

Генетика, геномика, фенотип

УДК 636.2:636.082.2:577.2

doi: 10.15389/agrobiology.2025.6.986rus

ПОИСК ОТПЕЧАТКОВ СЕЛЕКЦИИ В ГЕНОМАХ СОВРЕМЕННОГО СЕВЕРНОГО И СТЕПНОГО СКОТА В СРАВНЕНИИ С ОБРАЗЦАМИ КОНЦА XIX—НАЧАЛА XX ВЕКОВ***А.С. АБДЕЛЬМАНОВА[✉], Т.Е. ДЕНИСКОВА, А.В. ШАХИН, А.А. НИКОЛАЕВ,
Г.О. АБРАМОВ, О.И. БОРОНЕЦКАЯ, В.А. БАГИРОВ, В.И. ТРУХАЧЕВ,
Н.А. ЗИНОВЬЕВА**

Развитие технологии полногеномного секвенирования (WGS) открывает новые возможности в изучении изменчивости геномов сельскохозяйственных животных под воздействием естественного и искусственного отбора. Анализ WGS-данных позволяет выявлять ключевые гены и геномные регионы, ответственные за формирование адапционных свойств к специфическим природно-климатическим условиям разведения. Селекция и генетический дрейф — ключевые источники генетической дифференциации между представителями исходных и современных пород. Изучение динамики изменений генофонда местных пород актуально, поскольку именно местные породы — это носители редких и ценных генетических вариантов, которые могут стать приоритетными в случае смены требований рынка и при климатических изменениях. В настоящей работе впервые описаны результаты идентификации участков генома, закрепившихся вследствие отбора у контрастных групп крупного рогатого скота, приспособленных к значительно различающимся условиям внешней среды, проведена структурная и функциональная аннотация локализованных в них генов. Выявлены гены, общие для современных и предковых популяций скота исследуемых групп. Цель исследований состояла в сравнительном анализе геномов музейных и современных образцов крупного рогатого скота разных пород на основании данных полногеномного секвенирования для выявления участков генома, подвергшихся действию естественного и искусственного отбора. Материалом для изучения служили музейные образцы черепов крупного рогатого скота из Музея животноводства им. Е.Ф. Лискуна (РГАУ — МСХА им. К.А. Тимирязева, г. Москва), датированные концом XIX—началом XX века, а также образцы от племенных животных, хранящиеся в биологической коллекции Национального центра генетических ресурсов сельскохозяйственных животных ФГБНУ ФИЦ ВИЖ им. академика Л.К. Эрнста. Группа северного скота была представлена образцами холмогорской и ярославской пород, группа степного скота — образцами калмыцкого скота. Выборку музейных образцов ($n = 20$) составили 15 образцов от северного скота (холмогорский скот, $n = 8$; ярославский скот, $n = 7$) и 5 образцов степного (калмыцкого) скота. Выборка современных образцов включала соответственно 17, 11 и 12 образцов названных выше пород. Итоговый набор данных включал генотипы по 1 615 259 SNPs (Single Nucleotide Polymorphism). Выделение ДНК из музейных образцов проводили с использованием методики, описанной ранее, из современных образцов — с помощью набора ДНК-Экстран 2 (ЗАО «Синтол», г. Москва). Библиотеки для полногеномного секвенирования с использованием технологии NGS готовили с помощью наборов TruSeq DNA Nano Library Prep kit («Illumina, Inc.», США) и Accel-NGS® 2S Plus DNA Library Kit (IDT). Секвенирование выполняли на секвенаторе NovaSeq 6000 (2×150 bp). Для выявления участков генома, находящихся под давлением отбора, применили три метода: парное сравнение популяций на основании генетических дистанций F_{ST} с использованием скользящего окна, анализ межпопуляционной гомозиготности гаплотипов (XP-EHH) и определение регионов избыточной гомозиготности (островки ROH). Показано отсутствие в геноме изучаемых пород скота перекрывающихся геномных регионов, находящихся под давлением отбора по результатам сравнительного анализа F_{ST} и XP-EHH предковых и современных популяций для пар (холмогорская + ярославская)—калмыцкая, что указывает на существенные различия в целях и направлениях селекции в различные периоды формирования пород. Идентифицированы семь островков ROH, перекрывающихся в музейных и современных образцах холмогорского и ярославского скота, которые локализованы в известных QTL удоя, состава молока, энергии роста, размера тела и фертильности, что свидетельствует о длительном использовании этих показателей в качестве целевых параметров отбора при разведении пород. Выявлены гены, общие для современных и предковых популяций скота исследуемых групп. Идентифицированные нами геномные регионы, подверженные действию отбора в предковых популяциях и перекрывающиеся с «отпечатками» селекции в геноме современных представителей пород, могут быть использованы при отборе животных для программ сохранения генетических ресурсов и аутентичных геномных компонентов.

Ключевые слова: крупный рогатый скот, холмогорская порода, ярославская порода, калмыцкая порода, музейные образцы, полногеномное секвенирование, давление отбора, гены-кандидаты.

* Исследования выполнены при поддержке Российского научного фонда, проект № 19-76-20012.

Крупный рогатый скот (КРС) — ценная модель для изучения изменения генома, происходящего под действием таких процессов, как одомашнивание, адаптация, селекция и скрещивание (1, 2). Результатом совместного воздействия этих факторов стало формирование современных пород и популяций, значительно различающихся по молочной и мясной продуктивности, плодовитости, окраске шерсти, размерам тела, форме рогов или их отсутствию, а также по другим фенотипическим особенностям (3-5).

С помощью секвенирования последовательностей полных геномов (whole-genome sequencing, WGS) становится возможным изучать редкие генетические варианты, которые отсутствуют в коммерчески доступных ДНК-чипах (6), и охватить весь спектр изменчивости, присутствующий в геноме (7). Использование результатов секвенирования, полученных с глубоким покрытием генома, снижает вероятность ошибочного определения аллелей в целевых позициях (8). Более того, анализ WGS позволяет более точно идентифицировать короткие сегменты гомозиготности (runs of homozygosity, ROH) (9) и выявлять отпечатки селекции с более высоким разрешением, чем при применении микроматриц (10). Использование в исследованиях большого числа SNP (single nucleotide polymorphism) маркеров (более 1 млн SNPs) позволяет компенсировать малый размер выборки (11), что особенно актуально для музейных и археологических образцов, размер выборки которых зависит от числа имеющихся экспонатов и находок (12).

На основе WGS подтверждаются геномные регионы, находящиеся под давлением селекции и идентифицированные с помощью ДНК-чипов у различных пород КРС. Эти регионы содержат гены, ассоциированные с размерами тела и молочной продуктивностью (*PLAG1*, *NCAPG*, *LCORL*, *LAP3*), эффективностью использования корма и липидным метаболизмом (*R3HDM1*, *AOXI*), а также окраской шерстного покрова (*MC1R*, *KIT*) (9, 10).

Н.А. Mulim с соавт. (13), изучая полные геномы представителей европейских, зебувидных и кроссбредных пород скота, идентифицировали 1662 позиционных гена-кандидата, вовлеченных в регуляцию 400 биологических процессов и 319 молекулярных функций. Островки ROH в популяциях пород Gig и Brahman содержали гены, связанные с меланогенезом, а в популяции джерсейской породы — гены, влияющие на молочную продуктивность и задействованные в кальциевых сигнальных путях. Х. Ма с соавт. (14) на основе анализа полногеномных последовательностей 50 представителей мясной китайской породы Bohai Black выявили регионы, имеющие признаки селекции и содержащие гены, связанные с формированием мышечной ткани (*ITGA9*, *ENAH*, *CAPG*) и отложением жира (*TBC1D1*, *CYB5R4*, *TUSC3* и *EPS8*). S. Voitard с соавт. (15) изучали последовательности у представителей испанской породы Asturiana de los Valles, которая изначально использовалась как универсальная, а затем в течение нескольких поколений подверглась селекции на мясную продуктивность (производство говядины). В результате были выявлены гены, влияющие на качество туши (*MSTN*, *FLRT2*, *CRABP2*) и определяющие пригодность к разведению в условиях полунтенсивных технологий, включая гены, вовлеченные в иммунный ответ (*GIMAP7* и *TICAM1*), а также гены, участвующие в регуляции обонятельной функции (*OR2D2* и *OR2D3*). В геноме китайской мясной породы Pinan обнаружены регионы, которые содержат гены-кандидаты, ассоциированные с ростом, развитием, эмбриогенезом и влияющие на иммунитет (16). М.У. Nawaz с соавт. (17), изучая полные геномы специализированных мясных пород, установили, что отпечатки селекции в абердин-ангусской породе и породе Hanwoo различались. Так, у ангусов выявлены геномные регионы, находящиеся под давлением селекции и содержащие позиционные гены-

кандидаты, которые связаны с ростом и развитием, иммунитетом, репродуктивными качествами, эффективностью использования корма и адаптацией к окружающей среде. Гены-кандидаты, идентифицированные в породе Hanwoo, регулировали качество мяса, отложение жира, метаболизм холестерина и синтез липидов (17).

Анализ WGS-данных позволяет выявлять ключевые гены и геномные регионы, ответственные за формирование адаптаций к специфическим природно-климатическим условиям разведения. Изучая полные геномы креольского крупного рогатого скота, J.A. Ward с соавт. (18) обнаружили «отпечатки» селекции, связанные с множеством адаптивных признаков, — термостойчивостью, иммунитетом и различными вариантами окраса шерсти и кожи (18). M. Tiwari с соавт. (19) исследовали генетические механизмы, лежащие в основе приспособленности аборигенного крупного рогатого скота к гипоксии в горном районе Индии. В качестве контрастной группы сравнения авторы выбрали представителей породы Sahiwal, обитающей на равнинах Центральной Индии. Были использованы два биоинформационных подхода — поиск сегментов гомозиготности (ROH) и расчет индекса фиксации (Fst). Гены-кандидаты, связанные с адаптацией к высокогорью, включая *FIA*, *VEGFC*, *ZEB1* и *SENP2*, были идентифицированы с помощью обоих подходов. Дополнительно посредством расчета Fst был выявлен ген *EPAS1* — наиболее значимый функциональный кандидат.

Селекция и генетический дрейф — ключевые источники генетической дифференциации между представителями исходных и современных пород (20). Изучение динамики изменений генофонда местных пород актуально, поскольку именно местные породы — это носители редких и ценных генетических вариантов, которые могут стать приоритетными в случае смены требований рынка и при климатических изменениях (20). A. Moscarelli с соавт. (21) сообщили о наличии геномных регионов, которые различаются между исходными и современными популяциями бурого скота. I. Hulsegge с соавт. (20) исследовали генетическое разнообразие в геноме быков фризской породы, относящихся к предковой (1961–1989 годы) и современной (2003–2015 годы) популяциям, и выявили сокращение генетического разнообразия.

В настоящей работе впервые описаны результаты идентификации участков генома, закрепившихся в процессе отбора у контрастных групп крупного рогатого скота (северного и степного), приспособленных к значительно различающимся условиям внешней среды, проведена структурная и функциональная аннотация локализованных в этих участках генов. Выявлены гены, общие для современных и предковых популяций скота исследуемых групп.

Цель работы состояла в сравнительном анализе геномов музейных и современных образцов от крупного рогатого скота разных пород на основании данных полногеномного секвенирования для выявления участков генома, подвергшихся действию естественного и искусственного отбора.

Методика. Образцы черепов крупного рогатого скота, датированные концом XIX—началом XX века, были получены из Музея животноводства им. Е.Ф. Лискуна РГАУ — МСХА им. К.А. Тимирязева (г. Москва), образцы от современных племенных животных — из биологической коллекции Национального центра генетических ресурсов сельскохозяйственных животных ФГБНУ ФИЦ ВИЖ им. академика Л.К. Эрнста.

