



## Оценка генетической дифференциации популяций молекулярным дисперсионным анализом (аналитический обзор)

© 2021. В. М. Кузнецов ✉

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

В этом исследовании мы сравнили разные подходы к использованию анализа молекулярной дисперсии (Analysis of Molecular Variance, AMOVA) для оценки генетической дифференциации популяций. Были использованы данные по 11 микросателлитным локусам 84 быков семи пород. Сравнивали результаты по трём опциям модуля AMOVA программы GenAlEx 6.502: по матрице дистанций между аллелями (рассчитывалась  $F_{ST(W\&C)}$  ( $=\theta$ ) статистика – вариант AMOVA1); по матрице дистанций между генотипами ( $\Phi_{PT}$  – AMOVA2); по матрице различий в размерах аллелей ( $R_{ST}$  – AMOVA3). Получены сходные сводные оценки генетической дифференциации пород:  $F_{ST(W\&C)} = 0,108$ ,  $\Phi_{PT} = 0,115$ ,  $R_{ST} = 0,110$  (все с  $p_{perm} \leq 0,001$ ). Между попарными оценками  $F_{ST(W\&C)}$  и  $\Phi_{PT}$  коэффициент корреляции был 0,99 ( $p_{value} < 0,0001$ ); статистически значимых корреляций с  $R_{ST}$  не установлено. Обнаружена высокая корреляция  $F_{ST(W\&C)}$  и  $\Phi_{PT}$  с попарными оценками дифференциации по Нею (0,96). Иные, чем GenAlEx программы (Arlequin v.3.5, GenePop v.4.7.3, RST22), давали схожих AMOVA-оценок. Установлена негативная линейная зависимость  $F_{ST(W\&C)}$  и  $\Phi_{PT}$  оценок от уровня средней гетерозиготности породных выборок ( $R^2 = 0,6$ ,  $r_s = -0,75$  при  $p_{value} < 0,02$ ) и отсутствие таковой для  $R_{ST}$ -оценок ( $R^2 = 0,04$ ,  $r_s = -0,23$  при  $p_{value} = 0,47$ ). Стандартизация оценок  $F_{ST(W\&C)}$  и  $\Phi_{PT}$  по Хедрику устранила эту зависимость и повысила первоначальные оценки до 0,35 и 0,37 соответственно. Последние были сопоставимы с оценками, полученными методами Нея-Хедрика (0,364-0,375), Джосты (0,292) и Морисита-Хорна (0,308). Корреляции Мантеля между матрицами парных по породам генетических дистанций (GD), рассчитанными разными мерами, в большинстве случаев были  $>0,9$ . Проекция матриц GD в анализе главных координат (PCoA) на 2D плоскости были в общем сходными. PCoA выделил кластер голштинских «экотипов», кластер «красных» пород и ветку джерсейской породы. В двухфакторном AMOVA данных по кластерам (как двух «регионов») межрегиональная GD составила 0,357; дифференциация пород в пределах «регионов» не превышала 0,027. Моделирование объединения пород с близкими к нулю GD привело к увеличению числа аллелей на locus в «новых» породах на 29 % и повышению сводной оценки генетической дифференциации на 29-46 %. Полученные результаты могут быть использованы при разработке мероприятий по сохранению вытесняемых пород.

**Ключевые слова:** разведение животных, микросателлиты, разнообразие, генетическая дифференциация, генетическая дистанция, AMOVA, анализ главных координат, сохранение генофонда

**Благодарности:** работа выполнена при поддержке Минобрнауки РФ в рамках Государственного задания ФГБНУ «Федеральный аграрный научный центр Северо-Востока имени Н. В. Рудницкого» (тема № 0767-2019-0089).

Автор благодарит рецензентов за вклад в экспертную оценку работы.

**Конфликт интересов:** автор заявил об отсутствии конфликта интересов.

**Для цитирования:** Кузнецов В. М. Оценка генетической дифференциации популяций молекулярным дисперсионным анализом (аналитический обзор). Аграрная наука Евро-Северо-Востока. 2021;22(2):167-187. DOI: <https://doi.org/10.30766/2072-9081.2021.22.2.167-187>

