствии равновесию Харди-Вайнберга. Полученные результаты не противоречат предположению о селективной нейтральности данных локусов. Однако, соответствие равно-

весию Харди-Вайнберга может быть обусловлено и другими факторами, либо их комбинацией, такими как отсутствие мутаций, миграции, дрейфа генов и неслучайного скрещивания в отношении данных локусов. Необходимы дополнительные исследования для подтверждения селективной нейтральности. Применение точного теста Фишера, обладающего большей статистической мощностью при анализе данных с ограниченным объемом выборки, выявило статистически значимые отклонения от равновесия Харди-Вайнберга для локуса VM1818.

Исследование генетического разнообразия популяции скота Западной Сибири по микросателлитным профилям показало высокую степень полиморфизма, что является благоприятным признаком с точки зрения адаптивного потенциала и возможностей для селекции. Анализ 12 микросателлитных локусов выявил 47 аллельных вариантов, что свидетельствует о широком спектре генетических вариантов в популяции. Среднее значение эффективного числа аллелей (N_e), равное 3,09, подтверждает высокий уровень генетического разнообразия. Высокое генетическое разнообразие позволяет популяции лучше адаптироваться к изменениям окружающей среды, увеличивает вероятность нахождения благоприятных мутаций и способствует более эффективной селекции.

По локусам TGLA122, TGLA227 и TGLA53 обнаружена высокая степень полиморфизма среди исследованных микросателлитов в популяции скота Западной Сибири. Это свидетельствует о генетическом разнообразии и делает их особенно ценными для дальнейшего изучения генетической структуры популяции. С учётом полученных результатов дальнейшие исследования позволят более детально изучить генетическое родство оценить филогенез породы и степень внутривидовой генетической дифференциации. Это важно для оптимизации селекционных программ, снижения эффекта инбредной депрессии, изучения адаптивных свойств и выявления генетических маркеров с учётом паратипических факторов разных географических или экологических зон.

Низкий уровень гетерозиготности, обнаруженный по локусу SPS115, может быть как результатом отбора по признакам, которые могут быть связаны с этим локусом, так и следствием генетического дрейфа. Последний может возникнуть в

результате «бутылочного горлышка», когда популяция проходит через период резкого сокращения численности, что приводит к уменьшению генетического разнообразия. Материалы описанные в данном разделе расширяют исследование представленное в статье:

- Петров, А.Ф. Генетическая структура скота сибирского отродья по микросателлитным локусам / А.Ф. Петров, Е.В. Камалдинов // Вестник НГАУ. – 2024. – №3(72). – С. 230-239.

3.6 Кластеризация предприятий по используемым быкам и микросателлитам

Построены дендрограммы, которые визуализируют результаты кластеризации племенных предприятий Западной Сибири, основанные на двух различных критериях: сходстве по количеству общих быков-производителей (рисунок 26) и генетическом сходстве популяций скота по микросателлитным локусам. Номера присвоены предприятиям в соответствии с теми, которые были приняты при анализе заполненности рядов родословных по исходным базам данных (приложение А,Б), а отсутствующие номера в названиях хозяйств продиктованы тем, что не все предприятия имеют данные о микросателлитных профилях животных в своих базах данных. Важно было сохранить сквозную нумерацию по хозяйствам для разных разделов исследования и при этом принять для кластерного анализа единую группу общих предприятий, чтобы избежать искажений в полученных результатах.

Сравнение двух дендрограмм позволяет выявить интересные закономерности в структуре региональной популяции чёрно-пёстрого скота, пролить свет на взаимосвязь между племенными практиками и генетическим разнообразием, а также сделать выводы, важные для дальнейшей селекционно-племенной работы.

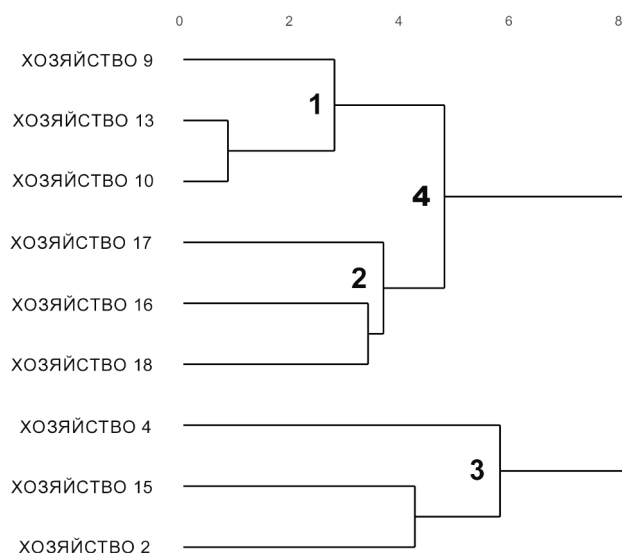


Рисунок 26 - Кластеризация предприятий по сходству использования быков (дистанция Канберра, метод «дальнего соседа» / complete)

Вследствие существенной вариабельности числа общих быков-производителей между анализируемыми хозяйствами совместное применение расстояния Канберры и метода дальнего соседа в проведенном кластерном анализе дали наиболее эффективную комбинацию, так как данное расстояние выполняет нормировку вклада каждой переменной. Метод дальнего соседа, в свою очередь способствовал формированию компактных, хорошо разделённых кластеров, что являлось ещё одним преимуществом данного сочетания. Результаты иерархического кластерного анализа, визуализированные посредством дендрограммы, выявили наличие нескольких основных кластеров, характеризующих специфику использования генетических ресурсов быков-производителей в исследуемых хозяйствах. Так, верхний кластер (кластер 1) объединяет хозяйства под номерами 9,10,13. Высокая степень сходства в наборе используемых быков-производителей может свидетельствовать о единой стратегии племенной работы и общих источниках приобретения племенного материала. Возможно данные хозяйства ориентируются на одни и те же селекционные тренды, используют семя быков одних и тех же линий и семейств или имеют общие каналы поставок племенной продукции, например, через региональные племпредприятия или дистрибьюторов. Не исключается влияние географической близости и исторически сложившихся межхо-

зяйственных связей, что могло способствовать обмену племенным материалом и формированию схожих селекционных подходов. Хозяйства 13 и 10 в ретроспективном аспекте используют единую стратегию племенной работы и общие источники племенного материала.

Центральный кластер (кластер 2) (рисунок 26) также демонстрируют значительное сходство в наборе используемых быков-производителей. Возможно одно из предприятий задает определенные тренды в селекционно-племенной работе и является источником племенного материала для многих хозяйств Западной Сибири. Включение двух других хозяйств в данный кластер может свидетельствовать о том, что они ориентируется на селекционные подходы, применяемые в одном из хозяйств данного кластера или использует племенной материал, полученный от него. Еще одним фактором, который возможно объединяет эти три хозяйства, выбор быков производителей одной командой специалистов, отвечающих за стратегию селекционно-племенной работы.

В нижний кластер (кластер 3) вошли хозяйства под номерами 2,4 и 15. Эти предприятия образуют отдельный кластер, отличающийся от других предприятий региона по набору используемых быков-производителей. Данный факт может указывать на собственную, отличную от других предприятий, стратегию племенной работы, ориентированную на специфические селекционные цели или источники племенного материала. Данные предприятия отличаются численностью поголовья и, как следствие, потенциалом для приобретения дорогостоящего племенного материала. Хозяйство под номером 4 имеет наименьшее количество общих быков по соотношению к другим хозяйствам и возможно ориентируется на более доступные или альтернативные источники поставок племенного материала.

Кластеризация предприятий по сходству использования быков наглядно демонстрирует структуру генетических связей между хозяйствами СФО по использованию быков-производителей. Выделенные кластеры, вероятно, отражают роль различных селекционных программ и/или пула используемого генетического материала. Наличие таких кластеров важно учитывать при планировании межхозяйственного обмена племенным материалом. Скрещивание животных из разных

кластеров может способствовать повышению гетерозиготности и проявлению эффекта гетерозиса. С другой стороны, слишком тесное родство внутри кластера (особенно если он небольшой) может привести к инбридингу и снижению продуктивности.



Рисунок 27 - Кластеризация предприятий по генетическому сходству на основе микросаттеллитов (дистанция Нея, метод Варда / Ward)

Комбинация генетической дистанции Нея и метода кластеризации Варда была выбрана для оценки генетического сходства на основе микросаттеллитов (рисунок 27) в популяции крупного рогатого скота, так как хорошо подходит для анализа микросателлитных данных. Дистанция Нея, рассчитанная по микросателлитам, учитывает их мутационную природу и дрейф генов, предоставляя интерпретируемые оценки генетической дивергенции, относительно устойчивые к размеру выборки. В отличие от некоторых других метрик, дистанция Нея не завышается искусственно в локусах с большим числом аллелей. Метод Варда, в свою очередь, минимизирует внутрикластерную дисперсию, формируя четкие и компактные кластеры, отображаемые в виде дендрограммы, которая отражает иерархию генетических отношений между популяциями. Комбинация этих методов позволяет эффективно выявлять генетическую структуру популяций крупного

рогатого скота. Дендрограмма на рисунке 27 отражает генетическое сходство популяций скота различных предприятий, оцененное по микросателлитным маркерам чёрно-пёстрого скота Западной Сибири. Анализ данной дендрограммы позволяет выявить группы предприятий, популяции скота, которые имеют высокое генетическое сходство.

На дендограмме на самом высоком уровне иерархии (генетическая дистанция около 6) наблюдается четкое разделение на два основных кластера. Первый кластер, объединяющий семь хозяйств демонстрирует более сложную структуру. На уровне генетической дистанции около 2 происходит разделение на два подкластера 1.1.1 и 1.2.1 кластера 1.1. Примечательно то, что частично данный кластер (1.1) перекликается с кластером 4 дендограммы, представленной на рисунке 26, объединяя хозяйства, входящие в его состав. Это может свидетельствовать о том, что сходство в использовании быков-производителей в данных предприятиях действительно отражается на генетическом сходстве их популяций. Однако, включение в данный кластер хозяйства под номером 15, которое не входило в кластер 1 дендограммы, кластеризующей предприятия по сходству использования быков, указывает на то, что генетическое сходство может быть обусловлено более глубокими историческими факторами, такими как общее происхождение маточного поголовья или наличие генетического обмена в прошлом.

Наиболее тесное генетическое родство, ожидаемо, демонстрируют два хозяйства 10 и 13 в подкластере 1.1.2, что подтверждает их принадлежность к одной популяции и как было отмечено ранее данные предприятия имеют родственное маточное поголовье. В подкластере 1.1.1 представлены уже 3 предприятия, что также указывает на генетическое сходство между популяциями их скота. Хозяйство под номером 15 генетически ближе к хозяйствам, которые входят в подкластер 1.1.2, что может быть следствием недавней интрогрессии генов (например, использование быков хозяйств 10 и 13 в хозяйстве 15). Можно также отметить, что хозяйство 16 и 9 являются наиболее генетически удаленным в кластере 1. Хозяйство 17 и 18 образуют отдельный подкластер, что указывает на их относительное генетическое сходство.

Второй основной кластер 2 включает хозяйства, отделяющихся от остальных. Это может указывать на длительную репродуктивную изоляцию этих хозяйств, использование генетически дивергентных родственных групп. Требуется дальнейший анализ родословных и исторических данных. Следует отметить, что и по набору используемых быков-производителей (рисунок 26) хозяйства 2 и 4 также располагались обособленно.

Сравнение двух дендрограмм выявляет сложную и неоднозначную картину взаимосвязей между племенными предприятиями Западной Сибири. С одной стороны, наблюдается определенная корреляция между сходством в использовании быков-производителей и генетическим сходством популяций скота. Предприятия, использующие схожие наборы быков, как правило, демонстрируют и более близкое генетическое родство, что подтверждает влияние современных племенных практик на генетическую структуру популяций. С другой стороны, существуют и заметные расхождения между кластеризацией по разным критериям, указывающие на то, что генетическое сходство популяций не всегда полностью определяется реализуемыми селекционными программами, и на него могут влиять исторические факторы, географическая близость, популяционная структура и другие эволюционные процессы, такие как генетический дрейф и адаптация к локальным условиям среды.

Кластеризация по генетическому сходству, вычисленному на основе анализа полиморфизма микросателлитных локусов, отражает более глубокие и долгосрочные исторические и биологические взаимосвязи между популяциями скота. Микросателлиты, как нейтральные маркеры, накапливают изменения (мутации) относительно медленно и отражают суммарный эффект генетического дрейфа, мутационного процесса и потока генов (миграции) на протяжении многих поколений.

Кластеризация по общим быкам – это отражение сегодняшнего дня, а кластеризация по генетическим дистанциям – это история, отражающая формирование генетической структуры популяций. Некоторое несовпадение выбранных подходов естественно, поскольку текущие стратегии разведения на отдельно взятых

предприятиях – лишь один из факторов, влияющих на генетическое разнообразие, наряду с более фундаментальными и долгосрочными процессами.

Кластеризация, основанная на генетическом сходстве, выявляет популяции с уникальным генофондом, представляющие ценный резерв генетического разнообразия для выращивания быков-производителей, адаптированных к условиям региона. Так популяции хозяйств 2 и 4, выделившиеся в отдельный кластер на обеих дендрограммах, заслуживают особого внимания как потенциально уникальные популяции.

Информация о генетической структуре популяций является основой для разработки научно обоснованных рекомендаций по обмену племенным материалом между хозяйствами. При планировании скрещиваний следует учитывать не только данные родословных, но и результаты генетической кластеризации, минимизируя риск инбридинга и поддерживая оптимальный уровень генетического разнообразия. Полученные результаты подчеркивают важность комплексного подхода к оценке генетических ресурсов и разработке селекционных программ, учитывающего как современные племенные технологии, так и исторические и эволюционные факторы, формирующие генофонд популяции. Дальнейшие исследования, направленные на углубленное изучение генетической структуры чёрно-пёстрого скота Западной Сибири с использованием SNP-маркеров и методов геномной селекции, позволят более эффективно использовать генетический потенциал региональной популяции и повысить конкурентоспособность отечественного молочного скотоводства.

3.7 Генетическая структура популяции по SNP-локусам, ассоциированными с признаками молочной продуктивности

Таблица 14 представляет собой анализ генетической изменчивости популяции чёрно-пестрого скота по 64 SNP-локусам, ассоциированным с признаками молочной продуктивности согласно литературным источникам (приложение В).

Таблица 14 – Генетическая изменчивость популяции по SNP

SNP	Частота аллеля 1	Частота аллеля 2	MAF	He	Ho	DG	PIС	Пропущенных аллелей, %	χ^2	p	RS	Позиция, п.н.	Ближайший ген
ARS-BFGL-BAC-2469	0,81±0,028	0,19±0,028	0,19	0,3	0,31	0,3	0,26	0	0,079	0,779	-	33433160	HEATR7B2
ARS-BFGL-BAC-27930	0,83±0,027	0,17±0,027	0,17	0,28	0,28	0,28	0,24	0	0,006	0,937	rs109774038	29373244	HCN1
ARS-BFGL-BAC-6525	0,35±0,034	0,65±0,034	0,35	0,45	0,4	0,46	0,35	0	1,461	0,226	rs110594909	92127288	NRXN3
ARS-BFGL-NGS-100459	0,64±0,034	0,36±0,034	0,36	0,46	0,44	0,46	0,35	0	0,204	0,651	-	22128569	ABCG8
ARS-BFGL-NGS-103091	0,12±0,023	0,88±0,023	0,12	0,21	0,2	0,21	0,19	0,01	0,265	0,606	rs110111162	74518588	RBFOX2
ARS-BFGL-NGS-107037	0,23±0,03	0,77±0,03	0,23	0,36	0,34	0,36	0,29	0,01	0,137	0,711	rs109394729	46486647	USP3
ARS-BFGL-NGS-107379	0,47±0,035	0,53±0,035	0,47	0,5	0,57	0,5	0,37	0,01	1,783	0,181	rs109350371	2054457	LOC786966
ARS-BFGL-NGS-107749	0,06±0,017	0,94±0,017	0,06	0,1	0,11	0,1	0,1	0,01	0,343	0,558	rs109830113	33986486	CDH16
ARS-BFGL-NGS-110497	0,47±0,035	0,53±0,035	0,47	0,5	0,42	0,5	0,37	0	2,464	0,116	rs110475419	45870133	ADAM12
ARS-BFGL-NGS-111443	0,28±0,032	0,72±0,032	0,28	0,4	0,48	0,4	0,32	0	3,628	0,056	rs109498193	94269370	DERA
ARS-BFGL-NGS-11319	0,48±0,035	0,52±0,035	0,48	0,5	0,54	0,5	0,37	0	0,668	0,413	rs109084201	6763227	SLC40A1
ARS-BFGL-NGS-113507	0,72±0,032	0,28±0,032	0,28	0,4	0,41	0,4	0,32	0,01	0,102	0,748	rs110440617	98407974	PTRH1
ARS-BFGL-NGS-113876	0,74±0,031	0,26±0,031	0,26	0,38	0,37	0,38	0,31	0	0,069	0,793	rs109392509	40644182	AGXT2L2
ARS-BFGL-NGS-118998	0,73±0,031	0,27±0,031	0,27	0,4	0,36	0,4	0,32	0,01	0,687	0,407	rs110482506	32030332	GHR
ARS-BFGL-NGS-14781	0,23±0,03	0,76±0,03	0,23	0,36	0,33	0,36	0,29	0	0,675	0,411	rs111027205	56766165	DDIT3
ARS-BFGL-NGS-18010	0,82±0,027	0,18±0,027	0,18	0,29	0,31	0,29	0,25	0	0,542	0,461	rs110643010	37101861	ADAMTS8
ARS-BFGL-NGS-24998	0,83±0,027	0,17±0,027	0,17	0,28	0,3	0,28	0,24	0	0,398	0,528	rs111003135	26721324	SAA3/SAA1
ARS-BFGL-NGS-29557	0,13±0,024	0,87±0,024	0,13	0,23	0,24	0,23	0,2	0,01	0,388	0,533	rs110614142	81612132	PTHLH
ARS-BFGL-NGS-29581	0,58±0,035	0,42±0,035	0,42	0,49	0,51	0,49	0,37	0,01	0,113	0,736	-	114000000	KCNH2
ARS-BFGL-NGS-39539	0,19±0,028	0,81±0,028	0,19	0,31	0,32	0,31	0,26	0	0,157	0,691	rs110558583	41457147	JARID2
ARS-BFGL-NGS-42950	0,69±0,033	0,31±0,033	0,31	0,43	0,46	0,43	0,34	0,01	0,636	0,425	rs110565416	49531815	TMEM132C
ARS-BFGL-NGS-49079	0,81±0,028	0,19±0,028	0,19	0,3	0,27	0,3	0,26	0	1,095	0,295	-	6574398	-
ARS-BFGL-NGS-4939	0,68±0,033	0,32±0,033	0,32	0,44	0,5	0,44	0,34	0	2,217	0,136	rs109421300	1801116	DGAT1
ARS-BFGL-NGS-53398	0,18±0,027	0,82±0,027	0,18	0,29	0,27	0,29	0,25	0,01	0,392	0,531	rs110836976	21953655	MAGEA11
ARS-BFGL-NGS-57820	0,66±0,033	0,34±0,033	0,34	0,45	0,51	0,45	0,35	0	1,65	0,198	rs109146371	1651311	FOXH1
ARS-BFGL-NGS-5790	0,88±0,023	0,12±0,023	0,12	0,22	0,21	0,22	0,19	0	0,16	0,689	rs110457668	26986116	NR4A1
ARS-BFGL-NGS-61452	0,54±0,035	0,46±0,035	0,46	0,5	0,48	0,5	0,37	0	0,114	0,735	-	75250982	HUS1

SNP	Частота аллеля 1	Частота аллеля 2	MAF	He	Ho	DG	PIC	Пропущенных аллелей, %	χ^2	p	RS	Позиция, п.н.	Ближайший ген
ARS-BFGL-NGS-69013	0,32±0,033	0,68±0,033	0,32	0,44	0,4	0,44	0,34	0	0,654	0,418	rs110252311	123717269	SESN2
ARS-BFGL-NGS-71395	0,5±0,035	0,5±0,035	0,5	0,5	0,58	0,5	0,38	0,01	2,273	0,131	rs42216771	24707157	SLC24A2
ARS-BFGL-NGS-72191	0,09±0,02	0,91±0,02	0,09	0,16	0,18	0,16	0,15	0	0,978	0,322	rs110840223	30279220	BoLA-DQB/H4
ARS-BFGL-NGS-85980	0,58±0,035	0,42±0,035	0,42	0,49	0,52	0,49	0,37	0,01	0,332	0,564	rs43392997	76091078	FAM71A
ARS-BFGL-NGS-89510	0,78±0,029	0,22±0,029	0,22	0,35	0,35	0,35	0,29	0	0,001	0,971	rs110970157	13936262	IDE
ARS-USMARC-Parent-DQ846692-rs29010281	0,56±0,035	0,44±0,035	0,44	0,49	0,59	0,49	0,37	0	3,781	0,051	rs29010281	80082923	RALYL
BTA-121739-no-rs	0,09±0,02	0,92±0,019	0,09	0,16	0,13	0,16	0,14	0	2,698	0,100	rs41622323	38063313	PKD2
BTA-39609-no-rs	0,8±0,028	0,2±0,028	0,2	0,33	0,33	0,33	0,27	0	0,015	0,901	-	0	-
BTA-48480-no-rs	0,19±0,028	0,81±0,028	0,19	0,31	0,34	0,31	0,26	0	1,094	0,295	rs41641868	95119968	ADAM23
BTA-50402-no-rs	0,75±0,031	0,25±0,031	0,25	0,38	0,42	0,38	0,3	0	1,44	0,230	rs41945918	34451383	-
BTA-68781-no-rs	0,91±0,02	0,09±0,02	0,09	0,17	0,18	0,17	0,15	0,01	0,99	0,319	rs43711373	30711619	HIST1H2AC
BTB-00231742	0,4±0,035	0,6±0,035	0,4	0,48	0,58	0,48	0,36	0	4,34	0,037	rs43436389	77095345	-
BTB-00285653	0,76±0,03	0,24±0,03	0,24	0,37	0,38	0,37	0,3	0,01	0,2	0,654	rs43499009	30036807	NFIB
BTB-00411816	0,07±0,018	0,93±0,018	0,07	0,13	0,14	0,13	0,12	0	0,567	0,451	rs43613632	11899342	THBS4
BTB-00777571	0,66±0,033	0,34±0,033	0,34	0,45	0,46	0,45	0,35	0,01	0,091	0,762	rs41938472	34017024	-
BTB-01052867	0,76±0,03	0,24±0,03	0,24	0,37	0,39	0,37	0,3	0	0,294	0,587	rs42215728	24734717	SLC24A2
BTB-01731924	0,6±0,035	0,4±0,035	0,4	0,48	0,51	0,48	0,36	0	0,45	0,502	rs42844632	75830763	GABRG2
BTB-01766447	0,64±0,034	0,36±0,034	0,36	0,46	0,55	0,46	0,35	0,02	3,92	0,047	rs42877704	17484533	RPL23A
Нармаp29482-BTA-146449	0,24±0,03	0,76±0,03	0,24	0,36	0,34	0,36	0,3	0	0,462	0,496	rs110728437	38804655	LAMA4
Нармаp30383-BTC-005848	0,52±0,035	0,48±0,035	0,48	0,5	0,53	0,5	0,37	0	0,393	0,530	rs109752439	1489496	C14H8orf33 – ZNF34
Нармаp38455-BTA-100999	0,4±0,035	0,6±0,035	0,4	0,48	0,44	0,48	0,36	0	0,694	0,404	rs41613779	76346736	OLIG3
Нармаp39690-BTA-82650	0,6±0,035	0,4±0,035	0,4	0,48	0,47	0,48	0,36	0	0,028	0,867	rs41590921	104829510	ZNF618
Нармаp39717-BTA-112973	0,14±0,025	0,86±0,025	0,14	0,23	0,25	0,23	0,21	0	0,496	0,481	rs41617243	26781358	KBTBD10
Нармаp40191-BTA-73919	0,76±0,03	0,23±0,03	0,23	0,36	0,41	0,36	0,29	0	1,969	0,160	rs41648982	71978791	SYN3
Нармаp42216-BTA-45665	0,32±0,033	0,68±0,033	0,32	0,44	0,51	0,44	0,34	0	2,637	0,104	rs41636126	45934555	GOSR2
Нармаp42263-BTA-60093	0,29±0,032	0,71±0,032	0,29	0,41	0,37	0,41	0,33	0,01	0,775	0,378	rs41567529	35342491	MIR2388

SNP	Частота аллеля 1	Частота аллеля 2	MAF	He	Ho	DG	PIС	Пропущенных аллелей, %	χ^2	p	RS	Позиция, п.н.	Ближайший ген
Нарmap47777-ВТА-91000	0,83±0,027	0,17±0,027	0,17	0,28	0,26	0,28	0,24	0	0,619	0,431	rs41659076	140892698	LOC785455
Нарmap48369-ВТА-50306	0,79±0,029	0,21±0,029	0,21	0,33	0,31	0,33	0,27	0,01	0,214	0,643	rs41639307	7627111	-
Нарmap48524-ВТА-92140	0,26±0,031	0,74±0,031	0,26	0,38	0,41	0,38	0,31	0	0,626	0,429	rs41663866	75684520	NCF4
Нарmap49309-ВТА-78604	0,65±0,034	0,35±0,034	0,35	0,45	0,48	0,45	0,35	0	0,302	0,582	rs41658012	24655689	GADD45B
Нарmap50621-ВТА-21320	0,83±0,027	0,17±0,027	0,17	0,28	0,28	0,28	0,24	0,01	0,003	0,954	rs41574875	64425164	-
Нарmap51303-ВТА-74377	0,16±0,026	0,84±0,026	0,16	0,28	0,31	0,28	0,24	0	1,563	0,211	rs41652649	83790390	ITPR2/PTHLH
Нарmap54188-rs29022489	0,28±0,032	0,72±0,032	0,28	0,4	0,47	0,4	0,32	0	3,193	0,073	rs29022489	75017253	-
Нарmap56889-rs29016097	0,39±0,034	0,61±0,034	0,39	0,48	0,44	0,48	0,36	0,01	0,474	0,490	rs29016097	31468943	HMGCS1
Нарmap59494-rs29020429	0,34±0,033	0,66±0,033	0,34	0,45	0,42	0,45	0,35	0	0,412	0,521	rs29020429	54062560	UFL1
Нарmap60955-rs29022431	0,63±0,034	0,37±0,034	0,37	0,47	0,48	0,47	0,36	0	0,088	0,767	rs29022431	21292766	-
UA-IFASA-6479	0,56±0,035	0,44±0,035	0,44	0,49	0,47	0,49	0,37	0,02	0,216	0,642	rs41566160	33201457	CACNB2

* MAF- частота минорных аллелей; He - ожидаемая гетерозиготность; Ho - наблюдаемая гетерозиготность; DG - генетическое разнообразие Нэя; PIС - уровень полиморфизма; Miss - процент пропущенных значений; χ^2 (хи-квадрат)- значение статистики хи-квадрат, используемой для проверки соответствия наблюдаемых частот генотипов ожидаемым при равновесии Харди-Вайнберга; p - уровень значимости (*p-value*) для теста хи-квадрат; RS - идентификатор SNP в базе данных dbSNP.

