

<https://doi.org/10.30766/2072-9081.2020.21.2.169-182>

УДК 636.082:575.174.015.3

Сравнение методов оценки генетической дифференциации популяций по микросателлитным маркерам

© 2020. В. М. Кузнецов ✉

ФГБНУ «Федеральный аграрный научный центр Северо-Востока имени Н. В. Рудницкого», г. Киров, Российская Федерация

Проведены сравнительные исследования восьми методов оценки генетической дифференциации по микросателлитным (STR) маркерам (11 локусов) ДНК на примере семи породных выборок ($N = 84$) молочного скота. Диапазон показателей генетического разнообразия выборок был: числа аллелей на локус – 3,5-6,2, числа эффективных аллелей на локус – 2,4-4,3, индекса Шеннона – 0,95-1,56, наблюдаемой и ожидаемой гетерозиготности – 0,56-0,97 и 0,53-0,75. Сводные оценки генетической дифференциации методами группы А (F_{ST} , G_{ST} и $G_{ST(NEI)}$) составили 13,4, 10,3 и 11,8 % ($p_{perm} < 0,001$); различия между оценками были статистически незначимыми. Оценки методами группы В ($G'_{ST(HED)}$, $G''_{ST(HED)}$, D_{EST}) составили 36,4, 37,5 и 29,2 % ($p_{perm} < 0,001$); различия между оценками также были статистически незначимыми. Оценки, полученные методами группы В, статистически значимо превышали оценки методами группы А почти в 3 раза. Методами групп А, В и С (G_{DN} и uG_{DN}) были рассчитаны парные по выборкам генетические дистанции. Несмотря на значительные различия в оценках, тест Мантеля показал высокую степень соответствия матриц генетических дистанций ($R_M \geq 0,97$; $p_{perm} < 0,001$), которая проявилась в проекциях генетических отношений породных выборок на плоскости 1 и 2 главных координат. В совокупности две первые главные координаты объясняли 97-99 % STR-изменчивости в матрицах генетических дистанций. Можно полагать, что при оценке текущей дифференциации популяций по STR-маркерам следует использовать методы группы В. Из них – D_{EST} -статистику, как не зависящую от уровня внутривидовой гетерозиготности. При исследовании пространственной ординации генофондов популяций, по всей вероятности, правомерно использовать любой метод.

Ключевые слова: микросателлиты, разнообразие, генная дифференциация, генетическая дистанция, анализ главных координат, животноводство

Благодарности: работа выполнена в рамках Государственного задания ФГБНУ «Федеральный аграрный научный центр Северо-Востока имени Н. В. Рудницкого» (тема № АААА-А19-119042290136-4).

Конфликт интересов: автор заявил об отсутствии конфликта интересов.

Для цитирования: Кузнецов В. М. Сравнение методов оценки генетической дифференциации популяций по микросателлитным маркерам. Аграрная наука Евро-Северо-Востока. 2020;21(2):169-182.

<https://doi.org/10.30766/2072-9081.2020.21.2.169-182>

Поступила: 22.01.2020

Принята к публикации: 03.03.2020

Опубликована онлайн: 21.04.2020

Comparison of methods for evaluating genetic differentiation of populations by microsatellite markers

© 2020. Vasily M. Kuznetsov ✉

Federal Agricultural Research Center of the North-East named N. V. Rudnitsky, Kirov, Russian Federation

Comparative studies of eight methods for evaluation of genetic differences by microsatellite (STR) markers (11 loci) DNA were carried out using the example of seven breed samples ($N = 84$) of dairy cattle. The range of indicators of genetic diversity of samples was as follows: the number of alleles per locus 3.5-6.2, the number of effective alleles per locus – 2.4-4.3, the Shannon index – 0.95-1.56, the observed and expected heterozygosity – 0.56-0.97 and 0.53-0.75. Evaluation summary of genetic differentiation by group A methods (F_{ST} , G_{ST} and $G_{ST(NEI)}$) were 13.4, 10.3, and 11.8% ($p_{perm} < 0.001$); the differences between the estimates were statistically insignificant. Estimates by group B methods ($G'_{ST(HED)}$, $G''_{ST(HED)}$, D_{EST}) were 36.4, 37.5 and 29.2% ($p_{perm} < 0.001$); differences between the estimates were also statistically insignificant. Estimates obtained by group B methods were statistically significantly higher than those obtained by group A methods by almost 3 times. Methods of groups A, B, and C (G_{DN} and uG_{DN}) were used to calculate paired genetic distances in the samples. Despite significant differences in estimates, the Mantel test showed a high degree of correspondence of the genetic distance matrices ($R_M \geq 0.97$; $p_{perm} < 0.001$), which was manifested in the projections of genetic relations of breed samples on plane 1 and 2 of the principal coordinates. The first two principal coordinates explained 97-99% of STR variation in the genetic distance matrices. It can be assumed that when assessing the actual differentiation of populations by STR-markers, the methods of group B should

be used and among them the statistics of D_{EST} as independent of the level of mean within-subpopulation heterozygosity. In the study of spatial ordination of gene pools of populations, in all probability, it is reasonably to use any method.

Keywords: *microsatellites, diversity, genetic differentiation, genetic distance, principal coordinate analysis, animal breeding*

Acknowledgement: the research was carried out within the state assignment of the Federal Agricultural Research Center of the North-East named N. V. Rudnitsky (theme No. AAAA-A19-119042290136-4).

Conflict of interest: the authors declare no conflict of interest.

For citation: Kuznetsov V. M. Comparison of methods for evaluating genetic differentiation populations by microsatellite markers. *Agrarnaya nauka Evro-Severo-Vostoka* = Agricultural Science Euro-North-East. 2020;21(2):169-182. (In Russ.). <https://doi.org/10.30766/2072-9081.2020.21.2.169-182>

Received: 22.01.2020

Accepted for publication: 03.03.2020

Published online: 21.04.2020

В популяциях большинства видов животных имеет место определенный уровень генетической дифференциации (структурирования), что может быть связано с различными факторами. Так, природные популяции дифференцируются вследствие эволюционных процессов (мутаций, миграции, дрейфа генов, естественного отбора), из-за эколого-географических барьеров и изоляции расстоянием (популяции в непосредственной близости генетически более сходны, чем отдаленные).

Популяции одомашненных животных могут относиться к разным породам, иметь разную численность и несхожие системы спаривания, с разной интенсивностью подвергаться давлению отбора и обмениваться генетическим материалом. Всяческие знания о генетической дифференциации популяций могут быть полезными при выборе улучшающей породы или пород для скрещивания, при принятии решения по сохранению генофонда той или иной вытесняемой породы. В случае чистопородного разведения селекционерам небезынтересно знать генетические различия между линиями, племенными стадами (хозяйствами), при групповом подборе – для минимизации коэффициента инбридинга. Поэтому поиск объективных методов контроля за уровнем генетического разнообразия как внутри, так и между популяциями (породами, линиями, стадами, группами животных) продолжает оставаться актуальным.

При исследовании генетического разнообразия популяций предпочтение отдают полиморфизму на уровне ДНК. Наиболее часто используют высокополиморфные локусы микросателлитной ДНК (Simple Tandem Repeats, STR – простые тандемные последовательности), которые нередко имеют «большое разнообразие аллелей» гена [1]. В последние годы – однонуклеотидный полиморфизм (Single Nucleotide Polymorphism, SNP – однонуклеотидные позиции в геномной ДНК). SNP-маркеры в основном имеют два аллельных

состояния гена. Несмотря на то, что SNP-маркеры считают перспективными для применения в популяционно-генетических анализах [2], STR-маркеры более информативны при изучении близкородственных групп [3]. Они обладают «беспрецедентной способностью» обнаруживать и описывать даже небольшие генетические различия между популяциями; вследствие наличия локусов с большим числом аллелей дают лучшие оценки генетических дистанций [4, 1]. Оба типа маркеров позволяют разложить генетическое разнообразие на внутри- и межпопуляционные компоненты.

В подобных работах, как правило, рассматриваются выборки из нескольких популяций (пород, линий, стад) и целый ряд локусов. Для измерения разнообразия между популяциями используются меры генетического сходства/различия. Оценка по группе популяций характеризует сводный (усреднённый) показатель *генетической дифференциации*. С помощью сводных показателей проясняются межпопуляционные взаимоотношения, скрытые под массой цифр. Мету генетических различий между парой популяций называют «Genetic Distance» – *генетическая дистанция*.

Для оценки генетической дифференциации применяется, как правило, индекс фиксации Райта [5], приспособленный для полиморфных локусов M. Nei [6, 7] – F_{ST} . С введением поправки на гармоничный средний размер выборок и родственное спаривание M. Nei назвал индекс «коэффициентом генной дифференциации» – G_{ST} [8]. В последующем был предложен несмещённый вариант G_{ST} – $G_{ST(NEI)}$, учитывающий число популяций [9].

