

## ИДЕНТИФИКАЦИЯ *Bos taurus* И *Bos grunniens* НА ОСНОВАНИИ SNP

© 2025 г. В. Н. Кипень<sup>1, \*</sup>, Ж. Т. Исакова<sup>2, 4</sup>, М. М. Патрин<sup>3</sup>, К. Б. Чекиров<sup>4</sup>,  
К. А. Айтбаев<sup>2</sup>, А. Р. Карыпова<sup>4</sup>, М. И. Ирсалиев<sup>2</sup>

<sup>1</sup>Институт генетики и цитологии Национальной академии наук Беларуси, Минск, 220072 Республика Беларусь

<sup>2</sup>Научно-исследовательский институт молекулярной биологии и медицины, Бишкек, 720040

Кыргызская Республика <sup>3</sup>ООО “Максим Медикал”, Москва, 123423 Россия

<sup>4</sup>Кыргызско-Турецкий Университет “Манас”, Бишкек, 720038 Кыргызская Республика

\*e-mail: v.kipen@igc.by

Поступила в редакцию 02.07.2024 г.

После доработки 01.08.2024 г.

Принята к публикации 08.08.2024 г.

В статье исследованы образцы домашнего яка и трех пород крупного рогатого скота для оценки дифференцирующего потенциала полиморфных вариантов Chr4:68609356G>T (ген *JAZF1*), Chr14:35695388G>T (ген *SLCO5A1*) и Chr19:63181970C>G (ген *CEP112*). Подтверждена высокая точность (99.67%), специфичность (100%) и чувствительность (100%) предложенной тест-модели, состоящей из этих трех полиморфизмов, для идентификации домашних яков и крупного рогатого скота. Разработан быстрый и простой метод идентификации на основе данной модели с использованием технологии конкурентной аллель-специфической ПЦР (KASP), что позволяет существенно сократить временные и финансовые затраты на молекулярно-генетический анализ, а также снизить риск перекрестного загрязнения образцов.

**Ключевые слова:** *Bos taurus*, *Bos grunniens*, однонуклеотидный полиморфизм, идентификация, конкурентная аллель-специфическая ПЦР (KASP, Kompetitive allele specific PCR), генотипирование *in silico*.

**DOI:** 10.31857/S0016675825010079 **EDN:** VEOPTT

Домашний як (*Bos grunniens*) представляет собой domesticiрованную форму дикого яка (*Bos mutus*). Эти животные известны своей способностью выживать в экстремальных условиях, таких как низкие температуры, гипоксия и нехватка пищи, что делает их важным объектом для исследования молекулярно-генетических механизмов выживаемости и адаптации к высокогорным районам. Долгая и сложная история домостикации яка, а также физико-географические особенности его местообитания оказывают значительное влияние на генетическое разнообразие современных популяций этого животного. В настоящее время идет активное изучение *Bos grunniens* по всему миру, включая исследования различий в популяциях, генетического многообразия, геногеографии и филогенетических связей [1].

В горах Кыргызстана активно развивается яководство, в то время как на более низких и средних высотах преобладает разведение крупного рогатого скота (КРС). В отличие от последнего, яки предпочитают использовать пастбищные корма низкого роста, что свидетельствует об их уникальных адаптационных возможностях. Домашний як имеет ограниченное географическое распространение

в высокогорных районах Центральной Азии, которое обусловлено его высокой адаптивностью к жизни в условиях гор и плато, подобно дикому родственнику – тибетскому яку.

Оценка генетического разнообразия яков и пород КРС, разводимых в Кыргызстане, в контексте возможной гибридизации между видами проводилась в нашем предыдущем исследовании [2]. Кроме того, в работе [3] были исследованы гибриды яка с КРС, по результатам которой в популяциях яка и гибридов первого поколения был выявлен видоспецифичный для яка паттерн из восьми ISSR-фрагментов. Также на основе микросателлитного анализа ранее был изучен аллелофонд яков и их гибридов с *Bos taurus*, который показал высокое генетическое разнообразие для гибридов первого поколения в сравнении с исходными видами [4].

Идентификация домашнего яка и КРС имеет важное значение для оценки возможной гибридизации между этими видами. Эта задача актуальна как для селекционеров, так и экспертов-криминалистов, особенно в условиях ограниченной доступности референсных данных и высоких затрат на анализ STR-локусов. Оценка частот

распространения видоспецифичных SNP (Single Nucleotide Polymorphism) также представляет интерес с точки зрения динамики эволюционных изменений у домашних видов животных по сравнению с их дикими сородичами, что может быть связано с искусственным отбором, проводимым человеком.

Цель настоящего исследования — использование методов биоинформатики для выявления SNP с высоким потенциалом дифференциации, необходимых для идентификации принадлежности образцов как к биологическому виду *Bos taurus*, так и *Bos grunniens*. На основании проведенного анализа предлагается разработать тест-систему, включающую несколько SNP для идентификации домашнего яка и представителей КРС.