Группа северного скота (North Cattle, NC) была представлена образцами холмогорской и ярославской пород, ведущих свое происхождение от северного великорусского скота и относящихся к краниологическому типу

Bos taurus primigenius, группа степного скота (Steppe Cattle, SC) — образцами калмыцкого скота, относящегося к краниологическому типу *Bos indicus Kuleschowi Sakowsky*. Выборка музейных образцов (M, $n = 20$) включала образцы NC ($n = 15$), в том числе 8 гол. холмогорского (KHLMH) и 7 гол. ярославского (YRSLH) скота, а также 5 гол. калмыцкого скота (KALMH), относящегося к группе SC. Выборка современных образцов включала соответственно 17 образцов холмогорского (KHLM), 11 образцов ярославского (YRSL) и 12 образцов калмыцкого скота (KALM).

ДНК из музейных образцов выделяли по методике, описанной ранее (12), из современных образцов — с помощью набора ДНК-Экстрэн 2 (ЗАО «Синтол», г. Москва). Библиотеки для полногеномного секвенирования с использованием технологии NGS готовили с помощью наборов TruSeq DNA Nano Library Prep kit («Illumina, Inc.», США) и Accel-NGS® 2S Plus DNA Library Kit («Integrated DNA Technologies, Inc.», США). Секвенирование выполняли на секвенаторе NovaSeq 6000 (2×150 bp). Полученные риды выравнивали на секвенаторе референсного генома KPC по сборке ARS_UMD 3.1.1 (https://www.ncbi.nlm.nih.gov/assembly/GCF_000003055.6/) с использованием инструментов bwa-mem2 и SAMtools. В анализ были включены только SNPs, генотипированные более чем у 80 % образцов (--mind 0.2) с частотой минорного аллеля более 5 % (--maf 0.05), и образцы, у которых было генотипировано не менее чем 90 % локусов (--geno 0.10). Используемые фильтры (--geno, --maf и --mind) применялись последовательно, чтобы сохранить максимальное число успешно генотипированных локусов, общих для большинства образцов. Итоговый набор данных включал генотипы по 1 615 259 SNPs.

Для выявления участков генома, находящихся под давлением отбора, применяли три метода: попарное сравнение популяций на основании генетических дистанций F_{ST} с использованием скользящего окна, которые отражают различия в частотах альтернативных аллелей, закрепившихся в каждой из исследуемых групп (22); анализ межпопуляционной гомозиготности гаплотипов (cross-population extended haplotype homozygosity, XP-EHH), который позволяет определить локусы, где исследуемый аллель приблизился или достиг фиксации в одной группе, но остается полиморфным в другой (23); определение регионов избыточной гомозиготности (островки ROH), что дает возможность выявлять участки с меньшей полиморфностью, чем в среднем по геному, возникшие в результате селекционного давления или событий инбридинга (24, 25).

Анализ F_{ST} при попарном сравнении пород проводили в программе VCFtools (26) на основании средних значений F_{ST} для скользящего окна, чтобы уменьшить вариации значений от локуса к локусу (27), поскольку изменения в процессе селекционного давления захватывают не только целевой регион, но и связанные с ним участки. Размер окна составлял 100 Kb с шагом 50 Kb (zF_{ST}) (--fst-window-size 100000, --fst-window-step 50000) и был выбран, исходя из предположения, что при большем размере окна некоторые участки, подвергнувшиеся давлению отбора, могут быть нивелированы из-за нарушения сцепления (28). Для предотвращения ложноположительных результатов и корректного сравнения значений между парами пород значения F_{ST} были стандартизированы и рассчитаны z -оценки. За участки, наиболее сильно подвергшиеся давлению отбора, были приняты регионы, содержащие SNPs, для которых значения zF_{ST} входили в 0,1 % максимальных значений, как было предложено ранее другими авторами (29, 30). Как отмечалось выше, небольшой размер групп, включенных в анализ, компенсировался значительным числом анализируемых локусов (11).

Анализ XP-EHH проводили с использованием R-пакета rehh (31)

для тех же пар пород, для которых определяли F_{st} , с целью корректного сравнения результатов и выявления общих регионов.

Для идентификации следов селекции в геноме на первом этапе проводили поиск сегментов ROH методом последовательных прогонов (32), реализованным в пакете R detectRUNS (33), с использованием следующих параметров: минимальная длина сегмента — 0,5 Mb, минимальное число SNPs — 456, допускается 5 не генотипированных и 1 гетерозиготный вариант. Островками ROH считали участки ROH, перекрывающиеся на 0,3 Mb и более для более чем 50 % образцов в каждой исследуемой популяции. Принимая во внимание снижение качества секвенирования музейных образцов, мы уменьшили порог перекрывающегося участка в островках ROH для них с 0,3 Mb до 0,003 Mb.

Для структурной аннотации отбирали области генома, локализованные в пределах окна 0,4 Mb (от 0,2 Mb downstream до 0,2 Mb upstream от выявленного SNP или блока из SNPs) для SNPs, выявленных с помощью метода F_{st} , и идентифицированные регионы XP-EHH и островки ROH. Идентификацию генов, полностью или частично локализованных внутри таких регионов, проводили с использованием инструментов VEP (34) и BioMart (35), реализованных на платформе ENSEMBL, выпуск 94 (36). Списки генов для каждой геномной области, полученные с помощью инструмента BioMart, были сопоставлены с опубликованными данными для определения основных генов-кандидатов и других генов, представляющих интерес. В случаях, когда возникала необходимость преобразовать координаты генома между ныне действующей референсной сборкой генома *Bos taurus* (ARS-UCD1.2) и более ранней версией (UMD3.1.1) для сравнения с другими исследованиями, применяли инструмент UCSC liftOver (37).

Для функциональной аннотации и анализа обогащения списка генов генами с определенной функцией в терминах генной онтологии (GO) использовали базу данных для аннотации, визуализации и интегрированного обнаружения (DAVID) v6.8 (38, 39). Значимыми считали кластеры аннотаций с оценкой обогащения более 1,12, что соответствует $p < 0,05$. Для более точного понимания связей между выявленными генами и хозяйственно полезными признаками провели поиск локусов количественных признаков, расположенных внутри генов, с использованием базы данных cattleQTLdb (<http://www.animalgenome.org/cgi-bin/QTLdb/index>) (40) по сборке генома крупного рогатого скота ARS_UCD 1.2 с помощью оригинального R-скрипта. Анализ фенотипов, потенциально связанных с идентифицированными генами-кандидатами, проводили на основании источников литературы, размещенных в базе данных PubMed Национального центра биотехнологической информации (<https://pubmed.ncbi.nlm.nih.gov/>; NCBI Resource Coordinators, 2016). Информацию для аннотации некоторых генов получали из Gene Cards (<http://www.genecards.org/>) (41) и UniProt (<https://www.uniprot.org/>; The UniProt Consortium, 2021).

Результаты. Как следует из полученных данных, мы не выявили перекрывающихся регионов zF_{st} как в современных, так и предковых популяциях северного и калмыцкого скота (рис. 1, табл. 1). У современного скота такие регионы были локализованы на ВТА4, ВТА6, ВТА14 и ВТА16, в то время как у предкового скота — на ВТА5, ВТА7, ВТА9, ВТА10, ВТА13, ВТА18, ВТА20 и ВТА24. Наибольшее число регионов zF_{st} у современного скота было локализовано на ВТА16 (16 регионов), у предкового — на ВТА18 (7 регионов) и ВТА9 (6 регионов).

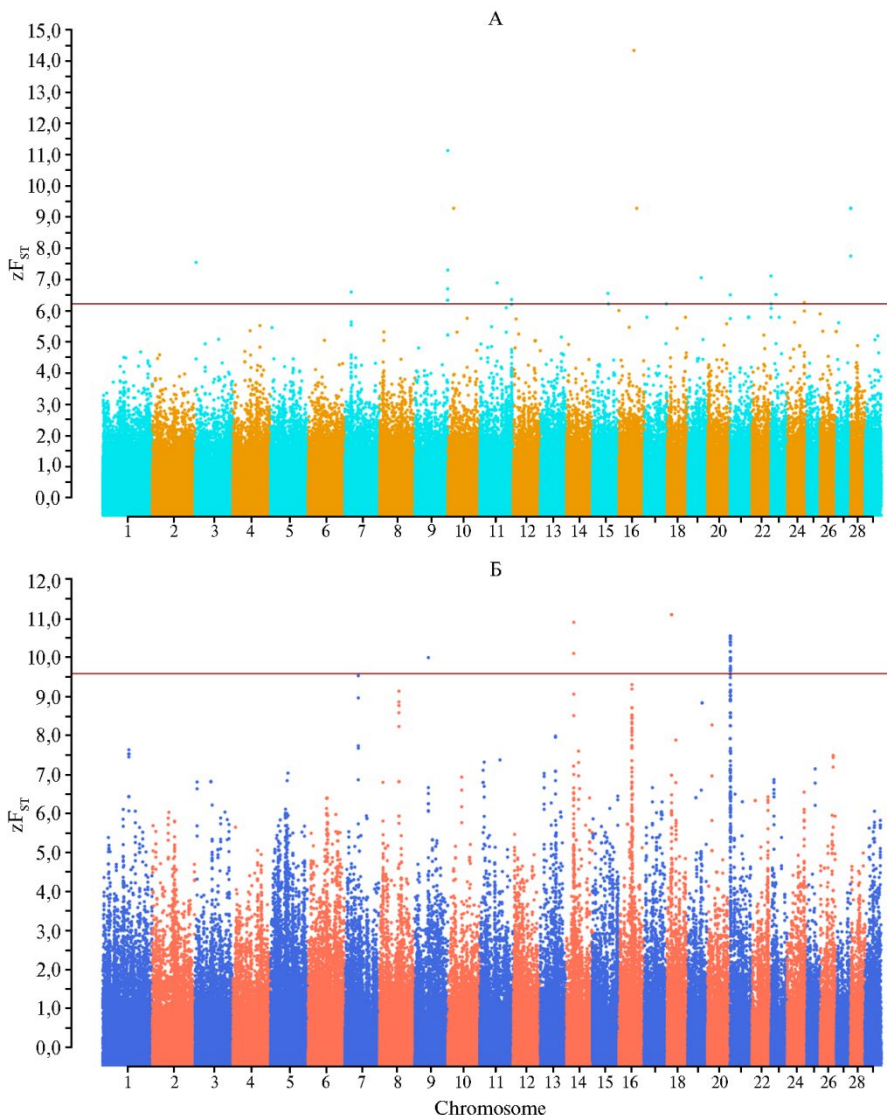


Рис. 1. Значения zF_{ST} при попарном сравнении северного и калмыцкого скота (*Bos taurus*), рассчитанные на основании генотипов по 1 615 259 SNPs: А — музейные образцы, Б — современные образцы. SNP локализованы по оси X в соответствии с их расположением на хромосоме; горизонтальная линия показывает порог, который был установлен на уровне 0,1 % блоков SNP с наиболее высокими значениями zF_{ST} . Описание образцов см. в разделе «Методика».