Вместе с тем было показано, что при анализе STR-маркеров уровень гетерозиготности в пределах популяций оказывает негативное влияние на F_{ST} (G_{ST} , $G_{ST(NEI)}$), сужая диапазон их оценок [10]. Были предложены различные способы устранения этого недостатка. Так, P. W. Hedrick ввёл стандартизацию G_{ST} на его возможное максимальное значение –

$G'_{ST(NEI)}$ [11]. Он также предложил стандартизировать $G_{ST(NEI)}$ и рассчитывать $G''_{ST(NEI)}$. Альтернативный метод разработал L. Jost [12]. Его статистика, D_{EST} – «истинная аллельная дифференциация», базируется на изменчивости эффективного числа аллелей (в отличие от предыдущих, в основе которых ожидаемая гетерозиготность). Для оценки генетических различий между двумя (суб)популяциями обычно используют стандартную (GDN) и/или несмещённую ($uGDN$) дистанции M. Nei, которые есть логарифмы нормированных коэффициентов идентичности аллельных профилей [6, 13].

Было показано, что только в 15% зарубежных публикаций использовались методы, предложенные P. W. Hedrick [14]. В российских работах в основном употребляются F_{ST} -статистика и дистанция Nei. Из последних публикаций можно отметить такие, как: исследование происхождения девяти российских (местных) пород крупного рогатого скота с использованием STR- [15] и SNP-маркеров [16], генетическая дифференциация 19 пород крупного рогатого скота России по SNP-маркерам [17], полногеномный SNP-анализ происхождения, селекции и адаптации российских пород скота [18], STR-дифференциация региональных популяций холмогорского скота [19], SNP-разнообразие и популяционная структура 25 местных пород овец [20], дифференциация региональных популяций романовских овец по STR-маркерам [21], дифференциация линий по STR-маркерам в чистокровной верховой породе лошадей [22], STR-дифференциация четырёх пород свиней [23], генетическое разнообразие по SNP-маркерам популяций domestцированных и диких оленей [24] и дифференциация дальневосточных пчёл по STR-маркерам [25].

Представленные выше методы можно условно объединить в три группы: «группа А» – F_{ST} , G_{ST} и $G_{ST(NEI)}$ статистики, «группа В» – статистики $G'_{ST(NEI)}$, $G''_{ST(NEI)}$ и D_{EST} , «группа С» – статистики GDN и $uGDN$. Возникают, естественно, вопросы: (а) в какой степени оценки, рассчитанные разными методами, соотносятся по величине и (б) какая статистика обеспечивает получение адекватных оценок текущей генетической дифференциации популяций?

Цель исследований – сравнительный анализ восьми методов оценки генетической дифференциации популяций по STR-маркерам.

Материал и методы. Были использованы данные по быкам-производителям, представленные в открытом доступе на сайте ВНИИплем¹. Файл включал 140 быков с различным количеством идентификационных данных и числом STR-локусов (до 18). После фильтрации (наличие породы, страны происхождения, одинакового набора локусов) были отобраны 84 быка, каждый с 11-ю локусами, именно: 10 быков айрширской породы (**AYR**), 10 быков джерсейской породы (**JER**), 10 – красной датской (**RDAT**), 9 – красной шведской (**RSH**), 45 – голштинской породы, из них 13 быков из Германии (**H-DEU**), 17 – из Нидерландов (**H-NLD**), 15 – из США (**H-USA**).

Для характеристики STR-разнообразия породных выборок рассчитывали среднее число аллелей на локус (N_a), число эффективных аллелей на локус (N_e), наблюдаемую (N_o) и ожидаемую (N_e) гетерозиготности, индекс Шеннона (I) и индекс фиксации (F). Сравнимые методы оценки генетической дифференциации популяций представлены в таблице 1.

В основе F_{ST} и G_{ST} -подобных мер – аддитивное разложение общей ожидаемой гетерозиготности на внутри- и межвыборочный компоненты. Высокий полиморфизм STR-маркеров сужает диапазон оценок, получаемых методами группы А. Этот недостаток устраняют методы группы В: $G'_{ST(NEI)}$ и $G''_{ST(NEI)}$ – это оценки G_{ST} и $G_{ST(NEI)}$, стандартизированные на их максимально возможные значения ($G_{ST(max)}$).

D_{EST} – альтернативная мера; в её основе мультипликативное разложение общей вариации эффективного числа аллелей на внутри- и межвыборочный компоненты. Считается, что повышенный полиморфизм STR-маркеров не влияет на значения D_{EST} -оценок. Методы групп А и В обеспечили получение сводных оценок дифференциации по всем породным выборкам (по L. Jost [12] методы группы А количественно определяют близость аллели к фиксации, D_{EST} – относительную степень аллельной дифференциации, а $G'_{ST(NEI)}$ и $G''_{ST(NEI)}$ – попытка сделать методы группы А «настоящими» мерами типа D_{EST}). Они также использовались, наравне с методами группы С, для оценки попарных генетических дистанций. Расчёты проводили по программе GenAlEx 6.502 [26, 27, 28].

¹База генетических данных быков-производителей/ Микросателлиты [Электронный ресурс]. Режим доступа: <http://www.vniiplm.ru/rus/files/Database/DNK/mikrosatellity.pdf> (дата обращения: 09.12.2018).

Таблица 1 – Группы мер генетической дифференциации пород / Table 1 – Groups of measures of genetic differentiation of breeds

| <i>Группа / Group</i> | <i>Статистика / Statistics</i> | <i>Диапазон / Range</i> | <i>Источник / Source</i> |
|-----------------------|--|---------------------------|---------------------------------|
| A | $F_{ST} = (H_T - H_S) / H_T$ | 0...(1 - H _S) | M. Nei, 1972 [6] |
| | $G_{ST} = 1 - cH_S / cH_T$ | 0...(1 - H _S) | M. Nei, R. K. Chesser, 1983 [8] |
| | $G_{ST(NEI)} = \frac{k(cH_T - cH_S)}{kcH_T - cH_S}$ | 0...(1 - H _S) | M. Nei, 1986 [9] |
| B | $G'_{ST(HED)} = \frac{G_{ST}(k-1+cH_S)}{(k-1)(1-cH_S)}$ | 0...1 | P. W. Hedrick, 2005 [11] |
| | $G''_{ST(HED)} = \frac{G_{ST(NEI)}}{(1-cH_S)}$ | 0...1 | P. W. Hedrick, 2005 [11] |
| | $D_{EST} = \left(\frac{k}{(k-1)} \right) \left(\frac{cH_T - cH_S}{(1-cH_S)} \right)$ | 0...1 | L. Jost, 2008 [12] |
| C | $GDN = -\ln \left(\frac{J_{XY}}{\sqrt{J_X J_Y}} \right)$ | 0...∞ | M. Nei, 1972, 1978 [6, 13] |
| | $uGDN = -\ln \left(\frac{J_{XY}}{\sqrt{cJ_X cJ_Y}} \right)$ | 0...∞ | M. Nei, 1972, 1978 [6, 13] |

Примечания: H_S и H_T – ожидаемые гетерозиготности в пределах выборок и в целом по объединённой выборке; cH_S и cH_T – скорректированные H_S и H_T; k – число выборок; J_X, J_Y, J_{XY} – вероятности идентичности аллелей в выборках X, Y и XY соответственно; cJ_X, cJ_Y – скорректированные на размеры выборок J_X, J_Y; ln – натуральный логарифм /

Notes: H_S and H_T – expected heterozygosity within sample and the heterozygosity of the combined sample; cH_S and cH_T – adjusted H_S and H_T; k – number of samples; J_X, J_Y, J_{XY} – probabilities of identity of alleles in samples X, Y and XY, respectively; cJ_X, cJ_Y – adjusted for sample sizes J_X, J_Y; ln – natural logarithm.

Результаты и их обсуждение. Описательная статистика STR-разнообразия породных выборок представлена в таблице 2. Среднее число аллелей на локус варьировало от 3,5 до 6,2, число эффективных аллелей – от 2,4

до 4,3. Это в среднем на треть меньше фактического, что указывало на наличие достаточно большого числа низкочастотных аллелей. Самые низкие показатели получены по JER-выборке, самые высокие – по H-NLD-выборке.

Таблица 2 – Показатели STR-разнообразия породных выборок / Table 2 – Indicators of STR-diversity of breed samples

| <i>Выборка / Sample</i> | <i>n_i</i> | <i>Na</i> | <i>Ne</i> | <i>I</i> | <i>Ho</i> | <i>He</i> | <i>F</i> |
|-------------------------|----------------------|-----------|-----------|----------|-----------|-----------|----------|
| AYR | 10 | 5,1 | 3,3 | 1,33 | 0,69 | 0,67 | -0,04 |
| JER | 10 | 3,5 | 2,4 | 0,95 | 0,56 | 0,53 | -0,09 |
| RDAT | 10 | 5,4 | 3,2 | 1,32 | 0,76 | 0,66 | -0,15 |
| RSH | 9 | 4,8 | 3,4 | 1,30 | 0,76 | 0,65 | -0,15 |
| H-DEU | 13 | 5,6 | 3,9 | 1,44 | 0,74 | 0,71 | -0,04 |
| H-NLD | 17 | 6,2 | 4,3 | 1,56 | 0,97 | 0,75 | -0,30 |
| H-USA | 15 | 4,7 | 3,2 | 1,23 | 0,70 | 0,64 | -0,07 |

Примечания: n_i – размер выборки; Na и Ne – число аллелей и эффективных аллелей на локус; I – индекс Шеннона; Ho и He – наблюдаемая и ожидаемая гетерозиготности; F – индекс фиксации /

Notes: n_i – sample size; Na, Ne – number of alleles and number of effective alleles per locus; I – Shannon index; Ho, He – observed and expected heterozygosity; F – fixation index.