## МАТЕРИАЛЫ И МЕТОДЫ

**Биологические образцы.** Для молекулярно-генетического исследования были собраны образцы крови от 56 домашних яков (*Bos grunniens*) из высокогорного региона Калмак-Ашуу (Кочкорский р-н, Нарынская обл., Кыргызская Республика). Эта группа обозначена как YAK. Также были взяты образцы крови у 146 коров (*Bos taurus*) трех пород: абердин-ангусской ( $n = 45$ , группа ABR), голштинской ( $n = 51$ , группа HOL) и алатауской ( $n = 50$ , группа ALA). Совокупность этих образцов составила выборку COW. Биологический материал был отобран сотрудниками опытных хозяйств и поступал в лабораторию для молекулярно-генетического анализа в пробирках типа Vacutainer (BD Vacutainer® Sodium Citrate Tubes), каждая из которых была снабжена биркой с описанием животного (видовая и породная принадлежность, возраст, пол).

**Выделение ДНК.** Образцы крови хранились при температуре  $-20^{\circ}\text{C}$  в течение 1–2 мес. либо при  $-80^{\circ}\text{C}$  при долговременном хранении. ДНК экстрагировали с использованием набора Blood-Animal-Plant DNA Preparation Kit (Jena Bioscience, Германия) в соответствии с инструкциями производителя. Количественная оценка выделенной ДНК проводилась с помощью NanoDrop 1000 (Thermo Fisher Scientific, США). Средняя концентрация ДНК составила  $72.4 \pm 23.3$  нг/мкл, при соотношении 260/280 —  $1.89 \pm 0.09$ .

**KASP.** Определение генотипа по SNP (версия генома *Bos taurus* UMD\_3.1.1, GCF\_000003055.6) Chr4:68609356G>T (*JAZF1*), Chr14:35695388G>T (*SLC05A1*), Chr19:63181970C>G (*CEP112*) проводилось с применением технологии, основанной на конкурентной аллель-специфической ПЦР (KASP, Kompetitive allele specific PCR). Генотипирование проводили с использованием KASP Assay mix (KASP by Design, KBD) и KASP Master mix (LGC Biosearch Technologies, Великобритания; ООО “Максим Медикал”, РФ) в объеме 10 мкл

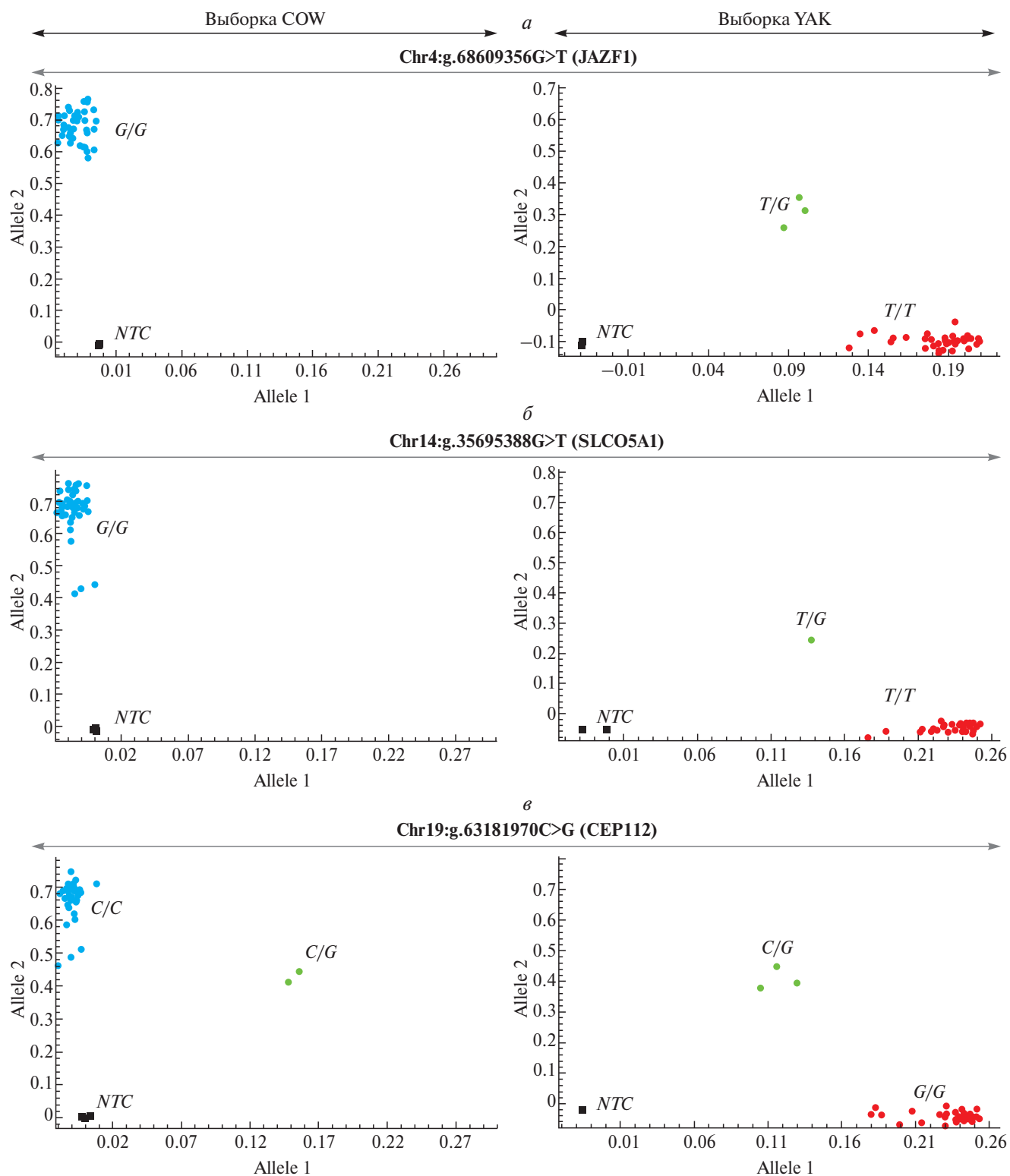
в термоциклере QuantStudio™ 5 Real-Time PCR System (Thermo FS, США) в соответствии с рекомендациями производителя. На рис. 1 представлены 2D-графики аллельной дискриминации для SNP Chr4:68609356G>T (*JAZF1*), Chr14:35695388G>T (*SLC05A1*), Chr19:63181970C>G (*CEP112*), использовавшихся в тестах для идентификации домашних яков и КРС.