1. Регионы, входящие в 0,1 % по среднему значению zF_{ST} , рассчитанному для SNPs, локализованных внутри скользящего окна длиной 50 Кб, при попарном сравнении северного и степного скота (*Bos taurus*) музейных образцов (М) и современных популяций (С)

ВТА	Группа	Позиция, Мб		Число SNPs в окне	F_{ST} (среднее)	z-Оценка
		начало	конец			
3	М	3,10	3,15	6	0,290	7,544
7	М	19,90	19,95	19	0,255	6,592
	С	42,74	42,79	2	0,684	9,999
9	М	104,12	104,17	3	0,245	6,333
		104,13	104,18	3	0,245	6,333
	104,14	104,19	3	0,245	6,333	
	104,15	104,20	4	0,258	6,696	
	104,52	104,57	1	0,424	11,129	
10	М	104,53	104,58	3	0,281	7,298
		18,77	18,82	1	0,355	9,278

11	M	55,48	55,53	13	0,266	6,887	
		101,50	101,55	20	0,246	6,353	
14	C	23,24	23,29	14	0,741	10,905	
		23,25	23,30	12	0,691	10,105	
15	M	49,92	49,97	1	0,253	6,549	
		51,42	51,47	1	0,241	6,216	
		51,43	51,48	1	0,241	6,216	
		48,15	48,20	1	0,544	14,339	
16	M	56,99	57,04	1	0,355	9,278	
		70,73	70,78	1	0,241	6,216	
17	M	70,74	70,79	1	0,241	6,216	
		14,66	14,71	1	0,753	11,101	
18	C	14,67	14,72	1	0,753	11,101	
		14,68	14,73	1	0,753	11,101	
19	M	43,24	43,29	2	0,272	7,050	
		0,56	0,61	28	0,251	6,503	
	C	0,83	0,88	8	0,694	10,152	
		0,84	0,89	8	0,694	10,152	
		0,85	0,90	8	0,709	10,404	
	21		0,86	0,91	8	0,679	9,914
			0,87	0,92	6	0,662	9,657
			0,92	0,97	1	0,709	10,394
			0,93	0,98	2	0,715	10,487
			0,94	0,99	2	0,715	10,487
			0,95	1,00	3	0,719	10,553
			0,96	1,01	3	0,719	10,553
			1,05	1,10	6	0,704	10,323
			1,06	1,11	5	0,709	10,397
			1,23	1,28	6	0,683	9,978
1,24			1,29	7	0,684	9,999	
1,25			1,30	5	0,681	9,949	
1,26			1,31	8	0,666	9,721	
1,27	1,32	6	0,667	9,722			
23	M	1,28	1,33	7	0,670	9,779	
		1,29	1,34	8	0,663	9,665	
		0,09	0,14	2	0,274	7,109	
		0,60	0,65	2	0,241	6,216	
		15,98	16,03	10	0,252	6,513	
		54,08	54,13	14	0,242	6,255	
		45,54	45,59	1	0,355	9,278	
27	M	45,55	45,60	1	0,355	9,278	
		45,56	45,61	2	0,298	7,747	

Примечание. ВТА — номер аутосомы крупного рогатого скота; Fst (среднее) — среднее значение Fst, рассчитанное для SNPs, расположенных внутри скользящего окна размером 100 Kb; z-оценка — стандартизированные значения Fst. Описание образцов см. в разделе «Методика».

Как и в случае zFst, мы не выявили перекрывающихся ХР-ЕНН регионов как в современных, так и в предковых популяциях при сравнении северного и калмыцкого скота (рис. 2, табл. 2). У современного скота такие регионы были локализованы на ВТА1, ВТА3, ВТА5, ВТА10, ВТА14, ВТА15, ВТА20, ВТА21 и ВТА22, у предкового скота — на ВТА5, ВТА8, ВТА11, ВТА15, ВТА18, ВТА20, ВТА22, ВТА23, ВТА25 и ВТА26. У современного скота два региона ХР-ЕНН локализовались на ВТА1, у предкового скота — на ВТА8 и ВТА25. На остальных хромосомах, на которых выявили регионы ХР-ЕНН, мы идентифицировали по одному такому региону.

Было показано наличие только одного перекрывающегося региона, идентифицированного обоими методами, при сравнении музейных образцов от северного и калмыцкого скота на ВТА11 (позиции 101,50-101,55 Мб для метода zFst и 100,30-102,20 Мб — для метода ХРЕНН). Структурная аннотация этого региона показала локализацию в нем 23 генов, в том числе два гена (*U6* и *SNORD62*) не относились к протеин-кодирующим, а были связаны с модификациями рибосомных РНК в ядрышке (табл. 3). Интересно отметить, что, согласно данным базы OMIA, ген *ASS1* связан с цитруллинемией (запись OMIA000194) — моногенным заболеванием, характеризующимся нарушением обмена мочевины и вызывающим гибель новорожденных телят на 3-и–5-е сут жизни.

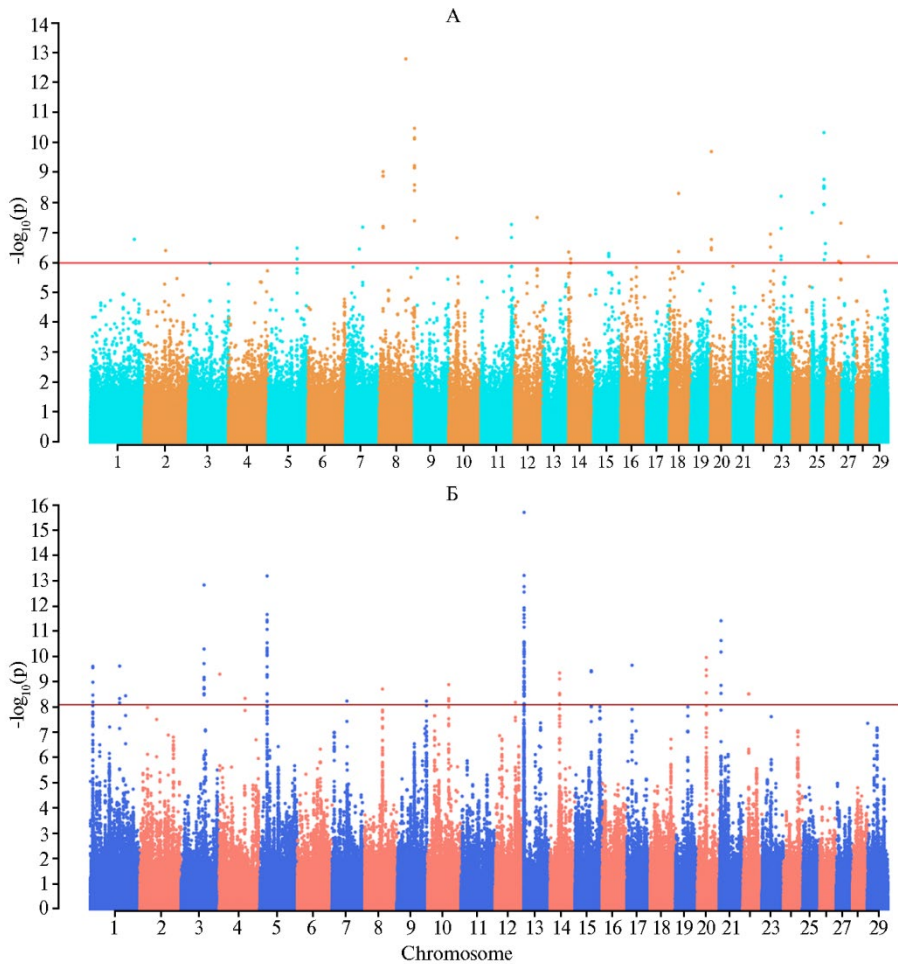


Рис. 2. Регионы ХР-ЕНН, выявленные при попарном сравнении северного и калмыцкого скота (*Bos taurus*) на основании анализа генотипов по 1 615 259 SNPs: А — музейные образцы, Б — современные образцы. SNPs локализованы по оси X в соответствии с их расположением на хромосоме; горизонтальная линия показывает порог, который был установлен на уровне 0,1 % блоков SNPs с наиболее высокими значениями $-\log_{10}(p)$. Описание образцов см. в разделе «Методика».

2. Характеристика регионов ХР-ЕНН, идентифицированных на основании 1 615 259 SNPs, при попарном сравнении северного и степного скота (*Bos taurus*) музейных образцов (М) и современных популяций (С)

ВТА	Группа	Позиция, Мб		Число SNPs	Среднее число SNPs	Число экстремальных SNPs	Доля экстремальных SNPs	Среднее число экстремальных SNPs
		начало	конец					
1	С	1,20	3,10	1193	1,12	5	0,42	8,96
		98,60	100,50	597	0,49	3	0,50	8,70
3	С	71,90	73,80	739	0,98	9	1,22	9,51
5	С	23,50	25,40	861	1,72	22	2,56	10,07
		90,70	92,60	444	0,55	2	0,45	6,29
8	М	9,30	11,20	645	0,58	5	0,78	8,23
		112,10	114,00	495	0,54	9	1,82	9,18
10	С	57,80	59,70	560	0,89	3	0,54	8,48
11	М	100,30	102,20	528	1,25	2	0,38	7,04
13	С	0,00	1,20	1710	1,61	51	2,98	9,88
14	С	22,30	24,20	669	2,02	5	0,75	8,72
15	М	44,90	46,80	541	0,61	3	0,55	6,24
		50,50	52,40	1197	0,55	3	0,25	9,41
18	М	29,20	31,10	417	0,46	2	0,48	7,33
20	М	1,40	3,30	385	0,49	4	1,04	7,34
		30,70	32,60	648	1,09	4	0,62	9,30

21	C	1,10	3,00	551	2,59	5	0,91	9,92
22	C	17,90	19,80	748	0,86	2	0,27	8,51
	M	49,50	51,40	356	0,74	2	0,56	6,72
23	M	16,90	18,80	358	0,58	4	1,12	6,91
25	M	37,90	39,80	621	0,98	9	1,45	8,34
		40,50	42,40	740	0,85	2	0,27	6,46
26	M	49,90	51,80	475	0,59	2	0,42	6,64

Примечание. ВТА — номер аутосомы крупного рогатого скота; число SNPs — число SNP внутри XP-EHH региона; среднее число SNPs — среднее число SNPs внутри региона; число экстремальных SNPs — число SNPs, превышающих пороговое значение (при $p \leq 10^{-4}$); доля экстремальных SNPs — доля SNPs, превышающих пороговое значение (при $p \leq 10^{-4}$). Описание образцов см. в разделе «Методика».

3. Гены, локализованные на ВТА11 внутри региона, идентифицированного методами zFST и XP-EHH, при сравнении музейных образцов северного и калмыцкого скота (*Bos taurus*)

Позиция, п.н.		Ген	Полное название гена	Тип гена
начало	конец			
100410491	100463233	<i>GPR107</i>	G protein-coupled receptor 107	ПРК
100490472	100534966	<i>NCS1</i>	Neuronal calcium sensor 1	ПРК
100580768	100757175	<i>HMCN2</i>	Hemicentin 2	ПРК
100770166	100822252	<i>ASS1*</i>	Argininosuccinate synthase 1	ПРК
100882266	100931850	<i>FUBP3</i>	Far upstream element binding protein 3	ПРК
100937424	100937524	<i>U6</i>	U6 spliceosomal RNA	мяРНК
100951292	100966190	<i>PRDM12</i>	PR/SET domain 12	ПРК
100975458	100985125	<i>EXOSC2</i>	Exosome component 2	ПРК
100988747	101131037	<i>ABL1</i>	ABL protoonco 1, non-receptor tyrosine kinase	ПРК
101136448	101136852	<i>QRFP</i>	Pyroglutamylated RFamide peptide	ПРК
101147884	101179089	<i>FIBCD1</i>	Fibrinogen C domain containing 1	ПРК
101225317	101291690	<i>LAMC3</i>	Laminin subunit gamma 3	ПРК
101296503	101321933	<i>AIF1L</i>	Allograft inflammatory factor 1 like	ПРК
101323401	101414345	<i>NUP214</i>	Nucleoporin 214	ПРК
101450754	101465047	<i>FAM78A</i>	Family with sequence similarity 78 member A	ПРК
101477364	101504765	<i>PLPP7</i>	Phospholipid phosphatase 7 (inactive)	ПРК
101536817	101618215	<i>PRRC2B</i>	Proline rich coiled-coil 2B	ПРК
101607876	101607961	<i>SNORD62</i>	Small nucleolar RNA SNORD62	мяРНК
101611878	101611963	<i>SNORD62</i>	Small nucleolar RNA SNORD62	мяРНК
101622925	101642291	<i>POMT1</i>	Protein O-mannosyltransferase 1	ПРК
101639654	101645280	<i>UCK1</i>	Uridine-cytidine kinase 1	ПРК
101652894	101678757	<i>PRRT1B</i>	Proline rich transmembrane protein 1B	ПРК
101696993	101834040	<i>RAPGEF1</i>	Rap guanine nucleotide exchange factor 1	ПРК
101945497	102165314	<i>MED27</i>	Mediator complex subunit 27	ПРК

Примечание. ВТА — номер аутосомы крупного рогатого скота, позиция — позиция на хромосоме (п.н. — пары нуклеотидов), мяРНК — ген малой ядрышковой РНК, ПРК — белок-кодирующий ген.