Анализ распределения частот минорных аллелей выявил существенную вариабельность между исследованными SNP. Частота минорного аллеля (MAF) варьирует в широком диапазоне: от 0,06 (ARS-BFGL-NGS-107749) до 0,5 (ARS-BFGL-NGS-71395) (рисунок 28), что свидетельствует о наличии в популяции как редких генетических вариантов, так и более распространенных полиморфизмов. Для большинства SNP значение MAF находилось в диапазоне от 0,15 до 0,30, что может указывать на адаптивную роль данных полиморфизмов или нейтральность по отношению к селективным силам.

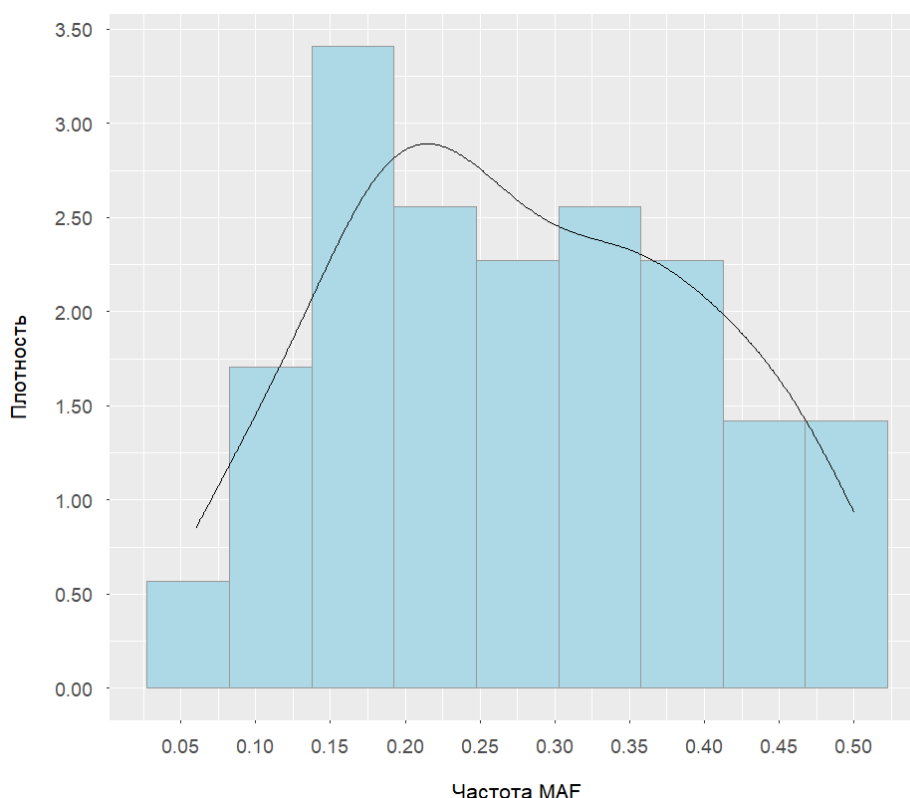


Рисунок 28 – Плотность распределения частот минорных аллелей (MAF)

Наличие нескольких однонуклеотидных полиморфизмов (SNP) с частотой минорного аллеля (MAF), приближающейся к 0,5 свидетельствует о потенциальном влиянии балансирующего отбора на поддержание генетического разнообразия в исследуемой популяции. Поддержание высокой MAF может быть обусловлено диверсифицирующим отбором в гетерогенной среде, где различные аллели предоставляют преимущества в различных условиях [171, 179]. Это особенно актуально для популяций, обитающих в условиях переменной среды, где разные

генотипы могут быть более адаптивными в разные периоды времени или в разных экосистемах. Для подтверждения гипотез о балансирующем и диверсифицирующем отборе необходимы дальнейшие исследования, включающие анализ функционального воздействия данных SNP на фенотип, проведение полногеномного поиска ассоциаций (GWAS) для выявления связей с хозяйственно-полезными признаками, и сравнительный анализ частот аллелей в различных популяциях того же вида для оценки адаптивной роли SNP. При этом важно учитывать и возможное влияние случайного генетического дрейфа, особенно в малочисленных популяциях. Комплексный анализ позволит более точно определить роль балансирующего отбора в поддержании высокой MAF для SNP (ARS-BFGL-NGS-71395 (MAF = 0,5) и Нармар30383-BTC-005848 (MAF = 0,48)) и оценить их вклад в адаптацию и генетическое разнообразие популяции чёрно-пёстрого скота Западной Сибири.

На рисунке 29 представлено распределение ранжированных разностей частот аллелей для 64 однонуклеотидных полиморфизмов (SNP) у животных в изучаемой популяции. Замечена значительная вариабельность частот аллелей. Наблюдаются как SNP с близкими значениями частотами обоих аллелей (например, ARS-BFGL-NGS-71395, Нармар30383-BTC-005848), так и SNP с преобладанием одного из аллелей (например, ARS-BFGL-NGS-107749, ВТВ-00411816). Такая вариабельность может отражать различные эволюционные процессы, влияющие на данные локусы, такие как селективное давление или генетический дрейф. SNP с близкими частотами аллелей (близкими к 0,5) могут свидетельствовать о балансирующем отборе, поддерживающем оба аллеля в популяции, в то время как SNP с преобладанием одного аллеля могут указывать на направленный отбор или эффект бутылочного горлышка.

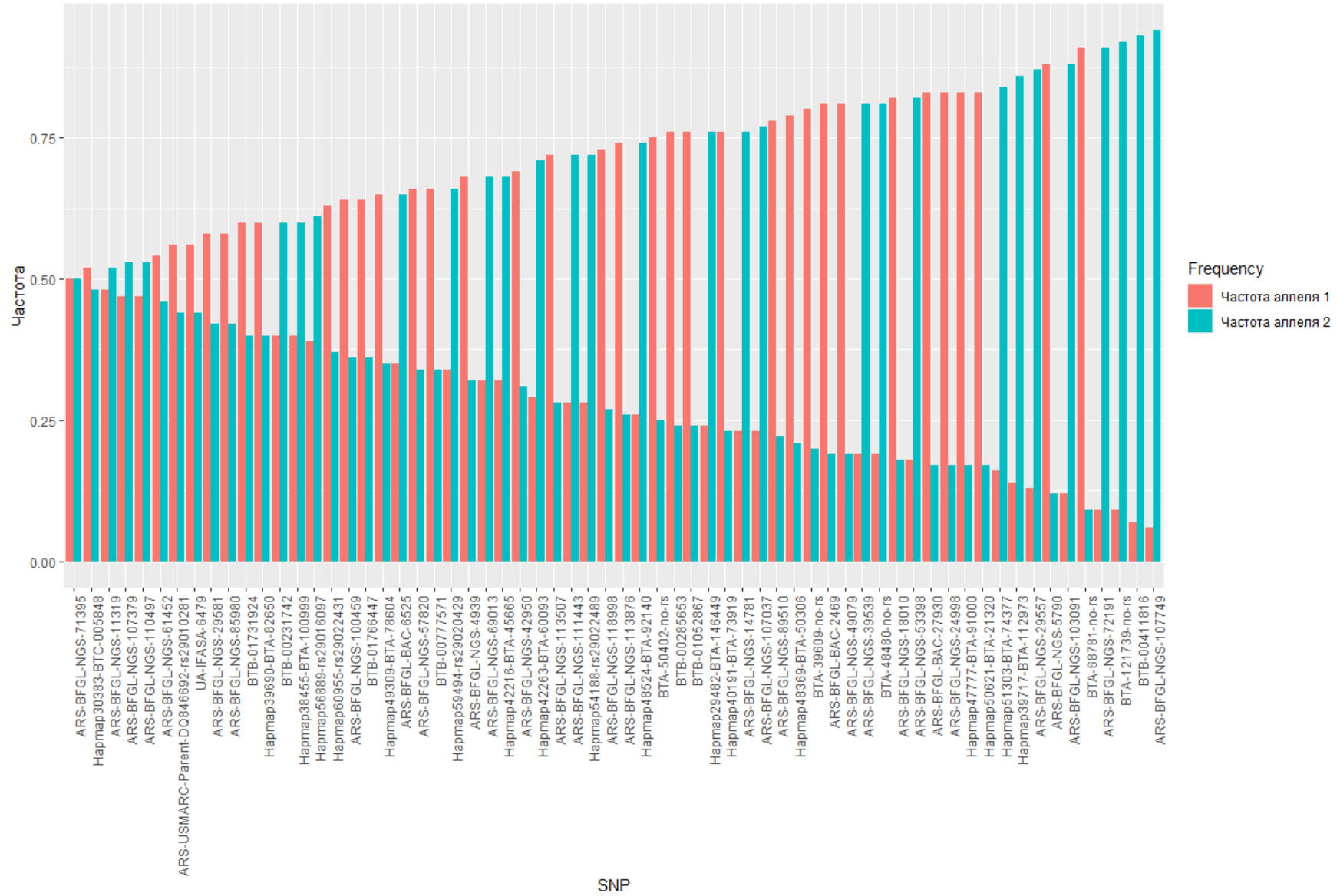


Рисунок 29 – Частоты аллелей по изучаемым SNP

Для оценки генетического разнообразия популяции были рассчитаны показатели ожидаемой (H_e) и наблюдаемой (H_o) гетерозиготности, а также уровень полиморфизма (PIC). Среднее значение ожидаемой гетерозиготности (H_e) составило $0,37 \pm 0,013$, что свидетельствует о достаточно высоком уровне генетического разнообразия в исследуемой популяции. Распределение значений H_e представлено на рисунке 30. Наибольшее отклонение в абсолютной разности наблюдаемой и ожидаемой гетерозиготности выявлено для SNP VTB-00231742 и ARS-USMARC-Parent-DQ846692-rs2901028. Важно отметить, что наблюдаемая гетерозиготность (H_o) для большинства SNP соответствовала ожидаемой, что указывает на отсутствие выраженных нарушений равновесия Харди-Вайнберга (HWE) в исследуемой популяции. Это, в свою очередь, может указывать на то, что влияние направленного отбора на эти локусы не является доминирующим, и генетическая структура популяции в отношении этих SNP достаточно стабильна. Анализируя значения p , полученные в результате теста хи-квадрат (χ^2) для каждого SNP на соответствие равновесию Харди-Вайнберга можно оценить, соответствует ли наблюдаемое распределение генотипов ожидаемому при равновесии Харди-Вайнберга. Учитывая порог значимости в 0,05 и p в таблице почти все SNP не показывают статистически значимого отклонения от HWE. Тем не менее, для двух SNP наблюдаются незначительные отклонения от равновесия, о чем свидетельствуют низкие значения p для критерия хи-квадрат. Например, для VTB-00231742 ($p = 0,037$) и VTB-01766447 ($p = 0,047$). Это может быть связано с селективным преимуществом гетерозигот, явлением, известным как сверхдоминирование, когда гетерозиготы имеют более высокую приспособленность по сравнению с гомозиготами.

Чтобы снизить вероятность ложноположительных результатов была сделана корректировка на множественное тестирование. В реальных исследованиях с большим количеством SNP корректировка практически обязательна. Для этих целей была принята корректировка Бонферрони, что позволило принять порог значимости как 0,0007. Все p в таблице больше 0,0007, поэтому с корректировкой Бонферрони ни один из SNP не показывает статистически значимого отклонения

от HWE.

Уровень полиморфизма (PIC) варьировал от 0,1 до 0,38, при среднем значении $0,29 \pm 0,008$. Разделив SNP по уровню полиморфизма на три группы с низким ($PIC < 0,25$), средним ($0,25 \leq PIC \leq 0,5$) и высоким уровнем ($PIC > 0,5$) можно утверждать, что из 64 SNP – 14 вошли в группу с низким полиморфизмом, 50 со средним, в группу с высоким ни один SNP не вошел. Это подтверждает то, что в целом уровень полиморфизма в исследуемой популяции скорее средний, чем высокий. Небольшое значение стандартной ошибки (0,008) говорит о том, что значения PIC распределены относительно равномерно вокруг среднего значения.

Наблюдаемый диапазон PIC может быть обусловлен несколькими факторами, связанными с историей и структурой исследуемой популяции. Один из таких факторов может проявляться в виде эффекта бутылочного горлышка или эффекта основателя, что привело к утрате редких аллелей и снижению общего генетического разнообразия. Селекционное давление, будь то очищающая селекция против вредных мутаций или направленная селекция на определенные признаки, также может снизить вариабельность в связанных локусах и, следовательно, PIC. Инбридинг, распространенный в некоторых популяциях, увеличивает гомозиготность и уменьшает гетерозиготность, что непосредственно влияет на снижение PIC. Наконец, некоторые участки генома могут быть изначально менее полиморфными из-за функциональных ограничений или других локус-специфических факторов, что также может внести свой вклад в наблюдаемый диапазон значений PIC. Важно учитывать все эти факторы при интерпретации результатов и планировании дальнейших исследований. Анализ значений генетического разнообразия Нэя (DG) показывает, что исследованные SNP демонстрируют достаточно высокий уровень генетического разнообразия от 0,1 до 0,5. Наличие SNP как с низким, так и с высоким GD подчеркивает сложную генетическую структуру изучаемой популяции чёрно-пёстрого скота. Среднее значение GD $0,37 \pm 0,013$ указывает на "умеренный" уровень генетического разнообразия в исследованной популяции в целом.

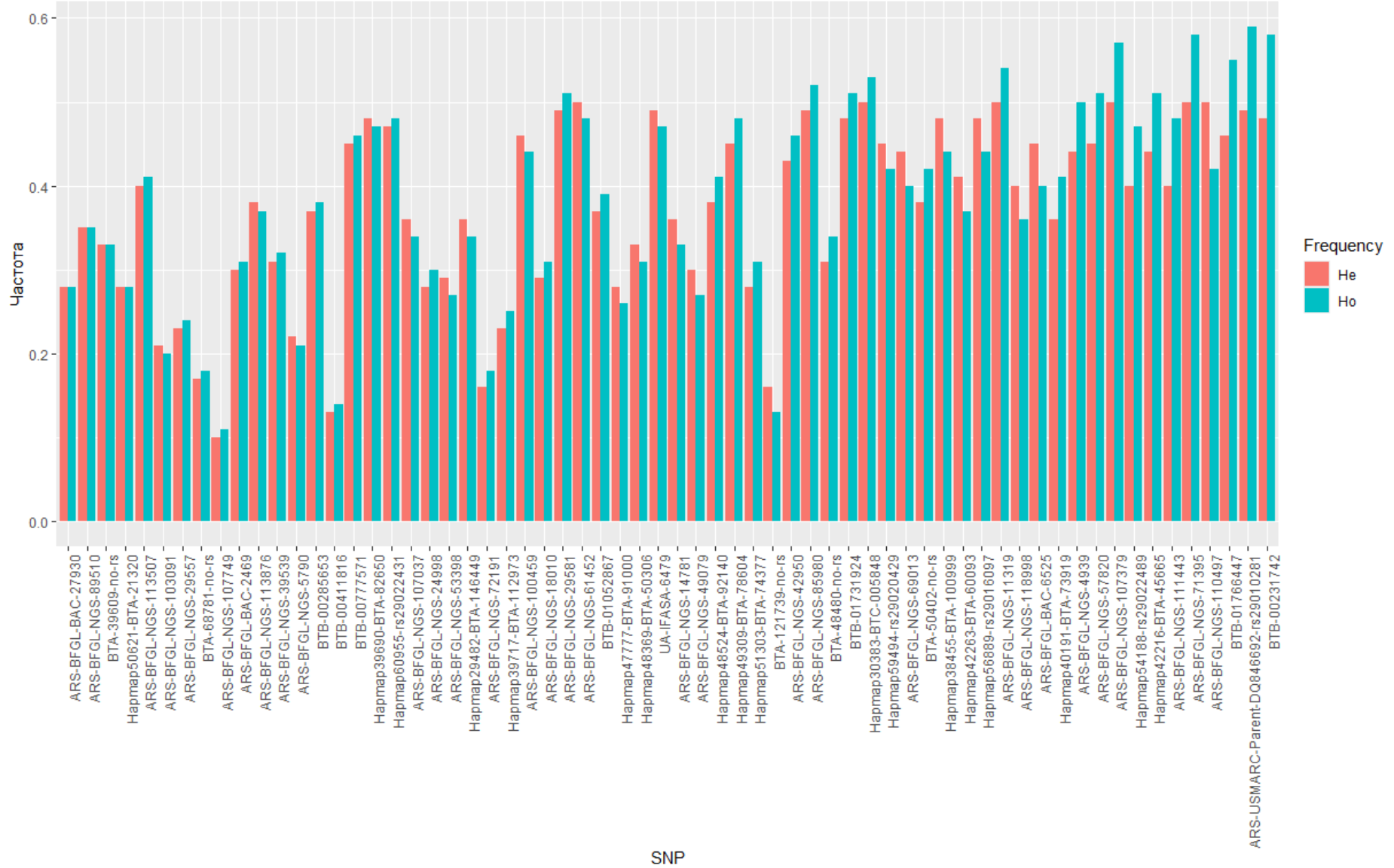


Рисунок 30 – Сравнение наблюдаемой и ожидаемой гетерозигонности

Рассматривая гены, ближайšie к исследуемым SNP (таблица 14, приложение В), можно предположить потенциальные механизмы влияния генетических вариантов на молочную продуктивность. Например, SNP ARS-BFGL-NGS-4939 расположен рядом с геном DGAT1 (диацилглицерол О-ацилтрансферазы 1). DGAT1 ассоциирован с различными признаками молочной продуктивности, такими как выход жира, удой, процент жира и процент белка. DGAT1 играет ключевую роль в синтезе триацилглицеридов, основных компонентов молочного жира. Вариации в гене DGAT1, отраженные через полиморфизм SNP ARS-BFGL-NGS-4939, вероятно, могут влиять на эффективность синтеза молочного жира и, следовательно, на жирномолочность скота. Интересно, что этот SNP демонстрирует относительно высокое значение GD (0,44) и PIC (0,34), что делает его потенциально важным маркером для селекции по признакам жирномолочности в данной популяции [166, 240].

Другие SNP также ассоциированы с генами, чьи функции могут быть связаны с молочной продуктивностью. Так, SNP ARS-BFGL-NGS-100459 расположен рядом с геном ABCG8 (ATP-binding cassette sub-family G member 8), который ассоциирован с процентом жира и процентом белка в молоке. Гены семейства ABC (ATP-binding cassette) часто участвуют в транспорте различных молекул через клеточные мембраны. В контексте молочной продуктивности, ABCG8 может быть вовлечен в транспорт липидов или других компонентов молока в молочную железу [251].

SNP ARS-BFGL-NGS-71395 ассоциирован с геном SLC24A2 (Solute Carrier Family 24 Member 2), который также связан с молочной продуктивностью. Гены семейства SLC кодируют белки-переносчики растворенных веществ. SLC24A2 может играть роль в транспорте ионов или других молекул, важных для секреции молока или поддержания осмотического баланса в молочной железе [155].

SNP ARS-BFGL-NGS-11319 ассоциирован с геном SLC40A1 (Solute Carrier Family 40 Member 1), также связанным с молочной продуктивностью. SLC40A1, известный как ферропортин, играет ключевую роль в транспорте железа из клеток. Хотя связь железа с молочной продуктивностью может показаться неочевид-

ной, железо необходимо для многих метаболических процессов, включая синтез белков и липидов, которые являются важными компонентами молока.

Следует отметить, что многие из исследуемых SNP ассоциированы с генами, функции которых еще не полностью изучены в контексте молочной продуктивности, как, например, HCN1, NRXN3, RBFOX2, USP3, CDH16, ADAM12, DERA, PTRN1, AGXT2L2, GHR, DDIT3, ADAMTS8, PTHLH, KCNH2, JARID2, TMEM132C, DGAT1, MAGEA11, FOXH1, NR4A1, HUS1, SESN2, SLC24A2, BoLA-DQB/H4, FAM71A, IDE, RALYL, PKD2, ADAM23, HIST1H2AC, THBS4, GABRG2, RPL23A, LAMA4, C14H8orf33-ZNF34, OLIG3, ZNF618, KBTBD10, SYN3, GOSR2, CACNB2, HMGC S1, UFL1 [142, 153, 155, 156, 165, 201, 253, 256, 260]. Дальнейшие исследования, направленные на изучение генных сетей и путей, в которых участвуют эти гены, а также их взаимодействия, необходимы для полного понимания генетических механизмов, лежащих в основе молочной продуктивности чёрно-пёстрого скота Западной Сибири.

Анализ генетической структуры популяции чёрно-пёстрого скота Сибири по SNP-локусам, ассоциированным с признаками молочной продуктивности, выявил умеренный уровень генетического разнообразия и наличие SNP, потенциально связанных с важными генами, участвующими в различных аспектах молочной продуктивности. Дальнейшие исследования с использованием более широкого спектра геномных данных и интеграция с фенотипическими данными позволят углубить понимание генетической архитектуры молочной продуктивности и разработать эффективные стратегии геномной селекции для чёрно-пёстрого скота Западной Сибири.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ

Настоящее диссертационное исследование, посвященное комплексному изучению генеалогической и генетической структуры чёрно-пёстрого скота Западной Сибири, представляет собой комплексную попытку в понимании генетического потенциала и особенностей селекционно-племенной работы с сибирской популяцией скота в специфических условиях Сибири. В отличие от предшествующих исследований, которые, как правило, ограничивались изучением отдельных хозяйств или небольших выборок животных, настоящая работа охватила популяцию чёрно-пёстрого скота на региональном уровне, что позволило получить более целостное и репрезентативное представление о её генетической структуре и изменчивости хозяйственно-полезных признаков.

Одним из наиболее важных и трудоемких этапов диссертационной работы стала разработка и апробация комплексной методологии верификации и валидации данных первичного зоотехнического учёта. В условиях российской практики, где базы данных племенных предприятий зачастую фрагментированы, разрознены и не всегда соответствуют единым стандартам, а проблема достоверности исходной информации приобретает критическое значение. Так фрагментарность генеалогической информации представляет собой существенный ограничивающий фактор для формирования высококачественных баз данных регионального и федерального уровня, препятствуя реализации эффективных программ по улучшению и сохранению генетического потенциала популяций сельскохозяйственных животных. Неполные родословные снижают точность оценки племенной ценности животных, что затрудняет выявление и отбор лучших особей для дальнейшего разведения. Кроме того, вариабельность генеалогической информации ограничивает возможности для эффективного подбора родительских пар, снижая эффективность селекционных мероприятий, направленных на повышение продуктивности и улучшение других хозяйственно-полезных признаков.