**Определение генома in silico.** Для проведения генотипирования *in silico* использовались геномы животных, доступные в базе данных NCBI, конвертированные в формат \*.fasta с помощью пакета SRA-Toolkit версии 2.11. Для определения генотипов использовалось оригинальное программное обеспечение GENIS, разработанное на языке Python версии 3.10. Подробности методологии описаны в статье [5].

В рамках исследования генотипирование было проведено для 316 особей домашнего яка и 385 особей КРС. Таким образом, в биоинформатическом анализе был задействован 701 отсеквенированный геном, файлы для которых расположены в базе Sequence Read Archive (SRA) [6]: PRJNA74739 (2012, Китайская Народная Республика — КНР), PRJNA217895 (2013, КНР), PRJNA285834 (2015, КНР), PRJEB18113 (2016, Швейцария), PRJNA508864 (2018, КНР), PRJNA431934 (2018, Австралия), PRJNA531398 (2019, КНР), PRJNA762180 (2021, Великобритания), PRJNA766811 (2021, КНР), PRJNA842787 (2022, КНР), PRJNA899924 (2022, КНР), PRJNA950586 (2023, КНР).

**Статистический анализ данных.** Для оценки потенциала SNP в качестве идентификационных маркеров использовался ROC-анализ (Receiver Operating Characteristic analysis) в программе SPSS версии 20.0. SNP полагался высокоэффективным маркером при условии, что нижняя граница асимптотического 95%-ного доверительного интервала (ДИ) для параметра площади под кривой (AUC) превышала 0.8.

Комплексную оценку дифференцирующего потенциала для совокупности SNP проводили с использованием программы MDR v.3.0.2 [7]. Вклад конкретного генотипа определялся величиной энтропии  $H$  (выраженной в %). При  $H = 100\%$  генотип способен однозначно дифференцировать, к какой группе относится образец. В программе MDR для определения оптимальной модели дифференциации использовались следующие высококонсервативные настройки: количество атрибутов (attribute count range) — от 1 до  $n$  (где  $n$  — количество переменных в модели); воспроизводимость модели (cross-validation count) — 100; анализ топ-моделей (track top models) — 1000; поиск конфигурации модели (search method configuration) — всесторонний (exhaustive); метод сравнения (ambiguous cell analysis) — точный тест Фишера (Fisher's exact test);



**Рис. 1.** График аллельной дискриминации результатов KASP для полиморфизмов: а – Chr4:68609356G>T (*JAZF1*); б – Chr14:35695388G>T (*SLCO5A1*); в – Chr19:63181970C>G (*CEP112*).

классификация ячеек (ambiguous cell assignment) – неклассифицированные (unclassified). Корректность модели оценивалась по значению сбалансированной точности (adj. Balanced Accuracy). Вероятность отнесения образца к одной из двух выборок – YAK или COW – рассчитывалась в SPSS v.20.0 с использованием логистической регрессии.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Отбор наиболее информативных SNP включал два этапа. На первом этапе были сформированы две тестовые группы, образцы в которые были включены случайным образом: BG-1 (домашний як,  $n = 56$ ) и BT-1 (КРС,  $n = 60$ ), для которых были определены генотипы по 947 SNP (перечень SNP представлен в файле “Дополнительные материалы”, табл. Д-1). В перечень потенциально информативных SNP вошли как ранее описанные SNP [8], так и фланкирующие их SNP из BovineHD BeadChip от Illumina© [9].