Функциональная аннотация выявленных генов, проведенная с помощью онлайн-ресурса DAVID, показала достоверное ($p < 0,05$) обогащение генами, связанными с процессами развития аксонов, обменом кальция, клеточным ответом на стимуляцию эпидермальным фактором роста (табл. 4).

4. Кластеризация генов на ВТА11, выявленных при сравнении музейных образцов северного и калмыцкого скота (*Bos taurus*)

Категория	Термины	n	%	p	Гены
Кластер 1 (степень насыщения: 1,71)					
GOTERM_CC_DIRECT	GO: 0005604~basement membrane	3	14,3	0,002	<i>LAMC3, HMCN2, AIF1L</i>
UP_KW_LIGAND	KW-0106~Calcium	5	23,8	0,005	<i>LAMC3, FIBCD1, NCS1, HMCN2, AIF1L</i>
GOTERM_BP_DIRECT	GO: 0007411~axon guidance	3	14,3	0,008	<i>LAMC3, HMCN2, AIF1L</i>
UP_KW_CELLULAR_COM					
PONENT	KW-0272~Extracellular matrix	3	14,3	0,011	<i>LAMC3, HMCN2, AIF1L</i>
SMART	SM00181: EGF	3	14,3	0,015	<i>LAMC3, HMCN2, AIF1L</i>
GOTERM_MF_DIRECT	GO:0005509~calcium ion binding	4	19,1	0,017	<i>LAMC3, NCS1, HMCN2, AIF1L</i>
INTERPRO	IPR000742: EGF-like_dom	3	14,3	0,018	<i>LAMC3, HMCN2, AIF1L</i>
UP_KW_CELLULAR_COM					
PONENT	KW-0964~Secreted	5	23,8	0,025	<i>LAMC3, FIBCD1, HMCN2, QRFP, AIF1L</i>
Кластер 2 (степень насыщения: 1,52)					
UP_KW_LIGAND	KW-0106~Calcium	5	23,8	0,005	<i>LAMC3, FIBCD1, NCS1, HMCN2, AIF1L</i>

UP_SEQ_FEATURE	DOMAIN: EF-hand	3	14,3	0,013	<i>LAMC3, NCSI, AIFIL</i>
INTERPRO	IPR002048: EF_hand_dom	3	14,3	0,015	<i>LAMC3, NCSI, AIFIL</i>
GOTERM_MF_DIRECT	GO: 0005509~calcium ion binding	4	19,1	0,017	<i>LAMC3, NCSI, HMCN2, AIFIL</i>
INTERPRO	IPR011992: EF-hand-dom_pair	3	14,3	0,021	<i>LAMC3, NCSI, AIFIL</i>

Ген *LAMC3* служит частью метаболического пути формирования внеклеточного матрикса, оказывает влияние на развитие тканей, процессы репарации, дифференцировки и миграции клеток (42). Полногеномный ассоциативный мета-анализ, проведенный на голштинском, монбельярдском и нормандском крупном рогатом скоте, обнаружил связь этого гена с удоем (43). Была показана связь SNP в гене *LAMC3* с содержанием бета-лактоглобулина у итальянского бурого швицкого скота (44), а также с показателями фертильности (доля стельных животных, степень оплодотворяемости телок и коров) у голштинского скота американской селекции (45). Ген *LAMC3* был локализован в регионе, находящимся под давлением отбора в популяции немецкого симментальского скота (46). Принимая во внимание связь мутаций в *LAMC3* с пороками развития затылочной коры головного мозга у человека (47), авторы предположили, что давление отбора на этот ген могло быть нацелено на признаки поведения, такие как сдержанный темперамент во время доместикации. У свиней *LAMC3* участвует в ремоделировании жировой ткани, воспалении и резистентности к инсулину, ассоциированных с ожирением (48). Экспрессия *LAMC3* в жировой ткани свиней коррелировала с показателями упитанности (49) и варьировала в мышцах в зависимости от содержания жира в рационе (50). Связь с метаболизмом липидов у свиней показана еще у одного идентифицированного нами гена — *NCSI* (51). Обнаружена ассоциация гена *FIBCD1* с признаками конституции у китайского голштинского скота (52). Ген *HMCN2* был связан с мраморностью мяса у японского скота Hanwoo (53), теплоустойчивостью у восьми китайских пород крупного рогатого скота и яков (54). Регион на BTA11 в области 101 Mb включал позиционные гены-кандидаты *LAMC3*, *ABLI*, *FIBCD1*, *QRFP* и другие, локализованные в непосредственной близости от SNP rs110144789 (BTA11), ассоциированного с устойчивостью к клещам у бразильских брафордов и герефордов (55).

По результатам анализа ROH в музейных образцах мы идентифицировали 17 островков ROH, в том числе 13 — в группе северного скота (5 — у KHLMH и 8 — у YRSLH) и 4 — у KALMH (табл. 5). В геноме KALMH островки ROH локализовались на BTA1, BTA5, BTA20 и BTA29, в геноме KHLMH — на BTA16 (2 островка) и BTA18 (3 островка), в геноме YRSLH — на BTA1, BTA3, BTA4, BTA5, BTA7 (2 островка), BTA10 и BTA24.

Из 7 островков ROH, выявленных в геноме YRSLH, 5 перекрывались с островками ROH в геноме YRSL, из которых 1 был обнаружен также в геноме KALMH, а 2 — в геноме KALM и KHLM. Анализ базы этих локусов количественных признаков (40) показал, что выявленные островки ROH локализовались в области QTL фертильности (BTA1), удоя (BTA7), состава молока, включая количество белка (BTA3, BTA5) и количество каппа-казеина (BTA1), а также скорости роста и живой массы (BTA7).

Из 5 островков ROH, обнаруженных в геноме KHLMH, 2 перекрывались с островками ROH в геноме KHLM, из которых 1 был также идентифицирован у YRSL. По результатам скрининга базы данных QTL (40) была установлена локализация этих островков ROH в области QTL роста, живой массы (BTA16), потребления корма (BTA16, BTA18). Один из четырех островков ROH в геноме KALMH на BTA5 перекрывался с островков

РОН у современного и предкового ярославского скота.

5. Островки РОН, идентифицированные в музейных образцах исследованных пород крупного рогатого скота (*Bos taurus*), и перекрывающиеся с ними островки РОН, выявленные в современных образцах

ВТА	Островки РОН					
	KALM	KALMH	KHLM	KHLMH	YRSL	YRSLH
1					<u>1,62-2,02</u> (282; 0,403)	<u>1,71-1,76</u> (21; 0,043)
3		<u>61,6-61,7</u> (31; 0,044)			<u>43,0-43,3</u> (145; 0,311)	<u>42,2-42,3</u> (25; 0,074)
4						<u>32,9-33,3</u> (23; 0,037)
5		<u>48,9-49,9</u> (41; 0,084)			<u>48,3-48,9</u> (142; 0,531)	<u>48,9-49,1</u> (104; 0,151)
7	<u>49,4-49,9</u> (60; 0,486) <u>49,9-50,6</u> (136; 0,643)		<u>49,4-51,2</u> (303; 1,701)		<u>49,4-50,6</u> (220; 1,147)	<u>50,1-50,3</u> (26; 0,139) <u>50,3-50,3</u> (20; 0,046)
10	<u>43,3-43,8</u> (150; 0,501)					<u>43,6-43,7</u> (13; 0,029)
16			<u>26,9-27,3</u> (215; 0,383)	<u>26,9-27,0</u> (41; 0,083) <u>75,5-75,6</u> (90; 0,111)		
18				<u>13,5-13,5</u> (7; 0,022) <u>13,5-13,6</u> (30; 0,028)		
20		<u>22,0-22,1</u> (35; 0,078)	<u>14,1-14,8</u> (274; 0,748)	<u>14,3-14,4</u> (30; 0,065)	<u>14,2-14,6</u> (155; 0,356)	
24						<u>28,9-29,0</u> (14; 0,041)
29		<u>38,1-38,2</u> (14; 0,030)				

Примечание. Над чертой — начало-конец, Мб, под чертой — (число SNPs; длина, Мб). KALM и KALMH — современные и музейные образцы от калмыцкого скота, KHLM и KHLMH — современные и музейные образцы холмогорского скота, YRSL и YRSLH — современные и музейные образцы ярославского скота; ВТА — номер аутосомы крупного рогатого скота.

6. Гены, локализованные внутри островков РОН, перекрывающихся для музейных и современных образцов у крупного рогатого скота (*Bos taurus*)

ВТА	Позиция, п.н.		Группа	Гены
	начало	конец		
1	1618395	2021023	YRSL	<i>ATP5PO, ITSN^h, bta-mir-12045, bta-mir-2285bs^a, CRYZL1, DONSON, SON, bta-mir-6501, GART^b</i>
3	1715056	1757597	YRSLH	<i>ITSN^h, bta-mir-2285bs^a</i>
3	42234489	42308800	YRSLH	<i>DPH5</i>
4	42998765	43309404	YRSL	<i>RTCA, DBT, LRRC39, TRMT13, SASS6, MFS14A, SLC35A3^b, U5</i>
4	70715789	70718893	YRSLH	<i>C4H7orf31</i>
5	48333100	48863644	YRSL	<i>LEMD3, WIF1, U6</i>
5	48888892	49039803	YRSLH	<i>TBC1D30^a</i>
5	48934026	49017630	KALMH	<i>TBC1D30^a</i>
7	49430074	50577564	YRSL	<i>WNT8A, NME5, BRD8, CDC23, KIF20A, GFRA3, CDC25C, SLBP2, FAM53C, bta-mir-2459, KDM3B, REEP2, EGR1, ETF1, HSPA9, SNORD63, CTNNA1, 5S_rRNA^a, LRRTM2^a, SIL^h</i>
7	49430074	49916441	KALM	<i>WNT8A, NME5, BRD8, CDC23, KIF20A, GFRA3, CDC25C, SLBP2, FAM53C, bta-mir-2459, KDM3B, REEP2, EGR1, ETF1, HSPA9, SNORD63</i>
7	49463093	51163997	KHLM	<i>WNT8A, NME5, BRD8, CDC23, KIF20A, GFRA3, CDC25C, SLBP2, FAM53C, bta-mir-2459, KDM3B, REEP2, EGR1, ETF1, HSPA9, SNORD63, CTNNA1, 5S_rRNA^a, LRRTM2^a, SIL^h, SNORA74, bta-mir-1949, MATR3, PAIP2, SLC23A1, MZB1, PROB1, SPATA24, DNAJC18, ECSCR, SMIM33, STING1, UBE2D2, CXXC5, PSD2, NRG2</i>
7	49989331	50632057	KALM	<i>5S_rRNA^a, LRRTM2^a, SIL^h, SNORA74, bta-mir-1949, MATR3, PAIP2</i>
7	50117105	50256308	YRSLH	<i>5S_rRNA^a, LRRTM2^a</i>
7	50265014	50310721	YRSLH	<i>SIL^h</i>