С целью улучшения качества генеалогических данных и для определения корректной региональной генеалогической структуры необходимо подключение

дополнительных зарубежных источников с информацией о предках животных, представленных в местных популяциях. Это объясняется тем, что на протяжении последних десятилетий маточное поголовье в Западной Сибири и России в целом формировалось преимущественно путём поглотительного скрещивания местных чёрно-пёстрых коров с иностранными быками голштинской породы. Также необходимо проводить комплексные меры по актуализации генеалогических данных на самих племенных предприятиях, с учётом важности заполнения в родословных информации о происхождении быков-производителей. Проведенный в диссертационном исследовании анализ генеалогических баз данных племенных предприятий Сибирского федерального округа наглядно продемонстрировал обоснованность всех изложенных опасений. Выявленная значительная вариабельность качества генеалогических данных между хозяйствами стала отправной точкой для разработки системы мер по повышению надежности исходной информации.

Глубокий анализ заполненности родословных выявил не только пропуски и неполноту данных, что само по себе является серьезной проблемой, но и наличие систематических ошибок, таких как заикливания и дублирования записей. Эти ошибки, на первый взгляд кажущиеся незначительными, на самом деле могут вносить существенные искажения в результаты популяционно-генетических исследований и, как следствие, приводить к принятию неоптимальных селекционных решений. Более того, в ходе верификации данных были обнаружены косвенные признаки, указывающие на потенциальную фальсификацию (искажение) данных, такие как присвоение одной корове непропорционально большого количества потомков или несоответствие дат рождения предков и потомков. Эти тревожные факты подчеркивают необходимость усиления контроля за достоверностью первичной учетной документации и внедрения механизмов, исключающих возможность манипуляций с данными.

Разработанные в рамках диссертационного исследования алгоритмы верификации и валидации, основанные на сопоставлении данных из различных источников (включая международные базы данных CDN и ВНИИплем, а также открытые источники генеалогической информации), графической визуализации генеа-

логических деревьев и применении методов многомерной статистики, позволили не только выявить широкий спектр ошибок и искажений, но и автоматизировать процесс их исправления. Созданное программное обеспечение, реализующее разработанные алгоритмы, представляет собой ценный практический инструмент, который может быть использован племенными предприятиями и региональными информационно-селекционными центрами для систематической очистки и улучшения качества генеалогических баз данных. Внедрение подобных систем верификации на федеральном уровне может помочь обеспечению прозрачности, объективности и эффективности племенной работы в масштабах всей страны. Только на основе достоверных и качественных данных возможно построение надежных моделей прогнозирования племенной ценности животных и принятие обоснованных селекционных решений, направленных на повышение продуктивности и конкурентоспособности отечественного молочного скота.

Валидация фенотипических данных, проведенная на примере признаков молочной продуктивности, также выявила ряд важных особенностей региональных баз данных. Использование критерия Андерсона-Дарлинга для оценки соответствия эмпирических распределений нормальному закону позволило количественно оценить качество фенотипических данных и выявить отклонения, свидетельствующие о потенциальных проблемах. Наличие асимметрии распределений и «тяжелых хвостов» для ряда признаков, особенно массовой доли жира, может быть обусловлено как объективными биологическими факторами, такими как генетическая гетерогенность популяции и влияние средовых факторов, так и субъективными факторами, связанными с ошибками измерений или некорректным ведением учета.

Особого внимания заслуживает выявленная двухвершинность распределения удоя для ряда хозяйств, что может указывать на наличие в стаде двух или более групп животных, существенно различающихся по генетическому потенциалу и уровню продуктивности, реализующих свой генетический потенциал в разной степени. Это связано с условиями кормления и содержания животных и тем, какие быки-производители используются на предприятии, что сопряжено с каче-

ством отбора. Генотип одного и того же производителя на предприятиях с разным уровнем продуктивности маточного поголовья будет неодинаково реализовываться в показателях продуктивности и здоровья своих дочерей. Такая проблема особенно актуальна для России, где уровень молочной продуктивности варьируется в широких пределах как в разных регионах, так и предприятиях. Это контрастирует с продуктивностью, например, в Америке или Канаде, где наблюдается большая консолидация стад по этому признаку, что создаёт основу для формирования однородного стада при относительно однородных условиях среды. Таким образом, следует рассмотреть возможность введения в России категориальной индексной системы оценки племенной ценности животных, при которой прогноз по признакам отбора по одному быку-производителю будет неодинаковым на предприятиях с разным уровнем производственной культуры в тех или иных природно-климатических условиях.

Генеалогическая структура популяции чёрно-пёстрого скота Западной Сибири, выявленная в ходе диссертационного исследования, является своеобразным отражением истории формирования и развития породы в регионе, а также современных тенденций в селекционно-племенной работе. Неравномерность заполнения родословных, с явным преобладанием полноты данных по отцовским линиям, является, с одной стороны, объективным следствием широкого распространения искусственного осеменения и доступности генеалогической информации о быках-производителях. Использование семени высокоценных быков иностранной селекции, информация о происхождении которых тщательно документируется и доступна в международных базах данных, стало обычной практикой в молочном скотоводстве России на протяжении последних десятилетий. В то же время, недостаточная полнота данных по материнским линиям может рассматриваться как негативная тенденция, ограничивающая возможности для комплексной оценки племенной ценности коров и формирования сбалансированных селекционных программ. В условиях современной селекции, все большее внимание уделяется материнским качествам, репродуктивному долголетию и общей адаптивности животных, что требует полноценной генеалогической информации по обоим роди-

тельским линиям.

Высокая доля пропущенных данных о происхождении быков-осеменителей (85.5% от общего числа пропусков), выявленная в исследовании, является тревожным сигналом, указывающим на потенциальные проблемы с контролем за происхождением и качеством племенного материала, используемого в регионе. Данный факт может быть обусловлен различными причинами, включая недостатки системы племенного учета, неполноту или недостоверность информации, предоставляемой поставщиками племенного материала, а также человеческий фактор, связанный с ошибками при внесении данных в базы данных предприятий. В любом случае, недостаточная информация о происхождении быков-осеменителей создает риски для селекционно-племенной работы, затрудняя оценку их племенной ценности и контроль за генетическим разнообразием популяции. Усиление контроля за происхождением и качеством племенного материала, используемого в регионе, должно стать одним из приоритетных направлений деятельности региональных органов управления АПК и племенных служб.

Кластерный анализ хозяйств по качеству генеалогической информации позволил выявить определенную стратификацию предприятий региона по уровню ведения племенного учета. Выделение групп хозяйств, демонстрирующих схожие характеристики полноты и достоверности генеалогических данных, может быть использовано для разработки дифференцированных стратегий поддержки и стимулирования улучшения племенного учета. В частности, для хозяйств, отнесенных к кластерам с низким качеством генеалогической информации, могут быть разработаны специальные программы обучения и методической помощи, направленные на повышение квалификации специалистов-зоотехников и внедрение современных технологий ведения племенного учета. В то же время, хозяйства-лидеры, демонстрирующие высокий уровень качества генеалогической информации, могут служить примером для других предприятий и стать площадками для распространения лучших практик и обмена опытом в области племенного дела.

Визуализация генеалогических деревьев, реализованная в рамках диссертационной работы, оказалась не только эффективным инструментом для выявления

и исправления ошибок в родословных, но и ценным средством для наглядного представления генеалогической структуры популяции. Графический подход позволяет специалистам-зоотехникам лучше ориентироваться в массивах генеалогической информации, выявлять наиболее распространенные линии и семейства, оценивать степень родства между животными и принимать более обоснованные решения в области селекции и племенного дела. Внедрение инструментов визуализации генеалогических деревьев в практику племенных предприятий может существенно повысить эффективность селекционно-племенной работы и способствовать сохранению и улучшению генетического потенциала чёрно-пёстрого скота Западной Сибири.

Генетический анализ изучаемой популяции чёрно-пёстрого скота, проведенный с использованием STR-локусов микросателлитов, выявил высокий уровень генетического разнообразия. Этот результат, на первый взгляд, может показаться неожиданным, учитывая многолетнюю историю интенсивной голштинизации, которая, как правило, приводит к гомогенизации генофонда и снижению генетической изменчивости. Однако, более глубокий анализ полученных данных позволяет предложить ряд объяснений для данного феномена.

Во-первых, высокий уровень генетического разнообразия по STR-локусам микросателлитов может отражать исторические особенности формирования чёрно-пёстрого скота в Сибири, включающие различные этапы интродукции и селекции, а также адаптацию к специфическим условиям среды. Сибирское отродье чёрно-пёстрого скота формировалось на протяжении длительного времени в условиях сурового континентального климата, что могло способствовать сохранению и развитию адаптивных генетических вариантов, не всегда связанных с признаками молочной продуктивности, которые являются основными целями селекции в голштинской породе.

Во-вторых, высокий уровень полиморфизма STR-локусов может быть обусловлен нейтральным характером данных маркеров по отношению к селекционному давлению. Микросателлиты, как правило, располагаются в некодирующих областях генома и не оказывают прямого влияния на хозяйственно-полезные при-

знаки, что позволяет им сохранять высокую изменчивость даже в условиях интенсивной селекции по продуктивности. В этом контексте, STR-локусы могут служить ценными маркерами генетического разнообразия популяции, отражая ее историческое происхождение и эволюционные процессы, но не обязательно коррелируя с признаками, представляющими непосредственный интерес для селекции.

В то же время, отклонения от равновесия Харди-Вайнберга, выявленные для ряда STR-локусов, указывают на действие эволюционных сил, влияющих на генетическую структуру популяции. В частности, дефицит гетерозигот по локусу TGLA53 может быть обусловлен селективным преимуществом гомозигот по данному локусу или эффектом Валунда, связанным с популяционной субструктурой и смешиванием генетически дифференцированных групп животных. Для более точной интерпретации отклонений от равновесия Харди-Вайнберга необходимы дальнейшие исследования, направленные на изучение популяционной структуры и характера отбора, действующего на отдельные локусы и аллели.

В будущих исследованиях анализ распределения частот аллелей может позволить выявить генетическую дифференциацию внутри популяции и определить наличие субпопуляций. Частоты аллелей могут быть использованы в качестве переменных в ассоциативных исследованиях для выявления генетических маркеров, связанных с продуктивностью, устойчивостью к заболеваниям и другими важными признаками.

Зная, что равновесие Харди-Вайнберга описывает теоретическое состояние популяции, в котором частоты аллелей и генотипов остаются постоянными из поколения в поколение при определенных условиях (отсутствие мутаций, миграции, селекции и случайного дрейфа генов, а также при случайном скрещивании), то рассмотрение количества наблюдаемых гомозигот по отношению к ожидаемым представляет особый интерес, поскольку оно позволяет выявить отклонения от этого равновесия.

Анализ генетической структуры популяции по SNP-локусам, ассоциированным с признаками молочной продуктивности, выявил несколько иную картину генетической изменчивости. Уровень генетического разнообразия по данным мар-

керам оказался умеренным, что, в целом, соответствует ожиданиям для породы, подвергавшейся интенсивной селекции на молочную продуктивность. Преобладание гомозигот и высокая частота мажорных аллелей для большинства SNP, ассоциированных с признаками молочной продуктивности, может быть обусловлено селекционным давлением, направленным на фиксацию аллелей, благоприятных для удоя, жирности и белковости молока. Искусственный отбор, осуществляемый на протяжении многих поколений, приводит к увеличению частоты желательных аллелей и снижению частоты нежелательных, что, в конечном итоге, ведет к гомогенизации генофонда по селективируемым признакам.

Вместе с тем, даже в SNP-локусах, подверженных селекционному давлению, сохраняется определенный уровень генетической изменчивости, о чем свидетельствует наличие SNP с частотой минорного аллеля (MAF) выше 0,23. Эти SNP могут представлять интерес для геномной селекции, поскольку они могут быть связаны с QTL (Quantitative Trait Loci), контролирующими хозяйственно-полезные признаки, и использованы для повышения точности оценки племенной ценности животных. Отрицательное значение среднего коэффициента инбридинга ($F = -0,03$), полученное на основе SNP-данных, свидетельствует об эффективности селекционных программ, направленных на снижение инбридинга в популяции. Поддержание низкого уровня инбридинга является важным фактором для сохранения жизнеспособности, продуктивного долголетия и адаптивного потенциала скота, особенно в условиях интенсивного использования геномной селекции, которая может способствовать ускоренному распространению генотипов отдельных выдающихся производителей и, как следствие, увеличению инбридинга в популяции.

Существенное преобладание одного аллеля, выявленное для ряда SNP, ассоциированных с признаками молочной продуктивности, может быть интерпретировано как результат действия как естественного, так и искусственного отбора. В условиях промышленного производства, отбор может быть направлен на закрепление аллелей, обеспечивающих высокую продуктивность и технологичность молока, что приводит к увеличению частоты мажорных аллелей и снижению генети-

ческого разнообразия по данным локусам. В то же время, нельзя исключать влияние естественного отбора, направленного на адаптацию животных к специфическим условиям среды сибирского региона. Для более точной интерпретации аллельных частот SNP необходимы дальнейшие исследования, направленные на изучение функциональной роли генов, в которых расположены данные SNP, и проведение ассоциативных исследований с фенотипическими признаками, характеризующими не только молочную продуктивность, но и адаптивность, устойчивость к заболеваниям и другие хозяйственно-полезные качества скота в сибирских условиях.

В целом, анализ генетической структуры чёрно-пёстрого скота Западной Сибири по STR- и SNP-локусам выявил сложную картину, отражающую баланс между генетическим разнообразием и селекционным давлением. Сохранение высокого уровня генетического разнообразия по микросателлитным маркерам, наряду с умеренным полиморфизмом SNP, ассоциированных с признаками молочной продуктивности, свидетельствует о наличии генетического потенциала для дальнейшей селекции и адаптации популяции к изменяющимся условиям среды. Дальнейшие исследования, направленные на расширение геномных исследований, идентификацию QTL и разработку региональных геномных индексов племенной ценности, позволят более эффективно использовать генетические ресурсы чёрно-пёстрого скота Западной Сибири для повышения продуктивности и конкурентоспособности молочного скотоводства в регионе.

Валидация зарубежных баз данных CDN путем кластеризации стран-лидеров молочного скотоводства по признакам молочной продуктивности стала важным этапом диссертационного исследования, позволившим оценить позиции России в мировом молочном скотоводстве и выявить перспективные направления для международного сотрудничества в области селекции и племенного дела. Кластерный анализ, проведенный на основе комплекса признаков молочной продуктивности, выявил существенную дифференциацию стран по селекционным стратегиям и моделям ведения молочного скотоводства. Выделение различных кластеров, ориентированных на интенсивное увеличение удоя, пастбищное содержание и ка-

чество молока, или сбалансированную селекцию, подчеркивает необходимость учета региональной специфики и хозяйственно-экономических условий при разработке селекционных программ. Лидирующие позиции США и Нидерландов в кластере стран с наиболее высокой молочной продуктивностью подтверждают эффективность интенсивных селекционных программ, направленных на максимальное увеличение удоя, но при этом выделение Новой Зеландии и Ирландии в отдельный кластер указывает на существование альтернативных, экономически и экологически обоснованных моделей, которые могут быть релевантны для регионов с ограниченными ресурсами и специфическими климатическими условиями, таких как Сибирь.

Интересным результатом кластеризации стало включение России и Беларуси во второй кластер, характеризующийся более сбалансированным подходом к селекции. Данный факт, с одной стороны, может свидетельствовать о наличии определенного потенциала для интенсификации селекционной работы и повышения продуктивности в отечественном молочном скотоводстве, а с другой – указывает на необходимость сохранения внимания к адаптивным качествам и устойчивости животных, традиционно важным для российских пород скота. Результаты валидации зарубежных баз данных CDN могут быть использованы для бенчмаркинга и международного сотрудничества в области селекции молочного скота, позволяя оценить позиции России на мировом рынке племенной продукции и выявить перспективные направления для обмена генетическим материалом и технологиями.

Разработка математической модели прогнозирования удоя чёрно-пёстрого скота с учетом паратипических и генетических факторов стала важным шагом на пути к созданию региональной системы геномной селекции. Полученная модель, основанная на смешанных регрессионных моделях и учитывающая влияние широкого спектра факторов, продемонстрировала достаточно высокую точность прогноза на верифицированных данных ($R^2 = 0.819$). Данный результат подтверждает значимость комплексного учета как генетических, так и средовых факторов в изменчивости удоя и открывает перспективы для повышения эффективности се-

лекции за счет использования математического моделирования и современных информационных технологий. Включение в модель фиксированных эффектов, таких как «Год отёла» и «Сезон отёла», а также случайных эффектов, связанных с индивидуальными особенностями животных, лактацией и влиянием отцов, позволило уточнить прогноз племенной ценности и учесть влияние различных источников изменчивости удоя. В то же время, необходимо отметить, что модель, разработанная на основе фенотипических данных, имеет определенные ограничения и может быть существенно улучшена за счет включения геномной информации. Дальнейшие исследования, направленные на интеграцию данных генотипирования SNP в модель прогнозирования удоя, позволят создать более точный и надежный инструмент для геномной селекции чёрно-пёстрого скота, адаптированный к региональным условиям и хозяйственно-экономическим реалиям.

Полученные в диссертации результаты имеют выраженную практическую и теоретическую значимость для молочного скотоводства Сибирского региона и отрасли в целом. В теоретическом плане, диссертационная работа вносит вклад в развитие популяционной генетики и селекции сельскохозяйственных животных, расширяя знания о генеалогической и генетической структуре чёрно-пёстрого скота, выявляя факторы, влияющие на изменчивость удоя, и развивая методологию валидации баз данных и математического моделирования в животноводстве. В практическом плане, диссертация предлагает конкретные инструменты и методические подходы, которые могут быть непосредственно внедрены в практику племенных предприятий и региональных информационно-селекционных центров. Разработанная единая региональная база данных, включающая верифицированные генеалогические и фенотипические данные, является ценным ресурсом для проведения дальнейших исследований и селекционно-племенной работы. Созданное программное обеспечение для верификации данных и визуализации генеалогических деревьев позволяет автоматизировать рутинные операции и повысить качество племенного учета. Разработанная математическая модель прогнозирования удоя может быть использована для оценки племенной ценности животных и оптимизации селекционных программ.

Результаты исследования могут быть использованы для совершенствования селекционно-племенной работы с чёрно-пёстрым скотом Западной Сибири, повышения точности оценки племенной ценности, разработки региональных программ селекции, направленных на повышение продуктивности и адаптивности животных в условиях Западной Сибири, и, в конечном итоге, для увеличения производства молока и повышения конкурентоспособности отечественного молочного скотоводства. Особую практическую значимость имеет снижение зависимости от импортных технологий и племенного материала, что является стратегически важной задачей в условиях современных геополитических вызовов.

Перспективы дальнейших исследований в данной области представляются весьма широкими и многообещающими. В первую очередь, необходимо расширение базы данных генотипирования SNP для проведения полногеномных ассоциативных исследований (GWAS) и идентификации генетических маркеров, ассоциированных не только с молочной продуктивностью, но и с адаптивными качествами, устойчивостью к заболеваниям и продуктивным долголетием скота в сибирских условиях. На основе результатов GWAS возможно разработка и внедрение региональных геномных индексов племенной ценности, учитывающих специфику генофонда чёрно-пёстрого скота Западной Сибири и региональные факторы среды. Совершенствование математических моделей прогнозирования продуктивности с использованием данных генотипирования SNP, биохимических показателей крови и других физиологических параметров позволит повысить точность и персонализацию селекционных решений, переходя от массового отбора к индивидуальному подходу, учитывающему генетические и фенотипические особенности каждого животного. Разработка специализированного программного обеспечения для автоматизации процессов верификации данных, генеалогического и генетического анализа, геномной селекции и прогнозирования продуктивности станет важным шагом на пути к внедрению современных технологий в практику племенных предприятий и региональных селекционно-информационных центров. Наконец, дальнейшее изучение адаптивного потенциала чёрно-пёстрого скота с использованием геномных технологий и методов фенотипирования, направленных на выяв-

ление генетических и физиологических механизмов адаптации к стрессовым факторам среды, позволит разработать селекционные программы, направленные на повышение устойчивости и продуктивного долголетия скота в условиях изменяющегося климата и возрастающих требований к экологической безопасности и устойчивости молочного скотоводства.

Следует подчеркнуть, что комплексный подход, реализованный в данном диссертационном исследовании, открывает новые перспективы для развития молочного скотоводства в Сибирском регионе и способствует повышению конкурентоспособности отечественной племенной продукции. Дальнейшее развитие исследований в данном направлении, с использованием современных геномных и биоинформатических технологий, позволит в полной мере раскрыть генетический потенциал чёрно-пёстрого скота и обеспечить устойчивое и эффективное производство молока в условиях Западной Сибири и других регионов России.

ЗАКЛЮЧЕНИЕ

В настоящей диссертационной работе впервые проведено комплексное региональное исследование генеалогической и генетической структуры чёрнопёстрого скота Западной Сибири, целенаправленно ориентированное на разработку научно-обоснованных основ для региональных программ крупномасштабной селекции. Для достижения поставленной цели был реализован принципиально новый подход, основанный на верификации и валидации данных первичного зоотехнического учета для получения несмещенных оценок и применении современных методов биоинформатики и математического моделирования. Ключевые выводы, полученные в ходе исследования, заключаются в следующем:

1. Разработана и практически внедрена уникальная методология верификации и валидации данных первичного зоотехнического учета, включающая многоступенчатые алгоритмы цифровой идентификации животных, методы многомерной статистики и сопоставление с международными базами данных. Особенность методологии заключается в комплексном подходе к обеспечению несмещенной оценки исходной информации, что позволило выявить и устранить критические ошибки в генеалогических базах данных племенных предприятий и повысить согласованность фенотипических сведений.

2. Выявлена значительная межхозяйственная вариабельность качества генеалогических данных в регионе, указывающая на децентрализацию племенного учета и дискретизацию региональной генеалогической информации. Это затрудняет получение комплексного несмещённого представления о генетической структуре популяции. Ранжирование хозяйств по качеству генеалогической информации позволило дифференцировать подходы к внедрению разработанной методологии верификации и валидации данных, ориентируясь на специфические потребности и уровень цифровизации региональных предприятий. Это позволяет формировать региональную популяцию скота с достаточным генеалогическим разнообразием для снижения негативных последствий возможного инбридинга.

3. Установлен высокий уровень генетического разнообразия чёрно-пёстрого скота Западной Сибири по STR-локусам микросателлитов на основе несмещённых генетических данных. Это свидетельствует о сохранении ценного генетического потенциала и адаптивности региональной популяции, несмотря на длительный процесс голштинизации. Выявленные отклонения от равновесия Харди-Вайнберга для ряда локусов (BM1824, BM2113, ETH10, ETH3, NRA023, TGLA122, TGLA227, TGLA53 и BM1818) обосновывают необходимость дальнейшего изучения региональной популяционной структуры для управления генетическим разнообразием в рамках программ крупномасштабной селекции.

4. Наблюдается умеренный уровень генетического разнообразия по SNP-локусам, ассоциированным с признаками молочной продуктивности, на основе несмещённых генетических оценок. Преобладание гомозигот и высокая частота мажорных аллелей для большинства SNP свидетельствует о селекционном давлении, направленном на увеличение молочной продуктивности. В то же время, наличие полиморфных SNP с высокой частотой минорных аллелей (rs42216771 (0,5), rs109752439 (0,48)) открывает перспективы для геномной селекции и выявления новых генетических маркеров, ассоциированных с хозяйственно-полезными признаками в региональных условиях.

5. Разработана математическая модель прогнозирования удоя чёрно-пёстрого скота Западной Сибири с учётом паратипических и генетических факторов, демонстрирующая высокую точность за счёт использования верифицированных данных первичного зоотехнического учёта. Модель позволяет количественно оценить вклад генетических и паратипических факторов (год и сезон отела, лактация, возраст, генетические особенности) в изменчивость удоя и предлагается в качестве инструмента для повышения точности несмещённой оценки племенной ценности и оптимизации селекционных программ в рамках отечественной крупномасштабной селекции.