Для оценки информативности SNP был проведен ROC-анализ, который позволил выделить 64 SNP с наибольшими значениями параметра AUC (area under ROC curve) – площади, ограниченной ROC-кривой и осью доли ложных положительных классификаций. Эти маркеры были отобраны для дальнейших исследований. На втором, расширенном этапе генотипы этих 64 SNP были определены для дополнительных 260 животных *Bos grunniens* (группа BG-2) и 325 животных *Bos taurus* (группа BT-2). Список отобранных для генотипирования *in silico* животных представлен в табл. Д-2 (доп. мат.). Повторный ROC-анализ был проведен для уточнения SNP с наибольшим дифференцирующим потенциалом для различения домашнего яка и крупного рогатого скота. Полученные результаты использовались для построения ROC-кривых и оценки показателя AUC. По результатам биоинформатического анализа, значения AUC для отобранных SNP находились в диапазоне от 0.899 до 0.999, что указывает на их высокий дифференцирующий потенциал (доп. мат., табл. Д-3).

Для исследования взаимодействий SNP был использован метод многомерного сокращения размерности (MDR v.3.0.2). На основе данных, полученных в ходе биоинформатического анализа SRA, были отобраны три SNP с наибольшим дифференцирующим потенциалом: Chr4:68609356G>T (*JAZF1*), Chr14:35695388G>T (*SLCO5A1*) и Chr19:63181970C>G (*CEP112*). Значения энтропии  $H$  для этих SNP составили 78.19, 77.79 и 77.07% соответственно. Далее, с использованием технологии KASP, были определены генотипы для 202 образцов домашнего яка и крупного рогатого скота из Кыргызстана. Повторный анализ с использованием метода MDR позволил рассчитать величину энтропии  $H$  для каждого SNP: для Chr4:68609356G>T (*JAZF1*)

$H = 85.16\%$ , для Chr14:35695388G>T (*SLCO5A1*)  $H = 85.16\%$ , для Chr19:63181970C>G (*CEP112*)  $H = 82.76\%$ . Результаты генотипирования представлены в табл. Д-4 (доп. мат.).

В результате ROC-анализа было установлено, что все три SNP обладают высочайшим дифференцирующим потенциалом для различения домашнего яка и крупного рогатого скота. Значения AUC для SNP составили: Chr4:68609356G>T (*JAZF1*) AUC = 1.0, 95%ДИ = [1.0–1.0],  $p = 4.17 \cdot 10^{-28}$ ; Chr14:35695388G>T (*SLCO5A1*) AUC = 1.0, 95%ДИ = [1.0–1.0],  $p = 4.17 \cdot 10^{-28}$ ; Chr19:63181970C>G (*CEP112*) AUC = 1.0, 95%ДИ = [0.999–1.0],  $p = 4.56 \cdot 10^{-28}$ . Частоты аллелей также подтвердили их значимость: аллель  $G$  для Chr4:68609356G>T в выборке COW составил 100%, аллель  $G$  для Chr14:35695388G>T – 100%, аллель  $C$  для Chr19:63181970C>G – 99.32%. В выборке YAK частоты аллелей были следующими: аллель  $T$  для Chr4:68609356G>T – 97.32%, аллель  $T$  для Chr14:35695388G>T – 99.11%, аллель  $G$  для Chr19:63181970C>G – 97.32%.