Индекс Шеннона (I) варьировал от 0,95 (JER) до 1,55 (H-NLD). Вообще, оценки I имеют диапазон от 0 до некоторого максимального значения – I_{max} . Величина I, близкая к нулю, говорит о небольшом разнообразии и концентрации наблюдений выборки в 1-2 аллелях. Высокие значения I свидетельствуют о наличии значительного аллельного разнообразия и слабой концентрации наблюдений, что указывает на возрастание энтропии (неоднородности) структуры анализируемой выборки. Величина I_{max} не является строго определённой. Поэтому интерпретировать индекс относительно «количества разнообразия» сложно. Можно только полагать, что «неоднородности» в H-NLD-выборке больше, чем в JER (Галинская и др. [1] отмечали: «...значения индекса Шеннона для микросателлитных маркеров считаются высокими, если они не менее 1,5»).

Распространённое мнение, что гетерозиготность является наиболее естественной мерой уровня генетического разнообразия. В JER-выборке наблюдаемая гетерозиготность (H_o) была наименьшей – 56%, самая большая, 97%, в H-NLD-выборке. В остальных пяти выборках H_o была на уровне 70-75%.

Наблюдаемая гетерозиготность подвержена математическому давлению системы спаривания ($H_o = H_e (1 - F)$, где F – индекс фиксации или коэффициент инбридинга). На ожидаемую гетерозиготность (H_e) такого давления нет – она есть функция только частот аллелей (p_i) в пуле генов выборки ($H_e = 1 - \sum p_i^2$). Независимо от системы спаривания, уровень H_e будет одним. Поэтому H_e более отвечает цели корректного сравнения выборок по уровню генетического разнообразия.

Изолированное разведение животных приводит к повышению гомозиготности, а миграция (поток генов) и скрещивание (межпородное, межлинейное) наоборот – к повышению гетерозиготности животных. При средней $H_e = 0,66$, низкая частота H_e в JER-выборке (0,53) указывает на большую «закрытость» породы, которую она представляет (табл. 2). С другой стороны, высокая частота H_e в выборках H-DEU (0,71) и H-NLD (0,75) свидетельствует о большей открытости этих популяций для «лучшего мирового генофонда» и, возможно, о проводимых дисассортативных спариваниях (отрицательно ассортативных (неслучайных), когда образуются пары, которые характеризуются дополнительно признаками). В выборках R DAT, RSH и H-NLD имел место избыток гетерозигот.

Результаты по гетерозиготности хорошо согласуются с данными литературы. Так, в работе [29] обобщены 145 оценок H_e по STR-маркерам, которые варьировали от 0,40 до 0,91, средняя – 0,66. По пяти выборкам джерсейской породы H_e была 0,57, в голштинской породе США по трём выборкам – 0,64, в айрширской породе по двум выборкам – 0,68, в красной датской – 0,57, в красной шведской – 0,60.

Сводные оценки генетической дифференциации. Все три оценки, рассчитанные методами группы А (табл. 3), были статистически значимые ($p_{perm} \leq 0,001$), но их 95%ДИ перекрывались. Следовательно, различия между оценками можно считать статистически незначимыми. Это можно объяснить достаточно хорошей выравненностью численности выборок. Усреднённая оценка была на уровне 12% с 95%ДИ 8...16%. Отметим, в пределах этого доверительного интервала были оценки генетической дифференциации девяти локальных пород России: при анализе молекулярной вариации (AMOVA) по 11 микросателлитным локусам – 11% [15], при оценке F_{ST} по 35874 SNP-маркерам – 9,8% [16].

Значения статистик группы А не могут быть выше уровня усреднённой по выборкам STR-гомозиготности ($1 - cH_s$) [11]. Чтобы обойти это ограничение диапазона оценок, P. W. Hedrick предложил относить G_{ST} к максимально возможному значению, $G_{ST(max)}$. Стандартизованная оценка, $G'_{ST(HEd)}$, составила $36,4 \pm 2,9\%$ при $p_{perm} \leq 0,001$ с 95%ДИ 31,0...41,5%. Подобное преобразование было применено и к оценке $G_{ST(NEd)}$, которая после стандартизации составила $G''_{ST(HEd)} = 37,5 \pm 2,9\%$ при $p_{perm} \leq 0,001$ с 95%ДИ 32,4...42,6%. Различие $G'_{ST(HEd)}$ и $G''_{ST(HEd)}$ оценок было небольшое и статистически незначимое. Усреднённая оценка была на уровне 37%, что в три раза больше, чем до стандартизации ($\approx 12\%$). 95%ДИ между нестандартизованными и стандартизованными оценками не перекрывались, следовательно, различия можно считать статистически значимыми.

Выборочная оценка «истинной аллельной дифференциации», D_{EST} , составила $29,2 \pm 2,6\%$ при $p_{perm} \leq 0,001$ с 95%ДИ 24,4...33,9%. Она была на 28% ниже стандартизованных оценок, но 95%ДИ перекрывались. Следовательно, нет оснований для отвержения нулевой гипотезы – методы группы В давали сходные сводные оценки генетической дифференциации породных выборок.

Таблица 3 – Уровень STR-дифференциации породных выборок / Table 3 – Level of STR-differentiation of breed samples

| Показатель / Indicator | Метод / Method | | | | | |
|------------------------|--------------------|--------------|---------------|--------------------|-----------------|--------------|
| | группа А / group A | | | группа В / group B | | |
| | F_{ST} | G_{ST} | $G_{ST(NEI)}$ | $G'_{ST(HED)}$ | $G''_{ST(HED)}$ | D_{EST} |
| Оценка / Est. | 0,134 | 0,103 | 0,118 | 0,364 | 0,375 | 0,292 |
| Ошибка / S.E. | 0,013 | 0,014 | 0,015 | 0,029 | 0,029 | 0,026 |
| 95% CIL | 0,111 | 0,080 | 0,093 | 0,310 | 0,324 | 0,244 |
| 95% CIU | 0,159 | 0,130 | 0,148 | 0,415 | 0,426 | 0,339 |
| p_{perm} | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 | 0,001 |
| Est. / Fst. | 1,0 | 0,8 | 0,9 | 2,7 | 2,8 | 2,2 |

Примечания: 95%CI (95%ДИ) – 95% доверительный интервал (L – нижняя граница, U – верхняя граница; рассчитывался bootstrap-методом); p_{perm} – достигнутый перестановочный уровень статистической значимости (jackknife- и bootstrap-оценки, permutation test рассчитан по 999 перестановкам/репликам); Est. – оценка; S.E. – стандартная ошибка / Notes: 95%CI (95%ДИ) – 95% confidence interval (L – lower limit, U – upper limit; calculated by bootstrap method); p_{perm} – achieved permutation level of statistical significance (jackknife- and bootstrap-estimates, permutation test calculated by 999 permutations/replicas); Est. – estimate; S.E. – standard error.

Для проверки влияния внутривыборочной гетерозиготности (sH_s) на оценки статистик группы А допустили, что имеющиеся данные состоят из 11 псевдонаборов, каждый с разным локусом. По каждому псевдонабору были рассчитаны sH_s и сводные оценки. Корреляции sH_s с F_{ST} , G_{ST} и $G_{ST(NEI)}$ оценками составили (все три) -0,82 при $p_{value} \leq 0,002$, с $G'_{ST(HED)}$ и $G''_{ST(HED)}$ – -0,34 и -0,37 при $p_{value} \geq 0,259$, а с D_{EST} – -0,01 при $p_{value} = 0,979$. Корреляционный анализ подтвердил негативную статистически значимую зависимость оценок генетической дифференциации методами группы А от sH_s и отсутствие таковой при использовании методов группы В.

95%ДИ оценок методами групп А и В не перекрывались. Следовательно, можно полагать, что методы группы А статистически значимо (почти в 3 раза) занижали степень генетической дифференциации породных выборок. Это противоречит мнению, что любую статистику (F_{ST} , G_{ST} , D_{EST}), в контексте STR-дифференциации, «использовать правомерно» [1].