Поступила: 16.02.2021

Принята к публикации: 02.04.2021

Опубликована онлайн: 19.04.2021

## Assessment of genetic differentiation of populations by analysis of molecular variance (analytical review)

© 2021. Vasily M. Kuznetsov ✉

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

Different approaches to using the analysis of molecular variance (AMOVA) to assess the genetic differentiation of populations have been compared in the research. Data on 11 microsatellite loci of 84 bulls of seven breeds were used. The results were compared for three options of the AMOVA module of the GenAlEx 6.502 program: the allele distance matrix (calculated  $F_{ST(W\&C)}$  ( $=\theta$ ) statistics – variant AMOVA1); the genotype distance matrix ( $\Phi_{PT}$  – AMOVA2); and the allele size difference matrix ( $R_{ST}$  – AMOVA3). Similar summary estimates of the genetic differentiation of breeds were obtained:  $F_{ST(W\&C)} = 0.108$ ,  $\Phi_{PT} = 0.115$ ,  $R_{ST} = 0.110$  (all with  $p_{perm} \leq 0.001$ ). Between the estimates of  $F_{ST(W\&C)}$  and  $\Phi_{PT}$  for each locus, the correlation coefficient was 0.99 ( $p_{value} < 0.0001$ ); no statistically significant correlations with  $R_{ST}$  were found. A high correlation of  $F_{ST(W\&C)}$  and  $\Phi_{PT}$  with the estimates of differentiation according to Nei's (0.96) was found. Programs other than GenAlEx (Arlequin v.3.5, GenePop v.4.7.3, RST22) gave similar AMOVA estimates. The negative linear dependence of  $F_{ST(W\&C)}$  and  $\Phi_{PT}$  on the level of the average heterozygosity of the breed samples was established ( $R^2 = 0.6$ ,  $r_s = -0.75$  for  $p_{value} < 0.02$ ) and the absence of such dependence for  $R_{ST}$  ( $R^2 = 0.04$ ,  $r_s = -0.23$  for  $p_{value} = 0.47$ ). The standardization of the

$F_{ST(W\&C)}$  and  $\Phi_{PT}$  estimates according to Hedrick's eliminated this dependence and raised the initial estimates to 0.35 and 0.37, respectively. The latter were comparable to the estimates obtained by the Nei-Hedrick's (0.364-0.375), Jost's (0.292), and Morisit-Horn's (0.308) methods. The Mantel correlations between the matrices of paired genetic distances (GD) calculated by different measures were  $>0.9$  in most cases. The projections of the GD matrices in the principal coordinate analysis (PCoA) on the 2D plane were generally similar. The PCoA identified a cluster of Holstein «ecotypes», a cluster of «Red» breeds, and a branch of the Jersey breed. In the two-factor AMOVA of data on clusters (as two «regions»), the interregional GD was 0.357; the differentiation of breeds within the «regions» did not exceed 0.027. Modeling the association of breeds with close to zero GD resulted in an increase in the number of alleles per locus in the «new» breeds by 29 %, and an increase in the combined estimate of genetic differentiation by 29-46 %. The results obtained can be used in the development of measures for the conservation of endangered breeds.

**Keywords:** animal breeding, microsatellites, diversity, genetic differentiation, genetic distance, AMOVA, principal coordinate analysis, gene pool conservation

**Acknowledgement:** The research was carried out under the support of the Ministry of Science and Higher Education of the Russian Federation within the state assignment of the Federal Agricultural Research Center of the North-East named N. V. Rudnitsky (theme No. 0767-2019-0089).

The author thank the reviewers for their contribution to the peer review of this work.

**Conflict of interest:** the author declare no conflict of interest.

**For citation:** Kuznetsov V. M. Assessment of genetic differentiation of populations by analysis of molecular variance (analytical review). *Agrarnaya nauka Evro-Severo-Vostoka* = Agricultural Science Euro-North-East. 2020;22(2):167-187. (In Russ.). DOI: <https://doi.org/10.30766/2072-9081.2021.22.2.167-187>

Received: 16.02.2021

Accepted for publication: 02.04.2021

Published online: 19.04.2021

Каждый вид сельскохозяйственных животных подразделяется на породы. Породы, в свою очередь, подразделяются на стада, линии, семейства. С развитием молекулярно-генетических методов и ДНК-технологий стало возможным оценивать не только фенотипические различия между породами и их структурными элементами, но и непосредственно генетические различия (дифференциацию). Последнее важно и необходимо знать для эффективного планирования разведения и селекции коммерческих, сохранения исчезающих пород.

Почти 80 лет назад Wright [1] определил генетическую дифференциацию как различия частот аллелей между субпопуляциями, отражающие значительные отклонения генотипических частот от равновесия Харди-Вайнберга (Hardy-Weinberg Equilibrium, HWE) в общей популяции. Для характеристики распределения генетической изменчивости (гетерозиготности) внутри и между популяциями Райт ввёл *индексы фиксации* [2, 3]. Допускалась бесконечная островная модель (Infinite Island Model) структуры популяции, в которой диаллельные однолокусные субпопуляции равного размера обмениваются мигрантами в равных пропорциях. Индексы фиксации Райта оценивали размеры эффектов, подобных инбридингу, в пределах субпопуляций ( $F_{IS}$ ), во всей популяции ( $F_{IT}$ ) и между субпопуляциями ( $F_{ST}$ ). Индексы не независимы и связаны отношением  $(1 - F_{IT}) = (1 - F_{IS})(1 - F_{ST})$ , из которого следует:  $F_{ST} = (F_{IT} - F_{IS}) / (1 - F_{IS})$ , что адекватно  $F_{ST} = \sigma^2_{q(ST)} / [q_T(1 - q_T)]$ , где  $q_T$  – частота аллеля A

в диаллельном локусе общей популяции,  $\sigma^2_{q(ST)}$  – дисперсия частот аллели по субпопуляциям.  $F_{ST}$  есть мера степени завершения процесса фиксации аллелей в субпопуляциях («демографическая дифференциация» [4]). Если во всех субпопуляциях аллели имеют равные частоты, то  $\sigma^2_{q(ST)} = 0$  и  $F_{ST} = 0$ . Если дисперсия равна биномиальному максимуму,  $\sigma^2_{q(ST)} = q_T(1 - q_T)$ , то  $F_{ST} = 1$ , т. е. разные аллели фиксированы по субпопуляциям (в смысле фиксации одного аллеля в локусе для формирования гомозиготы). Значения  $F_{ST}$  между этими двумя крайностями интерпретируют как «отображающие различные уровни структурирования» [5]: 0-0,05 – незначительная, 0,05-0,15 – умеренная, 0,15-0,25 – большая и  $>0,25$  – очень большая дифференциация субпопуляций [6].