ПРЕДЛОЖЕНИЯ ПРОИЗВОДСТВУ

Результаты исследований могут быть использованы для совершенствования селекционно-племенной работы с чёрно-пёстрым скотом в племенных хозяйствах и региональных информационно-селекционных центрах Западной Сибири. На основании полученных данных и сделанных выводов, предлагается следующий комплекс мер по внедрению результатов исследования в практику молочного скотоводства региона:

1. Внедрить разработанную методологию верификации и валидации данных первичного зоотехнического учета на всех племенных предприятиях региона. Обеспечить обязательное использование разработанного программного обеспечения для систематической проверки и очистки генеалогических и фенотипических баз данных. Разработать и внедрить единые региональные стандарты ведения племенного учета, обеспечивающие полноту, достоверность и унификацию данных, собираемых на различных предприятиях.
2. Внедрить в практику племенных предприятий разработанную математическую модель прогнозирования удоя. Провести адаптацию и калибровку модели для различных хозяйственно-климатических зон Западной Сибири. Рассмотреть возможность интеграции модели в региональную информационно-аналитическую систему племенных ресурсов.
3. Рассмотреть возможность введения в регионе категориальной индексной системы оценки племенной ценности животных, при которой прогноз по признакам отбора по одному быку-производителю будет дифференцироваться в зависимости от уровня производственной культуры и хозяйственно-климатических условий конкретного предприятия. Это позволит более адекватно оценивать племенную ценность животных в различных производственных средах и повысить эффективность селекции в региональном масштабе.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Алтухов, А.И. Развитие молочного скотоводства в стране: проблемы и возможные пути их решения / А.И. Алтухов // АПК: Экономика, управление. – 2018. – №9. – С. 86-92.
2. Амерханов, Х. Состояние и перспективы развития племенного животноводства в Российской Федерации / Х. Амерханов // Молочное и мясное скотоводство. – 2012. – №2. – С. 7-9.
3. Анищенко, А.Н. Территориальная дифференциация молочного скотоводства России / А.Н. Анищенко, Д.И. Усманов // Продовольственная политика и безопасность. 2021. – Т.8, № 2. – С. 179-188.
4. Бабина, С.И. Цифровые и информационные технологии в управлении предприятием: реальность и взгляд в будущее / С.И. Бабина // Креативная экономика. – 2019. – Т.13, №4. – С. 723-742.
5. Багиров, В. Генетические ресурсы животноводства / В. Багиров // Животноводство России. – 2008. – №2. – С. 10-12.
6. Балыкина, В.В. Использование межвидовых фенотипических дистанций Ив по морфометрических показателях листа для видовой диагностики / В.В. Балыкина // Молодой Вчений. – 2015. – №11-1(26). – С. 54-58.
7. Банникова, А.А. Молекулярные маркеры и современная филогенетика млекопитающих / А.А. Банникова // Журнал общей биологии. – 2004. – Т.65, №4. – С. 278-305.
8. Баранов, В.С. Генетический паспорт - основа индивидуальной и предиктивной медицины / В.С. Баранов, А.С. Глотов, Т.Э. Иващенко [и др.] – Санкт-Петербург: ООО «Издательство Н-Л», 2009. – 528 с.
9. Бегучев, А.П. Скотоводство / А.П. Бегучев, Т.И. Безенко, Л.Г. Боярский [и др.] – 3-е издание, переработанное и дополненное. – Москва: Агропромиздат, 1992. – 542 с.

10. Белоног, М.Г. Барабинский государственный племенной рассадник крупного рогатого скота Новосибирской области / М.Г. Белоног. – Новосибирск: Новосибирское областное государственное издательство, 1950. – 48 с.
11. БЕЛПЛЕМЖИВОБЪЕДИНЕНИЕ: Белорусское государственное объединение по племенному животноводству : [сайт]. Минск. URL: <https://www.belplem.by/> (дата обращения: 16.06.2024).
12. Березина, Т.И. Использование коров чёрно-пёстрой породы разных типов телосложения и кровности по голштинам для эффективного производства молока: 06.02.07 «Разведение, селекция и генетика сельскохозяйственных животных» : диссертация на соискание учёной степени кандидата сельскохозяйственных наук / Березина Т.И.; Самарская государственная сельскохозяйственная академия. – Кинель, 2016. – 143 с.
13. Богданова, О.В. Научно-теоретическое обоснование системы совершенствования селекционно-племенной работы в молочном скотоводстве Новосибирской области / О.В. Богданова, Е.В. Камалдинов, С.Г. Куликова [и др.] // Вестник НГАУ. – 2023. – №2(67). – С. 149-155.
14. Васильева, Е.Н. Современная генеалогическая структура ленинградской популяции молочного скота / Е.Н. Васильева, О.В. Тулинова // Молочное и мясное скотоводство. – 2022. – №5. – С. 30-35.
15. Васильева, Л.А. Статистические методы в биологии, медицине и сельском хозяйстве / Л.А. Васильева. – Новосибирск: Новосибирский государственный университет, 2007. – 124 с.
16. Воронкова, О.А. Продуктивные и воспроизводительные качества скота черно-пестрой породы различного происхождения: 06.02.10 «Частная зоотехния, технология производства продуктов животноводства» : диссертация на соискание учёной степени кандидата сельскохозяйственных наук / Воронкова О.А.; Федеральное государственное бюджетное образовательное учреждение высшего образования "Российский государственный аграрный университет - МСХА имени К.А. Тимирязева". – Москва, 2019. – 172 с.

17. Гареева, И.Т. Взаимосвязь полиморфных вариантов генов пролактина и β -лактоглобулина с молочной продуктивностью коров: 06.02.07 «Разведение, селекция и генетика сельскохозяйственных животных»: диссертация на соискание учёной степени кандидата биологических наук / Гареева И.Т.; Всероссийский научно-исследовательский институт генетики и разведения сельскохозяйственных животных. – Санкт-Петербург-Пушкин, 2012. – 117 с.
18. Гарт, В.В. Полиномиальная сопряжённая изменчивость признаков линейной оценки экстерьера и удоя высокопродуктивного голштинского скота / В.В. Гарт, С.Г. Куликова, О.В. Богданова [и др.] // Известия Тимирязевской сельскохозяйственной академии. – 2024. – №5. – С. 86-100.
19. Гарт, В.В. Характеристика линейного профиля первотёлок голштинской породы разной продуктивности в условиях промышленного комплекса / В.В. Гарт, О.В. Ефремова, С.Г. Куликова // Достижения науки и техники АПК. – 2024. – Т.38, №3. – С. 45-50.
20. Гашева, Н.А. Классификационно-диагностическая шкала рода *Salix* как возможность мониторинговых и таксационных эвм-тестирований / Н.А. Гашева // Вестник Оренбургского государственного университета. – 2006. – №4. – С. 23-27.
21. Гашева, Н.А. Математическое представление формы листа в исследованиях структуры биоразнообразия ив / Н.А. Гашева // Вестник Тверского государственного университета. Серия: Биология и экология. – 2008. – №9. – С. 42-46.
22. Гашева, Н.А. Межвидовые фенотипические дистанции ив по комплексам морфометрических признаков листа / Н.А. Гашева // Вестник экологии, лесоведения и ландшафтоведения. – 2008. – №8. – С. 116-123.
23. Генетическое расстояние [Электронный ресурс]: Википедия. Свободная энциклопедия. – URL: https://ru.wikipedia.org/wiki/Генетическое_расстояние (дата обращения: 03.09.2024).
24. Герчиков, Н.П. Крупный рогатый скот: учебное пособие для зоотехнических институтов и факультетов / Н.П. Герчиков. – Москва: Сельхозгиз, 1958. – 351 с.

25. Гладырь, Е.А. Изучение изменчивости микросателлитов при создании нового типа мясного скота Сибири / Е.А. Гладырь, Г.М. Гончаренко, П.В. Горелов [и др.] // Достижения науки и техники АПК. – 2011. – №10. – С. 30-32.
26. Гладырь, Е.А. Использование микросателлитов для характеристики аллелофонда популяций крупного рогатого скота Таджикистана / Е.А. Гладырь, Н.А. Зиновьева, В.А. Багиров [и др.] // Достижения науки и техники АПК. – 2012. – №8. – С. 58-62.
27. Гладырь, Е.А. Характеристика аллелофонда крупного рогатого скота некоторых мясных пород, разводимых на территории Южного Урала и Западной Сибири / Е.А. Гладырь, Н.А. Зиновьева, Д.Б. Косян [и др.] // Достижения науки и техники АПК. – 2013. – №3. – С. 61-63.
28. Глазко, Т.Т. ДНК-технологии для повышения мясной продуктивности / Т.Т. Глазко, А.Б. Комаров, Е.В. Борзаковская – Текст : непосредственный. // Известия Тимирязевской сельскохозяйственной академии. – 2008. – №1. – С. 75-80.
29. Говор, Е.М. Генетическая дивергенция самоопыленных линий кукурузы / Е.М. Говор, Л.П. Шиманский // Земледелие и селекция в Беларуси. – 2019. – №55. – С. 290-295.
30. Горелик, О. В. Оценка влияния голштинизации на продуктивные качества черно-пестрого скота / О.В. Горелик, С.Ю. Харлап, Н.А. Андрюшечкина // Вестник биотехнологии. – 2020. – №1(22). – С. 9.
31. Грашин, В.А. Молочная, мясная продуктивность и технологические свойства молока черно-пестроголштинских коров в зоне Поволжья: 06.02.04 «Частная зоотехния, технология производства продуктов животноводства» : диссертация на соискание учёной степени кандидата сельскохозяйственных наук / Грашин В.А.; Всероссийский научно-исследовательский институт племенного дела (ГНУ ВНИИплем). – п. Лесные Поляны, Московской области, 2002. – 135 с.
32. Гридина, С.Л. Выведение новых высокопродуктивных линий уральского типа черно-пестрого скота / С.Л. Гридина, О.И. Лешонок: в сборнике статей / «Вопросы повышения эффективности сельскохозяйственного производства на

Среднем Урале» // Уральский научно-исследовательский институт сельского хозяйства Россельхозакадемии. – Екатеринбург, 2003. – С. 342-345.

33. Денискова, Т.Е. Характеристика аллелофонда популяции симментальского и помесного скота Поволжья с использованием микросателлитов / Т.Е. Денискова, В.В. Волкова, О.В. Костюнина [и др.] // Вестник Алтайского государственного аграрного университета. – 2019. – №3(173). – С. 100-106.

34. Донник, И.М. Продуктивные характеристики коров в условиях интенсификации производства молока / И.М. Донник, О.С. Чеченихина. – Екатеринбург: Уральский государственный аграрный университет, 2024. – 316 с.

35. Евтеева, Е.В. Использование информационных технологий в системе управления предприятия / Е.В. Евтеева // Вестник Волжского университета им. В.Н. Татищева. – 2015. – №2(24). – С. 19-25.

36. Желтиков, А.И. Черно-пестрый скот Сибири / А.И. Желтиков, В.Л. Петухов, О.С. Короткевич [и др.] – Новосибирск: Общество с ограниченной ответственностью «Издательство Прометей», 2010. – 500 с.

37. Заварзин, А.В. Разработка и исследование иерархических агломеративных алгоритмов районирования / А.В. Заварзин, В.И. Месюра // Радиоэлектроника, информатика, управление. – 2003. – №1. – С. 76-81.

38. Зиннатов, Ф.Ф. Влияние некоторых полиморфных вариантов генов белкового и липидного обмена у джерсейской породы коров на показатели молочной продуктивности / Ф.Ф. Зиннатов, Т.М. Ахметов, С.В. Тюлькин [и др.] // Вестник биотехнологии и физико-химической биологии им. Ю.А. Овчинникова. – 2024. – Т.20, №2. – С. 68-75.

39. Зиннатов, Ф.Ф. Характерные особенности влияния генов белкового обмена на показатели качества молока коров джерсейской породы / Ф.Ф. Зиннатов, С.В. Тюлькин, Т.М. Ахметов [и др.] // Ученые записки Казанской государственной академии ветеринарной медицины им. Н.Э. Баумана. – 2024. – Т.257, №1. – С. 98-104.

40. Зиновьева, Н.А. Генетические ресурсы животных: развитие исследований аллелофонда Российских пород крупного рогатого скота – миниобзор / Н.А.

Зиновьева, А.А. Сермягин, А.В. Доцев [и др.] // Сельскохозяйственная биология. – 2019. – Т.54, №4. – С. 631-641.

41. Зиновьева, Н.А. Роль ДНК-диагностики в контроле и элиминации рецессивных наследственных аномалий у сельскохозяйственных животных / Н.А. Зиновьева, Е.А. Гладырь, О.В. Костюнина [и др.] // Достижения науки и техники АПК. – 2012. – №11. – С. 37-40.

42. Казаровец, Н.В. Организационно-методические приемы, используемые при формировании оптимальной структуры белорусской популяции черно-пестрого скота / Н.В. Казаровец, И.П. Шейко, Т.В. Павлова [и др.] // Весці нацыянальнай акадэміі навук Беларусі. Серыя аграрных навук. – 2019. – Т.57, №4. – С. 454–469.

43. Калашникова, Л.А. Селекция XXI века: использование ДНК-технологий / Л.А. Калашникова, И.М. Дунин, В.И. Глазко. – Лесные Поляны: Всероссийский НИИ племенного дела, 2000. – 31 с.

44. Калашникова, Л.А. Генетическая характеристика крупного рогатого скота с использованием микросателлитов / Л.А. Калашникова, Я.А. Хабибрахманова, Т.Б. Ганченкова [и др.] // Зоотехния. – 2016. – №2. – С. 9-11.

45. Камалдинов, Е.В. Современные проблемы животноводства / Е.В. Камалдинов, А.Ф. Петров: в сборнике статей / «Роль аграрных вузов в решении задач биологизации сельского хозяйства» // Воронежский государственный аграрный университет им. Императора Петра I. – Воронеж, 2021. – С. 53-55.

46. Камалдинов, Е.В. Генофонд пород крупного рогатого скота и свиней Западной Сибири: 06.02.07 «Разведение, селекция и генетика сельскохозяйственных животных», 03.01.04 «Биохимия»: диссертация на соискание учёной степени доктора биологических наук / Камалдинов Е.В.; ФГБОУ ВПО «Новосибирский государственный аграрный университет». – Новосибирск, 2013. – 445 с.

47. Камалдинов, Е.В. Достоверность данных первичного зоотехнического учёта в молочном скотоводстве / Е.В. Камалдинов, А.Ф. Петров, К.С. Шатохин [и др.] // Вестник НГАУ. – 2022. – №2(63). – С. 76-83.

48. Камалдинов, Е.В. Межпородные фенотипические дистанции по комплексу биохимических показателей у свиней Западной Сибири / Е.В. Камалдинов //

Вестник Алтайского государственного аграрного университета. – 2011. – №1(75). – С. 59-63.

49. Кантор В. Методы кластеризации: K-Means, агломеративная кластеризация, DBSCAN. Оценка качества кластеризации [Электронный ресурс]: Яндекс образование. - 2024. - URL: <https://education.yandex.ru/handbook/ml/article/klasterizaciya> (дата обращения: 28.10.2024).

50. Каратаева, М.Б. Характеристика генофонда крупного рогатого скота казахской белоголовой породы по микросателлитным ДНК: в сборнике статей / М.Б. Каратаева, Ж.М. Хамзина, Ж.А. Байдилдинова [и др.] // Сборник научных трудов по итогам III международной научно-практической конференции «Актуальные вопросы сельскохозяйственных наук в современных условиях развития страны». – Санкт-Петербург: Инновационный центр развития образования и науки, 2017. – С. 12-15.

51. Ким, Дж. О. Факторный, дискриминантный и кластерный анализ / Дж. О. Ким, Ч.У. Мьюллер, У.Р. Клекка [и др.] – Москва: Финансы и статистика, 1989. – 215 с.

52. Клименок, И.И. Продуктивные и племенные качества черно-пестрого скота сибирского региона / И.И. Клименок, С.Б. Яранцева, М.А. Шишкина // Генетика и разведение животных. – 2014. – №2. – С. 30-33.

53. Климов, Н.Н. Эффективность использования голштинизированного черно-пестрого скота для производства молока / Н.Н. Климов, С.И. Коршун // Молодой ученый. – 2016. – №24(128). – С. 143-146.

54. Ковальчук, Д.А. Эффективность использования голштинизированного скота черно-пестрой и бестужевской пород при поточно-цеховой технологии производства молока: 06.02.04 «Частная зоотехния, технология производства продуктов животноводства»: диссертация на соискание учёной степени кандидата сельскохозяйственных наук / Ковальчук Д.А.; Оренбургский государственный аграрный университет. – Оренбург, 2007. – 184 с.

55. Колчанов, Н.А. Интеграция генных сетей, контролирующих физиологические функции организма / Н.А. Колчанов, О.А. Подколотная, Е.В. Игнатьева [и др.] // Информационный вестник ВОГИС. – 2005. – Т.9, №2. – С. 179-198.

56. Кольцов, Д. Н. Характеристика аллелофонда сычевской породы крупного рогатого скота по ДНК микросателлитам / Д.Н. Кольцов, В.В. Волкова, Е.А. Гладырь [и др.] // Достижения науки и техники АПК. – 2012. – №8. – С. 56-57.
57. Костомахин, Н.М. Совершенствование скота черно-пестрой породы различного происхождения: монография / Н.М. Костомахин, О.А. Воронкова, М.А. Габедава. – Калуга: Индивидуальный предприниматель Стрельцов Илья Анатольевич, 2023. – 122 с.
58. Костомахин, Н.М. Породы крупного рогатого скота: учебное пособие для студентов высших учебных заведений, обучающихся по специальности «Зоотехния» / Н.М. Костомахин. – Москва: КолосС, 2011. – 118 с.
59. Костюнина, О.В. Влияние генотипов по ДНК-маркерам на показатели молочной продуктивности коров черно-пестрой породы / О.В. Костюнина, А.М. Бакай, Г.А. Бушева [и др.] // Достижения науки и техники АПК. – 2011. – № 10. – С. 33-34.
60. Косяченко, Н.М. Голштинская порода в создании улучшенных генотипов и внутрипородных типов крупного рогатого скота: монография / Н.М. Косяченко, М.В. Абрамова, А.В. Ильина [и др.] – Ярославль: ООО «Канцлер», 2020. – 157 с.
61. Крамаренко, Н.М. Опыт племенной работы с крупным рогатым скотом черно-пестрой породы / Н.М. Крамаренко, Н.П. Семенов, Л.К. Эрнст. – Москва: Россельхозиздат, 1965. – 80 с.
62. Красота, В.Ф. Разведение сельскохозяйственных животных: учебное пособие для студентов высших учебных заведений, обучающихся по специальности «Зоотехния» / В.Ф. Красота, Т.Г. Джапаридзе, Н.М. Костомахин; 5-е издание, переработанное и дополненное. – Москва: КолосС, 2005. – 423 с.
63. Крюков, В.И. ДНК-диагностика в селекции крупного рогатого скота / В.И. Крюков, О.А. Шалимова, Н.Г. Друшляк [и др.] // Вестник Орловского государственного аграрного университета. – 2012. – №1(34). – С. 62-67.
64. Кудинов, А.А. Модель геномной оценки племенной ценности молочного скота Ленинградской области: 06.02.07 «Разведение, селекция и генетика сельскохозяйственных животных»: диссертация на соискание учёной степени

- кандидата биологических наук / Кудинов А.А.; Всероссийский научно-исследовательский институт племенного дела. – Лесные Поляны, 2019. – 124 с.
65. Кузнецов, В.М. Методы Нея для анализа генетических различий между популяциями / В.М. Кузнецов // Проблемы биологии продуктивных животных. – 2020. – №1. – С. 91-110.
66. Куликова, С.Г. Связь между признаками линейной оценки экстерьера голштинских коров: в сборнике трудов конференции / С.Г. Куликова, О.В. Богданова, В.В. Гарт [и др.] // VII национальная (всероссийская) научная конференция с международным участием «Теория и практика современной аграрной науки». – Новосибирск: ФГБУ ВО «Новосибирский ГАУ», 2024. – С. 389-392.
67. Луконина, О.Н. Ежегодник по племенной работе в молочном скотоводстве в хозяйствах Российской Федерации (2023 год) / О.Н. Луконина, Г.Ф. Сафина, В.В. Чернов [и др.] – Москва: Издательство ФГБНУ ВНИИплем, 2024. – 242 с.
68. Ляшук, Р.Н. К вопросу о голштинизации черно-пестрого скота в Орловской области / Р.Н. Ляшук, А.И. Шендаков, М.В. Востров [и др.] // Вестник Орловского государственного университета. – 2007. – №1(4). – С. 26-28.
69. Мартынов, Н.А. Молекулярно-генетическая оценка распространённости дефицита фактора свёртываемости крови в уральской популяции голштинизированного чёрно-пёстрого скота / Н.А. Мартынов, М.В. Модоров, О.В. Соколова [и др.] // Проблемы биологии продуктивных животных. – 2022. – №1. – С. 54-61.
70. Меркурьева, Е.К. Биометрия в селекции и генетике сельскохозяйственных животных / Е.К. Меркурьева. – Москва: Колос, 1970. – 424 с.
71. Меркурьева, Е.К. Генетические основы селекции в скотоводстве / Е.К. Меркурьева. – Москва: Колос, 1977. – 239 с.
72. Модоров, М.В. Генетическая структура популяции голштинизированного черно-пестрого скота на территории Урала / М.В. Модоров, И.В. Ткаченко, А.А. Грин [и др.] // Генетика. – 2021. – Т.57, №4. – С. 437-444.

73. Моисеева, И.Г. Генофонды сельскохозяйственных животных. Генетические ресурсы животноводства России / И.Г. Моисеева, С.В. Уханов, Ю.А. Столповский [и др.] – Москва: Федеральное государственное унитарное предприятие «Академический научно-издательский, производственно-полиграфический и книгораспространительский центр «Наука», 2006. – 462 с.
74. Мороз, М.Т. Развитие кадрового потенциала молочного животноводства в условиях цифровой экономики России / М.Т. Мороз, А.Б. Люлин, Е.Н. Тюренкова // Современное образование: содержание, технологии, качество. – 2018. – Т.2. – С. 126-128.
75. Мороз, М.Т. Совершенствование управления животноводством на основе информационных технологий: в сборнике трудов конференции / М.Т. Мороз, Е.Н. Тюренкова, В.С. Грачёв // Сборник научных трудов по материалам международной научно-практической конференции, посвященной году науки и технологий «Приоритеты развития АПК в условиях цифровизации и структурных изменений национальной экономики». – Санкт-Петербург: Санкт-Петербургский государственный аграрный университет, 2021. – С. 107-112.
76. Мымрин, В.С. Использование геномных индексов для отбора быков-производителей / В. Мымрин, О. Ткачук, Н.Е. Шавшукова // Молочное и мясное скотоводство. – 2012. – №3. – С. 4-7.
77. Мымрин, В.С. Результаты голштинизации черно-пестрого скота в уральском регионе / В.С. Мымрин, С.Л. Гридина, В.Ф. Гридин // Генетика и разведение животных. – 2014. – №2. – С. 17-20.
78. Нардид, А.В. Селекционные аспекты совершенствования коров черно-пестрой породы по продуктивным и технологическим признакам в условиях промышленных ферм: 06.02.07 «Разведение, селекция и генетика сельскохозяйственных животных», 06.02.10 «Частная зоотехния, технология производства продуктов животноводства»: диссертация на соискание учёной степени кандидата сельскохозяйственных наук / Нардид А.В.; Всероссийский научно-исследовательский институт животноводства. – п. Дубровицы Московской области, 2011. – 120 с.

79. Нардин, Д. С. Функциональные возможности программного продукта для зоотехнического и племенного учета «Селэкс - Молочный» / Д.С. Нардин, А.И. Малинина // Электронный научно-методический журнал Омского ГАУ. – 2015. – №3(3). – С. 10-15.
80. Нацгенофонд: база данных национального генофонда крупного рогатого скота молочного направления продуктивности / ФГБНУ «Всероссийский научно-исследовательский институт племенного дела», 2024 : [база данных], данные в формате XLSX. URL: https://vniiplem.ru/l/gisc/bd_ng/db_ng_mol_krs/ (дата обращения: 16.11.2024).
81. Николаенко, О. Генеалогическая структура стада, племенная ценность и возрастные параметры стада крупного рогатого скота голштинской породы: в сборнике статей по материалам конференции / О. Николаенко, Н. Брехин, Р. Епифанов // Всероссийская национальная научно-практическая конференция для студентов, аспирантов и молодых ученых «Перспективные разработки молодых ученых в области производства и переработки сельскохозяйственной продукции». – Ставрополь: ФГБОУ ВО «Ставропольский ГАУ», 2021. – С. 58-63.
82. Нурбаев, С.Д. Описание аллелофонда крупного рогатого скота Жетысуского типа породы санта-гертруда по микросателлитным ДНК / С.Д. Нурбаев, А.М. Омбаев, Т.Н. Карымсаков [и др.] // Известия национальной академии наук республики Казахстан. 2015. – Т.6, №30. – С. 18-22.
83. Панов, Б.Л. Проблемы селекции сельскохозяйственных животных / Б.Л. Панов, В.Л. Петухов, Л.К. Эрнст [и др.] – Новосибирск: Сибирское издательско-полиграфическое и книготорговое предприятие «Наука» РАН, 1997. – 283 с.
84. Паронян, И.А. Сохранение и использование отечественного генофонда животных важнейшая задача животноводства России / И.А. Паронян, О.П. Юрченко, С.А. Шабанова [и др.] // Достижения науки и техники АПК. – 2010. – №4. – С. 70-71.
85. Патюков, А.Г. Статистические методы анализа в биологии и медицине / А.Г. Патюков. – Омск: Вариант-Омск, 2012. – 230 с.