Наши результаты согласуются с теоретическими расчётами [11, 12]. Как косвенное подтверждение можно привести результаты сравнительного исследования информативности STR- и SNP-маркеров для внутривидовой и межвидовой дифференциации рода *Ovis* [30]. Оценки F_{ST} , рассчитанные по STR-маркерам, были в диапазоне 2,8-11,5% (усреднённая 8,4%), по SNP-маркерам – 15,0-53,7% (39,6%). Считается, что при использовании SNP-маркеров нет необходимости в коррекции F_{ST} -оценок на уровень sH_s . Если допустить, как предполагают авторы, «...потенциальное преимуще-

ство SNP-маркеров в выявлении генетических различий между видами, экотипами и другими мелкими таксономическими единицами» [30; с. 93], то F_{ST} -оценки по STR-маркерам были ниже (занижены) почти в 5 раз! Наряду с этим следует отметить, что оценки F_{ST} по SNP-маркерам были примерно в два раза ниже, чем по STR-маркерам (28-43% против 73-78%). Возникает вопрос, именно: отражали ли оценки F_{ST} по SNP-маркерам подлинное внутривидовое генетическое разнообразие рода *Ovis*?

Райт [цит. по 31] для F_{ST} -оценок предложил классификацию, которая, как представляется, применима и к иным мерам:

| Величина F_{ST} | Генетическая дифференциация популяций |
|-------------------|---------------------------------------|
| $\leq 0,05$ | - незначительная (little); |
| 0,06-0,15 | - умеренная (moderate); |
| 0,16-0,25 | - большая (great); |
| $> 0,25$ | - очень большая (very great). |

По оценкам, полученным методами группы А, генетическую дифференциацию породных выборок можно классифицировать как «умеренная». С другой стороны, по оценкам методами группы В – «очень большая». При подобных противоречивых заключениях решения по разведению/сохранению популяций могут существенно различаться и даже быть противоположными. Это указывает на важность правильного выбора метода для оценки генетической дифференциации пород, линий, стад.

Из-за чувствительности F_{ST} , G_{ST} и $G_{ST(NEI)}$ -статистик к уровню sH_s их оценки могут приближаться к нулю, даже при отсутствии

в субпопуляциях общих аллелей [12]. Поэтому для получения адекватных оценок *текущей* генетической дифференциации популяций представляется целесообразным использовать D_{EST} -статистику L. Jost, расчёт которой базируется на числе эффективных аллелей и оценка не зависит от уровня s_N . Стандартизированные оценки по P. W. Hedrick ($G'_{ST(HEDE)}$ и $G''_{ST(HEDE)}$) более корректны, чем оценки методами группы А (F_{ST} , G_{ST} и $G_{ST(NEI)}$), т. к. при высоких значениях s_N приближаются к таковым по D_{EST} ($r = 0,93$; $p_{value} < 0,0001$). Считается, что $G'_{ST(HEDE)}$

и $G''_{ST(HEDE)}$ статистики могут быть полезны при сравнении генетической дифференциации популяций разных видов, с разными эффективными размерами, при использовании генетических маркеров, отличающихся по темпам мутирования, как STR и SNP [32].

Генетические дистанции между выборками. Сводные оценки характеризовали усреднённую дифференциацию, нивелируя генетические отношения между каждой парой породных выборок. Для выяснения последних были рассчитаны генетические дистанции (**GD**; табл. 4).

Таблица 4 – Треугольные матрицы парных генетических дистанций / Table 4 – Triangular matrices of paired genetic distances

| Выборка / Sample | AYR | JER | RDAT | RSH | H-DEU | H-NLD | H-USA |
|---|--------------|--------|---------------|---------------|--------------|---------------|---------------|
| F _{ST} -дистанции / F _{ST} -distances | | | | | | | |
| AYR | - | 0,0001 | 0,4022 | 0,2628 | 0,0001 | 0,0001 | 0,0001 |
| JER | 0,147 | - | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
| RDAT | 0,025 | 0,166 | - | 0,3028 | 0,0001 | 0,0001 | 0,0001 |
| RSH | 0,030 | 0,183 | 0,027 | - | 0,0003 | 0,0001 | 0,0001 |
| H-DEU | 0,082 | 0,124 | 0,086 | 0,087 | - | 0,0935 | 0,2355 |
| H-NLD | 0,064 | 0,109 | 0,070 | 0,073 | 0,019 | - | 0,0003 |
| H-USA | 0,086 | 0,143 | 0,085 | 0,085 | 0,021 | 0,029 | - |
| G _{ST(NEI)} -дистанции / G _{ST(NEI)} -distances | | | | | | | |
| AYR | - | 0,0001 | 0,4022 | 0,2601 | 0,0001 | 0,0001 | 0,0001 |
| JER | 0,217 | - | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
| RDAT | 0,001 | 0,250 | - | 0,3028 | 0,0001 | 0,0001 | 0,0001 |
| RSH | 0,008 | 0,274 | 0,005 | - | 0,0003 | 0,0001 | 0,0001 |
| H-DEU | 0,113 | 0,185 | 0,121 | 0,120 | - | 0,1057 | 0,2369 |
| H-NLD | 0,089 | 0,168 | 0,101 | 0,106 | 0,008 | - | 0,0003 |
| H-USA | 0,123 | 0,219 | 0,123 | 0,121 | 0,006 | 0,031 | - |
| G'' _{ST(HEDE)} -дистанции / G'' _{ST(HEDE)} -distances | | | | | | | |
| AYR | - | 0,0001 | 0,4034 | 0,2599 | 0,0001 | 0,0001 | 0,0001 |
| JER | 0,590 | - | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
| RDAT | 0,005 | 0,663 | - | 0,3028 | 0,0001 | 0,0001 | 0,0001 |
| RSH | 0,026 | 0,728 | 0,017 | - | 0,0003 | 0,0001 | 0,0001 |
| H-DEU | 0,401 | 0,524 | 0,417 | 0,416 | - | 0,1072 | 0,2362 |
| H-NLD | 0,335 | 0,501 | 0,371 | 0,388 | 0,032 | - | 0,0001 |
| H-USA | 0,387 | 0,565 | 0,377 | 0,372 | 0,019 | 0,107 | - |
| D _{EST} -дистанции / D _{EST} -distances | | | | | | | |
| AYR | - | 0,0001 | 0,4034 | 0,2599 | 0,0001 | 0,0001 | 0,0001 |
| JER | 0,477 | - | 0,0001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
| RDAT | 0,003 | 0,551 | - | 0,3028 | 0,0001 | 0,0001 | 0,0001 |
| RSH | 0,019 | 0,626 | 0,012 | - | 0,0003 | 0,0001 | 0,0001 |
| H-DEU | 0,325 | 0,417 | 0,337 | 0,336 | - | 0,1081 | 0,2356 |
| H-NLD | 0,270 | 0,401 | 0,300 | 0,316 | 0,024 | - | 0,0003 |
| H-USA | 0,302 | 0,443 | 0,289 | 0,285 | 0,014 | 0,079 | - |

Примечания: под диагональю оценки, над диагональю p_{perm} – достигнутый уровень статистической значимости ($\alpha_{Bonf} = 0,05/21 = 0,0024$) / Notes: under the evaluation diagonal, above the p_{perm} diagonal - the reached level of statistical significance ($\alpha_{Bonf} = 0,05/21 = 0,0024$).

С вероятностью ошибки $\alpha_{\text{Bonf}} = 0,0024$ (множественный тест Бонферрони) из 21 парных сравнений 16 оценок GD были статистически значимыми. GD между породными выборками AYR с RDAT, AYR с RSH, RDAT с RSH и H-DEU с двумя другими голштинскими выборками были статистически незначимыми по всем мерам. Возможной причиной подобия генофондов быков красных пород могло быть наличие достаточно интенсивного обмена генетическим материалом; между голштинскими выборками – интродукция генов голштинской породы Северной Америки.

Похожие результаты приводил A. V. Dotsev с коллегами²: оценки SNP/ F_{ST} -дистанций между выборкой голштинских быков из Германии ($n_1 = 5$) и таковыми из Нидерландов (4), Дании (3) и Канады (4) были, можно сказать, нулевыми: -0,01, 0,006 и -0,001 соответственно.

GD по F_{ST} -статистике были в диапазоне от 0,019 до 0,183 (усреднённая 0,083), по $G_{ST(NEI)}$ – от 0,001 до 0,274 (0,114), по $G''_{ST(HED)}$ – от 0,005 до 0,728 (0,345), по D_{EST} – от 0,003 до 0,626 (0,277). Если за

«истинную дифференциацию» взять сводную D_{EST} -оценку (табл. 3), то F_{ST} -мера занижала оценки генетических дистанций в среднем на $(1 - 0,134/0,292)100 = 54\%$, $G_{ST(NEI)}$ -мера – на 60%, т. е. более чем в 2 раза!

Отметим, STR/ F_{ST} -дистанции между девятью локальными российскими породами скота были в диапазоне 0,022-0,106 [15], между региональными популяциями холмогорской породы – 0,016-0,041 [19], между стадами романовских овец – 0,038-0,059 [21], дифференциация линий в чистокровной верховой породе лошадей была на уровне 0,056 [22], популяций дальневосточных пчёл – 0,049-0,149 [30]. Если исходить из полученных нами результатов, то можно полагать, что эти оценки также являются значительно заниженными.