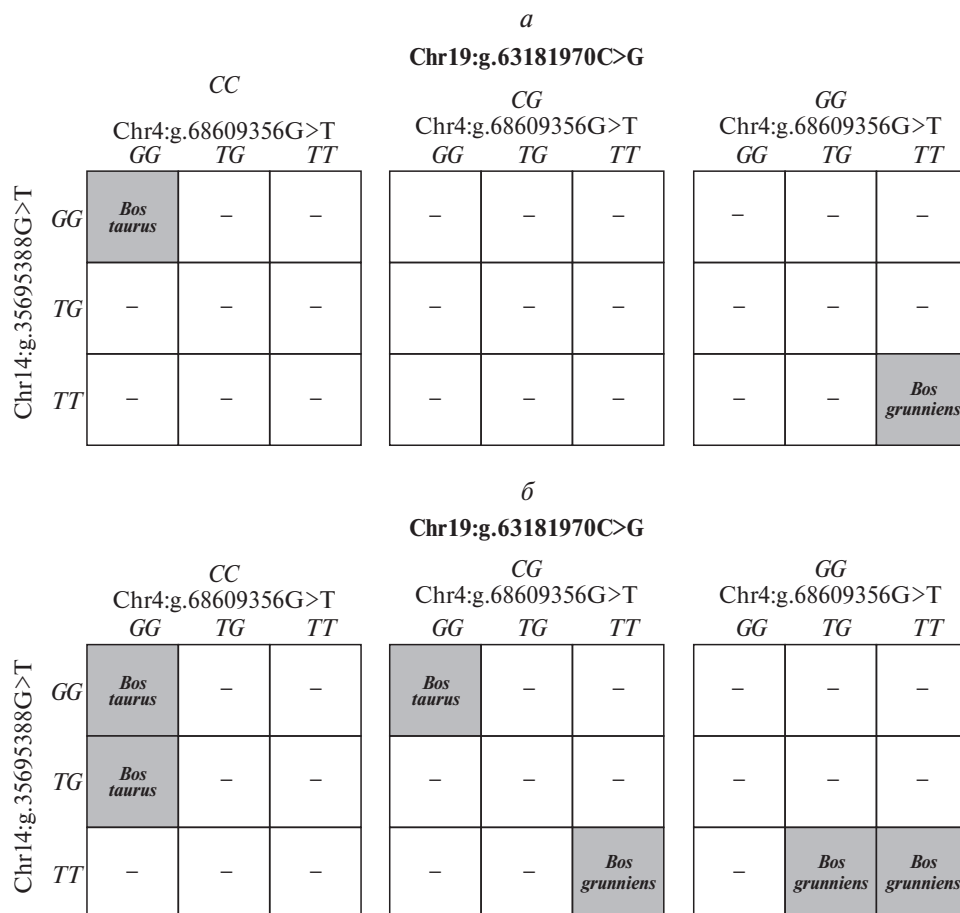
Таким образом, в тест-систему были включены три SNP, графическая интерпретация модели на основании анализа 202 образцов представлена на рис. 2, а. Согласно полученным результатам, сбалансированная точность идентификации, (adj. Balanced accuracy) домашнего яка и КРС при анализе 202 образцов по SNP Chr4:68609356G>T (*JAZF1*), Chr14:35695388G>T (*SLCO5A1*) и Chr19:63181970C>G (*CEP112*) составила 98.87% (специфичность модели – 100%, чувствительность – 100%). При объединении двух массивов данных – образцов из Кыргызстана и образцов из NCBI-SRA, для которых имелись данные по трем SNP, с последующим анализом в MDR определено, что сбалансированная точность идентификации (adj. Balanced accuracy) домашнего яка и КРС составила 99.67% (специфичность модели – 100%, чувствительность – 100%). Графическая интерпретация модели на основании анализа 611 образцов представлена на рис. 2, б.

При оценке точности отнесения образца к одной из двух групп: домашний як или КРС, с использованием логистической регрессии, получены следующие данные (доп. мат., табл. Д-5):

при наличии генотипа  $CC$  (Chr19:63181970C>G) /  $GG$  (Chr4:68609356G>T) /  $GG$  (Chr14:35695388G>T) вероятность отнесения образца к выборке COW составляет 100% (в совокупности распространенность особей с данными генотипами в группе COW – 94.79%);

при наличии генотипа  $CG$  (Chr19:63181970C>G) /  $GG$  (Chr4:68609356G>T) /  $GG$  (Chr14:35695388G>T) или  $CC$  (Chr19:63181970C>G) /  $GG$  (Chr4:68609356G>T) /  $TG$  (Chr14:35695388G>T) вероятность отнесения образца к выборке COW





**Рис. 2.** Графическое представление модели из трех полиморфизмов для идентификации домашнего яка и КРС: *a* – 202 образца (образцы из Кыргызстана); *б* – 611 образцов (образцы из Кыргызстана и данные SRA).

составляет 100% (в совокупности распространенность особей с данными генотипами в группе COW – 4.22%);

при наличии генотипа *GG* (Chr19:63181970C>G) / *TT* (Chr4:68609356G>T) / *TT* (Chr14:35695388G>T) вероятность отнесения образца к выборке YAK составляет 100% (в совокупности распространенность особей с данными генотипами в группе COW – 90.87%);

при наличии генотипа *CG* или *GG* (Chr19:63181970C>G) / *TT* или *TG* (Chr4:68609356G>T) / *TT* (Chr14:35695388G>T) вероятность отнесения образца к выборке YAK составляет 100% (в совокупности распространенность особей с данными генотипами в группе COW – 7.21%);

в 1.31% случаев особь не удалось отнести ни к одному кластеру с заявленным в MDR уровнем точности в 99.0%.

Таким образом, в настоящей работе мы показали, что сбалансированная точность дифференциации *Bos taurus* и *Bos grunniens* составила более

99%, что вполне достаточно для решения большинства задач популяционной генетики. Однако в 1.31% случаев (согласно модели, рис. 2, б) образцы домашнего яка и КРС не могут быть корректно дифференцированы с уровнем значимости  $p < 0,01$ . Это обусловлено относительно небольшим количеством особей с редкими генотипами, такими как *CC* (Chr19:63181970C>G) / *TG* (Chr4:68609356G>T) / *GG* (Chr14:35695388G>T) – для КРС, или *GG* (Chr19:63181970C>G) / *TT* (Chr4:68609356G>T) / *TG* (Chr14:35695388G>T) – для домашнего яка. Если же использовать уровень значимости  $p < 0.05$  при MDR-анализе, то все 611 образцов дифференцируются корректно. Аналогичная ситуация показана для результатов дифференциации при использовании логистической регрессии.