86. Петров, А.Ф. Генетическая структура скота сибирского отродья по микросателлитным локусам / А.Ф. Петров, Е.В. Камалдинов // Вестник НГАУ. – 2024. – №3(72). – С. 230-239.
87. Петров, А.Ф. Моделирование объема туловища по линейным признакам скота ирменского типа / А.Ф. Петров, Е.В. Камалдинов, О.Д. Панферова [и др.] // Сибирский вестник сельскохозяйственной науки. – 2021. – Т.50, №6(227). – С. 106-114.
88. Петухов, В.Л. Генетические основы селекции животных / В.Л. Петухов, Л.К. Эрнст, И.И. Гудилин [и др.] – Москва: Росагропромиздат, 1989. – 448 с.
89. Племяшов, К. Геномная селекция - будущее животноводства / К. Племяшов // Животноводство России. – 2014. – № 5. – С. 2-4.
90. Половикова, О.Н. Использование евклидова и манхэттенского расстояний в качестве меры близости для решения задачи классификации / О.Н. Половикова, В.В. Фокина // Известия Алтайского государственного университета. – 2010. – №1-1(65). – С. 101-102.
91. Попов, П.П. Кластерные группы, структура и относительное положение популяций ели на востоке Европы / П.П. Попов // Экология. – 2015. – №2. – С. 95-102.
92. Попов, П.П. Фенотипическое положение популяций ели в европейско-сибирской части ареала / П.П. Попов // Вестник экологии, лесоведения и ландшафтоведения. – 2008. – №8. – С. 108-115.
93. Прожерин, В.П. Подходы в оценке и сохранении генофонда отечественной холмогорской породы крупного рогатого скота в Архангельской области / В.П. Прожерин, И.В. Селькова // Вестник российской сельскохозяйственной науки. – 2024. – №5. – С. 101-104.
94. Прохоренко, П.Н. Межпородное скрещивание в молочном скотоводстве / П. Н. Прохоренко, Ж. Г. Логинов. – Москва: Россельхозиздат, 1986. – 191 с.
95. Прохоренко, П.Н. Генетические ресурсы крупного рогатого скота Российской Федерации / П.Н. Прохоренко, И.А. Паронян // Розведення і генетика тварин. – 2006. – №40. – С. 141-151.

96. Прохоренко, П.Н. Прошлое, настоящее и будущее генетики и селекции в животноводстве / П.Н. Прохоренко // Зоотехния. – 2008. – №1. – С. 8-10.
97. Республиканская система животноводства: Информационно-аналитическая система : [сайт]. Астана. URL: <http://www.plem.kz/> (дата обращения: 17.05.2024). Режим доступа: для зарегистрированных пользователей.
98. Решение Коллегии Евразийской экономической комиссии от 02.06.2020 №74 «Об утверждении Положения о проведении молекулярной генетической экспертизы племенной продукции государств – членов Евразийского экономического союза» // Официальный сайт Евразийского экономического союза. URL: <https://docs.eaeunion.org/documents/247/5175/> (дата обращения: 21.05.2024).
99. Решение Коллегии Евразийской экономической комиссии от 08.09.2020 №108 «Об утверждении Порядка определения породы (породности) племенных животных» // Официальный сайт Евразийского экономического союза. URL: <https://docs.eaeunion.org/documents/247/5306/> (дата обращения: 22.06.2024).
100. Российская Федерация. Президент. О Стратегии научно-технологического развития Российской Федерации : указ Президента Российской Федерации от 28.02.2024 №145 // ГАРАНТ: информационно-правовая система. – URL: <https://base.garant.ru/408618353/> (дата обращения: 01.05.2024).
101. Руденко, О.В. Влияние кровности по голштинской породе на продуктивное долголетие и пожизненную молочную продуктивность чёрно-пёстрых коров / О.В. Руденко, С.П. Еремин // Вестник Ульяновской государственной сельскохозяйственной академии. – 2015. – №2(30). – С. 132-136.
102. Ружевский, А.Б. Черно-пестрый скот. – Москва: Сельхозгиз, 1959. – 336 с.
103. Рукин, И.В. Геномная селекция - будущее в разведении животных / И.В. Рукин, Е.С. Пантюх, Д.С. Груздев // Зоотехния. – 2013. – №7. – С. 8-9.
104. Салах, Э.С.Р.А.В. Результаты использования поглотительного скрещивания черно-пестрого скота с производителями голштинской породы: 06.02.01 «Разведение, селекция, генетика и воспроизводство сельскохозяйственных животных»: диссертация на соискание учёной степени кандидата

сельскохозяйственных наук / Салах Э.С.Р.А.В.; Всероссийский научно-исследовательский институт генетики и разведения сельскохозяйственных животных. – Санкт-Петербург-Пушкин, 1993. – 123 с.

105. Санников, С.Н. Генетическая и фенотипическая дифференциация притобольских и европейских популяций *Calluna vulgaris* (L.) Hull / С.Н. Санников, И.В. Петрова, О.С. Дымшакова [и др.] // Генетика. – 2014. – Т.50, №9. – С. 1050.

106. Санова, З.С. Конституционально-продуктивные особенности популяции голштинизированного скота в Калужской области / З.С. Санова, Н.Е. Джумаева // Эффективное животноводство. – 2022. – №1(176). – С. 82-84.

107. Свидетельство о государственной регистрации базы данных №2024625179 Российская Федерация. «База данных генеалогической структуры быков голштинской породы канадского происхождения»: №2024624918 : заявл. 01.11.2024 : опубл. 14.11.2024 / А.Ф. Петров, Е.В. Камалдинов, К.Н. Нарожных; заявитель и правообладатель ФГБУ ВО Новосибирский ГАУ.

108. Свидетельство о государственной регистрации базы данных №2024625178 Российская Федерация. «Генетическое разнообразие скота породы Герефорд Западной Сибири по микросателлитным локусам»: №2024624910 : заявл. 01.11.2024 : опубл. 14.11.2024 / М.А. Барсукова, А.Ф. Петров, К.Н. Нарожных; заявитель и правообладатель ФГБУ ВО Новосибирский ГАУ.

109. Свидетельство о государственной регистрации базы данных №2024622398 Российская Федерация. «Комплексная линейная оценка вымени популяции скота канадской селекции»: №2024622072 : заявл. 16.05.2024 : опубл. 31.05.2024 / А.Ф. Петров, Е.В. Камалдинов, К.Н. Нарожных [и др.]; заявитель и правообладатель ФГБУ ВО Новосибирский ГАУ.

110. СЕЛЭКС. Молочный скот [Электронный ресурс] : информационно-аналитическая система / разработчик: Региональный центр информационного обеспечения племенного животноводства «ПЛИНОР». – Санкт-Петербург : РЦ «ПЛИНОР», 2024. – URL: <https://plinor.ru/selexdairycattle> (дата обращения: 01.03.2024).

111. Сермягин, А.А. Перспективы использования оценки геномной племенной ценности в селекции молочного скота / А.А. Сермягин, Е.Н. Нарышкина, Т.В. Карпушкина [и др.] // Молочное и мясное скотоводство. – 2015. – №7. – С. 2-5.
112. Скворцова, Е.Г. Влияние доли кровности по голштинской породе на продуктивное долголетие черно-пестрого скота / Е.Г. Скворцова // Вестник биотехнологии. – 2020. – №1(22). – С. 15.
113. Смарагдов, М.Г. 9-й Всемирный конгресс по прикладной генетике животных / М.Г. Смарагдов // Генетика. – 2011. – Т.47, №5. – С. 715-717.
114. Смарагдов, М.Г. Геномная селекция молочного скота в мире. Пять лет практического использования / М.Г. Смарагдов // Генетика. – 2013. – Т.49, №11. – С. 1251.
115. Смарагдов, М.Г. Полногеномный анализ межстадной F_{st} - гетерогенности голштинизированного скота / М.Г. Смарагдов, Е.И. Сакса, А.А. Кудинов [и др.] // Генетика. – 2016. – Т.52, №2. – С. 198.
116. Смарагдов, М.Г. Тотальная геномная селекция с помощью SNP как возможный ускоритель традиционной селекции / М.Г. Смарагдов // Генетика. – 2009. – Т.45, №6. – С. 725-728.
117. Смирнова, Ю.М. Влияние генотипа на долголетие и пожизненную продуктивность коров / Ю.М. Смирнова, А.В. Платонов // Агробиотехника. – 2019. – Т.2, №3. – С. 1.
118. Снедекор, Д.У. Статистические методы в применении к исследованиям в сельском хозяйстве и биологии / Д.У. Снедекор. – Москва: Сельхозиздат, 1961. – 503 с.
119. Солошенко, В.А. О возможности использования генетических маркеров в селекции мясного скота для повышения качественных показателей мяса / В.А. Солошенко, Г.И. Гончаренко, А.А. Дворяткин [и др.] // Вестник мясного скотоводства. – 2013. – №1(79). – С. 37-41.
120. Столповский, Ю.А. Популяционно-генетические основы сохранения генофондов domesticированных видов животных / Ю.А. Столповский // Вавиловский журнал генетики и селекции. – 2013. – Т.17, №4-2. – С. 900-915.

121. Сулимова, Г.Е. ДНК-маркеры в генетических исследованиях: типы маркеров, их свойства и области применения / Г.Е. Сулимова // Успехи современной биологии. – 2004. – Т.124, №3. – С. 260-271.
122. Тюренкова, Е.Н. Управление племенным животноводством в регионе с применением информационных технологий / Е.Н. Тюренкова // Техника и оборудование для села. – 2012. – №10. – С. 41-45.
123. Тяпугин, С.Е. Племенные и продуктивные качества молочного скота Вологодского типа / С.Е. Тяпугин, Г.А. Симонов, В.С. Зотеев [и др.] // Горное сельское хозяйство. – 2020. – №1. – С. 96-101.
124. Уханов, С.В. Генетические ресурсы крупного рогатого скота: редкие и исчезающие отечественные породы / С.В. Уханов, Ю.А. Столповский, Л.В. Банникова. – Москва: Наука, 1993. – 169 с.
125. ФГИАС ПР (Федеральная государственная информационно-аналитическая система племенных ресурсов) [электронный ресурс] // URL: <https://www.tadviser.ru/index.php/>
- Продукт:ФГИАС_ПР_(Федеральная_государственная_информационно-аналитическая_система_племенных_ресурсов) (дата обращения: 22.02.2024).
126. Фенченко, Н.Г. История создания и генеалогия черно-пестрой породы крупного рогатого скота / Н.Г. Фенченко, Н.И. Хайруллина, Ф.Х. Сиразетдинов. – Уфа: БНИИСХ, 2002. – 332 с.
127. Харисова, Ч.А. Генеалогическая структура татарстанской популяции черно-пестрого скота по принадлежности к перспективным ветвям / Ч.А. Харисова, Р.Р. Шайдуллин, Т.М. Ахметов // Известия Кабардино-Балкарского государственного аграрного университета им. В.М. Кокова. – 2022. – №2(36). – С. 61-67.
128. Чижова, Л.Н. Межпородные особенности полиморфизма генов соматотропин, пролактин у коров молочного направления продуктивности: в сборнике трудов / Л.Н. Чижова, Е.С. Суржикова, Г.П. Ковалёва [и др.] // Сборник научных трудов всероссийского научно-исследовательского института овцеводства и козоводства. – 2017. – Т.2, №10. – С. 108-113.

129. Шабунин, Л.А. Влияние голштинизации на количество и качество молочной продуктивности коров черно-пестрой породы / Л.А. Шабунин, В.Г. Кахикало, О.Г. Назарченко // Вестник Красноярского государственного аграрного университета. – 2015. – №5(104). – С. 164-167.
130. Шайдуллин, Р.Р. Генетический потенциал черно-пестрого скота разных линий и ветвей / Р.Р. Шайдуллин, Ч.А. Харисова, Т.М. Ахметов // Агробиотехнологии и цифровое земледелие. – 2022. – №3(3). – С. 53-57.
131. Шайхутдинова, А.А. Методы оценки биоразнообразия: методические указания. Методы оценки биоразнообразия / А.А. Шайхутдинова. – Оренбург: Оренбургский государственный университет, 2019. – 37 с.
132. Шиманский, Л.П. Генетическая дивергенция самоопыленных линий подсолнечника / Л.П. Шиманский, О.А. Туровец, Е.М. Говор // Земледелие и селекция в Беларуси. – 2021. – №57. – С. 434-442.
133. Эрнст, Л.К. Биологические проблемы животноводства в XXI веке: монография / Л.К. Эрнст, Н.А. Зиновьева. – Москва: Всероссийский научно-исследовательский институт животноводства имени академика Л.К. Эрнста, 2008. – 508 с.
134. Эрнст, Л.К. Изучение влияния прилития крови голштинского скота на изменение генофонда крупного рогатого скота отечественных пород с использованием ДНК-микросателлитов / Л.К. Эрнст, Н.А. Зиновьева, Е.Н. Коновалова [и др.] // Зоотехния. – 2007. – №12. – С. 2-4.
135. Эрнст, Л. К. Черно-пестрая порода – золотой фонд молочного скотоводства страны / Л.К. Эрнст, А.П. Калашникова, Н.Г. Дмитриева [и др.] // Зоотехния. – 1990. – №2. – С. 2-3.
136. Юшкова, И.В. Влияние кровности по улучшающей породе на продолжительность и эффективность использования коров / И.В. Юшкова, М.Ю. Петрова, С.В. Борисенко // Вестник Омского государственного аграрного университета. – 2016. – №4(24). – С. 58-64.

137. Яковлев, А.Ф. Геномная селекция и прогнозирование качества потомства животных / А.Ф. Яковлев // Вестник Российской академии наук. – 2018. – Т.88, №10. – С. 946-950.
138. Aguilar, I. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of holstein final score / I. Aguilar, I. Misztal, D. L. Johnson [et al.] // Journal of Dairy Science. – 2010. – V.93, №2. – P. 743-752.
139. Akanno, E. Genomic prediction of breed composition and heterosis effects in angus, charolais and hereford crosses using 50k genotypes / E. Akanno, L. Chen, M. Abo-Ismael [et al.] // Canadian Journal of Animal Science. – 2017. – V.97. – P. 431-438.
140. Altukhov, Iu. P. DNA polymorphism in population genetics / Iu.P. Altukhov, E.A. Salmenkova // Genetika. – 2002. – V.38, №9. – P. 1173-1195.
141. Barendse, W. A genetic linkage map of the bovine genome / W. Barendse, S.M. Armitage, L.M. Kossarek [et al.] // Nature Genetics. – 1994. – V.6, №3. – P. 227-235.
142. Bekele, R. Genomic regions and candidate genes associated with milk production traits in holstein and its crossbred cattle: a review / R. Bekele, M. Taye, G. Abebe [et al.] // International Journal of Genomics. – 2023. – V.2023, №1. – P. 8497453.
143. Bishop, M.D. A genetic-linkage map for cattle / M.D. Bishop, S.M. Kappes, J.W. Keele [et al.] // Genetics. – 1994. – V.136, №2. – P. 619-639.
144. Bourdon, R.M. Understanding animal breeding / R.M. Bourdon. – 2nd ed. – NJ: Pearson Higher Education & Professional Group, 2014. – 520 p.
145. Bovine Genome Sequencing Project (BGSP) - Genome assembly Btau_3.1 / Human Genome Sequencing Center Baylor College of Medicine, 2024 : [база данных], datasets in JSON. URL: https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000003205.1/ (дата обращения: 11.07.2024).
146. Buitenhuis, B. Estimation of genetic parameters and detection of chromosomal regions affecting the major milk proteins and their post translational modifications in danish holstein and danish jersey cattle / B. Buitenhuis, N.A. Poulsen, G. Gebreyesus [et al.] // BMC Genetics. – 2016. – V.17, №1. – P. 114-125.

147. Buitenhuis, B. Estimation of genetic parameters and detection of quantitative trait loci for minerals in danish holstein and danish jersey milk / B. Buitenhuis, N.A. Poulsen, L.B. Larsen [et al.] // BMC Genetics. – 2015. – V.16, №1. – P. 52-59.
148. Burstin, J. Relationship between phenotypic and marker distances: theoretical and experimental investigations / J. Burstin, A. Charcosset // Heredity. – 1997. – V.79, №5. – P. 477–483.
149. Canadian Dairy Network : [сайт]. - URL: <https://www.cdn.ca/home.php> (дата обращения: 14.04.2024).
150. Council on Dairy Cattle Breeding (CDCB) : [сайт]. - URL: <https://uscddb.com> (дата обращения: 27.02.2024).
151. Christensen, O. Genomic prediction when some animal are not genotyped / O. Christensen, M.S. Lund // Genetics Selection Evolution. – 2010. – V.42. – P. 2-9.
152. Coghetto, R. Chebyshev Distance / R. Coghetto // Formalized Mathematics. – 2016. – V.24, №2. – P. 121-141.
153. Cole, J.B. Additional file 5 / J.B. Cole, G.R. Wiggans, L. Ma [et al.] [электронный ресурс] : ResearchGate. - 2011. - URL: https://www.researchgate.net/publication/293933955_Additional_file_5 (дата обращения: 15.09.2024).
154. Cole, J.B. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary u.s. holstein cows / J.B. Cole, G.R. Wiggans, L. Ma [et al.] // BMC Genomics. – 2011. – V.12, №1. – P. 408-424.
155. Cruz, A. Single-nucleotide polymorphism variations associated with specific genes putatively identified enhanced genetic predisposition for 305-day milk yield in the girolando crossbreed / A. Cruz, D.C. Silva, L.B. Minasi [et al.] // Frontiers in Genetics. – 2021. – V.11. – A. 573344.
156. Cui, X. Transcriptional profiling of mammary gland in holstein cows with extremely different milk protein and fat percentage using RNA sequencing / X. Cui, Y. Hou, S. Yang [et al.] // BMC Genomics. – 2014. – V.15, №1. – P. 226-240.
157. Czerneková, V. Genetic diversity between seven central european cattle breeds as revealed by microsatellite analysis / V. Czerneková, T. Kott, Dudkova [et al.] // Czech Journal of Animal Science. – 2006. – V.51, №1. – P. 1-7.

158. Darvishzadeh, R. Phenotypic and molecular marker distance as a tool for prediction of heterosis and F1 performance in sunflower (*helianthus annuus* L.) under well-watered and water-stressed conditions / R. Darvishzadeh // Australian Journal of Crop Science. – 2012. – V.6, №4. – P. 732-738.
159. Doormaal, B.V. Current status on haplotypes and genetic recessives / B.V. Doormaal, L. Beavers [Электронный ресурс]: Canadian Dairy Network. - 2019. - URL: <https://www.cdn.ca/document.php?id=522> (дата обращения: 01.05.2024).
160. Dugué, M. Genomic correlations between the gaits of young horses measured by accelerometry and functional longevity in jumping competition / M. Dugué, B.D.S. Priest, H. Crichan [et al.] // Frontiers in Genetics. – 2021. – V.12. – A. 619947.
161. Edwards, A.W.F. Distances between populations on the basis of gene frequencies / A.W.F. Edwards // Biometrics. – 1971. – V. 27, №4. – P. 873-881.
162. Eurogenomics: A European network for a reliable cattle breeding : [сайт]. - URL: <https://www.eurogenomics.com> (дата обращения: 15.08.2024).
163. European Federation for Animal Science (EAAP): Cattle Study Commission: [сайт]. - URL: <https://eaap.org/study-commissions/cattle> (дата обращения: 07.09.2024).
164. Fan, J.B. Highly parallel SNP genotyping / J.B. Fan, A. Oliphant, R. Shen [et al.] // Cold Spring Harbor Symposia on Quantitative Biology. – 2003. – V.68. – P. 69-78.
165. Fang, M. A multiple-SNP approach for genome-wide association study of milk production traits in chinese holstein cattle / M. Fang, W. Fu, D. Jiang [et al.] // PLOS One. – 2014. – V.9. №8. – P. e99544.
166. Faraj, S.H. DGAT1 gene polymorphism and its relationships with cattle milk yield and chemical composition / S.H. Faraj, A. Yheia, D.K. Seger // Periodico Tche Quimica. – 2020. – V.17, №35. – P. 174-180.
167. Firebird : официальный сайт / Firebird Foundation : [сайт]. - URL: <https://firebirdsql.org/> (дата обращения: 09.01.2024).
168. Gałeccki, A. Linear mixed-effects models using R: a step-by-step approach / A. Gałeccki, T. Burzykowski. – New York: Springer Science & Business Media, 2013. – 542 p.

169. Gebreyesus, G. Combining multi-population datasets for joint genome-wide association and meta-analyses: The case of bovine milk fat composition traits / G. Gebreyesus, A.J. Buitenhuis, N.A. Poulsen [et al.] // *Journal of Dairy Science*. – 2019. – V.102, №12. – P. 11124–11141.
170. Georges, M. Microsatellite mapping of a gene affecting horn development in *Bos taurus* // *Nature Genetics*. – 1993. – V.4, №2. – P. 206-210.
171. Gillespie, J.H. *Population genetics: a concise guide* / J.H. Gillespie. – Baltimore, Md: The Johns Hopkins University Press, 2004. – 214 p.
172. Goldstein, D.B. An evaluation of genetic distance for use with microsatellite loci / D.B. Goldstein, A.R. Linares, L.L. Cavalli-Sforza [et al.] // *Genetics*. – 1994. – V.139, №1. – P. 463-471.
173. González-José, R. Patterns of phenotypic covariation and correlation in modern humans as viewed from morphological integration / R. González-José, S. Van der Molen, E. González-Pérez [et al.] // *American Journal of Physical Anthropology*. – 2004. – V.123, №1. – P. 69-77.
174. Granato, I.S.C. snpReady: a tool to assist breeders in genomic analysis / I.S.C. Granato, G. Galli, E.G. de Oliveira Couto [et al.] // *Molecular Breeding*. – 2018. – V.38, №8. – A. 102.
175. Grover, A. Development and use of molecular markers: past and present / A. Grover, P.C. Sharma // *Critical Reviews in Biotechnology*. – 2014. – V.36, №2. – P. 290–302.
176. Guo, S.W. Performing the exact test of Hardy-Weinberg proportion for multiple alleles / S.W. Guo, E.A. Thompson // *Biometrics*. – 1992. – V.48, №2. – P. 361–372.
177. Gupta, S.K. Phenotypic and molecular diversity-based prediction of heterosis in pearl millet (*Pennisetum glaucum* L. (R.) Br.) / S.K. Gupta, T. Nepolean, C.G. Shaikh [et al.] // *The Crop Journal*. – 2018. – V.6, №3. – P. 271-281.
178. Hayes, B.J. Invited review: genomic selection in dairy cattle: progress and challenges / B.J. Hayes, P.J. Bowman, A.J. Chamberlain [et al.] // *Journal of Dairy Science*. – 2009. – V.92, №2. – P. 433–443.

179. Hedrick, P.W. Genetics of Populations / P.W. Hedrick – 4th ed. – Sudbury, Mass.: Jones and Bartlett, 2011. – 675 p.
180. Hill, E.W. The contribution of myostatin (MSTN) and additional modifying genetic loci to race distance aptitude in thoroughbred horses racing in different geographic regions / E.W. Hill, B.A McGivney, M. Rooney [et al.] // Equine Veterinary Journal. – 2019. – V.51, №5 – P. 625-633.
181. Ilie, D.E. Genome-wide association studies for milk somatic cell score in romanian dairy cattle / D.E. Ilie, A. Mizeranschi, C. V. Mihali [et al.] // Genes. – 2021. – V.12, №10. – P. 1495-1511.
182. Irish Cattle Breeding Federation (ICBF): [сайт]. - URL: <https://www.icbf.com> (дата обращения: 16.11.2024).
183. Iung, L.H.S. Genome-wide association study for milk production traits in a brazilian holstein population / L.H.S. Iung, J. Petrini, J. Ramírez-Díaz [et al.] // Journal of Dairy Science. – 2019. – V.102, №6. – P. 5305-5314.
184. Jaiswal, P. Gramene: development and integration of trait and gene ontologies for rice / P. Jaiswal, D. Ware, J. Ni [et al.] // Comparative and Functional Genomics. – 2002. – V.3. – P. 132–136.
185. Jarne, P. Microsatellites, from molecules to populations and back / P. Jarne, P.J. Lagoda // Trends in Ecology & Evolution. – 1996. – V.11, №10. – P. 424-429.
186. Jiang, J. A Large-scale genome-wide association study in u.s. holstein cattle / J. Jiang, L. Ma, D. Prakapenka [et al.] // Frontiers in Genetics. – 2019. – V.10. – A. 412.
187. Johanssen, W. The genotype conception of heredity / W. Johanssen // International Journal of Epidemiology. – 2014. – V.43, №4. – P. 989–1000.
188. Jombart, T. adegenet: a R package for the multivariate analysis of genetic markers / T. Jombart // Bioinformatics. – 2008. – V.24, №11. – P. 1403-1405.
189. Kalish, J.M. Triplex-induced recombination and repair in the pyrimidine motif / J.M. Kalish, M.M. Seidman, D.L. Weeks [et al.] // Nucleic Acids Research. – 2005. – V.33, №11. – P. 3492-3502.
190. Kamaldinov E.V. Assessment of the variability of reproductive abilities of a black and white cattle using genealogical data and paratypical factors / E.V. Kamaldinov,

- O.D. Panferova, O.V. Efremova [et al.] // *Data in Brief*. – 2021. – V.35, №2. – A. 106842.
191. Kamaldinov E.V. Modelling the variability of milk fat content in cattle / E.V. Kamaldinov, A.F. Petrov, U. Shergaziev [et al.] // Conference «Conference Proceedings: Global Food Forum 2021. Dialogue without borders». – Geneva: EurAsian Scientific Editions SA; Hong Kong: EurAsian Scientific Editions Ltd; Tallinn: EurAsian Scientific Editions OÜ, 2022. – P. 152-160.
192. Kang, H. Variance component model to account for sample structure in genome-wide association studies / H.M. Kang, J.H. Sul, S. K. Service [et al.] // *Nature Genetics*. – 2010. – V.42, №4. – P. 348–354.
193. Kantanen, J. Genetic diversity and population structure of 20 north european cattle breeds / J. Kantanen, I. Olsaker, L.E. Holm [et al.] // *The Journal of Heredity*. – 2000. – V.91, №6. – P. 446-457.
194. Keenan, K. Application DiveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors / K. Keenan, P. McGinnity, T. Cross [et al.] // *Methods in Ecology and Evolution*. – 2013. – V.4. – P. 782–788.
195. Kennedy, B.W. Estimation of effects of single genes on quantitative traits / B.W. Kennedy, M. Quinton, J.A.M. van Arendonk // *Journal of Animal Science*. – 1992. – V.70, №7. – P. 2000-2012.
196. Kim, E. Genome-wide scan of twinning rate QTL using linkage disequilibrium / E. Kim, P. Berger, B.W. Kirkpatrick // *Animal Genetics*. – 2009. – V.40(3). – P. 300–307.
197. Kolbehdari, D. A Whole-genome scan to map quantitative trait loci for conformation and functional traits in canadian holstein bulls / D. Kolbehdari, Z. Wang, J. R. Grant [et al.] // *Journal of Dairy Science*. – 2008. – V.91, №7. – P. 2844–2856.
198. Kumar, V. Deficiency of uridine monophosphate synthase: a recessive disorder in holstein friesland cattle / V. Kumar, R.K. Singh, A. Sharma // *Veterinary World*. – 2010. – V.3, №11. – P. 523-525.