10	43302482	43803928	KALM	<i>MAP4K5, AT11, SAV1, NIN^a, ABHD12B, PYGL</i>
	43630770	43659331	YRSLH	<i>NIN^a</i>
16	26899584	27282406	KHLM	<i>CAPN8^a, CAPN2, TP53BP2</i>
	26919607	27002553	KHLMH	<i>CAPN8^a</i>
	75493141	75604057	KHLMH	<i>bta-mir-29b-2, bta-mir-29c, CD46</i>
18	14056029	14803960	KHLM	<i>ACSF3, CDH15, SLC22A31, ANKRD11^a, SPG7, RPL13, SNORD68, CPNE7, DPEP1, CHMP1A, SPATA33, CDK10, SPATA2L, VPS9D1, ZNF276, FANCA, SPIRE2, TCF25, MC1R^b, TUBB3, DEF8, DBNDD1, GAS8, URAH</i>
	14244681	14600983	YRSL	<i>CDH15, SLC22A31, ANKRD11^a, SPG7, RPL13, SNORD68, CPNE7, DPEP1, CHMP1A, SPATA33, CDK10, SPATA2L, VPS9D1, ZNF276</i>
	14320926	14386043	KHLMH	<i>ANKRD11^a</i>
20	22016727	22094725	KALMH	<i>GPBP1</i>

Примечание. KALM и KALMH — современные и музейные образцы калмыцкого скота, KHLM и KHLMH — современные и музейные образцы холмогорского скота, YRSL и YRSLH — современные и музейные образцы ярославского скота; BTA — номер аутосомы крупного рогатого скота; ^a — гены, общие для музейных и современных групп, ^b — гены, для которых есть записи в базе данных OMIA. Описание образцов см. в разделе «Методика».

Островок ROH в области 49,3-52,4 Мб на BTA7, который в наших исследованиях был выявлен у музейных образцов ярославского скота и всех трех изученных пород современного скота, также обнаружен при исследовании 18 альпийских пород крупного рогатого скота (56) и 8 китайских пород скота (57). Структурная аннотация показала локализацию в островках ROH 102 генов на девяти аутосомах (BTA1, BTA3, BTA4, BTA5, BTA7, BTA10, BTA16, BTA18 и BTA20). Наибольшее число генов было выявлено на BTA7 и BTA18 (соответственно 39 и 26 генов) (табл. 6).

В современных образцах ярославской породы было выявлено 54 гена, в то время как в музейных — всего 9 генов. Причем 5 выявленных генов (*ITSN1* и *bta-mir-2285bs* на BTA1; *5S_rRNA*, *LRRTM2* и *SIL1* на BTA7) совпадали для YRSL и YRSLH. В современной холмогорской породе было идентифицировано 63 гена, в то время как в музейных образцах — всего 5 генов, 2 из которых (*CAPN8* на BTA16 и *ANKRD11* на BTA18) были выявлены в KHLM. У современного калмыцкого скота обнаружили 29 генов, а в музейных образцах — всего 2 гена, причем ни один не совпал с выявленными в современных образцах. Для музейных образцов ярославской и калмыцкой породы выявлен общий ген *TBC1D30*, локализованный на BTA5. Интересно отметить, что обнаруженные у YRSL гены *GART* и *SLC35A3* имеют записи в базе данных признаков животных с менделевским типом наследования (OMIA) (<https://www.omia.org/home/>). Рецессивная мутация в гене *GART* (OMIA 001826) связана с гаплотипом фертильности голштинского скота НН4 и приводит к прерыванию стельности (58). Была показана связь этой мутации в гене *GART* с репродуктивными качествами голштинского скота российской селекции, включая продолжительность сервис-периода и интервал между отелами (59). Ген *SLC35A3* (OMIA 001340) связан с комплексным пороком развития позвоночника (complex vertebral malformation, CVM) — рецессивной мутацией, приводящей к гибели плода (60). В KHLM выявлен ген *MC1R*, для которого в базе данных OMIA есть две записи. Одна запись (OMIA 001199) связана с окрасом шерсти (61, 62), вторая (OMIA001544) — с гипотрихозом и ослаблением окраски шерсти (синдром крысиного хвоста) (63).

Функциональная аннотация указанных генов показала формирование десяти кластеров, из которых только для двух (табл. 7) было выявлено значимое насыщение (уровень насыщения $>1,13$, $p \leq 0,05$). В первый кластер вошли гены *EGRI*, *CDC23*, *STING1*, *CTNNA1*, *ATP5PO*, *RPL13*, *ETF1*, *PAIP2*, *SIL1*, для которых отмечено участие в процессах регуляции транскрипции, опосредованном ответе на гипоксию и ишемию, регуляции био-

синтеза лютеинизирующего гормона, регуляции пролиферации и апоптоза клеток, регуляции клеточного цикла. Во второй кластер вошли гены *HSPA9*, *STING1*, *TUBB3*, *RTCA*, *UBE2D2*, *ATL1*, *CDK10*, *KIF20A*, *PYGL*, *GART*, *ACSF3*, которые связаны с регуляцией физиологических процессов, индукцией иммунного ответа, развитием нервной ткани.

7. Кластеризация генов, локализованных внутри островков ROH, перекрывающихся для музейных и современных образцов у северного и калмыцкого скота (*Bos taurus*)

Категория	Термины	n	%	p	Гены
UP_SEQ_FEATURE	Кластер 1 (степень насыщения: 1,27) CROSSLNK: Glycyl lysine isopeptide (Lys-Gly) (interchain with G-Cter in SUMO2)	5	5,95	0,027	<i>EGR1</i> , <i>CDC23</i> , <i>CTNNA1</i> , <i>RPL13</i> , <i>ETFF1</i>
UP_KW_PTM	KW-0832~Ubl conjugation	9	10,71	0,073	<i>EGR1</i> , <i>CDC23</i> , <i>STING1</i> , <i>CTNNA1</i> , <i>ATP5PO</i> , <i>RPL13</i> , <i>ETFF1</i> , <i>PAIP2</i> , <i>SIL1</i>
UP_KW_PTM	KW-1017~Isopeptide bond	7	8,33	0,078	<i>EGR1</i> , <i>CDC23</i> , <i>STING1</i> , <i>TUBB3</i> , <i>CTNNA1</i> , <i>RPL13</i> , <i>ETFF1</i>
UP_KW_LIGAND	Кластер 2 (степень насыщения: 1,14) KW-0547~Nucleotide-binding	11	13,09	0,039	<i>HSPA9</i> , <i>STING1</i> , <i>TUBB3</i> , <i>RTCA</i> , <i>UBE2D2</i> , <i>ATL1</i> , <i>CDK10</i> , <i>KIF20A</i> , <i>PYGL</i> , <i>GART</i> , <i>ACSF3</i>
GOTERM_MF_DIRECT	GO: 0005524~ATP binding	10	11,90	0,073	<i>HSPA9</i> , <i>RTCA</i> , <i>UBE2D2</i> , <i>CDK10</i> , <i>SPG7</i> , <i>KIF20A</i> , <i>PYGL</i> , <i>MAP4K5</i> , <i>GART</i> , <i>ACSF3</i>
UP_KW_LIGAND	KW-0067~ATP-binding	8	9,52	0,134	<i>HSPA9</i> , <i>RTCA</i> , <i>UBE2D2</i> , <i>CDK10</i> , <i>KIF20A</i> , <i>MAP4K5</i> , <i>GART</i> , <i>ACSF3</i>

Из общего списка идентифицированных генов семь находились в островках ROH, выявленных как у предкового, так и у современного северного скота, включая гены *ITSN1*, *bta-mir-2285bs* на ВТА1, *5S rRNA*, *LRRTM2*, *SIL1* на ВТА7 — у ярославского скота и *CAPN8* на ВТА16 и *ANKRD11* на ВТА18 — у холмогорского скота. Ген *CAPN8* находился под давлением отбора у бурого швицкого скота (64). У тайского молочного скота выявлена ассоциация гена *CAPN8* с содержанием соматических клеток в молоке (65). Ген *LRRTM2* связан с персистентностью лактации у буйволов (66). Ген *ANKRD11* подвержен действию отбора у скота породы Sahiwal (3). Ген *TBC1D30*, локализованный в островке ROH у предкового северного (ярославский скот) и степного скота, ассоциировался с возрастом первого отела и долей двоен у голштинского скота (67, 68), а также со способностью длительно сохранять репродуктивный потенциал (продолжительность продуктивного использования) у крупного рогатого скота породы Nelore (69). У итальянской локальной породы скота Reggiana была показана связь гена *TBC1D30* с пигментацией (интенсивностью окраски шерсти в градации красного) (70). Интересно, что у человека этот ген ассоциируется с цветом волос (71).

Таким образом, исследование северного (холмогорского и ярославского) и степного (калмыцкого) скота показало, что давлению отбора в большинстве случаев были подвержены различные участки генома. Это может быть следствием различий в целях и направлениях селекции в предковых и современных популяциях изучаемых пород. Интересно отметить, что для предковых популяций ярославской и калмыцкой пород выявлен общий ген *TBC1D30*, ассоциированный с репродуктивными качествами и пигментацией. Этот факт может свидетельствовать о том, что, несмотря на разное происхождение и условия обитания, в исследованных популяциях могли сформироваться схожие генетические механизмы приспособления к выживанию в различных климатических зонах — северной и степной. Эти данные подчеркивают конвергентную эволюцию ключевых генов в ответ на

влияние экологических факторов. Вместе с тем в геноме музейных образцов северного скота обнаружено 7 островков РОН (в том числе 5 — у ярославского скота, 2 — у холмогорского скота), перекрывающихся с одноименными современными популяциями. Выявленные островки РОН локализовались в области QTL удоя и состава молока, включая содержание белка и каппа-казеина, скорости роста и живой массы, а также фертильности. Полученные данные дают основание полагать, что указанные признаки были приоритетными в селекции как предкового, так и современного холмогорского и ярославского скота.

ФГБНУ Федеральный исследовательский центр животноводства — ВИЖ им. академика Л.К. Эрнста,
142132 Россия, Московская обл., г.о. Подольск, пос. Дубровицы, 60,
e-mail: abdelmanova@vij.ru ✉, horarka@yandex.ru, alexshakhin@vij.ru,
alexandraces@yandex.ru, g0work@mail.ru, liskun@rgau-msha.ru,
vugarbagirov@mail.ru, rector@rgau-msha.ru, n_zinovieva@mail.ru

Поступила в редакцию
29 мая 2025 года
Принята к публикации
22 июля 2025 года

Sel'skokhozyaistvennaya biologiya [Agricultural Biology], 2025, V. 60, № 6, pp. 986-1004

SEARCH FOR SELECTION SIGNATURES IN GENOMES OF NORTHERN AND STEPPE MODERN CATTLE IN RUSSIA AS COMPARED TO SAMPLES OF XIX-XX CENTURIES

A.S. Abdelmanova✉, *T.E. Deniskova*, *A.V. Shakhin*, *A.A. Nikolaev*, *G.O. Abramov*,
O.I. Boronetskaya, *V.A. Bagirov*, *V.I. Truchachev*, *N.A. Zinovieva*

Ernst Federal Research Center for Animal Husbandry, 60, pos. Dubrovitsy, Podolsk District, Moscow Province, 142132 Russia, e-mail abdelmanova@vij.ru (✉ corresponding author) horarka@yandex.ru, alexshakhin@vij.ru, alexandraces@yandex.ru, g0work@mail.ru, liskun@rgau-msha.ru, vugarbagirov@mail.ru, rector@rgau-msha.ru, n_zinovieva@mail.ru

ORCID:

Abdelmanova A.S. orcid.org/0000-0003-4752-0727

Deniskova T.E. orcid.org/0000-0002-5809-1262

Shakhin A.V. orcid.org/0000-0003-4959-878X

Nikolaev A.A. orcid.org/0000-0001-9355-3285

Abramov G.O. orcid.org/0009-0000-7239-3227

The authors declare no conflict of interests

Acknowledgements:

Supported financially from the Russian Science Foundation, project No. 19-76-20012.