В перспективе для увеличения точности модели, особенно для криминалистических приложений, может быть рекомендовано увеличение выборки исследуемых образцов за счет включения новых образцов с заведомо известной видовой принадлежностью. Увеличение числа редких генотипов при сохранении их относительной частоты

в выборках домашнего яка и КРС приведет к снижению процента образцов, которые не могут быть корректно дифференцированы.

#### Характеристика исследуемых генов

Ген *JAZF1* (*JAZF* zinc finger 1, NCBI Gene ID 616701) расположен на 4-й хромосоме (NC\_037331.1 (67992971..68321145), ARS-UCD2.0 (GCF\_002263795.3)). Ортолог данного гена у человека кодирует ядерный белок с тремя цинковыми пальцами типа C2H2 и действует как репрессор транскрипции. Этот ген участвует в липидном обмене, подавляя липогенез и усиливая липолиз, что приводит к уменьшению накопления липидов в жировой ткани. *JAZF1* также задействован в гомеостазе глюкозы, улучшая метаболизм глюкозы и чувствительность к инсулину. Белки *JAZF1* у КРС и человека имеют схожий размер. В исследовании Eusebi G.P. et al. изучались профили экспрессии генов в префронтальной коре агрессивных пород КРС, включая *JAZF1*, который показал пониженную экспрессию у испанской породы Lidia, известной своим агонистическим поведением [10]. Jin M. et al. обнаружили, что *JAZF1* участвует в адаптации овец *Ovis aries* к высокогорным условиям [11]. Zhao F. et al. выделили *JAZF1* как один из ключевых генов, влияющих на продуктивные характеристики КРС [12]. Исследователи установили, что полиморфизм *JAZF1* ассоциирован с ростом у людей [13].

Ген *SLCO5A1* (solute carrier organic anion transporter family member 5A1, NCBI Gene ID 535202) расположен на 14-й хромосоме (NC\_037341.1 (33483001..33770931), ARS-UCD2.0 (GCF\_002263795.3)). Ортолог данного гена у человека отвечает за активность трансмембранного переносчика органических анионов, независимую от натрия. Белок *SLCO5A1* расположен во внутриклеточных мембраносвязанных органеллах и плазматической мембране. Размер белков у КРС и человека различается незначительно – 846 и 848 аминокислот соответственно. При ассоциативном анализе репродуктивных признаков у *Sus scrofa domestica* и генных сетей на основании полногеномных исследований было показано, что ген *SLCO5A1* связан с транскрипционными факторами, участвующими в функционировании молочной железы и скелетных мышц [14]. В работе Gaddis K.L.P. et al. была обнаружена связь гена *SLCO5A1* с восприимчивостью к кетозу у джерсейской породы КРС [15]. В другом исследовании был определен регион на 14-й хромосоме генома *Bos taurus* (абердин-ангусы) с потенциально высокой ассоциацией с массой при рождении, где также расположен ген *SLCO5A1* [16].

Ген *CEP112* (centrosomal protein 112, NCBI Gene ID 617266) расположен на 19-й хромосоме (NC\_037346.1 (62280043..62600661), ARS-UCD2.0

(GCF\_002263795.3)). Ортолог этого гена у человека кодирует белок со спиральным доменом, который принадлежит к семейству эффекторных белков, контролирующих клеточное деление. *CEP112* был идентифицирован как компонент centrosомы человека. Для этого гена характерен альтернативный сплайсинг, что приводит к образованию множества вариантов транскрипта. В исследовании Kim S. и соавт. установлено, что полиморфизм rs42699274 гена *CEP112* может служить генетическим маркером при оценке генетического разнообразия и дивергенции среди корейских пород КРС [17]. В другом исследовании, направленном на повышение точности геномного прогнозирования характеристик телосложения у корейского КРС голштинской породы, было показано, что полиморфизм гена *CEP112* является важным маркером для понимания генетических основ этих признаков и может служить надежной основой для прогнозов, основанных на геномных данных [18].

Известно, что прямое перенесение результатов, полученных для одного биологического вида, на другой вид некорректно и требует дополнительных исследований. Тем не менее полученные результаты могут быть полезны для дальнейшего изучения эволюции филогеографической структуры популяций домашнего яка и пород КРС в Кыргызстане. Они также могут способствовать сохранению генетических ресурсов для будущих селекционных исследований и расширить понимание того, как животные могут адаптироваться к изменениям климата. Все три SNP, включенные в тест-систему, расположены в интронных областях генов и не должны оказывать влияние на функционирование экспрессируемых белков. Однако различные события одомашнивания, адаптация к различным климатическим зонам и дивергентный отбор по продуктивным признакам сформировали геномные различия между исследуемыми видами.