199. Legarra, A. A relationship matrix including full pedigree and genomic information / A. Legarra, I. Aguilar, I. Misztal // *Journal of Dairy Science*. – 2009. – V.92, №9. – P. 4656-4663.
200. Legarra, A. Single step, a general approach for genomic selection / A. Legarra, O.F Christensen, I. Aguilar [et al.] // *Livestock Science*. – 2014. – V.166, №1. – P. 54-65.
201. Li, Q. Identification of candidate genes for milk production traits by RNA sequencing on bovine liver at different lactation stages / Q. Li, R. Liang, Y. Li [et al.] // *BMC Genetics*. – 2020. – V.21, №1. – P. 72-83.
202. Liu, S. Hadamard, khatri-rao, kronecker and other matrix products / S. Liu, G. Trenkler // *International Journal of Information and Systems Sciences*. – 2008. – V.4, №1. – P. 160-177.
203. Ludden, T.M. Comparison of the akaike Information criterion, the schwarz criterion and the F test as guides to model selection / T.M. Ludden, S.L. Beal, L.B. Sheiner // *Journal of Pharmacokinetics and Biopharmaceutics*. – 1994. – V.22, №5. – P. 431-445.
204. Majumdar, A. Semiparametric allelic tests for mapping multiple phenotypes: binomial regression and mahalanobis distance / A. Majumdar, J.S. Witte, S. Ghosh // *Genetic Epidemiology*. - 2015. – V.39, №8. – P. 635-650.
205. Makvandi-Nejad, S. Four loci explain 83% of size variation in the horse / S. Makvandi-Nejad, G.E. Hoffman, J.J. Allen [et al.] // *PLOS One*. – 2012. – V.7, №7. – P. e39929.
206. Marete, A. Genome-wide association study for milking speed in french holstein cows / A. Marete, G. Sahana, S. Fritz [et al.] // *Journal of Dairy Science*. – 2018. – V.101, №7. – P. 6205-6219.
207. Matukumalli, L.K. Development and characterization of a high density SNP genotyping assay for cattle / L.K. Matukumalli, C.T. Lawley, R.D. Schnabel [et al.] // *PLOS One*. – 2009. – V.4, №4. – P. e5350.
208. McCue, M.E. A high density SNP array for the domestic horse and extant perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies / M.E. McCue, D.L. Bannasch, J.L. Petersen [et al.] // *PLoS Genetics*. – 2012. – V.8, №1. – P. e1002451.

209. Meredith, B.K. Genome-wide associations for milk production and somatic cell score in holstein-friesian cattle in Ireland / B.K. Meredith, F.J. Kearney, E.K. Finlay [et al.] // BMC Genetics. – 2012. – V.13, №1. – P. 21-31.
210. Meuwissen, T. Genomic selection: A paradigm shift in animal breeding / T. Meuwissen, B. Haeyes, M. Goddard // Animal Frontiers. – 2016. – V.6, №1. – P. 6-14.
211. Meuwissen, T.H.E. Prediction of total genetic value using genome-wide dense marker maps / T.H.E. Meuwissen, B.J. Hayes, M.E. Goddard // Genetics. – 2001. – V.157, №4. – P. 1819-1829.
212. Misztal, I. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information / I. Misztal, A. Legarra, I. Aguilar // Journal of Dairy Science. – 2009. – V.92, №9. – P. 4648-4655.
213. Moser, G. Accuracy of direct genomic values in holstein bulls and cows using subsets of SNP markers / G. Moser, M.S. Khatkar, B.J. Hayes [et al.] // Genetics Selection Evolution. – 2010. – V.42, №1. – P. 42-51.
214. Mrode, R.A. Linear models for the prediction of animal breeding values / R.A. Mrode, R. Thompson. – Wallingford: CAB International, 2005. – 344 p.
215. Mrode, R.A. Genomic selection and use of molecular tools in breeding programs for indigenous and crossbred cattle in developing countries: current status and future prospects / R.A. Mrode, J.M.K. Ojango, A.M. Okeyo [et al.] // Frontiers in Genetics. – 2019. – V.9. – A. 694.
216. Murtagh, F. Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? / F. Murtagh, P. Legendre // Journal of Classification. – 2014. – V.31, №3. – P. 274-295.
217. MySQL: The world's most popular open source database : [сайт]. – URL: <https://www.mysql.com> (дата обращения: 11.12.2023).
218. Narozhnykh, K. Prediction of milk productivity based on conformation traits in cows / K.N. Narozhnykh, E.V. Kamaldinov, O.V. Bogdanova [et al.] // International Journal of Chemical and Biochemical Sciences (IJCBS). – 2023. – V.24, №6. – P. 521-529.

219. Nayeri, S. Genome-wide association for milk production and female fertility traits in canadian dairy holstein cattle / S. Nayeri¹, M. Sargolzaei, M.K. Abo-Ismael [et al.] // BMC Genetics. – 2016. – V.17, №1. – P. 75-85.
220. Nei, M. Genetic distance between populations / M. Nei // The American Naturalist. – 1972. – V.106, №949. – P. 283-292.
221. Nielsen, R. Correcting for ascertainment biases when analyzing SNP data: applications to the estimation of linkage disequilibrium / R. Nielsen, J. Signorovitch // Theoretical Population Biology. – 2003. – V.63, №3. – P. 245-255.
222. Oliveira, H. Genome-wide association for milk production traits and somatic cell score in different lactation stages of ayrshire, holstein, and jersey dairy cattle / H. Oliveira, J.P. Cant, L.F. Brito [et al.] // Journal of Dairy Science. – 2019. – V.102, №9. – P. 8159-8174.
223. Peakall, R. GenAIEx V5: genetic analysis in excel. populations genetic software for teaching and research / R. Peakall, P.E. Smouse // Bioinformatics. – 2012. – V.28, №19, – P. 2537-2539.
224. Pedrosa, V.B. Genomewide association analyses of lactation persistency and milk production traits in holstein cattle based on imputed whole-genome sequence data / V.B. Pedrosa, F.S. Schenkel, S.-Y. Chen [et al.] // Genes. – 2021. – V.12, №11, P. 1830-1858.
225. Pellegrino, I. Development of SNP markers for population structure and phylogeography characterization in little owl (*athene noctua*) using a genotyping-bysequencing approach / I. Pellegrino, L. Boatti, M. Cucco [et al.] // Conservation Genetics Resources. – 2015. – V.8, №1. – P. 1-4.
226. Birteeb, P.T. Multivariate characterisation of the phenotypic traits of djallonke and sahel sheep in northern ghana / P.T. Birteeb, S.O. Peters, A. Yakubu [et al.] // Tropical Animal Health and Production. – 2012. – V.45, №1. – P. 267-274.
227. Petrov, A.F. Clustering of countries based on dairy productivity characteristics of holstein cattle for breeding material selection / A.F. Petrov, O.V. Bogdanova, K.N. Narozhnykh // Veterinary World. – 2024. – V.17, №5. – P. 1108-1118.

228. PHP: Hypertext Preprocessor : [сайт]. - URL: <https://www.php.net> (дата обращения: 12.01.2024).
229. PLINK: Whole genome data analysis toolset : [сайт]. - URL: <https://zzz.bwh.harvard.edu/plink/download.shtml> (дата обращения: 09.07.2024).
230. Prevosti, A. Distances between populations of drosophila subobscura, based on chromosome arrangement frequencies / A. Prevosti, J. Ocaña, G. Alonso // *Theoretical and Applied Genetics*. – 1975. – V.45, №6. – P. 231–241.
231. Pucket, E.E. Variability in total project and per sample genotyping costs under varying study designs including with microsatellites or SNPs to answer conservation genetic questions / E.E. Pucket // *Conservation Genetics Resources*. – 2017. – V.9. – P. 289-304.
232. Queirós, J. Effect of microsatellite selection on individual and population genetic inferences: an empirical study using cross-specific and species-specific amplifications / J. Queirós, R. Godinho, S. Lopes [et al.] // *Molecular Ecology Resources*. – 2015. – V.15, №4. – P. 747-760.
233. R: The R project for statistical computing : [сайт]. - URL: <https://www.r-project.org> (дата обращения: 11.03.2024).
234. Reynolds, J. Estimation of the coancestry coefficient: basis for a short-term genetic distance / J. Reynolds, B. Weir, C.C. Cockerham // *Genetics*. – 1983. – V.105, №3. – P. 767-779.
235. Ritchie, M.E. Comparing genotyping algorithms for illumina's infinium wholegenome SNP BeadChips / M.E. Ritchie, R.C. Liu, B. Carvalho [et al.] // *BMC Bioinformatics*. – 2011. – V.12, №1. – P. 68-79.
236. Robinson, J.L. Low incidence of citrullinemia carriers among dairy cattle of the united states / J.L. Robinson, J.L. Burns, C.E. Magura [et al.] // *Journal of Dairy Science*. – 1993. – V.76, №3. – P. 853-858.
237. Robinson, J.L. Testing for deficiency of uridine monophosphate synthase among holstein-friesian cattle of north america and europe / J.L. Robinson, R.G. Popp, R. Shanks [et al.] // *Livestock Production Science*. – 1993. – V.36, №4. – P. 287-298.

238. Rogers, J.S. Measures of genetic similarity and genetic distance / J.S. Rogers, N. Dogan // *Turkiye Klinikleri Journal of Biostatistics*. – 2016. – V.8, №1. – P. 87-93.
239. Salem, M.M.I. Genome-wide association study for milk-fat yield in portuguese holstein cattle / M.M.I. Salem, G. Thompson, S. Chen [et al.] // Conference «10th World Congress on Genetics Applied to Livestock Production». – Vancouver, BC Canada, 2014. – P. 1-3.
240. Samuel, B. Genetic diversity of DGAT1 gene linked to milk production in cattle populations of Ethiopia / B. Samuel, H. Dinka, H. Dadi [et al.] // *BMC Genomic Data*. – 2022. – V.23, №64. – P. 1-10.
241. Schopen, G.C.B. Whole-genome association study for milk protein composition in dairy cattle / G.C.B. Schopen, M.H.P.W. Visker, P.D. Koks [et al.] // *Journal of Dairy Science*. – 2011. – V.94, №6. – P. 3148–3158.
242. Smouse, P.E. Converting quadratic entropy to diversity: Both animals and alleles are diverse, but some are more diverse than others / P.E. Smouse, S.C. Banks, R. Peakall // *PLOS One*. – 2017. – V.12, №10. – P. e0185499.
243. Steffen, P. Isolation and mapping of polymorphic microsatellites in cattle / P. Steffen, A. Eggen, A.B. Dietz [et al.] // *Animal Genetics*. – 1993. – V.24, №2. – P. 121-124.
244. Sunden, S.L.F. A highly polymorphic bovine microsatellite locus: BM2113 / S.L.F. Sunden, R.T. Stone, M.D. Bishop [et al.] // *Animal Genetics*. – 1993. – V.24, №1. – P. 69.
245. Sunnucks, P. Efficient genetic markers for population biology / P. Sunnucks // *Trends in Ecology & Evolution*. – 2000. – V.15, №5. – P. 199-203.
246. Teklewold, A. Comparison of phenotypic and molecular distances to predict heterosis and F1 performance in ethiopian mustard (*brassica carinata* A. Braun) / A. Teklewold, H. Becker // *Theoretical and Applied Genetics*. – 2006. – V.112, №4. – P. 752-759.
247. Toldo, S.S. Physically mapped, cosmid-derived microsatellite markers as anchor loci on bovine chromosomes / S.S. Toldo, R. Fries, P. Steffen [et al.] // *Mammalian Genome*. – 1993. – V.4, №12. – P. 720-727.

248. Vaiman, D. A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism / D. Vaiman, D. Mercier, K. Goudarzi [et al.] // *Mammalian Genome*. – 1994. – V.5, №5. – P. 288-297.
249. Tassell, C.P.V. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries / C.P.V. Tassell, T. Smith, L. K Matukumalli [et al.] // *Nature Methods*. – 2008. – V.5, №3. – P. 247-252.
250. VanRaden, P.M. Efficient methods to compute genomic predictions / P.M. VanRaden // *Journal of Dairy Science*. – 2008. – V.91, №11. – P. 4414-4423.
251. Viturro, E. Identification, sequence analysis and mRNA tissue distribution of the bovine sterol transporters ABCG5 and ABCG8 / E. Viturro, C. Farke, H.H.D. Meyer [et al.] // *Journal of Dairy Science*. – 2006. – V.89, №2. – P. 553-561.
252. WAAP: World Association for Animal Production : [сайт]. - URL: <https://waap.it> (дата обращения: 15.09.2024).
253. Wang, P. Genome-wide association analysis of milk production, somatic cell score, and body conformation traits in Holstein cows / P. Wang, X. Li, Y. Zhu [et al.] // *Frontiers in Veterinary Science*. – 2022. – V.9. – P. 932034.
254. Willet, C. From the phenotype to the genotype via bioinformatics / C.E Willet, C.M. Wade // *Methods in Molecular Biology*. – 2014. – V.1168. – P. 1-16.
255. Yakubu, A. Multivariate analysis of phenotypic differentiation in bunaji and sokoto gudali cattle / A. Yakubu, K.O. Idahor, H.S. Haruna [et al.] // *Acta Agriculturae Slovenica*. – 2010. – V.96, №2. – P. 75-80.
256. Yodklaew, P. Genome-wide association study for lactation characteristics, milk yield and age at first calving in a thai multibreed dairy cattle population / P. Yodklaew, S. Koonawootrittriron, M. Elzo [et al.] // *Agriculture and Natural Resources*. – 2017. – V.51, №3. – P. 223-230.
257. Yu, G.X. An anchored AFLP- and retrotransposon-based map of diploid *Avena* / G.X. Yu, R.P Wise // *Genome*. – 2000. – V.43, №5. – P. 736-749.
258. Zatoń-Dobrowolska, M. Genetic distance between the polish red, czech red and german red cattle estimated based on selected loci of protein coding genes and DNA mi-

microsatellite sequences / M. Zatoń-Dobrowolska, J. Čitek [et al.] // *Animal Science Papers and Reports*. – 2007. – V.25, №1. – P. 45-54.

259. Zhou, G.L. Genetic diversity analysis of five cattle breeds native to china using microsatellites / G.L. Zhou, H.G. Jin, Q. Zhu [et al.] // *Journal of Genetics*. – 2005. – V.84, №1. – P. 77-80.

260. Zhou, J. Genome-wide association study of milk and reproductive traits in dual-purpose xinjiang brown cattle / J. Zhou, L. Liu, J. Chen [et al.] // *BMC Genomics*. – 2019. – V.20, №1. – P. 827.

ПРИЛОЖЕНИЯ

Приложение А: Анализ качества входных генеалогических баз данных племенных предприятий Западной Сибири

Хозяйство	Число строк в базе	Всего рядов предков (по 4 ряда на проб.)	Группа	Нет предков, всего	О	ОО	ОМ	ООО	ООМ	ОМО	ОММ	М	ММ	МО	МММ	ММО	МОМ	МОО
Хозяйство 1	6067	84938	Быков-осеменителей	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
			Всех быков	5211	230	342	351	417	417	425	430	225	348	341	425	427	416	417
			Коров	561	0	15	3	46	44	129	14	1	5	39	16	69	75	105
			Тёлок	488	1	30	8	31	17	48	19	2	11	34	24	35	112	116
Хозяйство 2	8810	123340	Быков-осеменителей	12	0	1	1	1	1	1	1	0	1	1	1	1	1	1
			Всех быков	3997	154	238	264	316	320	343	349	157	265	246	348	346	323	328
			Коров	666	0	7	15	44	28	250	0	8	0	253	33	27	1	0
			Тёлок	232	0	0	20	2	28	4	65	0	19	0	67	4	23	0
Хозяйство 3	6939	97146	Быков-осеменителей	24	1	1	1	2	2	2	2	1	1	1	3	3	3	1
			Всех быков	3997	154	238	264	316	320	343	349	157	265	246	348	346	323	328
			Коров	666	0	0	7	15	44	28	250	0	8	0	253	33	27	1
			Тёлок	2524	9	63	27	400	311	63	355	8	162	63	308	63	292	400
Хозяйство 4	12716	178024	Быков-осеменителей	26	1	2	2	2	2	2	2	1	2	2	2	2	2	2
			Всех быков	3539	141	212	242	268	285	283	307	159	245	227	307	287	288	288
			Коров	530	0	0	9	0	57	0	105	0	42	0	270	0	47	0
			Тёлок	55	0	0	0	0	0	0	9	0	0	0	46	0	0	0
Хозяйство 5	7337	102718	Быков-осеменителей	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
			Всех быков	3650	153	230	238	296	293	299	312	155	239	231	312	300	294	298
			Коров	148	0	0	0	17	6	18	33	0	0	0	33	18	6	17
			Тёлок	12	0	0	0	0	0	0	6	0	0	0	6	0	0	0

Хозяйство	Число строк в базе	Всего рядов предков (по 4 ряда на проб.)	Группа	Нет предков, всего	О	ОО	ОМ	ООО	ООМ	ОМО	ОММ	М	ММ	МО	МММ	ММО	МОМ	МОО
			Быков-осеменителей	19	1	1	1	1	1	2	2	1	2	1	2	2	1	1
Хозяйство 6	15154	212156	Всех быков	12224	594	784	858	920	965	986	1018	592	847	780	1013	984	964	919
			Коров	839	2	100	18	125	119	67	91	5	17	3	105	33	39	115
			Тёлок	922	6	121	6	121	80	111	23	6	8	104	26	181	8	121
			Быков-осеменителей	8	0	0	0	1	1	1	1	0	0	0	1	1	1	1
Хозяйство 7	9340	130760	Всех быков	3097	144	204	214	251	249	247	253	139	211	204	249	246	244	242
			Коров	800	0	89	21	89	39	0	112	0	55	0	270	0	36	89
			Тёлок	922	6	121	6	121	80	111	23	6	8	104	26	181	8	121
			Быков-осеменителей	31	0	0	0	4	4	4	3	0	0	0	4	4	4	4
Хозяйство 8	8292	116088	Всех быков	2572	97	149	177	196	209	221	219	108	177	158	221	222	213	205
			Коров	848	31	35	22	77	27	79	173	2	22	35	177	77	27	64
			Тёлок	317	13	13	2	16	5	17	98	0	3	13	100	16	5	16
			Быков-осеменителей	2846	8	75	234	211	275	258	351	7	239	81	353	264	279	211
Хозяйство 9	29372	411208	Всех быков	29283	1135	1631	2193	2181	2381	2382	2668	1154	2210	1671	2680	2416	2394	2187
			Коров	8151	0	48	0	825	313	1717	1055	0	0	53	1037	1793	380	930
			Тёлок	1037	0	23	0	108	28	304	5	0	0	24	3	307	32	203
			Быков-осеменителей	12	0	0	0	1	1	0	2	0	0	0	2	1	2	3
Хозяйство 10	32157	450198	Всех быков	46037	1784	2571	3520	3115	3858	3555	4267	2029	3610	2724	4313	3642	3888	3161
			Коров	3287	0	30	145	66	363	223	555	0	216	42	812	199	467	169
			Тёлок	965	0	86	0	86	23	95	142	0	0	86	227	97	33	90
			Быков-осеменителей	10	0	0	0	0	0	0	4	0	0	0	6	0	0	0
Хозяйство 11	55239	773346	Всех быков	85016	3491	4441	7107	4900	7470	6152	8381	3797	7276	4641	8443	6354	7542	5021
			Коров	8375	0	12	6	104	877	1224	1696	0	24	19	1851	1345	977	240
			Тёлок	116	0	0	1	4	12	29	5	0	0	0	16	28	16	5

Хозяйство	Число строк в базе	Всего рядов предков (по 4 ряда на проб.)	Группа	Нет предков, всего	О	ОО	ОМ	ООО	ООМ	ОМО	ОММ	М	ММ	МО	МММ	ММО	МОМ	МОО	
Хозяйство 12	45870	642180	Быков-осеменителей	16	0	0	0	1	1	0	4	0	0	0	4	1	3	2	
			Всех быков	86987	3345	4434	7537	4943	7785	6405	8405	3598	7733	4697	8485	6761	7825	5034	
			Коров	5699	0	10	30	108	467	1014	1077	0	31	14	1184	1089	500	175	
			Тёлок	7431	0	13	1	199	604	1444	1356	0	3	13	1412	1506	646	234	
Хозяйство 13	15068	210952	Быков-осеменителей	44	1	1	1	3	3	4	7	1	1	1	9	4	3	5	
			Всех быков	35757	1427	2080	2752	2462	2976	2687	3233	1580	2823	2178	3274	2772	2997	2516	
			Коров	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
			Тёлок	154	1	5	1	8	5	31	23	1	1	5	27	32	6	8	
Хозяйство 14	110419	1545866	Быков-осеменителей	55	0	1	0	5	4	5	9	0	0	1	11	7	5	7	
			Всех быков	5699	0	10	30	108	467	1014	1077	0	31	14	1184	1089	500	175	
			Коров	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
			Тёлок	5014	0	25	0	217	139	1818	208	0	0	25	297	1844	192	249	
Хозяйство 15	7542	105588	Быков-осеменителей	241	11	12	11	20	20	20	23	11	13	12	25	21	22	20	
			Всех быков	9043	372	524	649	658	748	714	798	451	644	553	794	716	749	673	
			Коров	366	0	45	0	141	0	45	0	0	0	45	0	45	0	45	
			Тёлок	3050	1	490	1	502	27	491	1	1	1	491	1	525	27	491	
Хозяйство 16	38634	540876	Быков-осеменителей	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
			Всех быков	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
			Коров	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
			Тёлок	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Хозяйство 17	8585	120190	Быков-осеменителей	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
			Всех быков	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
			Коров	2721	0	2	1	10	24	403	666	0	1	58	669	434	336	117	
			Тёлок	119	0	8	0	8	1	18	0	0	0	13	0	18	43	10	

Хозяйство	Число строк в базе	Всего рядов предков (по 4 ряда на проб.)	Группа	Нет предков, всего	Нет предков, всего													
					О	ОО	ОМ	ООО	ООМ	ОМО	ОММ	М	ММ	МО	МММ	ММО	МОМ	МОО
			Быков-осеменителей	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Хозяйство 18	12366	173124	Всех быков	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
			Коров	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
			Тёлок	6	0	1	0	1	0	1	0	0	0	1	0	1	0	1
Всего	429907	6018698		393130	13291	19480	26746	24838	32511	35838	40231	14341	27561	20448	41692	36908	33321	25924
Всего по быкам-осеменителям				3344	23	94	251	252	315	299	411	22	259	100	423	311	326	258
Всего по быкам				336109	13221	18088	26396	21347	28743	26056	32066	14301	26924	18911	32396	26908	28960	21792
Всего по коровам				32991	33	393	270	1652	2364	5169	5577	16	413	561	6457	5129	2891	2066
Всего по телкам				23364	37	999	73	1824	1360	4585	2338	24	216	976	2586	4838	1443	2065
Средняя по всем категориям				5584,1	193,0	283,7	391,3	363,6	475,7	523,7	589,0	208,2	403,2	297,8	610,4	539,4	487,6	10950,7
Средняя по быкам-осеменителям				196,7	1,4	5,9	15,7	15,8	19,7	18,7	25,7	1,4	16,2	6,3	26,4	19,4	20,4	389,6
Средняя по быкам				18383,9	764,2	1043,9	1532,1	1231,2	1666,3	1507,8	1861,0	828,0	1563,4	1092,4	1880,7	1557,8	1679,1	36591,8
Средняя по коровам				2060,7	10,4	34,2	29,9	107,6	149,1	300,6	342,3	9,6	37,8	42,7	391,2	302,2	175,9	3994,2
Средняя по телкам				1307,9	2,0	54,2	4,4	102,1	76,2	265,9	128,8	1,7	11,4	66,4	144,2	268,3	74,0	2507,5