Final revision received May 29, 2025

doi: 10.15389/agrobiology.2025.6.986eng

Accepted July 22, 2025

Boronetskaya O.I. orcid.org/0000-0001-8389-5572

Bagirov V.A. orcid.org/0000-0001-5398-8815

Truchachev V.I. orcid.org/0000-0003-4676-5407

Zinovieva N.A. orcid.org/0000-0003-4017-6863

Abstract

The development of whole genome sequencing (WGS) technology opens new opportunities in studying the genomic variability of farm animals under the influence of natural and artificial selection. Analysis of WGS data allows us to identify key genes and genomic regions responsible for the formation of adaptive properties to specific natural and climatic breeding conditions. Selection and genetic drift are key sources of genetic differentiation between representatives of original and modern breeds. Studying the dynamics of changes in the gene pool of local breeds is relevant, since local breeds are carriers of rare and valuable genetic variants that may become a priority in the event of changing market demands and climate change. This paper describes for the first time the results of identifying genomic regions fixed as a result of selection in contrasting groups of cattle adapted to significantly different environmental conditions, and provides a structural and functional annotation of the genes localized in them. Genes common to modern and ancestral cattle populations of the studied groups were identified. The aim of the study was to compare the genomes of historical and modern cattle of different breeds based on WGS data to identify genomic regions affected by natural and artificial selection. The study material included museum cattle skull specimens from the E.F. Liskun Livestock Museum (Timiryazev Russian State Agrarian University - Moscow Agricultural Academy, Moscow), dating from the late 19th to early 20th centuries, as well as samples from breeding animals stored in the biological collection of the National Center for Farm Animal Genetic Resources of the L.K. Ernst Federal Research Center for Animal Husbandry. The northern cattle group was represented by samples of the Kholmogor and Yaroslavl breeds, and the steppe cattle group was represented by samples of Kalmyk cattle. The sample of historical specimens (n=20) included 15 samples of northern cattle (Kholmogor cattle, n=8; Yaroslavl cattle, n=7) and 5 samples of steppe (Kalmyk) cattle. The sample of modern specimens included 17, 11 and 12 samples of the above-mentioned breeds, respectively.

The final data set included the genotypes for 1,615,259 SNPs. DNA extraction from museum specimens was performed using the previously described method, and from modern specimens, using the DNA-Extran 2 kit (Sintol JSC, Moscow). Libraries for whole-genome sequencing using NGS technology were prepared using the TruSeq DNA Nano Library Prep kit (Illumina, Inc., USA) and the Accel-NGS® 2S Plus DNA Library Kit (IDT). Sequencing was performed on a NovaSeq 6000 sequencer (2.5150 bp). Three methods were used to identify genomic regions under selection pressure: pairwise comparison of populations based on F_{ST} genetic distances using a sliding window, analysis of cross-population extended haplotype homozygosity (XP-EHH), and identification of regions of excess homozygosity (ROH islands). The absence of overlapping genomic regions under selection pressure in the genome of the historical and modern populations of studied cattle breeds based on the results of a comparative analysis of F_{ST} and XP-EHH for pairs (Kholmogor + Yaroslavl) – Kalmyk indicates significant differences in the breeding goals in different periods of breed formation. Seven ROH islands have been identified, overlapping in museum and modern samples of Kholmogor and Yaroslavl cattle, localized in the known QTL for milk yield, milk composition, growth, body size and fertility, which indicates the long-term use of the above-mentioned indices as target parameters at breeding these breeds. The genes common to modern and historical cattle populations of the studied groups have been identified. The identified genomic regions, which are under selection pressure in the historical populations of the studied breeds and overlap with the signature of selection in the genome of modern representatives of the same breeds, can be used in the selection of animals for inclusion in genetic resource conservation programs in order to preserve authentic genomic components.

Keywords: cattle, Kholmogor breed, Yaroslavl breed, Kalmyk breed, museum specimens, whole genome sequencing, selection pressure, candidate genes.

REFERENCES

1. Mastrangelo S., Sardina M.T., Tolone M., Di Gerlando R., Sutura A.M., Fontanesi L., Portolano B. Genome-wide identification of runs of homozygosity islands and associated genes in local dairy cattle breeds. *Animal*, 2018, 12(12): 2480–2488 (doi: 10.1017/S1751731118000629).
2. Barbato M., Hailer F., Upadhyay M., Del Corvo M., Colli L., Negrini R., Kim E.S., Crooijmans R.P.M.A., Sonstegard T., Ajmone-Marsan P. Adaptive introgression from indicine cattle into white cattle breeds from Central Italy. *Scientific Reports*, 2020, 10(1): 1279 (doi: 10.1038/s41598-020-57880-4).
3. Illa S.K., Mukherjee S., Nath S., Mukherjee A. Genome-wide scanning for signatures of selection revealed the putative genomic regions and candidate genes controlling milk composition and coat color traits in Sahiwal cattle. *Frontiers in Genetics*, 2021, 12: 699422 (doi: 10.3389/fgene.2021.699422).
4. Chen Q., Huang B., Zhan J., Wang J., Qu K., Zhang F., Shen J., Jia P., Ning Q., Zhang J., Chen N., Chen H., Lei C. Whole-genome analyses identify loci and selective signals associated with body size in cattle. *Journal of Animal Science*, 2020, 98(3): skaa068 (doi: 10.1093/jas/skaa068).
5. Ulimbashev M.B., Konik N.V., Krasnova O.A., Kapitonova E.A., Tletseruk I.R., Tatueva O.V., Sannikova N.A., Golembovsky V.V. Use and preservation of local dairy cattle breeds (review). *Sel'skokhozyaystvennaya biologiya [Agricultural Biology]*, 2024, 59(6): 1055–1075 (doi: 10.15389/agrobiology.2024.6.1055rus).
6. Eynard S.E., Windig J.J., Hiemstra S.J., Calus M.P. Whole-genome sequence data uncover loss of genetic diversity due to selection. *Genetics, Selection, Evolution: GSE*, 2016, 48: 33 (doi: 10.1186/s12711-016-0210-4).
7. Eusebi P.G., Martinez A., Cortes O. Genomic tools for effective conservation of livestock breed diversity. *Diversity*, 2020, 12(1): 8 (doi: 10.3390/d12010008).
8. Günther T., Nettelblad C. The presence and impact of reference bias on population genomic studies of prehistoric human populations. *PLoS Genetics*, 2019, 15(7): e1008302 (doi: 10.1371/journal.pgen.1008302).
9. Bhati M., Kadri N.K., Crysnanto D., Pausch H. Assessing genomic diversity and signatures of selection in Original Braunvieh cattle using whole-genome sequencing data. *BMC Genomics*, 2020, 21(1): 27 (doi: 10.1186/s12864-020-6446-y).
10. Boitard S., Boussaha M., Capitan A., Rocha D., Servin B. Uncovering adaptation from sequence data: lessons from genome resequencing of four cattle breeds. *Genetics*, 2016, 203(1): 433–450 (doi: 10.1534/genetics.115.181594).
11. Willing E.M., Dreyer C., van Oosterhout C. Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. *PLoS ONE*, 2012, 7(8): e42649 (doi: 10.1371/journal.pone.0042649).
12. Abdelmanova A.S., Volkova V.V., Kharzinova V.R., Fornara M.S., Chinarov R.Yu., Boronetskaya O.I., Trukhachev V.I., Brem G., Zinovieva N.A. Determination of consensus genotypes by microsatellites for museum accessions of cattle (*Vos taurus*). *Sel'skokhozyaystvennaya biologiya, [Agricultural Biology]*, 2023, 58(6): 1035–1045 (doi: 10.15389/agrobiology.2023.6.1035rus).