Ранее нами было показано, что при анализе пяти STR-локусов (*BM1818*, *BM1824*, *BM2113*, *CSSM66* и *ILSTS006*) точность классификации для *Bos grunniens* составила  $98.8 \pm 3.4\%$ , для *Bos taurus* –  $99.1 \pm 1.2\%$  [2]. При использовании же трех SNP (Chr4:68609356G>T, Chr14:35695388G>T и Chr19:63181970C>G) точность классификации составила 99.67% при максимально возможных значениях специфичности и чувствительности. При этом образцы YAK\_107, YAK\_117 и YAK\_131, для которых по результатам STR-анализа точность отнесения к своему кластеру составила 70.3, 75.5 и 72.3% соответственно, по результатам SNP-анализа были однозначно (100% точность) классифицированы как домашние яки.

Итак, в настоящем исследовании нами предложена тест-система для дифференциации домашнего яка (*Bos grunniens*) и крупного рогатого скота (*Bos taurus*), основанная на анализе трех SNP

в генах *JAZF1*, *SLCO5A1* и *CEP112*. С использованием методов биоинформатики были определены SNP с высоким дифференцирующим потенциалом. Эти полиморфные варианты подтвердили свою эффективность на практике. Модель, основанная на трех SNP (Chr4:68609356G>T в гене *JAZF1*, Chr14:35695388G>T в гене *SLCO5A1* и Chr19:63181970C>G в гене *CEP112*), продемонстрировала высокую сбалансированную точность (не менее 99.67%) при анализе 611 образцов. Применение данного подхода позволило достичь 100% чувствительности и 100% специфичности в дифференциации двух видов. Метод KASP, используемый в молекулярно-генетическом анализе, отличается одностадийностью, что снижает риск кросс-контаминации и уменьшает трудовые затраты благодаря исключению этапов рестрикции и электрофореза. Дальнейшее расширение базы данных образцов с помощью генотипирования, особенно *in silico*, остается актуальной задачей для совершенствования и применения данной тест-системы.

Программное обеспечение GENIS разработано в рамках НИР “Биоинформатический подход к анализу данных полногеномного секвенирования для поиска однонуклеотидных замен, способных дифференцировать близкородственные биологические виды” (БРФФИ, 2023–2025 гг., Б23-060, рег. № 20231076).

Сбор биологических образцов домашнего яка и КРС выполнен в рамках правительственного задания Министерства образования и науки Кыргызской Республики (договор № 30-21 от 02.15.2021, № 129/1).

Молекулярно-генетические исследования выполнены при частичной финансовой поддержке ООО “МАКСИМ МЕДИКАЛ” в рамках частной научной инициативы ANALYSIS OF SRA DATA USING BIOINFORMATICS METHODS.

Исследование одобрено Этическим комитетом Научно-исследовательского института молекулярной биологии и медицины (Бишкек, Кыргызская Республика, 12.06.2024, протокол № 6).

Все применимые международные, национальные и/или институциональные принципы ухода и использования животных были соблюдены.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Wang X., Pei J., Xiong L. et al. Genetic diversity, phylogeography, and maternal origin of yak (*Bos grunniens*) // BMC Genomics. 2024. V. 15. № 25(1). P. 481. P. 1–13. <https://doi.org/10.1186/s12864-024-10378-z>
2. Chekirov K.B., Isakova Zh.T., Kipen V.N. et al. Differentiation of *Bos grunniens* and *Bos taurus* based on STR locus polymorphism // Vavilov J. of Genet. and Breeding. 2023. V. 27(5). P. 488–494. <https://doi.org/10.18699/VJGB-23-59>
3. Stolpovsky Y.A., Kol N.V., Evsyukov A.N. et al. Comparative analysis of ISSR marker polymorphism in population of YAK (*Bos mutus*) and in F1 hybrids between yak and cattle in the Sayzn-Altai region // Russ. J. Genet. 2014. V. 50. № 10. P. 1163–1176. <https://doi.org/10.7868/S0016675814100142>
4. Аль-Кейси Т.В. Сравнительное исследование аллелофонда яков и их гибридов с крупным рогатым скотом с использованием микросателлитов: дис. ... канд. биол. наук. М.: ВИЖ, 2011. 97 с.
5. Kipen V.N., Snytkov E.V. GENIS – methodological approach for *in silico* genotyping (validation on *Sus scrofa* sequencing) // Mathematical Biol. and Bioinf. 2024. V. 19(1). P. 36–51. <https://doi.org/10.17537/2024.19.36>
6. Sequence Read Archive (SRA) [электронный ресурс]. URL: <https://www.ncbi.nlm.nih.gov/sra>. Дата доступа: 20.06.2024.
7. Ritchie M.D., Hahn L.W., Roodi N. et al. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer // Am. J. Hum. Genet. 2001. V. 69(1). P. 138–147. <https://doi.org/10.1086/321276>
8. Kalbfleisch T., Petersen J.L., Tait Jr. R.G. et al. Using triallelic SNPs for determining parentage in North American yak (*Bos grunniens*) and estimating cattle (*B. taurus*) introgression // F1000Research. 2020. V. 9(1096). P. 1–26. <https://doi.org/10.12688/f1000research.25803.2>
9. BovineHD DNA Analysis Kit [электронный ресурс]. Режим доступа: <https://www.illumina.com/products/by-type/microarray-kits/bovinehd.html>. Дата доступа 21.06.2024.
10. Eusebi P.G., Sevane N., O'Rourke T. et al. Gene expression profiles underlying aggressive behavior in the prefrontal cortex of cattle // BMC Genomics. 2021. V. 22(245). P. 1–14. <https://doi.org/10.1186/s12864-021-07505-5>
11. Jin M., Wang H., Liu G. et al. Whole-genome resequencing of Chinese indigenous sheep provides insight into the genetic basis underlying climate adaptation // Genet. Sel. Evol. 2024. V. 2. № 56(1). P. 1–14. <https://doi.org/10.1186/s12711-024-00880-z>
12. Zhao F., McParland S., Kearney F. et al. Detection of selection signatures in dairy and beef cattle using high-density genomic information // Genet. Select. Evol. 2015. V. 47(1). № 49. P. 1–12. <https://doi.org/10.1186/s12711-015-0127-3>
13. Johansson A., Marroni F., Hayward C., Franklin C.S. et al. Common variants in the JAZF1 gene associated with height identified by linkage and genome-wide association analysis // Hum. Mol. Genet. 2009.