Приложение Б: Ранжирование предприятий по качеству генеалогической информации

Хозяйство	Всего рядов предков (по 4 ряда на пробанда)	Группа	Отсутствующие предки	Всего отсутствующих предков (без учёта быков-осеменителей)	Отсутствующих предков, %	Всего отсутствующих предков (без учёта быков-осеменителей), %	Ранги по группе	Ранг
Хозяйство 1	84938	Быков-осеменителей	-	6260	-	7,4	1	11
		Всех быков	5211		6,14		12	
		Коров	561		0,66		11	
		Тёлок	488		0,57		14	
Хозяйство 2	123340	Быков-осеменителей	12	4895	0,01	4	11	9
		Всех быков	3997		3,24		9	
		Коров	666		0,54		9	
		Тёлок	232		0,19		8	
Хозяйство 3	97146	Быков-осеменителей	24	7187	0,02	7,4	13	12
		Всех быков	3997		4,11		10	
		Коров	666		0,69		10	
		Тёлок	2524		2,6		16	
Хозяйство 4	178024	Быков-осеменителей	26	4124	0,01	2,3	11	4
		Всех быков	3539		1,99		6	
		Коров	530		0,3		6	
		Тёлок	55		0,03		5	
Хозяйство 5	102718	Быков-осеменителей	0	3810	0	3,7	1	6
		Всех быков	3650		3,55		8	
		Коров	148		0,14		5	
		Тёлок	12		0,01		3	

Хозяйство	Всего рядов предков (по 4 ряда на пробанда)	Группа	Отсутствующие предки	Всего отсутствующих предков (без учёта быков-осеменителей)	Отсутствующих предков, %	Всего отсутствующих предков (без учёта быков-осеменителей), %	Ранги по группе	Ранг
Хозяйство 6	212156	Быков-осеменителей	19	13985	0,01	6,6	9	10
		Всех быков	12224		5,76		8	
		Коров	839		0,4		6	
		Тёлок	922		0,43		10	
Хозяйство 7	130760	Быков-осеменителей	8	4819	0,01	3,7	8	7
		Всех быков	3097		2,37		7	
		Коров	800		0,61		6	
		Тёлок	922		0,71		11	
Хозяйство 8	116088	Быков-осеменителей	31	3737	0,03	3,2	9	6
		Всех быков	2572		2,22		6	
		Коров	848		0,73		7	
		Тёлок	317		0,27		8	
Хозяйство 9	411208	Быков-осеменителей	2846	38471	0,69	9,4	11	13
		Всех быков	29283		7,12		6	
		Коров	8151		1,98		9	
		Тёлок	1037		0,25		7	
Хозяйство 10	450198	Быков-осеменителей	12	50289	0	11,2	6	14
		Всех быков	46037		10,23		7	
		Коров	3287		0,73		6	
		Тёлок	965		0,21		6	

Хозяйство	Всего рядов предков (по 4 ряда на пробанда)	Группа	Отсутствующие предки	Всего отсутствующих предков (без учёта быков-осеменителей)	Отсутствующих предков, %	Всего отсутствующих предков (без учёта быков-осеменителей), %	Ранги по группе	Ранг
Хозяйство 11	773346	Быков-осеменителей	10	93507	0	12,1	4	16
		Всех быков	85016		10,99		7	
		Коров	8375		1,08		7	
		Тёлок	116		0,01		3	
Хозяйство 12	642180	Быков-осеменителей	16	100117	0	15,6	4	17
		Всех быков	86987		13,55		7	
		Коров	5699		0,89		6	
		Тёлок	7431		1,16		7	
Хозяйство 13	210952	Быков-осеменителей	44	35911	0,02	17	5	18
		Всех быков	35757		16,95		7	
		Коров	0		0		1	
		Тёлок	154		0,07		3	
Хозяйство 14	1545866	Быков-осеменителей	55	10713	0	0,7	4	3
		Всех быков	5699		0,37		5	
		Коров	0		0		1	
		Тёлок	5014		0,32		4	
Хозяйство 15	105588	Быков-осеменителей	241	12459	0,23	11,8	4	15
		Всех быков	9043		8,56		5	
		Коров	366		0,35		3	
		Тёлок	3050		2,89		5	
Хозяйство 16	540876	Быков-осеменителей	0	0	0	0	1	1
		Всех быков	0		0		1	
		Коров	0		0		1	
		Тёлок	0		0		1	

Хозяйство	Всего рядов предков (по 4 ряда на пробанда)	Группа	Отсутствующие предки	Всего отсутствующих предков (без учёта быков-осеменителей)	Отсутствующих предков, %	Всего отсутствующих предков (без учёта быков-осеменителей), %	Ранги по группе	Ранг
Хозяйство 17	120190	Быков-осеменителей	0	2840	0	2,4	1	5
		Всех быков	0		0		1	
		Коров	2721		2,26		2	
		Тёлок	119		0,1		2	
Хозяйство 18	173124	Быков-осеменителей	0	6	0	0	1	2
		Всех быков	0		0		1	
		Коров	0		0		1	
		Тёлок	6		0		1	
Всего	6018698		393130		6,53			
Всего по осеменителям			3344		0,06			
Всего по быкам			336109		5,58			
Всего по коровам			32991		0,55			
Всего по телкам			23364		0,39			
Средняя по всем категориям			5584,1					
Средняя по быкам-осеменителям			196,7					
Средняя по быкам			18383,9					
Средняя по коровам			2060,7					
Средняя по телкам			1307,9					

Приложение В: SNP ассоциированные с признаками молочной продуктивности

ИlluminaBovineChip	SNP	Замена нуклеотида	Хромосома	Позиция	Ближайший ген	Функция*	Источник
ARS-BFGL-BAC-2469	-	-	20	33433160	HEATR7B2	PP	[156]
ARS-BFGL-BAC-27930	rs109774038	T/C	20	29373244	HCN1	FP	[253]
ARS-BFGL-BAC-6525	rs110594909	T/C	10	92127288	NRXN3	PY	[165]
ARS-BFGL-NGS-100459	-	-	11	22128569	ABCG8	FY, PY, PP	[201]
ARS-BFGL-NGS-103091	rs110111162	A/G	5	74518588	RBFOX2	MY	[165]
ARS-BFGL-NGS-107037	rs109394729	T/C	10	46486647	USP3	PP	
ARS-BFGL-NGS-107379	rs109350371	T/C	14	2054457	LOC786966	FP	[153]
ARS-BFGL-NGS-107749	rs109830113	A/G	18	33986486	CDH16	FP, PP	[156]
ARS-BFGL-NGS-110497	rs110475419	T/C	26	45870133	ADAM12	PY	[165]
ARS-BFGL-NGS-111443	rs109498193	A/G	5	94269370	DERA	FP, PP	
ARS-BFGL-NGS-11319	rs109084201	T/C	2	6763227	SLC40A1	MY	
ARS-BFGL-NGS-113507	rs110440617	T/G	11	98407974	PTRH1	FP	
ARS-BFGL-NGS-113876	rs109392509	T/C	7	40644182	AGXT2L2	MY	[256]
ARS-BFGL-NGS-118998	rs110482506	T/C	20	32030332	GHR	FP	[165]
ARS-BFGL-NGS-14781	rs111027205	A/G	5	56766165	DDIT3	FP	[152]
ARS-BFGL-NGS-18010	rs110643010	T/C	29	37101861	ADAMTS8	MY	[256]
ARS-BFGL-NGS-24998	rs111003135	A/G	29	26721324	SAA3/SAA1	PP	[152]
ARS-BFGL-NGS-29557	rs110614142	T/C	5	81612132	PTHLH	FP,PP	
ARS-BFGL-NGS-29581	-	-	4	114000000	KCNH2	PY	[165]
ARS-BFGL-NGS-39539	rs110558583	A/G	23	41457147	JARID2	PY	
ARS-BFGL-NGS-42950	rs110565416	A/G	17	49531815	TMEM132C	MY	[256]
ARS-BFGL-NGS-49079	-	-	9	6574398	-	MY	[165]
ARS-BFGL-NGS-4939	rs109421300	A/G	14	1801116	DGAT1	MY, FY, PY, FP, PP	
ARS-BFGL-NGS-53398	rs110836976	A/G	X	21953655	MAGEA11	PP	
ARS-BFGL-NGS-57820	rs109146371	T/C	14	1651311	FOXH1	FP	[153]
ARS-BFGL-NGS-5790	rs110457668	T/G	5	26986116	NR4A1	FP, PP	[152]
ARS-BFGL-NGS-61452	-	-	4	75250982	HUS1	PP	[165]
ARS-BFGL-NGS-69013	rs110252311	T/C	2	123717269	SESN2	FP, PP	[152]
ARS-BFGL-NGS-71395	rs42216771	A/G	8	24707157	SLC24A2	MY	[155]

llumniaBovineChip	SNP	Замена нуклеотида	Хромосома	Позиция	Ближайший ген	Функция*	Источник
ARS-BFGL-NGS-72191	rs110840223	T/C	23	30279220	BoLA-DQB/H4	FP, PP	[152]
ARS-BFGL-NGS-85980	rs43392997	T/G	16	76091078	FAM71A	FP,PP	
ARS-BFGL-NGS-89510	rs110970157	A/C	26	13936262	IDE	MY	
ARS-USMARC-Parent-DQ846692-rs29010281	rs29010281	A/G	14	80082923	RALYL	MY	[256]
BTA-121739-no-rs	rs41622323	T/C	6	38063313	PKD2	PP	[165]
BTA-39609-no-rs	-	-	0	0	-	PP	
BTA-48480-no-rs	rs41641868	T/A	2	95119968	ADAM23	PP	
BTA-50402-no-rs	rs41945918	A/G	20	34451383	NA	PP	
BTA-68781-no-rs	rs43711373	A/C	23	30711619	HIST1H2AC	FP,PP	[152]
BTB-00231742	rs43436389	C/G	5	77095345	-	FP	[165]
BTB-00285653	rs43499009	T/C	8	30036807	NFIB	FP	
BTB-00411816	rs43613632	T/C	10	11899342	THBS4	FP,PP	[152]
BTB-00777571	rs41938472	A/C	20	34017024	-	FP	[165]
BTB-01052867	rs42215728	A/G	8	24734717	SLC24A2	MY	[142]
BTB-01731924	rs42844632	A/G	7	75830763	GABRG2	PY	[260]
BTB-01766447	rs42877704	T/C	11	17484533	RPL23A	FP,PP	[152]
Hapmap29482-BTA-146449	rs110728437	T/A	9	38804655	LAMA4	MY	[256]
Hapmap30383-BTC-005848	rs109752439	A/G	14	1489496	C14H8orf33 – ZNF34	FP	[153]
Hapmap38455-BTA-100999	rs41613779	A/C	9	76346736	OLIG3	PP	[165]
Hapmap39690-BTA-82650	rs41590921	A/C	8	104829510	ZNF618	MY, IY, PY, PS, AFC	[256]
Hapmap39717-BTA-112973	rs41617243	T/G	2	26781358	KBTBD10	FP	[165]
Hapmap40191-BTA-73919	rs41648982	A/C	5	71978791	SYN3	FY	
Hapmap42216-BTA-45665	rs41636126	T/C	19	45934555	GOSR2	PP	
Hapmap42263-BTA-60093	rs41567529	T/C	25	35342491	MIR2388	FY	
Hapmap47777-BTA-91000	rs41659076	A/G	X	140892698	LOC785455	MY	

ИlluminaBovineChip	SNP	Замена нуклеотида	Хромосома	Позиция	Ближайший ген	Функция*	Источник
Нарmap48369-ВТА-50306	rs41639307	A/G	1	7627111	-	MY	[165]
Нарmap48524-ВТА-92140	rs41663866	T/C	5	75684520	NCF4	PP	
Нарmap49309-ВТА-78604	rs41658012	A/G	7	24655689	GADD45B	PP	
Нарmap50621-ВТА-21320	rs41574875	A/G	6	64425164	-	PP	
Нарmap51303-ВТА-74377	rs41652649	A/G	5	83790390	ITPR2/PTHLH	FP	
Нарmap54188-rs29022489	rs29022489	A/G	6	75017253	-	PP	
Нарmap56889-rs29016097	rs29016097	C/G	20	31468943	HMGCS1	MY	[256]
Нарmap59494-rs29020429	rs29020429	A/G	9	54062560	UFL1	MY	
Нарmap60955-rs29022431	rs29022431	T/G	23	21292766	-	MY	[165]
UA-IFASA-6479	rs41566160	A/G	13	33201457	CACNB2	MY	[256]

* где MY - удой, в кг.; FY – жир в кг.; FP - жир в %; PY - белок в кг; PP - белок в %.

Приложение Г: Оценка изменчивости и распределения входных данных без предварительной обработки

Показатель	Хозяйство / номер лактации	n	Mean	SE	Me	Min	Max	Range	Q1	Q3	IQR	SD	Cv	AD	AD.p	SF	SF.p
Жир за 305, %	Хозяйство 4 (1 лакт.)	1492	3,91	0,002	3,92	3,6	4,3	0,7	3,89	3,95	0,06	0,083	2,1	47,031	<0,001	0,903	<0,001
Жир за 305, %	Хозяйство 11 (1 лакт.)	3732	3,83	0,002	3,82	3,6	4,47	0,87	3,75	3,9	0,15	0,118	3,1	35,55	<0,001	0,952	<0,001
Жир за 305, кг	Хозяйство 3 (2 лакт.)	693	206,2	1,94	191,4	129,8	374,5	244,7	170,2	225,5	55,3	51,2	24,8	26,98	<0,001	0,885	<0,001
Жир за 305, %	Хозяйство 4 (2 лакт.)	855	3,92	0,002	3,92	3,61	4,25	0,64	3,9	3,95	0,05	0,069	1,8	24,882	<0,001	0,906	<0,001
Удой за 305, кг	Хозяйство 3 (2 лакт.)	693	5294	46,8	4957	3503	9341	5838	4412	5797	1385	1231	23,3	23,214	<0,001	0,899	<0,001
Жир за 305, кг	Хозяйство 3 (1 лакт.)	1003	209,7	1,52	199,5	130,7	388,8	258,1	174,9	235,8	60,9	48,2	23	16,252	<0,001	0,936	<0,001
Жир за 305, %	Хозяйство 7 (1 лакт.)	763	4,19	0,005	4,18	3,61	4,75	1,14	4,14	4,26	0,12	0,125	3	15,962	<0,001	0,907	<0,001
Жир за 305, %	Хозяйство 15 (2 лакт.)	313	3,82	0,009	3,79	3,61	5	1,39	3,74	3,86	0,12	0,156	4,1	15,877	<0,001	0,722	<0,001
Жир за 305, %	Хозяйство 8 (2 лакт.)	551	3,88	0,006	3,85	3,6	4,68	1,08	3,8	3,92	0,12	0,13	3,4	15,539	<0,001	0,881	<0,001
Удой за 305, кг	Хозяйство 3 (1 лакт.)	1003	5466	39,2	5213	3506	9741	6235	4552	6172	1620	1241	22,7	15,519	<0,001	0,942	<0,001
Жир за 305, %	Хозяйство 8 (1 лакт.)	889	3,86	0,004	3,84	3,6	4,35	0,75	3,79	3,91	0,12	0,115	3	15,009	<0,001	0,933	<0,001
Жир за 305, %	Хозяйство 15 (1 лакт.)	512	3,79	0,006	3,77	3,6	4,78	1,18	3,71	3,84	0,126	0,128	3,4	14,228	<0,001	0,794	<0,001
Жир за 305, %	Хозяйство 5 (1 лакт.)	523	3,98	0,007	3,95	3,6	4,86	1,26	3,88	4,05	0,17	0,168	4,2	13,833	<0,001	0,913	<0,001
Жир за 305, %	Хозяйство 4 (3 лакт.)	435	3,92	0,003	3,92	3,6	4,09	0,49	3,9	3,94	0,04	0,059	1,5	13,471	<0,001	0,844	<0,001
Жир за 305, кг	Хозяйство 7 (1 лакт.)	763	199,8	1,12	196,3	128,7	436,5	307,8	181,4	212,5	31,1	30,8	15,4	13,028	<0,001	0,898	<0,001
Жир за 305, %	Хозяйство 5 (2 лакт.)	368	4,02	0,011	3,96	3,67	5,19	1,52	3,87	4,11	0,24	0,219	5,4	12,983	<0,001	0,874	<0,001
Жир за 305, кг	Хозяйство 3 (3 лакт.)	407	196,4	2,09	187,3	134,3	357,5	223,2	168,8	210,6	41,8	42,2	21,5	12,719	<0,001	0,897	<0,001
Удой за 305, кг	Хозяйство 3 (3 лакт.)	407	5039	52	4828	3501	9014	5513	4342	5403	1061	1050	20,8	11,8	<0,001	0,903	<0,001
Жир за 305, %	Хозяйство 15 (3 лакт.)	160	3,82	0,015	3,78	3,6	5,04	1,44	3,72	3,85	0,13	0,188	4,9	11,273	<0,001	0,717	<0,001
Удой за 305, кг	Хозяйство 7 (1 лакт.)	763	4764	25,8	4703	3504	9938	6434	4323	5080	756,3	711,8	14,9	11,088	<0,001	0,908	<0,001
Жир за 305, %	Хозяйство 10 (1 лакт.)	1936	3,82	0,002	3,81	3,64	4,18	0,54	3,77	3,87	0,1	0,066	1,7	10,435	<0,001	0,977	<0,001
Жир за 305, %	Хозяйство 12 (1 лакт.)	915	3,82	0,003	3,81	3,63	4,23	0,6	3,76	3,88	0,12	0,089	2,3	8,651	<0,001	0,955	<0,001
Жир за 305, %	Хозяйство 9 (1 лакт.)	3257	3,9	0,002	3,9	3,6	4,36	0,76	3,8	4	0,2	0,137	3,5	8,609	<0,001	0,99	<0,001
Жир за 305, %	Хозяйство 16 (1 лакт.)	2584	3,94	0,004	3,93	3,6	4,67	1,07	3,8	4,08	0,28	0,192	4,9	7,763	<0,001	0,983	<0,001
Жир за 305, %	Хозяйство 11 (2 лакт.)	1786	3,84	0,003	3,83	3,6	4,45	0,85	3,74	3,92	0,18	0,129	3,4	7,713	<0,001	0,976	<0,001
Жир за 305, %	Хозяйство 3 (2 лакт.)	693	3,88	0,004	3,9	3,6	4,24	0,64	3,83	3,96	0,13	0,117	3	7,643	<0,001	0,971	<0,001
Жир за 305, %	Хозяйство 4 (4 лакт.)	229	3,92	0,004	3,93	3,66	4,12	0,46	3,9	3,95	0,05	0,057	1,5	7,132	<0,001	0,885	<0,001
Жир за 305, %	Хозяйство 5 (3 лакт.)	258	4,07	0,015	4,01	3,61	4,92	1,31	3,89	4,19	0,301	0,247	6,1	7,078	<0,001	0,9	<0,001

Показатель	Хозяйство / номер лактации	n	Mean	SE	Me	Min	Max	Range	Q1	Q3	IQR	SD	Cv	AD	AD.p	SF	SF.p
Жир за 305, кг	Хозяйство 7 (2 лакт.)	469	208	1,82	202,1	147,1	358,9	211,8	179,1	229,5	50,4	39,5	19	6,687	<0,001	0,923	<0,001
Жир за 305, %	Хозяйство 10 (2 лакт.)	1295	3,81	0,002	3,8	3,6	4,11	0,51	3,76	3,84	0,08	0,06	1,6	6,535	<0,001	0,971	<0,001
Жир за 305, %	Хозяйство 8 (3 лакт.)	320	3,85	0,007	3,83	3,6	4,8	1,2	3,78	3,9	0,12	0,125	3,3	6,459	<0,001	0,87	<0,001
Удой за 305, кг	Хозяйство 7 (2 лакт.)	469	4929	42,4	4805	3505	8366	4861	4248	5418	1170	917,5	18,6	6,385	<0,001	0,925	<0,001
Жир за 305, кг	Хозяйство 9 (2 лакт.)	1720	414,2	1,27	418,2	156,3	592,7	436,4	381,6	451	69,4	52,5	12,7	5,563	<0,001	0,985	<0,001
Удой за 305, кг	Хозяйство 11 (2 лакт.)	1786	9583	31	9707	4369	13963	9594	8772	10477	1705	1310	13,7	5,215	<0,001	0,988	<0,001
Жир за 305, кг	Хозяйство 3 (4 лакт.)	186	196	3,17	187,9	135,3	382,4	247,1	166,8	211,6	44,8	43,2	22,1	4,982	<0,001	0,895	<0,001
Жир за 305, кг	Хозяйство 9 (1 лакт.)	3257	378,3	0,86	380,4	199,7	606	406,3	348,6	412,2	63,6	49,1	13	4,905	<0,001	0,993	<0,001
Жир за 305, %	Хозяйство 6 (1 лакт.)	983	3,88	0,005	3,87	3,6	4,57	0,97	3,77	3,98	0,208	0,158	4,1	4,865	<0,001	0,972	<0,001
Жир за 305, %	Хозяйство 6 (2 лакт.)	488	3,84	0,007	3,82	3,6	4,47	0,87	3,74	3,93	0,19	0,16	4,2	4,831	<0,001	0,949	<0,001
Жир за 305, %	Хозяйство 3 (1 лакт.)	1003	3,84	0,003	3,83	3,6	4,3	0,7	3,77	3,9	0,13	0,111	2,9	4,785	<0,001	0,971	<0,001
Удой за 305, кг	Хозяйство 3 (4 лакт.)	186	5036	79,5	4844	3502	9507	6005	4300	5415	1115	1084	21,5	4,548	<0,001	0,902	<0,001
Удой за 305, кг	Хозяйство 6 (2 лакт.)	488	8570	66,3	8826	3784	11742	7958	7857	9631	1774	1464	17,1	4,447	<0,001	0,968	<0,001
Жир за 305, кг	Хозяйство 11 (1 лакт.)	3732	312,3	0,744	314,4	142	472,1	330,1	283,7	343,7	60	45,5	14,6	4,321	<0,001	0,994	<0,001
Жир за 305, %	Хозяйство 12 (2 лакт.)	523	3,84	0,003	3,83	3,67	4,29	0,62	3,79	3,88	0,09	0,079	2,1	4,12	<0,001	0,946	<0,001
Жир за 305, %	Хозяйство 5 (4 лакт.)	165	4,09	0,022	4,03	3,66	5,06	1,4	3,86	4,28	0,417	0,279	6,8	4,068	<0,001	0,921	<0,001
Жир за 305, кг	Хозяйство 6 (2 лакт.)	488	328,3	2,33	333,2	150,5	450,6	300,1	302,1	363,4	61,3	51,4	15,7	4,05	<0,001	0,97	<0,001
Жир за 305, кг	Хозяйство 7 (3 лакт.)	240	210,3	2,73	201,8	149,9	339,3	189,4	177	240	63	42,3	20,1	4,029	<0,001	0,938	<0,001
Удой за 305, кг	Хозяйство 11 (1 лакт.)	3732	8162	20,9	8228	3534	12706	9172	7361	9035	1674	1276	15,6	4,012	<0,001	0,995	<0,001
Жир за 305, %	Хозяйство 15 (4 лакт.)	89	3,83	0,017	3,77	3,61	4,53	0,92	3,73	3,9	0,173	0,163	4,3	3,788	<0,001	0,817	<0,001
Удой за 305, кг	Хозяйство 9 (2 лакт.)	1720	10809	35,7	10922	3958	15559	11601	9878	11885	2007	1481	13,7	3,661	<0,001	0,99	<0,001
Жир за 305, %	Хозяйство 9 (2 лакт.)	1720	3,84	0,003	3,84	3,6	4,22	0,62	3,75	3,93	0,18	0,125	3,3	3,629	<0,001	0,989	<0,001
Удой за 305, кг	Хозяйство 7 (3 лакт.)	240	4983	63	4774	3545	7878	4333	4260	5609	1349	975,4	19,6	3,401	<0,001	0,947	<0,001
Удой за 305, кг	Хозяйство 4 (2 лакт.)	855	6647	44,2	6578	3754	12756	9002	5699	7503	1804	1294	19,5	3,362	<0,001	0,977	<0,001
Жир за 305, кг	Хозяйство 12 (2 лакт.)	523	351	2,73	347,5	170,7	559,6	388,9	313,7	380,5	66,8	62,3	17,8	3,357	<0,001	0,973	<0,001
Жир за 305, кг	Хозяйство 11 (2 лакт.)	1786	367,4	1,2	370,1	165,1	549,1	384	336,3	403,4	67,1	50,6	13,8	3,307	<0,001	0,992	<0,001
Жир за 305, %	Хозяйство 8 (4 лакт.)	148	3,87	0,013	3,85	3,6	4,75	1,15	3,78	3,93	0,146	0,159	4,1	3,251	<0,001	0,884	<0,001
Удой за 305, кг	Хозяйство 9 (1 лакт.)	3257	9718	23,4	9763	4770	14849	10079	8891	10615	1724	1338	13,8	3,147	<0,001	0,996	<0,001
Жир за 305, кг	Хозяйство 4 (1 лакт.)	1492	247,6	1,04	249,4	143,3	390,7	247,4	217,6	275,9	58,2	40,4	16,3	3,085	<0,001	0,994	<0,001
Удой за 305, кг	Хозяйство 16 (2 лакт.)	1445	12169	40,9	12298	6987	17091	10104	11220	13169	1950	1554	12,8	3,02	<0,001	0,992	<0,001