Стандартные GD Nei (табл. 5) варьировали от 0,086 до 1,052 (усреднённая 0,435), несмещённые оценки GD – от 0,000 до 0,962 (усреднённая 0,341), т. е. в среднем на 22% ниже. Отметим, оценки GD Nei несравнимы с таковыми таблицы 4, т. к. являются продуктом логарифмирования с диапазоном от 0 до ∞ .

Таблица 5 – Матрицы генетических дистанций Nei (GDN и uGDN) / Table 5 – Nei genetic distance matrices (GDN and uGDN)

| Выборка / Sample | AYR | JER | RDAT | RSH | H-DEU | H-NLD | H-USA |
|------------------|-------|-------|-------|-------|-------|-------|-------|
| AYR | - | 0,629 | 0,000 | 0,013 | 0,392 | 0,293 | 0,354 |
| JER | 0,716 | - | 0,781 | 0,962 | 0,509 | 0,448 | 0,576 |
| RDAT | 0,107 | 0,865 | - | 0,000 | 0,402 | 0,325 | 0,333 |
| RSH | 0,128 | 1,052 | 0,111 | - | 0,401 | 0,347 | 0,327 |
| H-DEU | 0,499 | 0,591 | 0,505 | 0,510 | - | 0,007 | 0,004 |
| H-NLD | 0,397 | 0,527 | 0,425 | 0,454 | 0,106 | - | 0,050 |
| H-USA | 0,442 | 0,639 | 0,418 | 0,417 | 0,086 | 0,130 | - |

Примечания: под диагональю GDN – стандартные оценки; над диагональю uGDN – несмещённые оценки. Между AYR и RDAT оценка uGDN = -0,002; по рекомендации Nei она приравнена к нулю. $R_M = 0,999$ /

Notes: below the GDN diagonal are standard estimates; above the uGDN diagonal are unbiased estimates. Between AYR and RDAT, the estimate uGDN = -0.002; according to Nei recommendation, it is equated to zero. $R_M = 0.999$.

Степень соответствия (correspondence) матриц GD была определена корреляцией Мантеля (R_M) с рандомизированной перму-

тационной (random permutation) проверкой статистической значимости. R_M составили:

| Матрица | F_{ST} | G_{ST} | $G_{ST(NEI)}$ | $G'_{ST(HED)}$ | $G''_{ST(HED)}$ | uGDN |
|-----------|----------|----------|---------------|----------------|-----------------|--------|
| G_{EST} | 0,970 | 0,976 | 0,983 | 1,000 | 0,998 | 0,986 |
| uGDN | 0,985 | 0,985 | 0,984 | 0,982 | 0,977 | 1,000. |

²Dotsev A. V., Sermyagin A. A., Shakhin A. V., Reyer H., Wimmers K., Brem G., Zinovieva N. A. Breed purity of Holstein bulls born in Russia and imported from different countries. ASAS-CSAS Annual Meeting and Trade Show. 2018 Vancouver, Canada [poster]. [Электронный ресурс].

URL: https://www.vij.ru/images/conf-18/05-12jul_Vancouver/Dotsev.ArsenV_3.pdf (дата обращения: 26.12.2018).

Все R_M были очень высокие и статистически значимые ($p_{perm} \leq 0,02$). R_M есть мера линейной зависимости. Квадрат R_M (R_M^2) – это коэффициент детерминации, который представляет долю вариации, общую для двух матриц, или степень связанности элементов матриц. Диапазон R_M^2 был 0,941-0,998, т. е. неучтённая вариация составляла 0,2-5,9%. Несмотря на значительные различия в оценках парных дистанций, полученных разными методами, степень соответствия матриц GD была очень высокой.

Анализ главных координат. Даже такая небольшая матрица, как 4×4 , достаточно сложна для восприятия и интерпретации. Матрицы больших размеров анализировать практически невозможно. Поэтому в популяционно-генетических исследованиях используют методы многомерной статистики (факторный анализ, кластерный анализ и др.) для представления матриц GD в наглядном, доступном для интерпретации, графическом виде. Нами для визуализации генетических отношений между породными выборками быков (табл. 4 и 5), использован «анализ главных координат» (Principal Coordinate Analysis, PCoA). Концептуально PCoA аналогичен «анализу главных компонент» (Principal Components Analysis, PCA); часто PCoA отождествляют с «метрическим многомерным шкалированием» (metric Multi-Dimensional Scaling, mMDS).

PCoA – это процедура, которая каждому элементу матрицы дистанций между n объектами (выборками) назначает местоположение в теоретическом пространстве небольшой размерности в виде конфигурации из n точек, минимизируя при этом потери информации. Другими словами, PCoA проецирует информацию матрицы дистанций между объектами в достаточно адекватный 2D (или 3D) график ординации (взаимном расположении точек при \min возможном искажении расстояния между ними).

Процедура PCoA создает в теоретическом пространстве с помощью дискриминантных (разделяющих) функций набор некоррелированных (ортогональных) осей (шкал) координат для обобщения разброса (дисперсии) между объектами в матрице дистанций. Каждая ось имеет собственное значение (eigenvalue), величина которого указывает размер дисперсии, зафиксированной на этой оси. Отношение данного собственного значения к сумме всех собственных значений пока-

зывает относительную «важность» каждой оси (используется как мера вклада объектов в заданную ось PCoA). PCoA считается успешным, если первые 2-3 оси «улавливают» не менее 80% общей дисперсии в матрице дистанций (60% считается минимально приемлемым уровнем). Вклад остальных осей оказывается на уровне погрешности анализа (при равномерном распределении вкладов по осям PCoA считается неудачным).

В PCoA все выборки имеют «оценки» (score) вдоль каждой оси. Оценки выборки представляют её координаты на графике ординации. Чем ближе точки-выборки на графике, тем больше общего в составе их аллелофонов. Если две точки-выборки на графике далеки друг от друга, то между соответствующими выборками имеет место значительная дифференциация. Отдалённость определяется значениями элементов матрицы GD, используемой в качестве входных данных.

На рисунке представлены результаты PCoA, основанного на F_{ST} -, $G''_{ST(HEG)}$ -, D_{EST} - и $uGDN$ -матрицах GD. Первая главная координата (Coor.1) объясняла от 66 до 80% межвыборочной дисперсии в структуре данных. При этом по «классическим» мерам (F_{ST} и $uGDN$) дисперсия была выше (74 и 80%), чем по «новым» ($G''_{ST(HEG)}$ и D_{EST} – 66 и 69%). На вторую главную координату (Coor.2) приходилось соответственно 23 и 19, 32 и 30% STR-дисперсии между выборками. Суммарно две главные координаты «улавливали» в F_{ST} -, $G''_{ST(HEG)}$ -, D_{EST} - и $uGDN$ -матрицах GD 96,8, 97,0, 97,8 и 99,1% STR-изменчивости между выборками. При существенно отличающихся парных оценках F_{ST} -, $G''_{ST(HEG)}$ -, D_{EST} - и $uGDN$ -статистик (табл. 4 и 5) их PCoA-информативность была почти равной (различия не более 2,3 процентных пункта) и высокой.

Первая главная координата отделила JER-выборку от остальных (можно сказать, «островной» генофонд от «материкового»). Вторая главная координата отделила голштинские выборки от выборок быков других пород («американский» генофонд от «европейского»). В данном случае смысл (интерпретация) каждой из осей не имеет значения; главным является взаимное расположение выборок. В целом, все четыре 2D-графика демонстрировали группировку выборок и степень их дифференциации на три кластера – «островков дивергенции». Внутри кластеров выборки характеризовались большим генетическим подобием, между кластерами – большим несходством.