- V. 18(2). P. 373–380.  
<https://doi.org/10.1093/hmg/ddn350>
14. Lima Verardo L.L. Gene networks from genome wide association studies for pig reproductive traits: Tese (Doutorado em Zootecnia). Viçosa: Universidade Federal de Viçosa, 2015. 170 p.  
<http://www.locus.ufv.br/handle/123456789/6773>
  15. Gaddis K.L.P., Megonigal J.H., Clay J.S., Wolfe C.W. Genome-wide association study for ketosis in US Jerseys using producer-recorded data // J. Dairy Sci. 2018. V. 101(1). P. 413–424.  
<https://doi.org/10.3168/jds.2017-13383>
  16. Zeng X. Angus cattle at high altitude: Pulmonary arterial pressure, estimated breeding value and genome-wide association study: Diss. ... PHD. Colorado: Department of Animal Sci., Colorado State Univ., 2016. 259 p.
  17. Kim S., Cheong H.S., Shin H.D. et al. Genetic diversity and divergence among Korean cattle breeds assessed using a BovineHD single-nucleotide polymorphism chip // Asian-Australas J. Anim. Sci. 2018. V. 31(11). P. 1691–1699.  
<https://doi.org/10.5713/ajas.17.0419>
  18. Lee J., Mun H., Koo Y. et al. Enhancing genomic prediction accuracy for body conformation traits in Korean Holstein Cattle // Animals. 2024. V. 14(7). P. 1–14.  
<https://doi.org/10.3390/ani14071052>

## Identification of *Bos taurus* and *Bos grunniens* Based on SNP

V. N. Kipen<sup>1, \*</sup>, Zh. T. Isakova<sup>2, 4</sup>, M. M. Patrin<sup>3</sup>, K. B. Chekirov<sup>4</sup>, K. A. Aitbaev<sup>2</sup>,  
 A. R. Karypova<sup>4</sup>, M. I. Irsaliev<sup>2</sup>

<sup>1</sup>*Institute of Genetics and Cytology of the National Academy of Sciences of Belarus, Minsk, 220072 Republic of Belarus*

<sup>2</sup>*Research Institute of Molecular Biology and Medicine, Bishkek, 720040, Kyrgyz Republic*

<sup>3</sup>*Maxim Medical LLC, Moscow, 123423 Russia*

<sup>4</sup>*Kyrgyz-Turkish Manas University, Bishkek, 720038 Kyrgyz Republic*

\*e-mail: v.kipen@igc.by

The article examined samples of domestic yak and three breeds of cattle to assess the differentiating potential of the polymorphic variants Chr4:68609356G>T (*JAZF1* gene), Chr14:35695388G>T (*SLC05A1* gene) and Chr19:63181970C>G (*CEP112* gene). The high accuracy (99.67%), specificity (100%) and sensitivity (100%) of the proposed test model consisting of these three polymorphisms for the identification of domestic yaks and cattle were confirmed. A fast and simple identification method has been developed based on this model using competitive allele-specific PCR (KASP) technology, which can significantly reduce the time and financial costs of molecular genetic analysis, as well as reduce the risk of cross-contamination of samples.

**Keywords:** *Bos taurus*, *Bos grunniens*, single nucleotide polymorphism, identification, kompetitive allele specific PCR (KASP), in silico genotyping.