Показатель	Хозяйство / номер лактации	n	Mean	SE	Me	Min	Max	Range	Q1	Q3	IQR	SD	Cv	AD	AD.p	SF	SF.p
Удой за 305, кг	Хозяйство 12 (2 лакт.)	523	9146	72,3	9051	4430	14557	10127	8099	9989	1890	1654	18,1	2,785	<0,001	0,977	<0,001
Жир за 305, кг	Хозяйство 3 (5 лакт.)	64	210,9	7,69	186,9	137,1	381,4	244,3	161,7	254,7	93,1	61,5	29,2	2,735	<0,001	0,89	<0,001
Удой за 305, кг	Хозяйство 4 (1 лакт.)	1492	6330	26	6392	3650	10073	6423	5573	7003	1430	1003	15,8	2,693	<0,001	0,993	<0,001
Жир за 305, %	Хозяйство 10 (3 лакт.)	733	3,8	0,002	3,8	3,61	4,07	0,46	3,76	3,84	0,08	0,067	1,8	2,679	<0,001	0,987	<0,001
Удой за 305, кг	Хозяйство 16 (3 лакт.)	736	12496	64,2	12602	5440	16570	11130	11459	13698	2239	1741	13,9	2,629	<0,001	0,983	<0,001
Удой за 305, кг	Хозяйство 4 (3 лакт.)	435	6648	62,1	6500	3600	11120	7520	5736	7420	1684	1294	19,5	2,547	<0,001	0,975	<0,001
Удой за 305, кг	Хозяйство 3 (5 лакт.)	64	5384	184,2	4930	3556	9536	5980	4227	6380	2153	1473	27,4	2,522	<0,001	0,898	<0,001
Жир за 305, %	Хозяйство 5 (5 лакт.)	87	4,05	0,026	3,97	3,68	5,01	1,33	3,89	4,17	0,28	0,243	6	2,47	<0,001	0,878	<0,001
Жир за 305, кг	Хозяйство 4 (2 лакт.)	855	260,5	1,74	259	146,2	493	346,8	223	294	71,1	51	19,6	2,448	<0,001	0,982	<0,001
Жир за 305, %	Хозяйство 12 (3 лакт.)	266	3,83	0,005	3,82	3,66	4,22	0,56	3,77	3,88	0,11	0,082	2,1	2,43	<0,001	0,943	<0,001
Жир за 305, %	Хозяйство 6 (3 лакт.)	209	3,84	0,011	3,82	3,6	4,36	0,76	3,7	3,95	0,25	0,16	4,2	2,41	<0,001	0,956	<0,001
Жир за 305, %	Хозяйство 7 (2 лакт.)	469	4,22	0,005	4,21	3,76	4,63	0,87	4,15	4,29	0,14	0,109	2,6	2,401	<0,001	0,974	<0,001
Удой за 305, кг	Хозяйство 6 (3 лакт.)	209	8497	116,7	8636	3900	12453	8553	7362	9865	2503	1687	19,9	2,391	<0,001	0,967	<0,001
Жир за 305, %	Хозяйство 15 (6 лакт.)	34	3,85	0,029	3,83	3,66	4,54	0,88	3,75	3,91	0,16	0,171	4,4	2,316	<0,001	0,727	<0,001
Жир за 305, кг	Хозяйство 12 (1 лакт.)	915	309,2	1,57	313,8	150,1	456,2	306,1	280	336,4	56,4	47,4	15,3	2,3	<0,001	0,991	<0,001
Жир за 305, кг	Хозяйство 4 (3 лакт.)	435	260,4	2,45	253,5	129,5	432,6	303,1	223,6	290,7	67,1	51,1	19,6	2,27	<0,001	0,978	<0,001
Жир за 305, %	Хозяйство 15 (5 лакт.)	56	3,84	0,024	3,81	3,61	4,61	1	3,73	3,92	0,182	0,177	4,6	2,249	<0,001	0,801	<0,001
Жир за 305, кг	Хозяйство 16 (3 лакт.)	736	503,5	2,71	510,3	195,6	728,5	532,9	458	552	94	73,6	14,6	2,14	<0,001	0,988	<0,001
Жир за 305, кг	Хозяйство 6 (3 лакт.)	209	324,8	4,14	334,3	148,2	455,3	307,1	281	367,3	86,3	59,9	18,4	2,098	<0,001	0,969	<0,001
Жир за 305, %	Хозяйство 9 (3 лакт.)	559	3,83	0,005	3,82	3,6	4,22	0,62	3,75	3,92	0,17	0,12	3,1	2,09	<0,001	0,985	<0,001
Жир за 305, кг	Хозяйство 7 (4 лакт.)	150	207,9	3,14	203,2	148,3	343	194,7	180,4	225,6	45,2	38,5	18,5	2,006	<0,001	0,947	<0,001
Жир за 305, кг	Хозяйство 8 (1 лакт.)	889	212,1	1,29	210,9	131,1	341,5	210,4	182	238,4	56,5	38,4	18,1	1,955	<0,001	0,989	<0,001
Жир за 305, %	Хозяйство 10 (5 лакт.)	99	3,84	0,009	3,84	3,69	4,01	0,32	3,76	3,92	0,158	0,089	2,3	1,912	<0,001	0,951	0,002
Жир за 305, %	Хозяйство 16 (2 лакт.)	1445	4,02	0,006	4,02	3,6	4,74	1,14	3,87	4,16	0,29	0,216	5,4	1,871	<0,001	0,99	<0,001
Жир за 305, %	Хозяйство 4 (5 лакт.)	84	3,94	0,007	3,93	3,73	4,13	0,4	3,92	3,97	0,05	0,06	1,5	1,849	<0,001	0,926	<0,001
Жир за 305, %	Хозяйство 12 (5 лакт.)	44	3,85	0,015	3,9	3,66	4,04	0,38	3,76	3,91	0,146	0,096	2,5	1,833	<0,001	0,913	0,004
Удой за 305, кг	Хозяйство 7 (4 лакт.)	150	4925	72	4770	3552	7771	4219	4311	5392	1081	882	17,9	1,786	<0,001	0,954	<0,001
Удой за 305, кг	Хозяйство 8 (3 лакт.)	320	5989	78,7	5816	3509	9879	6370	4828	6955	2127	1408	23,5	1,721	<0,001	0,979	<0,001
Удой за 305, кг	Хозяйство 8 (1 лакт.)	889	5509	34,5	5474	3505	8703	5198	4735	6219	1484	1028	18,7	1,698	<0,001	0,988	<0,001
Жир за 305, кг	Хозяйство 10 (1 лакт.)	1936	318,1	1,28	321,2	148,2	564,6	416,4	281,9	356,3	74,4	56,4	17,7	1,65	<0,001	0,997	<0,001

Показатель	Хозяйство / номер лактации	n	Mean	SE	Me	Min	Max	Range	Q1	Q3	IQR	SD	Cv	AD	AD.p	SF	SF.p
Жир за 305, %	Хозяйство 8 (5 лакт.)	67	3,86	0,015	3,84	3,62	4,26	0,64	3,8	3,91	0,108	0,12	3,1	1,638	<0,001	0,93	0,002
Жир за 305, %	Хозяйство 11 (3 лакт.)	354	3,84	0,007	3,85	3,6	4,31	0,71	3,73	3,93	0,2	0,136	3,5	1,614	<0,001	0,977	<0,001
Жир за 305, кг	Хозяйство 11 (5 лакт.)	85	365,5	5,77	377,8	234,4	460,2	225,8	330,3	407,4	77	53,2	14,5	1,603	<0,001	0,947	0,002
Удой за 305, кг	Хозяйство 16 (4 лакт.)	327	12626	96,4	12848	4970	16535	11565	11608	13764	2157	1744	13,8	1,588	<0,001	0,972	<0,001
Жир за 305, %	Хозяйство 3 (3 лакт.)	407	3,89	0,006	3,9	3,6	4,2	0,6	3,81	3,98	0,17	0,114	2,9	1,57	<0,001	0,987	0,002
Удой за 305, кг	Хозяйство 4 (5 лакт.)	84	5889	117,8	5666	4314	8778	4464	4982	6826	1844	1080	18,3	1,556	<0,001	0,944	0,002
Жир за 305, %	Хозяйство 12 (4 лакт.)	130	3,85	0,008	3,86	3,67	4,07	0,4	3,8	3,9	0,1	0,086	2,2	1,552	0,001	0,97	0,007
Жир за 305, кг	Хозяйство 8 (3 лакт.)	320	230,1	2,96	226,1	133,7	377,8	244,1	187,2	267,2	80	53	23	1,551	0,001	0,981	0,001
Жир за 305, %	Хозяйство 11 (4 лакт.)	185	3,83	0,01	3,82	3,6	4,45	0,85	3,72	3,92	0,2	0,14	3,7	1,53	0,001	0,947	<0,001
Удой за 305, кг	Хозяйство 15 (3 лакт.)	160	9676	124,7	9921	5342	13960	8618	8408	10675	2267	1577	16,3	1,51	0,001	0,982	0,033
Жир за 305, кг	Хозяйство 4 (5 лакт.)	84	232,1	4,64	224,8	171,4	345,1	173,7	193,7	269	75,3	42,5	18,3	1,498	0,001	0,944	0,002
Жир за 305, кг	Хозяйство 1 (1 лакт.)	318	249,3	2,39	253,8	133,4	372,9	239,5	219,4	276,2	56,8	42,5	17,1	1,486	0,001	0,989	0,02
Жир за 305, %	Хозяйство 5 (7 лакт.)	17	4,1	0,084	4,04	3,82	5,18	1,36	3,86	4,15	0,29	0,346	8,4	1,477	0,001	0,733	0,001
Жир за 305, %	Хозяйство 10 (6 лакт.)	26	3,82	0,013	3,8	3,74	3,99	0,25	3,77	3,83	0,063	0,064	1,7	1,469	0,001	0,853	0,003
Жир за 305, кг	Хозяйство 16 (2 лакт.)	1445	489,3	1,75	491,9	272,7	694,9	422,2	447,4	533,9	86,5	66,7	13,6	1,451	0,001	0,997	0,003
Удой за 305, кг	Хозяйство 12 (1 лакт.)	915	8093	42,3	8161	3555	12270	8715	7265	8866	1601	1280	15,8	1,405	0,001	0,993	<0,001
Удой за 305, кг	Хозяйство 12 (4 лакт.)	130	8522	143,4	8323	4734	12266	7532	7678	9244	1566	1635	19,2	1,389	0,001	0,968	0,005
Жир за 305, кг	Хозяйство 12 (4 лакт.)	130	328	5,42	320,6	192,6	473,8	281,2	289,4	357,9	68,5	61,8	18,8	1,359	0,002	0,969	0,006
Жир за 305, кг	Хозяйство 12 (3 лакт.)	266	347,3	3,81	344	186,8	508,5	321,7	307,2	380,1	72,9	62,1	17,9	1,352	0,002	0,983	0,004
Удой за 305, кг	Хозяйство 13 (2 лакт.)	18	8381	559,8	7501	5675	11639	5964	6118	10822	4704	2375	28,3	1,323	0,001	0,848	0,01
Жир за 305, кг	Хозяйство 6 (1 лакт.)	983	307,5	1,5	309,7	166,1	445	278,9	276,6	340,8	64,2	47,1	15,3	1,274	0,003	0,996	0,006
Жир за 305, кг	Хозяйство 13 (2 лакт.)	18	325,6	23,3	301,2	207	457,6	250,6	230,7	422	191,3	98,7	30,3	1,247	0,002	0,858	0,013
Удой за 305, кг	Хозяйство 16 (1 лакт.)	2584	10338	25,1	10365	5644	14424	8780	9517	11191	1673	1276	12,3	1,224	0,003	0,998	0,008
Жир за 305, кг	Хозяйство 9 (3 лакт.)	559	413,1	2,41	417,7	192,5	580,5	388	381	452,8	71,8	56,9	13,8	1,224	0,003	0,988	0
Жир за 305, кг	Хозяйство 12 (6 лакт.)	20	357,7	8,03	355,9	268,7	442,9	174,2	339	371	32	35,9	10	1,215	0,003	0,847	0,006
Удой за 305, кг	Хозяйство 6 (4 лакт.)	75	8573	197,3	8936	4881	11154	6273	7617	9818	2201	1708	19,9	1,214	0,003	0,95	0,007
Жир за 305, кг	Хозяйство 13 (3 лакт.)	12	309,6	26,1	351,4	178,2	405,6	227,4	201,8	368,9	167,1	90,6	29,3	1,202	0,002	0,808	0,013
Удой за 305, кг	Хозяйство 15 (4 лакт.)	89	9763	175,2	10124	5227	12968	7741	8646	11011	2365	1652	16,9	1,168	0,004	0,959	0,008
Удой за 305, кг	Хозяйство 12 (3 лакт.)	266	9064	100,1	8938	4927	13522	8595	7999	10175	2176	1633	18	1,144	0,005	0,983	0,005
Удой за 305, кг	Хозяйство 6 (1 лакт.)	983	7953	43,1	7997	3842	11662	7820	7038	8917	1879	1350	17	1,138	0,006	0,996	0,016

Показатель	Хозяйство / номер лактации	n	Mean	SE	Me	Min	Max	Range	Q1	Q3	IQR	SD	Cv	AD	AD.p	SF	SF.p
Удой за 305, кг	Хозяйство 1 (1 лакт.)	318	6392	59,1	6514	3528	8974	5446	5716	7088	1372	1053	16,5	1,136	0,006	0,99	0,028
Удой за 305, кг	Хозяйство 13 (3 лакт.)	12	8022	691,1	9104	4451	10783	6332	5359	9329	3970	2394	29,8	1,091	0,005	0,829	0,021
Жир за 305, %	Хозяйство 7 (3 лакт.)	240	4,22	0,007	4,22	3,79	4,53	0,74	4,15	4,29	0,14	0,109	2,6	1,09	0,007	0,978	0,001
Удой за 305, кг	Хозяйство 10 (1 лакт.)	1936	8325	33,6	8402	3931	14988	11057	7350	9353	2002	1479	17,8	1,081	0,008	0,997	0,004
Жир за 305, %	Хозяйство 8 (6 лакт.)	26	3,84	0,025	3,82	3,6	4,2	0,6	3,78	3,86	0,084	0,125	3,3	1,071	0,007	0,894	0,014
Удой за 305, кг	Хозяйство 15 (6 лакт.)	34	9157	334,4	9698	3805	12287	8482	7268	10674	3406	1950	21,3	1,064	0,007	0,923	0,021
Жир за 305, кг	Хозяйство 7 (7 лакт.)	14	190,2	9,91	176,4	154,3	294,4	140,1	169	194,8	25,8	37,1	19,5	1,055	0,006	0,779	0,004
Жир за 305, кг	Хозяйство 6 (4 лакт.)	75	327	6,95	333,6	187,9	439,8	251,9	293,2	367,1	73,9	60,2	18,4	1,032	0,01	0,963	0,027
Удой за 305, кг	Хозяйство 12 (6 лакт.)	20	9261	213,4	9219	7030	10989	3959	8670	9779	1109	954,4	10,3	1,005	0,009	0,879	0,019
Жир за 305, кг	Хозяйство 16 (4 лакт.)	327	514,5	4,15	520,3	201,6	712,7	511,1	465,9	563,2	97,3	75	14,6	1	0,012	0,983	0,001
Жир за 305, кг	Хозяйство 7 (5 лакт.)	79	207,2	4,35	204	149,2	326,9	177,7	178,4	226,2	47,8	38,6	18,7	1	0,012	0,945	0,003
Удой за 305, кг	Хозяйство 8 (2 лакт.)	551	5845	51,3	5867	3508	8917	5409	4910	6665	1755	1205	20,6	0,993	0,013	0,989	0,001
Удой за 305, кг	Хозяйство 7 (7 лакт.)	14	4488	240,9	4131	3587	6981	3394	3894	4747	853,1	901,4	20,1	0,952	0,011	0,797	0,006
Удой за 305, кг	Хозяйство 8 (4 лакт.)	148	5483	99,6	5406	3501	8797	5296	4430	6292	1862	1212	22,1	0,945	0,016	0,974	0,007
Удой за 305, кг	Хозяйство 9 (3 лакт.)	559	10804	67,5	10891	5193	14679	9486	9823	11851	2028	1596	14,8	0,944	0,017	0,99	0,001
Жир за 305, кг	Хозяйство 15 (6 лакт.)	34	351,6	12	370,4	146,3	450,6	304,3	281,2	412	130,8	69,9	19,9	0,944	0,015	0,916	0,015
Жир за 305, кг	Хозяйство 8 (2 лакт.)	551	226,1	1,94	224,7	129,1	365	235,9	190	257,9	67,8	45,4	20,1	0,943	0,017	0,991	0,002
Удой за 305, кг	Хозяйство 7 (5 лакт.)	79	4896	99,4	4821	3582	7653	4071	4297	5332	1035	883,1	18	0,935	0,017	0,946	0,003
Удой за 305, кг	Хозяйство 10 (2 лакт.)	1295	9424	50,6	9517	3979	14383	10404	8220	10673	2453	1822	19,3	0,929	0,018	0,997	0,009
Жир за 305, %	Хозяйство 16 (3 лакт.)	736	4,03	0,008	4,03	3,6	4,65	1,05	3,88	4,17	0,29	0,208	5,2	0,925	0,019	0,992	0,001
Жир за 305, %	Хозяйство 11 (5 лакт.)	85	3,8	0,014	3,79	3,61	4,23	0,62	3,71	3,88	0,173	0,125	3,3	0,92	0,019	0,938	0,001
Жир за 305, кг	Хозяйство 6 (5 лакт.)	37	314,9	12,8	326,7	170,3	468,7	298,4	282,3	362	79,7	78	24,8	0,917	0,018	0,948	0,077
Жир за 305, %	Хозяйство 10 (7 лакт.)	11	3,79	0,018	3,78	3,74	3,93	0,19	3,76	3,8	0,04	0,059	1,6	0,903	0,014	0,803	0,015
Жир за 305, кг	Хозяйство 10 (2 лакт.)	1295	358,5	1,91	362,1	157,6	543,3	385,7	312	405,2	93,2	68,9	19,2	0,888	0,023	0,997	0,013
Жир за 305, кг	Хозяйство 12 (5 лакт.)	44	358,2	10,1	347	229,2	499,3	270,1	314,6	392,7	78,1	66,8	18,6	0,886	0,021	0,944	0,035
Жир за 305, %	Хозяйство 1 (4 лакт.)	59	3,9	0,021	3,87	3,66	4,45	0,79	3,77	4,02	0,247	0,16	4,1	0,865	0,025	0,94	0,008
Удой за 305, кг	Хозяйство 5 (2 лакт.)	368	6251	50,8	6406	3765	8657	4892	5547	6923	1376	974,4	15,6	0,858	0,027	0,992	0,051
Удой за 305, кг	Хозяйство 4 (4 лакт.)	229	6580	85,5	6572	4433	11805	7372	5605	7382	1778	1294	19,7	0,854	0,028	0,973	<0,001
Жир за 305, %	Хозяйство 6 (5 лакт.)	37	3,85	0,03	3,84	3,62	4,3	0,68	3,7	3,97	0,273	0,185	4,8	0,85	0,026	0,925	0,018
Удой за 305, кг	Хозяйство 11 (5 лакт.)	85	9626	158,3	9898	5963	12066	6103	8704	10664	1960	1460	15,2	0,841	0,029	0,968	0,032

Показатель	Хозяйство / номер лактации	n	Mean	SE	Me	Min	Max	Range	Q1	Q3	IQR	SD	Cv	AD	AD.p	SF	SF.p
Жир за 305, кг	Хозяйство 15 (3 лакт.)	160	369	4,45	376,1	203,1	528	324,9	330,7	403,2	72,4	56,3	15,3	0,841	0,03	0,989	0,208
Жир за 305, %	Хозяйство 3 (4 лакт.)	186	3,89	0,007	3,89	3,69	4,19	0,5	3,82	3,95	0,13	0,092	2,4	0,84	0,03	0,979	0,009
Жир за 305, %	Хозяйство 3 (5 лакт.)	64	3,9	0,013	3,91	3,73	4,16	0,43	3,8	3,98	0,176	0,108	2,8	0,839	0,029	0,963	0,054
Жир за 305, %	Хозяйство 3 (6 лакт.)	14	3,96	0,018	3,95	3,89	4,15	0,26	3,92	3,98	0,061	0,066	1,7	0,825	0,024	0,812	0,009
Жир за 305, кг	Хозяйство 15 (4 лакт.)	89	372,4	6,12	383,8	198,6	487,3	288,7	335,7	415,3	79,7	57,7	15,5	0,805	0,036	0,967	0,024
Жир за 305, %	Хозяйство 6 (4 лакт.)	75	3,83	0,017	3,83	3,6	4,21	0,61	3,73	3,92	0,188	0,149	3,9	0,793	0,038	0,955	0,012
Жир за 305, кг	Хозяйство 16 (1 лакт.)	2584	407,7	1,05	408,1	221,3	607,4	386,1	372	442,8	70,8	53,6	13,1	0,791	0,04	0,998	0,004

Приложение Д: Уникальные генотипы по локусам

Локус	Аллели	N	Локус	Генотипы	Количество	Локус	Генотипы	N	Локус	Генотипы	N
BM1818	262/266	3846	TGLA122	143/149	1460	BM2113	125/135	1913	TGLA227	89/91	1232
	266/266	2276		143/143	1027		127/135	1350		89/97	1181
	262/262	1449		143/163	994		125/127	1292		91/97	991
	264/266	873		143/161	696		135/139	910		89/89	478
	262/264	595		143/151	613		125/139	828		83/89	432
	264/264	57		149/163	546		135/135	729		91/91	430
BM1824	178/188	2018		143/171	391	127/139	712	87/89		403	
	182/188	1808		149/161	322	125/125	674	87/91		363	
	188/188	1772		149/149	304	127/127	481	83/91		347	
	180/188	1142		149/151	300	139/139	202	97/97		346	
	178/182	812		161/163	296	127/125	1	87/97		334	
	178/180	558		149/171	238	148/150	3167	89/103		327	
	180/182	505	151/163	235	148/148	1805	91/103	309			
	178/178	365	163/171	204	150/150	1385	83/97	299			
	182/182	305	163/163	194	140/148	849	97/103	273			
	180/180	174	151/161	142	140/150	783	83/87	123			
SPS115	248/248	4769	161/171	107	140/140	101	87/103	108			
	248/252	2595	151/171	91	219/219	1427	83/103	91			
	252/252	298	151/151	49	217/219	1361	87/87	46			
INRA023	206/214	1903	161/161	45	209/219	1174	103/103	37			
	206/210	1496	171/171	19	209/217	503	83/83	29			
	210/214	1396	162/162	221	213/219	322	117/117	1295			
	214/214	885	160/162	133	217/217	242	117/125	1012			
	206/206	838	162/168	121	213/217	135	117/127	657			
	202/214	726	160/160	101	209/209	115	125/127	210			
			160/168	84	209/213	103	125/125	107			

Локус	Аллели	N	Локус	Генотипы	Количество	Локус	Генотипы	N	Локус	Генотипы	N
INRA023	202/206	701	TGLA53	158/162	80	ETH10	213/213	16	ETH3	127/127	36
	202/210	567		160/176	76		219/213	1		127/125	1
	210/210	487		162/176	76						
	202/202	127		168/168	68						
	214/210	1		168/176	46						
TGLA126	115/117	3900		158/160	43						
	117/117	2736		158/158	42						
	115/115	1190		158/176	31						
	117/121	742		158/168	30						
	115/121	465		176/176	27						
	121/121	44		162/160	1						

* N – количество особей в популяции