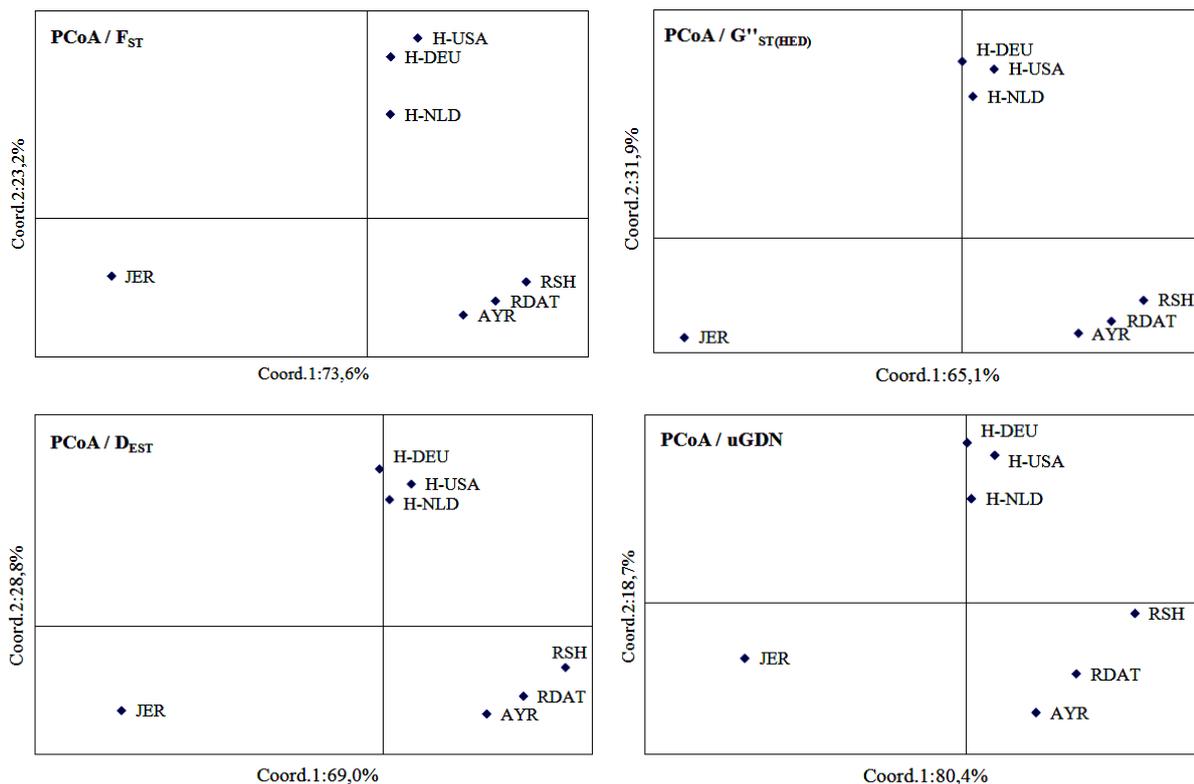


Рис. Проекция выборок быков на плоскости 1 и 2 главных координат с F_{ST} -, $G''_{ST(HEd)}$ -, D_{EST} - и $uGDN$ -матриц генетических дистанций. На осях даны проценты объяснённой дисперсии (числа показывают, какая доля общего разброса точек приходится на каждую из координат) /

Fig. Projection of bull samples on plane 1 and 2 of the main coordinates with F_{ST} -, $G''_{ST(HEd)}$ -, D_{EST} - and $uGDN$ -matrices of genetic distances. On the axes, the percentages of the explained variance are given (the numbers show how much of the total scatter of points falls on each of the coordinates).

На всех графиках ординация породных выборок была очень схожей. Первый кластер объединил голштинские выборки быков (H-DEU, H-NLD, H-USA), что вполне ожидаемо. На значительном расстоянии от первого кластера расположен второй – выборки быков трёх «красных» пород (AYR, RDAT, RSH), что также представляется логичным. В частности, в 2010 г. Финляндия, Дания, Швеция объединили айрширскую, красную датскую, шведскую красно-пёструю породы и улучшают по единой программе селекции VikingRed международной компании Viking Genetics (детали истории слияния пород даны в [33]). Обособлено, на достаточно большом и примерно равном расстоянии от первых двух кластеров, расположилась выборка быков джерсейской породы (JER), образуя, тем самым, третий кластер.

JER-выборку характеризовали повышенные GD с остальными породными выборками (табл. 4). Самыми высокими были оценки по $G''_{ST(HEd)}$ -мере, которые указывали на то, что дифференциация JER-выборки с остальными составляла от 50 до 73% от максимально

возможной при учёте фактической внутривыборочной гетерозиготности (cH_s).

Усреднённые GD JER-выборки с остальными выборками были: по F_{ST} -матрице – 0,145, $G''_{ST(HEd)}$ – 0,595, D_{EST} – 0,486, по $uGDN$ – 0,651. Это достаточно большие значения. Например, усреднённые GD AYR-выборки с остальными были в 2 и более раз ниже соответственно 0,072, 0,291, 0,233 и 0,280; объединённой выборки голштинских быков – 0,080, 0,385, 0,307, 0,353 (без JER). Подобная тенденция имела место в работе [16; Add. file 2: tabl. S2], именно: при полногеномном SNP/ F_{ST} -анализе усреднённая GD по девяти локальным российским породам с джерсейской породой составляла 0,156, с айрширской – 0,107, с голштинской – 0,085. Также в работе [34], где усреднённая GD джерсейской породы с 17 породами из разных частей мира по SNP/ F_{ST} -анализу была 0,208, аналогичная оценка по голштинской породе – 0,164.

Родина джерсейской породы – остров Джерзей; произошла от «слияния» нормандского и бретонского скота с участием альдер-

нейского скота [35], по некоторым данным с использованием швицев и зебу [36]. В 1789 году был издан закон, запрещающий импорт племенного скота, и после этого порода поддерживалась в «чистоте» с широким применением инбридинга³. В результате, как представляется, в породе и сформировался генофонд, отличающийся от такового материковых пород скота (проявилось и в пониженном генетическом разнообразии JER-выборки; см. табл. 2).

В общем, несмотря на значительные различия в оценках F_{ST} -, $G''_{ST(HEd)}$ -, D_{EST} - и $uGDN$ -статистик (табл. 4 и 5), все четыре 2D-графика имели очень схожие ординации выборок, наглядно демонстрируя секторальную структуру *объединённого* STR-генофонда быков. Эти результаты дают основание полагать, что для анализа пространственного местоположения популяций правомерно использовать любую меру GD (в этом контексте наше мнение согласуется с таковым в [1]).

Заключение. Вариация числа аллелей на микросателлитный локус (3,5-6,2) и гетерозиготности (56-97%) по породным выборкам свидетельствовала о наличии их дифференциации. Сводные оценки последней методами группы А (F_{ST} , G_{ST} , $G_{ST(NEI)}$) были в диапазоне 10...13% ($p_{perm} \leq 0,001$), методами группы В ($G'_{ST(HEd)}$, $G''_{ST(HEd)}$, D_{EST}) – от 29 до 37% ($p_{perm} \leq 0,001$). Оценки методами группы В статистически значимо превосходили оценки методами группы А почти в 3 раза. Было установлено воздействие повышенной внутривыборочной гетерозиготности на снижение оценок методами группы А ($r = -0,82$ при $p_{value} \leq 0,002$) и отсутствие такового при использовании стандартизации P. W. Hedrick

($r = -0,34$ и $-0,37$ при $p_{value} \geq 0,259$) и метода L. Jost ($r = -0,01$ при $p_{value} = 0,979$). Этот факт ставит под сомнение адекватность оценок, опубликованных ранее [15, 19, 21, 22, 30], реальной STR-дифференциации пород, линий, стад. Оценки парных генетических дистанций, рассчитанные разными методами, заметно отличались (по F_{ST} -матрице: 0,019-0,183, $G''_{ST(HEd)}$: 0,005-0,728, D_{EST} : 0,003-0,626, $uGDN$: 0,000-0,962). Однако между матрицами имело место высокая степень соответствия ($R_M \geq 0,97$), которая проявилась и в проекциях породных выборок на 2D-графиках PCoA. Это даёт основание считать, что рассматриваемые методы не оказывали существенного влияния на пространственную локализацию выборок и единообразно воспроизводили реальную структуру генофонда объединённой выборки быков. На основании результатов, полученных по имеющимся данным, можно полагать, что для оценки *текущей* STR-дифференциации популяций целесообразно использовать статистику L. Jost (D_{EST}) как меру «истинной аллельной дифференциации», на которую не влияет уровень внутривыборочной гетерозиготности. Стандартизированные оценки по P. W. Hedrick ($G'_{ST(HEd)}$ и $G''_{ST(HEd)}$) более корректны, чем оценки по методам группы А (F_{ST} , G_{ST} и $G_{ST(NEI)}$), т. к. при высокой внутривыборочной гетерозиготности приближаются к таковым по D_{EST} -статистике ($r = 0,90$ и $0,88$ при $p_{value} \leq 0,0003$). В то же время для исследования пространственной ординации генофондов популяций, по всей видимости, подходит любая из рассмотренных мер генетической дистанции.