13. Mulim H.A., Brito L.F., Pinto L.F.B., Ferraz J.B.S., Grigoletto L., Silva M.R., Pedrosa V.B. Characterization of runs of homozygosity, heterozygosity-enriched regions, and population structure in cattle populations selected for different breeding goals. *BMC Genomics*, 2022, 23(1): 209 (doi: 10.1186/s12864-022-08384-0).
14. Ma X., Cheng H., Liu Y., Sun L., Chen N., Jiang F., You W., Yang Z., Zhang B., Song E., Lei C. Assessing genomic diversity and selective pressures in Bohai Black cattle using whole-genome sequencing data. *Animals*, 2022, 12(5): 665 (doi: 10.3390/ani12050665).
15. Boitard S., Paris C., Sevane N., Servin B., Bazi-Kabbaj K., Dunner S. Gene banks as reservoirs to detect recent selection: the example of the Asturiana de los Valles bovine breed. *Frontiers in Genetics*, 2021, 12: 575405 (doi: 10.3389/fgene.2021.575405).
16. Zhang S., Yao Z., Li X., Zhang Z., Liu X., Yang P., Chen N., Xia X., Lyu S., Shi Q., Wang E., Ru B., Jiang Y., Lei C., Chen H., Huang Y. Assessing genomic diversity and signatures of selection in Pinan cattle using whole-genome sequencing data. *BMC Genomics*, 2022, 23(1): 460 (doi: 10.1186/s12864-022-08645-y).
17. Nawaz M.Y., Savegnago R.P., Lim D., Lee S.H., Gondro C. Signatures of selection in Angus and Hanwoo beef cattle using imputed whole genome sequence data. *Frontiers in Genetics*, 2024, 15: 1368710 (doi: 10.3389/fgene.2024.1368710).
18. Ward J.A., Ng'ang'a S.I., Randhawa I.A.S., McHugo G.P., O'Grady J.F., Flórez J.M., Browne J.A., Pérez O'Brien A.M., Landaeta-Hernández A.J., Garcia J.F., Sonstegard T.S., Frantz L.A.F., Salter-Townshend M., MacHugh D.E. Genomic insights into the population history and adaptive traits of Latin American Criollo cattle. *Royal Society Open Science*, 2024, 11(3): 231388 (doi: 10.1098/rsos.231388).
19. Tiwari M., Sodhi M., Chanda D., Kataria R.S., Niranjana S.K., Singh I., Bharti V.K., Iqbal M., Rabgais S., Amarjeet, Vivek P., Kumari P., Mukesh M. Deciphering genomic basis of unique adaptation of Ladakhi cattle to Trans-Himalayan high-altitude region of Leh-Ladakh in India. *Gene*, 2025, 942: 149251 (doi: 10.1016/j.gene.2025.149251).
20. Hulsege I., Oldenbroek K., Bouwman A., Veerkamp R., Windig J. Selection and drift: a comparison between historic and recent Dutch Friesian cattle and recent Holstein Friesian using WGS data. *Animals*, 2022, 12(3): 329 (doi: 10.3390/ani12030329).
21. Moscarelli A., Sardina M.T., Cassandro M., Ciani E., Pilla F., Senczuk G., Portolano B., Mastrangelo S. Genome-wide assessment of diversity and differentiation between original and modern Brown cattle populations. *Animal Genetics*, 2021, 52(1): 21-31 (doi: 10.1111/age.13019).
22. Holsinger K.E., Weir B.S. Genetics in geographically structured populations: defining, estimating and interpreting F(ST). *Nature Reviews Genetics*, 2009, 10(9): 639-650 (doi: 10.1038/nrg2611).
23. Sabeti P.C., Varilly P., Fry B., Lohmueller J., Hostetter E., Cotsapas C., Xie X., Byrne E.H., McCarroll S.A., Gaudet R., Schaffner S.F., Lander E.S., International HapMap Consortium, Frazer K.A., Ballinger D.G., Cox D.R., Hinds D.A., Stuve L.L., Gibbs R.A., Belmont J.W., Boudreau A., Hardenbol P., Leal S.M., Pasternak S., Wheeler D.A., Willis T.D., Yu F., Yang H., Zeng C., Gao Y., Hu H., Hu W., Li C., Lin W., Liu S., Pan H., Tang X., Wang J., Wang W., Yu J., Zhang B., Zhang Q., Zhao H., Zhao H., Zhou J., Gabriel S.B., Barry R., Blumenstiel B., Camargo A., Defelice M., Faggart M., Goyette M., Gupta S., Moore J., Nguyen H., Onofrio R.C., Parkin M., Roy J., Stahl E., Winchester E., Ziaugra L., Altshuler D., Shen Y., Yao Z., Huang W., Chu X., He Y., Jin L., Liu Y., Shen Y., Sun W., Wang H., Wang Y., Wang Y., Xiong X., Xu L., Wayne M.M., Tsui S.K., Xue H., Wong J.T., Galver L.M., Fan J.B., Gunderson K., Murray S.S., Oliphant A.R., Chee M.S., Montpetit A., Chagnon F., Ferretti V., Leboeuf M., Olivier J.F., Phillips M.S., Roumy S., Sallé C., Verner A., Hudson T.J., Kwok P.Y., Cai D., Koboldt D.C., Miller R.D., Pawlikowska L., Taillon-Miller P., Xiao M., Tsui L.C., Mak W., Song Y.Q., Tam P.K., Nakamura Y., Kawaguchi T., Kitamoto T., Morizono T., Nagashima A., Ohnishi Y., Sekine A., Tanaka T., Tsunoda T., Deloukas P., Bird C.P., Delgado M., Dermitzakis E.T., Gwilliam R., Hunt S., Morrison J., Powell D., Stranger B.E., Whittaker P., Bentley D.R., Daly M.J., de Bakker P.I., Barrett J., Chretien Y.R., Maller J., McCarroll S., Patterson N., Pe'er I., Price A., Purcell S., Richter D.J., Sabeti P., Saxena R., Schaffner S.F., Sham P.C., Varilly P., Altshuler D., Stein L.D., Krishnan L., Smith A.V., Tello-Ruiz M.K., Thorisson G.A., Chakravarti A., Chen P.E., Cutler D.J., Kashuk C.S., Lin S., Abecasis G.R., Guan W., Li Y., Munro H.M., Qin Z.S., Thomas D.J., McVean G., Auton A., Bottolo L., Cardin N., Eyheramendy S., Freeman C., Marchini J., Myers S., Spencer C., Stephens M., Donnelly P., Cardon L.R., Clarke G., Evans D.M., Morris A.P., Weir B.S., Tsunoda T., Johnson T.A., Mullikin J.C., Sherry S.T., Feolo M., Skol A., Zhang H., Zeng C., Zhao H., Matsuda ., Fukushima Y., Macer D.R., Suda E., Rotimi C.N., Adebamowo C.A., Ajayi I., Aniagwu T., Marshall P.A., Nkwodimma C., Royal C.D., Leppert M.F., Dixon M., Peiffer A., Qiu R., Kent A., Kato K., Niikawa N., Adewole I.F., Knoppers B.M., Foster M.W., Clayton E.W., Watkins J., Gibbs R.A., Belmont J.W., Muzny D., Nazareth L., Sodergren E., Weinstock G.M., Wheeler D.A., Yakub I., Gabriel S.B., Onofrio R.C., Richter D.J., Ziaugra L., Birren B.W., Daly M.J., Altshuler D., Wilson R.K., Fulton L.L., Rogers J., Burton J., Carter N.P., Clee C.M., Griffiths M., Jones M.C., McLay K., Plumb R.W., Ross M.T., Sims S.K., Willey D.L., Chen Z., Han H., Kang L., Godbout M., Wallenburg J.C., L'Archevêque P., Bellemare G., Saeki K.,

- Wang H., An D., Fu H., Li Q., Wang Z., Wang R., Holden A.L., Brooks L.D., McEwen J.E., Guyer M.S., Wang V.O., Peterson J.L., Shi M., Spiegel J., Sung L.M., Zacharia L.F., Collins F.S., Kennedy K., Jamieson R., Stewart J. Genome-wide detection and characterization of positive selection in human populations. *Nature*, 2007, 449(7164): 913-918 (doi: 10.1038/nature06250).
24. Dixit S.P., Singh S., Ganguly I., Bhatia A.K., Sharma A., Kumar N.A., Dang A.K., Jayakumar S. Genome-wide runs of homozygosity revealed selection signatures in *Bos indicus*. *Frontiers in Genetics*, 2020, 11: 92 (doi: 10.3389/fgene.2020.00092).
 25. Chen C., Zhu B., Tang X., Chen B., Liu M., Gao N., Li S., Gu J. Genome-wide assessment of runs of homozygosity by whole-genome sequencing in diverse horse breeds worldwide. *Genes*, 2023, 14(6):1211 (doi: 10.3390/genes14061211).
 26. Danecek P., Auton A., Abecasis G., Albers C.A., Banks E., DePristo M.A., Handsaker R.E., Lunter G., Marth G.T., Sherry S.T., McVean G., Durbin R., 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics*, 2011, 27(15): 2156-2158 (doi: 10.1093/bioinformatics/btr330).
 27. Rafiepour M., Ebrahimie E., Vahidi M.F., Salekdeh G.H., Niazi A., Dadpasand M., Liang D., Si J., Ding X., Han J., Zhang Y., Qanbari S. Whole-genome resequencing reveals adaptation prior to the divergence of buffalo subspecies. *Genome Biology and Evolution*, 2021, 13(1): evaa231 (doi: 10.1093/gbe/evaa231).
 28. Rahimadad S., Ghaffari M., Mokhber M., Williams J.L. Linkage disequilibrium and effective population size of buffalo populations of Iran, Turkey, Pakistan, and Egypt using a medium density SNP array. *Frontiers in Genetics*, 2021, 12: 608186 (doi: 10.3389/fgene.2021.608186).
 29. Kijas J.W., Lenstra J.A., Hayes B., Boitard S., Neto L.R.P., Cristobal M.S., Servin B., McCulloch R., Whan V., Gietzen K., Paiva S., Barendse W., Ciani E., Raadsma H., McEwan J., Dalrymple B., International Sheep Genomics Consortium Members. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology*, 2012, 10(2): e1001258 (doi: 10.1371/journal.pbio.1001258).
 30. Zhao F., McParland S., Kearney F., Du L., Berry D.P. Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genetics Selection Evolution*, 2015, 47: 49 (doi: 10.1186/s12711-015-0127-3).
 31. Gautier M., Klassmann A., Vitalis R. RehH 2.0: a reimplement of the R package rehH to detect positive selection from haplotype structure. *Molecular Ecology Resources*, 2017, 17(1): 78-90 (doi: 10.1111/1755-0998.12634).
 32. Marras G., Gaspa G., Sorbolini S., Dimauro C., Ajmone-Marsan P., Valentini A., Williams J.L., Macciotta N.P.P. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. *Animal Genetics*, 2015, 46(2): 110-121 (doi: 10.1111/age.12259).
 33. Biscarini F., Cozzi P., Gaspa G., Marras G. *DetectRUNS: Detect runs of homozygosity and runs of heterozygosity in diploid genomes: R Package Version 0.9.5* Available: <https://cran.r-project.org/web/packages/detectRUNS/index.html>. Accessed: 01/02/2025.
 34. McLaren W., Gil L., Hunt S.E., Riat H.S., Ritchie G.R., Thormann A., Flicek P., Cunningham F. The ensemble variant effect predictor. *Genome Biology*, 2016, 17(1): 122 (doi: 10.1186/s13059-016-0974-4).
 35. Smedley D., Haider S., Durinck S., Pandini L., Provero P., Allen J., Arnaiz O., Awedh M.H., Baldoock R., Barbiera G., Bardou P., Beck T., Blake A., Bonierbale M., Brookes A.J., Bucci G., Buetti I., Burge S., Cabau C., Carlson J.W., Chelala C., Chrysostomou C., Cittaro D., Collin O., Cordova R., Cutts R.J., Dassi E., Di Genova A., Djari A., Esposito A., Estrella H., Eyraes E., Fernandez-Banet J., Forbes S., Free R.C., Fujisawa T., Gadaleta E., Garcia-Manteiga J.M., Goodstein D., Gray K., Guerra-Assunção J.A., Haggarty B., Han D.J., Han B.W., Harris T., Harshbarger J., Hastings R.K., Hayes R.D., Hoede C., Hu S., Hu Z.L., Hutchins L., Kan Z., Kawaji H., Keliet A., Kerhornou A., Kim S., Kinsella R., Klopp C., Kong L., Lawson D., Lazarevic D., Lee J.H., Letellier T., Li C.Y., Lio P., Liu C.J., Luo J., Maass A., Mariette J., Maurer T., Merella S., Mohamed A.M., Moreews F., Nabihoudine I., Ndegwa N., Noirot C., Perez-Llamas C., Primig M., Quattrone A., Quesneville H., Rambaldi D., Reecy J., Riba M., Rosanoff S., Saddiq A.A., Salas E., Sallou O., Shepherd R., Simon R., Sperling L., Spooner W., Staines D.M., Steinbach D., Stone K., Stupka E., Teague J.W., Dayem Ullah A.Z., Wang J., Ware D., Wong-Erasmus M., Youens-Clark K., Zadissa A., Zhang S.J., Kasprzyk A. The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Research*, 2015, 43(W1): W589-W598 (doi: 10.1093/nar/gkv350).
 36. Zerbino D.R., Achuthan P., Akanni W., Amode M.R., Barrell D., Bhai J., Billis K., Cummins C., Gall A., Giryn C.G., Gil L., Gordon L., Haggerty L., Haskell E., Hourlier T., Izuogu O.G., Janacek S.H., Juettemann T., To J.K., Laird M.R., Lavidas I., Liu Z., Loveland J.E., Maurel T., McLaren W., Moore B., Mudge J., Murphy D.N., Newman V., Nuhn M., Ogeh D., Ong C.K., Parker A., Patricio M., Riat H.S., Schuilenburg H., Sheppard D., Sparrow H., Taylor K., Thormann A., Vullo A., Walts B., Zadissa A., Frankish A., Hunt S.E., Kostadima M., Langridge N., Martin F.J., Muffato M., Perry E., Ruffier M., Staines D.M., Trevanion S.J., Aken B.L., Cunningham F., Yates A., Flicek P. Ensembl 2018. *Nucleic Acids Research*, 2018, 46(D1): D754-