References

1. Галинская Т. В., Щепетов Д. М., Лысенков С. Н. Предубеждения о микросателлитных исследованиях и как им противостоять. Генетика. 2019;55(6):1-16. DOI: <https://doi.org/10.1134/S0016675819060043>
Galinskaya T. V., Shchepetov D. M., Lysenkov S. N. *Predubezhdeniya o mikrosatellitnykh issledovaniyakh i kak im protivostoyat'*. [Prejudices against microsatellite studies and the ways to resist them]. *Genetika = Russian Journal of Genetics*. 2019;55(6):1-16. (In Russ.). DOI: <https://doi.org/10.1134/S0016675819060043>
2. Зиновьева Н. А., Сермягин А. А., Доцев А. В., Бороонецкая О. И., Петрикеева Л. В., Абдельманова А. С., Врем Г. Генетические ресурсы животных: развитие исследований аллелофонда российских пород крупного рогатого скота – Миниобзор. Сельскохозяйственная биология. 2019;54(4):631-641. DOI: <https://doi.org/10.15389/agrobiology.2019.4.631rus>
Zinov'eva N. A., Sermyagin A. A., Dotsev A. V., Boronetskaya O. I., Petrikeeva L. V., Abdel'manova A. S., Brem G. *Geneticheskie resursy zhivotnykh: razvitie issledovaniy allelofonda rossiyskikh porod krupnogo rogatogo skota – Miniobzor*. [Animal genetic resources: developing the research of allele pool of Russian cattle breeds. Mini review]. *Sel'skokhozyaystvennaya biologiya = Agricultural Biology*. 2019;54(4):631-641. (In Russ.). DOI: <https://doi.org/10.15389/agrobiology.2019.4.631rus>

³Джерсейская (порода коров). Режим доступа: [https://ru.wikipedia.org/wiki/Джерсейская_\(порода_коров\)](https://ru.wikipedia.org/wiki/Джерсейская_(порода_коров)) (дата обращения: 12.09.2018).

3. Карафет Т. М., Зегура С. Л., Хаммер М.Ф. Историческое освоение человеком новых территорий: роль древних популяций Азии в заселении Америки. Информационный вестник ВОГиС. 2006;10(1):7-23. Режим доступа: <https://elibrary.ru/item.asp?id=9573134>

Karafet T. M., Zegura S. L., Khammer M. F. *Istoricheskoe osvoenie chelovekom novykh territoriy: rol' drevnikh populyatsiy Azii v zaselenii Ameriki*. [Historical development of new lands by a man: an ancient link between Asia and America]. *Informatsionnyy vestnik VOGiS*. 2006;10(1):7-23. (In Russ.). URL: <https://elibrary.ru/item.asp?id=9573134>

4. Kalinowski S. T. Evolutionary and statistical properties of three genetic distances. *Molec. Ecol.* 2002;11(8):1263-1273. DOI: <https://doi.org/10.1046/j.1365-294X.2002.01520.x>

5. Wright S. The genetical structure of populations. *Annals of Eugenics*. 1951;15(4):323-354. DOI: <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>

6. Nei M. Genetic distance between populations. *Amer. Natur.* 1972;106(949):283-292. URL: <https://www.jstor.org/stable/2459777>

7. Кузнецов В. М. F-статистики Райта: оценка и интерпретация. Проблемы биологии продуктивных животных. 2014;4:80-104. Режим доступа: <https://elibrary.ru/item.asp?id=22833217>

Kuznetsov V. M. *F-statistiki Rayta: otsenka i interpretatsiya*. [Wright's F-statistics: estimation and interpretation]. *Problemy biologii produktivnykh zhivotnykh = Problems of Productive Animal Biology*. 2014;4:80-104. (In Russ.). URL: <https://elibrary.ru/item.asp?id=22833217>

8. Nei M., Chesser R. K. Estimation of fixation indexes and gene diversities. *Ann. Hum. Genet.* 1983;47(3):253-259. DOI: <https://doi.org/10.1111/j.1469-1809.1983.tb00993.x>

9. Nei M. Definition and estimation of fixation indices. *Evolution*. 1986;40(3):643-645. DOI: <https://doi.org/10.1111/j.1558-5646.1986.tb00516.x>

10. Hedrick P. W. Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution*. 1999;53(2):313-318. DOI: <https://doi.org/10.1111/j.1558-5646.1999.tb03767.x>

11. Hedrick P. W. A standardized genetic differentiation measure. *Evolution*. 2005;59(8):1633-1638. URL: <https://www.jstor.org/stable/3449070>

12. Jost L. G_{ST} and its relatives do not measure differentiation. *Molec. Ecol.* 2008;17(18):4015-4026. DOI: <https://doi.org/10.1111/j.1365-294X.2008.03887.x>

13. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*. 1978;89(3):583-590. URL: <https://www.genetics.org/content/89/3/583>

14. Heller R., Siegmund H. R. Relationship between three measures of genetic differentiation G_{ST} , D_{EST} and G'_{ST} : how wrong have we been? *Molec. Ecol.* 2009;18(10):2080-2083. DOI: <https://doi.org/10.1111/j.1365-294x.2009.04185.x>

15. Волкова В. В., Денискова Т. Е., Костюнина О. В., Амерханов Х. А., Добрынина Т. И., Зиновьева Н. А. Характеристика аллелофонда локальных пород крупного рогатого скота России по микросателлитным маркерам. Генетика и разведение животных. 2018;1:3-10. DOI: <https://doi.org/10.31043/2410-2733-2018-1-3-10>

Volkova V. V., Deniskova T. E., Kostyunina O. V., Amerkhanov Kh. A., Dobrynina T. I., Zinov'eva N. A. *Kharakteristika allelofonda lokal'nykh porod krupnogo rogatogo skota Rossii po mikrosatelitnym markeram*. [Characteristic of allele pool of local cattle breeds of Russia based on microsatellite markers]. *Genetika i razvedenie zhivotnykh = Genetics and breeding animals*. 2018;1:3-10. (In Russ.). DOI: <https://doi.org/10.31043/2410-2733-2018-1-3-10>

16. Serymyagin A. A., Dotsev A. V., Gladyr E. A., Traspov A. A., Deniskova T. E., Kostyunina O. V., Reyer H., Wimmers K., Barbato M., Paronyan I. A., Plemashov K. V., Sölkner J., Popov R. G., Brem G., Zinovieva N. A. Whole-genome SNP analysis elucidates the genetic structure of Russian cattle and its relationship with Eurasian taurine breeds. *Genet. Sel. Evol.* 2018;50(37):1-13. DOI: <https://doi.org/10.1186/s12711-018-0408-8>

17. Yurchenko A., Yudin N., Aitnazarov R., Plyusnina A., Brukhin V., Soloshenko V., Lhasaranov B., Popov R., Paronyan I. A., Plemashov K. V., Larkin D. M. Genome-wide genotyping uncovers genetic profiles and history of the Russian cattle breeds. *Heredity*. 2018;120:125-137. DOI: <https://doi.org/10.1038/s41437-017-0024-3>

18. Юдин Н. С., Ларкин Д. М. Происхождение, селекция и адаптация российских пород крупного рогатого скота по данным полногеномных исследований. Вавиловский журнал генетики и селекции. 2019;23(5):559-568. DOI: <https://doi.org/10.18699/VJ19.525>

Yudin N. S., Larkin D. M. *Proiskhozhdenie, selektsiya i adaptatsiya rossiyskikh porod krupnogo rogatogo skota po dannym polnogenomnykh issledovaniy*. [Whole genome studies of origin, selection and adaptation of the Russian cattle breeds]. *Vavilovskiy zhurnal genetiki i selektsii = Vavilov Journal of Genetics and Breeding*. 2019;23(5):559-568. (In Russ.). DOI: <https://doi.org/10.18699/VJ19.525>

19. Волкова В. В., Романенкова О. С., Денискова Т. Е., Мишина А. И., Костюнина О. В., Зиновьева Н. А. Характеристика аллелофонда холмогорской породы крупного рогатого скота с использованием STR-маркеров. Молочное и мясное скотоводство. 2019;(7):3-7. Режим доступа: <https://elibrary.ru/item.asp?id=41852279>

Volkova V. V., Romanenkova O. S., Deniskova T. E., Mishina A. I., Kostyunina O. V., Zinov'eva N. A. *Kharakteristika allelofonda kholmogor-skoy porody krupnogo rogatogo skota s ispol'zovaniem STR-markerov*. [Assessment of the allele pool of the Kholmogory cattle breed using STR-markers]. *Molochnoe i myasnoe skotovodstvo = Dairy and Beef Cattle Breeding*. 2019;(7):3-7. (In Russ.). URL: <https://elibrary.ru/item.asp?id=41852279>

20. Deniskova T. E., Dotsev A. V., Selionova M. I., Kunz E., Medugorac I., Reyer H., Wimmers K., Barbato M., Traspov A. A., Brem G., Zinov'eva N. A. Population structure and genetic diversity of 25 Russian sheep breeds based on whole-genome genotyping. *Genet. Sel. Evol.* 2018;50(29). DOI: <https://doi.org/10.1186/s12711-018-0399-5>

21. Денискова Т. Е., Костюнина О. В., Соловьева А. Д., Зиновьева Н. А. Изучение генетического разнообразия и дифференциации региональных популяций романовских овец по микросателлитным маркерам. *Аграрная наука Евро-Северо-Востока*. 2018;64(3):75-80. DOI: <https://doi.org/10.30766/2072-9081.2018.64.3.75-80>

Deniskova T. E., Kostyunina O. V., Solov'eva A. D., Zinov'eva N. A. *Izuchenie geneticheskogo raznoobraziya i differentsiatsii regional'nykh populyatsiy romanovskikh ovets po mikrosatellitnym markeram*. [Study of genetic diversity and differentiation of regional populations of Romanov sheep using microsatellite markers]. *Agrarnaya nauka Evro-Severo-Vostoka = Agricultural Science Euro-North-East*. 2018;64(3):75-80. (In Russ.). DOI: <https://doi.org/10.30766/2072-9081.2018.64.3.75-80>

22. Khrabrova L. A., Blohina N. V., Suleymanov O. I., Rozhdestvenskaya G. A., Pustovoy V. F. Assessment of line differentiation in the Thoroughbred horse breed using DNA microsatellite loci. *Вавиловский журнал генетики и селекции*. 2019;23(5):569-574. DOI: <https://doi.org/10.18699/VJ19.526>

Khrabrova L. A., Blohina N. V., Suleymanov O. I., Rozhdestvenskaya G. A., Pustovoy V. F. Assessment of line differentiation in the Thoroughbred horse breed using DNA microsatellite loci. *Vavilovskiy zhurnal genetiki i seleksii = Vavilov Journal of Genetics and Breeding*. 2019;23(5):569-574. (In Russ.). DOI: <https://doi.org/10.18699/VJ19.526>

23. Харзинова В. Р., Костюнина О. В., Карпушкина Т. В., Быкова О. А., Зиновьева Н. А. Изучение популяционной структуры и генетического разнообразия свиней породы венгерская мангалица на основе анализа микросателлитов. *Аграрный вестник Урала*. 2019;7(186):77-81. DOI: https://doi.org/10.32417/article_5d52b081b3e348.43320197

Kharzinova V. R., Kostyunina O. V., Karpushkina T. V., Bykova O. A., Zinov'eva N. A. *Izuchenie populyatsionnoy struktury i geneticheskogo raznoobraziya sviney porody vengerskaya mangalitsa na osnove analiza mikrosatellitov*. [The study of the population structure and genetic diversity of Hungarian Mangalica breed of pigs based on microsatellites analysis]. *Agrarnyy vestnik Urala = Agrarian Bulletin of the Urals*. 2019;7(186):77-81. (In Russ.). DOI: https://doi.org/10.32417/article_5d52b081b3e348.43320197

24. Kharzinova V. R., Dotsev A. V., Deniskova T. E., Solovieva A. D., Fedorov V. I., Layshev K. A., Romanenko T. M., Okhlopov I. M., Wimmers K., Reyer H., Brem G., Zinov'eva N. A. Genetic diversity and population structure of domestic and wild reindeer (*Rangifer tarandus* L. 1758): A novel approach using BovineHD BeadChip. *PLoS ONE*. 2018. DOI: <https://doi.org/10.1371/journal.pone.0207944>

25. Форнара М. С., Крамаренко А. С., Шаров М. А., Зиновьева Н. А. Исследование аллелофонда и генетической дифференциации дальневосточных пчел. *Достижения науки и техники АПК*. 2016;30(10):101-104. Режим доступа: <https://elibrary.ru/item.asp?id=27175140>

Fornara M. S., Kramarenko A. S., Sharov M. A., Zinov'eva N. A. *Issledovanie allelofonda i geneticheskoy differentsiatsii dal'nevostochnykh pchel*. [Study of allele pool and genetic differentiation of the Far-Eastern honey bees]. *Dostizheniya nauki i tekhniki APK = Achievements of Science and Technology of AIC*. 2016;30(10):101-104. (In Russ.). URL: <https://elibrary.ru/item.asp?id=27175140>

26. Peakall R., Smouse P.E. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molec. Ecol.* 2006;6(1):288-295. DOI: <https://doi.org/10.1111/j.1471-8286.2005.01155.x>

27. Peakall R., Smouse P. GenAIEx Tutorial 1: Introduction to population genetic analysis. Australian National University. 2012. 57 p. URL: <https://mafiadoc.com/genalex-tutorial-1-introduction-to-population-genetic-597ef8441723dd6ae3e07272.html>

28. Peakall R., Smouse P. E. GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics*. 2012b;28(19):2537-2539. DOI: <https://doi.org/10.1093/bioinformatics/bts460>

29. Кузнецов В. М. Снижает ли кроссбридинг генетическое разнообразие? или Разведение и сохранение пород молочного скота в России. Киров: НИИСХ Северо-Востока, 2017. 104 с. Режим доступа: http://vm-kuznetsov.ru/files/book/Pap2017_he.pdf

Kuznetsov V. M. *Snizhaet li krossbridging geneticheskoe raznoobraziye? ili Razvedenie i sokhranenie porod molochnogo skota v Rossii*. [Whether crossbreeding reduces genetic diversity or Breeding and preservation of dairy cattle breeds in Russia]. Киров: NIISKh Severo-Vostoka, 2017. 104 p. URL: http://vm-kuznetsov.ru/files/book/Pap2017_he.pdf

30. Денискова Т. Е., Сермягин А. А., Багиров В. А., Охлопков И. М., Гладырь Е. А., Иванов Р. В., Брем Г., Зиновьева Н. А. Сравнительное исследование информативности STR и SNP маркеров для внутривидовой и межвидовой дифференциации рода *Ovis*. Генетика. 2016;52(1): 90-96. DOI: <https://doi.org/10.7868/S0016675816010021>

Deniskova T. E., Sermyagin A. A., Bagirov V. A., Okhlopkov I. M., Gladyr' E. A., Ivanov R. V., Brem G., Zinov'eva N. A. *Sravnitel'noe issledovanie informativnosti STR i SNP markerov dlya vnutrividovoy i mezhvidovoy differentsiatsii roda Ovis*. [Comparative analysis of the effectiveness of STR and SNP markers for intraspecific and interspecific differentiation of the genus *Ovis*]. *Genetika = Russian Journal of Genetics*. 2016;52(1): 90-96. (In Russ.). DOI: <https://doi.org/10.7868/S0016675816010021>

31. Hartl D. L., Clark A. G. Principles of Population Genetics. Third Edition. Sunderland, Massachusetts: Sinauer Associates, Inc., 1997. 542 p. URL: <https://archive.org/details/B-001-001-062/page/n13>

32. Meirmans P. G. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*. 2006;60(11):2399-2402. DOI: <https://doi.org/10.1111/j.0014-3820.2006.tb01874.x>

33. Кузнецов А. В., Тулинова О. В. Характеристика породных компонент айрширского стада ПАО «ПЗ им. В.И. Чапаева» Краснодарского края. Генетика и разведение животных. 2019;2:24-29. DOI: <https://doi.org/10.31043/2410-2733-2019-2-24-29>

Kuznetsov A. V., Tulina O. V. *Kharakteristika porodnykh komponent ayrshirskogo stada PAO «PZ im. V. I. Chapaeva» Krasnodarskogo kraya*. [Characteristics of breed component of Ayrshire herd of PC «BP of V. I. Chapaeva» Krasnodar territory]. *Genetika i razvedenie zhivotnykh*. 2019;2:24-29. (In Russ.). DOI: <https://doi.org/10.31043/2410-2733-2019-2-24-29>

34. Gautier M., Laloë D., Moazami-Goudarzi K. Insights into the Genetic History of French Cattle from Dense SNP Data on 47 Worldwide Breeds. *PLoS ONE*. 2010;5(9): e13038. DOI: <https://doi.org/10.1371/journal.pone.0013038>

35. Плёмб Ч. Типы и породы сельскохозяйственных животных. С.-Петербург, 1913. 560 с.

Plemb Ch. *Tipy i porody sel'skokhozyaystvennykh" zhivotnykh"*. [Types and breeds of farm animals]. S.-Peterburg", 1913. 560 p.

36. Ружевский А. Б., Рубан Ю. Д., Бердник П. П. Породы крупного рогатого скота. М.: «Колос»; 1980. 246 с.

Ruzhevskiy A. B., Ruban Yu. D., Berdnik P. P. *Porody krupnogo rogatogo skota*. [Breeds of cattle]. Moscow: «Kolos», 1980. 246 p.

Сведения об авторе

✉ Кузнецов Василий Михайлович, доктор с.-х. наук, профессор, зав. лабораторией популяционной генетики в животноводстве, ФГБНУ «Федеральный аграрный научный центр Северо-Востока имени Н. В. Рудницкого», ул. Ленина, д. 166а, г. Киров, Российская Федерация, 610007, e-mail: priemnaya@fanc-sv.ru, ORCID: <https://orcid.org/0000-0002-2219-805X>, e-mail: vm-kuznetsov@mail.ru

Information about the author

✉ Vasily M. Kuznetsov, DSc in Agricultural Science, professor, Head of the Laboratory of Population Genetics in Animal Husbandry, Federal Agricultural Research Center of the North-East named N. V. Rudnitsky, Lenin str., 166a, Kirov, Russian Federation, 610007, e-mail: priemnaya@fanc-sv.ru, ORCID: <https://orcid.org/0000-0002-2219-805X>, e-mail: vm-kuznetsov@mail.ru

✉ – Для контактов / Corresponding author