- D761 (doi: 10.1093/nar/gkx1098).
37. Kuehn L.A., Keele J.W., Bennett G.L., McDanel T.G., Smith T.P., Snelling W.M., Sonstegard T.S., Thallman R.M. Predicting breed composition using breed frequencies of 50,000 markers from the US Meat Animal Research Center 2,000 Bull Project. *Journal of Animal Science*, 2011, 89(6): 1742-1750 (doi: 10.2527/jas.2010-3530).
 38. Huang D.W., Sherman B.T., Lempicki R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 2009, 4(1): 44-57 (doi: 10.1038/nprot.2008.211).
 39. Huang D.W., Sherman B.T., Lempicki R.A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, 2009, 37(1): 1-13 (doi: 10.1093/nar/gkn923).
 40. Hu Z.-L., Park C.A., Reedy J.M., Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services, *Nucleic Acids Research*, 2022, 50(D1): D956-D961 (doi: 10.1093/nar/gkab1116).
 41. Safran M., Rosen N., Twik M., BarShir R., Stein T.I., Dahary D., Fishilevich S., Lancet D. The GeneCards Suite. In: *Practical guide to life science databases*. I. Abugessaisa, T. Kasukawa (eds.). Springer, Singapore, 2021: 27-56 (doi: 10.1007/978-981-16-5812-9_2).
 42. Ruiz-Ojeda F.J., Méndez-Gutiérrez A., Aguilera C.M., Plaza-Díaz J. Extracellular matrix remodeling of adipose tissue in obesity and metabolic diseases. *International Journal of Molecular Sciences*, 2019, 20(19): 4888 (doi: 10.3390/ijms20194888).
 43. Marete A.G., Guldbrandtsen B., Lund M.S., Fritz S., Sahana G., Boichard D. A meta-analysis including pre-selected sequence variants associated with seven traits in three french dairy cattle populations. *Frontiers in Genetics*, 2018, 9: 522 (doi: 10.3389/fgene.2018.00522).
 44. Pegolo S., Yu H., Morota G., Bisutti V., Rosa G.J.M., Bittante G., Cecchinato A. Structural equation modeling for unraveling the multivariate genomic architecture of milk proteins in dairy cattle. *Journal of Dairy Science*, 2021, 104(5): 5705-5718 (doi: 10.3168/jds.2020-18321).
 45. Parker Gaddis K.L., Null D.J., Cole J.B. Explorations in genome-wide association studies and network analyses with dairy cattle fertility traits. *Journal of Dairy Science*, 2016, 99(8): 6420-6435 (doi: 10.3168/jds.2015-10444).
 46. Qanbari S., Pausch H., Jansen S., Somel M., Strom T.M., Fries R., Nielsen R., Simianer H. Classic selective sweeps revealed by massive sequencing in cattle. *PLoS Genetics*, 2014, 10(2): e1004148 (doi: 10.1371/journal.pgen.1004148).
 47. Barak T., Kwan K.Y., Louvi A., Demirebilek V., Saygi S., Tüysüz B., Choi M., Boyacı H., Dörschner K., Zhu Y., Kaymakçalan H., Yılmaz S., Bakırcıoğlu M., Çağlayan A.O., Öztürk A.K., Yasuno K., Brunken W.J., Atalar E., Yalçınkaya C., Dinçer A., Bronen R.A., Mane S., Özçelik T., Lifton R.P., Šestan N., Bilgüvar K., Günel M. Recessive *LAMC3* mutations cause malformations of occipital cortical development. *Nature Genetics*. 2011, 43(6): 590-594 (doi: 10.1038/ng.836).
 48. Liu X., Wang L.G., Luo W.Z., Li Y., Liang J., Yan H. A genome-wide SNP scan in a porcine Large White \times Minzhu intercross population reveals a locus influencing muscle mass on chromosome 2. *Journal of Animal Science*, 2014, 85(12): 969-975 (doi: 10.1111/asj.12230).
 49. Ropka-Molik K., Pawlina-Tyszko K., Żukowski K., Tyra M., Derebecka N., Wesoly J., Szmatoła T., Piórkowska K. Identification of molecular mechanisms related to pig fatness at the transcriptome and miRNAome levels. *Genes*, 2020, 11(6): 600 (doi: 10.3390/genes11060600).
 50. Montenegro M.C., Peraza P., Balemian N., Carballo C., Barlocco N., Barrios P.G., Mernies B., Saadoun A., Castro G., Guimarães S.F., Llambí S. Gene expression analysis by RNA-sequencing of Longissimus dorsi muscle of pigs fed diets with differing lipid contents. *Genetics and Molecular Research*, 2019, 18(4): GMR18307.
 51. Passols M., Llobet-Cabau F., Sebastià C., Castelló A., Valdés-Hernández J., Criado-Mesas L., Sánchez A., Folch J.M. Identification of genomic regions, genetic variants and gene networks regulating candidate genes for lipid metabolism in pig muscle. *Animal*, 2023, 17(12): 101033 (doi: 10.1016/j.animal.2023.101033).
 52. Li S., Ge F., Chen L., Liu Y., Chen Y., Ma Y. Genome-wide association analysis of body conformation traits in Chinese Holstein Cattle. *BMC Genomics*, 2024, 25(1): 1174 (doi: 10.1186/s12864-024-11090-8).
 53. Srikanth K., Lee S.H., Chung K.Y., Park J.E., Jang G.W., Park M.R., Kim N.Y., Kim T.H., Chai H.H., Park W.C., Lim D. A Gene-set enrichment and protein-protein interaction network-based GWAS with regulatory SNPs identifies candidate genes and pathways associated with carcass traits in Hanwoo cattle. *Genes*, 2020, 11(3): 316 (doi: 10.3390/genes11030316).
 54. Freitas P.H.F., Wang Y., Yan P., Oliveira H.R., Schenkel F.S., Zhang Y., Xu Q., Brito L.F. Genetic diversity and signatures of selection for thermal stress in cattle and other two *Bos* species adapted to divergent climatic conditions. *Frontiers in Genetics*, 2021, 12: 604823 (doi: 10.3389/fgene.2021.604823).
 55. Sollero B.P., Junqueira V.S., Gomes C.C.G., Caetano A.R., Cardoso F.F. Tag SNP selection for prediction of tick resistance in Brazilian Braford and Hereford cattle breeds using Bayesian methods. *Genetics Selection Evolution*, 2017, 49(1): 49 (doi: 10.1186/s12711-017-0325-2).

56. Signer-Hasler H., Casanova L., Barenco A., Maitre B., Bagnato A., Vevey M., Berger B., Simčič M., Boichon D., Capitan A., Medugorac I., Bennewitz J., Mészáros G., Sölkner J., Drögemüller C., Flury C. Genomic regions underlying positive selection in local, Alpine cattle breeds. *Animal Genetics*, 2023, 54(3): 239-253 (doi: 10.1111/age.13295).
57. Xu L., Zhao G., Yang L., Zhu B., Chen Y., Zhang L., Gao X., Gao H., Liu G.E., Li J. Genomic patterns of homozygosity in Chinese local cattle. *Scientific Reports*, 2019, 9(1): 16977 (doi: 10.1038/s41598-019-53274-3).
58. Ask-Gullstrand P., Strandberg E., Bege R., Rius-Vilarrasa E., Berglund B. The effect of genetic defects on pregnancy loss in Swedish dairy cattle. *Journal of Dairy Science*, 2024, 107(5): 2999-3005 (doi: 10.3168/jds.2023-24159).
59. Romanenkova O.S., Volkova V.V., Kostyunina O.V., Gladyr' E.A., Naryshkina E.N., Serm-yagin A.A., Zinovieva N.A. The distribution for LoF mutations in the *FANCI*, *APAF1*, *SMC2*, *GART*, and *APOB* genes of the Russian Holstein cattle population. *Journal of Animal Science*, 2017, 95(Suppl_4): 83 (doi: 10.2527/asasann.2017.168).
60. Khan M.Y.A., Dai D., Su X., Tian J., Zhou J., Ma L., Wang Y., Wen W., Zhang Y. Multiplex fluorescent amplification-refractory mutation system PCR method for the detection of 10 genetic defects in Holstein cattle and its comparison with the KASP genotyping assay. *Animal Genetics*, 2024, 55(3): 457-464 (doi: 10.1111/age.13432).
61. Bertolini F., Moscatelli G., Schiavo G., Bovo S., Ribani A., Ballan M., Bonacini M., Prandi M., Dall'Olio S., Fontanesi L. Signatures of selection are present in the genome of two close autochthonous cattle breeds raised in the North of Italy and mainly distinguished for their coat colours. *Journal of Animal Breeding and Genetics*, 2022, 139(3): 307-319 (doi: 10.1111/jbg.12659).
62. Hauser M., Signer-Hasler H., Küttel L., Capitan A., Guldbrandtsen B., Hinrichs D., Flury C., Seefried F.R., Drögemüller C. Identification of two new recessive MC1R alleles in red-coloured Evolener cattle and other breeds. *Animal Genetics*, 2022, 53(3): 427-435 (doi: 10.1111/age.13206).
63. Knaust J., Hadlich F., Weikard R., Kuehn C. Epistatic interactions between at least three loci determine the "rat-tail" phenotype in cattle. *Genetics Selection Evolution*, 2016, 48: 26 (doi: 10.1186/s12711-016-0199-8).
64. Signer-Hasler H., Burren A., Neuditschko M., Frischknecht M., Garrick D., Stricker C., Gredler B., Bapst B., Flury C. Population structure and genomic inbreeding in nine Swiss dairy cattle populations. *Genetics Selection Evolution*, 2017, 49(1): 83 (doi: 10.1186/s12711-017-0358-6).
65. Buaban S., Lengnudum K., Boonkum W., Phakdeedindan P. Genome-wide association study on milk production and somatic cell score for Thai dairy cattle using weighted single-step approach with random regression test-day model. *Journal of Dairy Science*, 2022, 105(1): 468-494 (doi: 10.3168/jds.2020-19826).
66. Lázaro S.F., Tonhati H., Oliveira H.R., Silva A.A., Scalez D.C.B., Nascimento A.V., Santos D.J.A., Stefani G., Carvalho I.S., Sandoval A.F., Brito L.F. Genetic parameters and genome-wide association studies for mozzarella and milk production traits, lactation length, and lactation persistency in Murrah buffaloes. *Journal of Dairy Science*, 2024, 107(2): 992-1021 (doi: 10.3168/jds.2023-23284).
67. Daetwyler H.D., Schenkel F.S., Sargolzaei M., Robinson J.A. A genome scan to detect quantitative trait loci for economically important traits in Holstein cattle using two methods and a dense single nucleotide polymorphism map. *Journal of Dairy Science*, 2008, 91(8): 3225-3236 (doi: 10.3168/jds.2007-0333).
68. Cruickshank J., Dentine M., Berger P.J., Kirkpatrick B.W. Evidence of quantitative trait loci affecting twinning rate in North American Holstein cattle. *Animal Genetics*, 2004, 35(3): 206-212 (doi: 10.1111/j.1365-2052.2004.01138.x).
69. Barreto Amaral Teixeira D., Alves Fernandes Júnior G., Beraldo Dos Santos Silva D., Bernal Costa R., Takada L., Gustavo Mansan Gordo D., Bresolin T., Carvalheiro R., Baldi F., Galvão de Albuquerque L. Genomic analysis of stayability in Nelore cattle. *PLoS ONE*, 2017, 12(6): e0179076 (doi: 10.1371/journal.pone.0179076).
70. Bovo S., Schiavo G., Kazemi H., Moscatelli G., Ribani A., Ballan M., Bonacini M., Prandi M., Dall'Olio S., Fontanesi L. Exploiting within-breed variability in the autochthonous Reggiana breed identified several candidate genes affecting pigmentation-related traits, stature and udder defects in cattle. *Animal Genetics*, 2021, 52(5): 579-597 (doi: 10.1111/age.13109).
71. Kichaev G., Bhatia G., Loh P.R., Gazal S., Burch K., Freund M.K., Schoech A., Pasaniuc B., Price A.L. Leveraging polygenic functional enrichment to improve GWAS power. *American Journal of Human Genetics*, 2019, 104(1): 65-75 (doi: 10.1016/j.ajhg.2018.11.008).