Nei [7, 8, 9] представил «коэффициент генной дифференциации»  $G_{ST}$  как адаптированный для использования в мультиаллельных мультилокусных популяциях аналог  $F_{ST}$  Райта:  $G_{ST} = (H_T - H_S) / H_T$ , где  $H_T$  – генное разнообразие в общей популяции,  $H_S$  – усреднённая гетерозиготность по субпопуляциям при HWE («генное разнообразие»). Статистика  $G_{ST}$  выведена из аддитивного разложения гетерозиготности.  $G_{ST}$  измеряет уровень вариации между субпопуляциями по отношению к общей вариации в популяции (не устанавливая идентичность участвующих аллелей). В отличие от  $F_{ST}$  для диаллельного локуса, величина  $G_{ST}$  зависит от уровня вариации полиморфных генов. Так, при анализе микросателлитов  $H_S$  и  $H_T$  могут приближаться к 1, а оценки  $G_{ST}$  ограничиваться диапазоном от 0 до  $(1 - H_S)$ ,

т. е. усреднённым уровнем гомозиготности субпопуляций.

Hedrick [10] предложил стандартизовать  $G_{ST}$  на её максимальное значение:  $G'_{ST} = G_{ST}/G_{ST(max)} = G_{ST}(k - 1 + H_S)/[(k - 1)(1 - H_S)]$ , где  $k$  – число субпопуляций.  $G'_{ST}$  – это мера дифференциации субпопуляций относительно максимально возможной и, следовательно, она позволяет сравнивать результаты по локусам с разными уровнями вариации, типам маркеров, видам животных и исследованиям. Модификация  $G'_{ST}$  при небольшом числе анализируемых субпопуляций была предложена в [11]:  $G''_{ST} = k(H_T - H_S)/[(k \times H_T - H_S)(1 - H_S)]$ . Оценки  $G'_{ST}$  и  $G''_{ST}$  имеют диапазон от 0 до 1; их особо рекомендуют для расчёта парных генетических дистанций.

Альтернативный метод предложил Джост (Jost) [12]. Его статистика,  $D_{EST}$  – «истинная дифференциация», является мультипликативной и базируется на изменчивости эффективного числа аллелей. В терминах гетерозиготности  $D_{EST} = [(H_T - H_S)/(1 - H_S)] \times [k/(k - 1)]$ . Уровень гетерозиготности в субпопуляциях не влияет на значения  $D_{EST}$ ; их диапазон от 0 до 1. Как отмечал Уиллок (Whitlock) [13], « $F_{ST}$  измеряет отклонения от панмиксии, в то время как  $D_{EST}$  измеряет отклонения от полной аллельной дифференциации» (цит. по [14]).

Cockerham [15, 16], Weir и Cockerham [17] для анализа генетической структуры популяции представили метод моментов (для нужных параметров находят несмещённые статистики, которые используют в качестве оценок). Были предложены: линейная модель для генов разных подразделений популяции, набор параметров для описания корреляций частот аллелей ( $f$ ,  $F$  и  $\theta$ ), процедура их оценивания по дисперсиям между субпопуляциями, между индивидами в субпопуляции и между гаметам индивида. Weir [18] дал формулы декомпозиции молекулярной дисперсии в терминах и виде, обычных для классического дисперсионного анализа (Analyses of Variance, ANOVA). Принцип ANOVA: если дисперсия (варианса) между субпопуляциями такая же, как и дисперсия во всей популяции, то популяционной структуры нет. Оценки компонентов дисперсии внутри и между субпопуляциями используют для расчёта  $f$ ,  $F$  и  $\theta$  (тета) статистик, аналогичных индексам фиксации Райта, именно:  $f$  ( $F_{IS}$ ) – отклонение от HWE в пределах субпопуляций или корреляция между аллелями индивида относительно собственной

субпопуляции (инбридинг в субпопуляциях);  $F$  ( $F_{IT}$ ) – отклонение от HWE в популяции или корреляция между аллелями индивида относительно всей популяции (инбридинг в общей популяции);  $\theta$  ( $F_{ST}$ ) – доля генетического разнообразия, обусловленная различиями в частотах аллелей среди субпопуляций или корреляция между аллелями индивидов субпопуляции по отношению ко всей популяции (дифференциация субпопуляций). Оценка  $\theta$  является количественной мерой дивергенции в группе из  $k$  субпопуляций; для двух субпопуляций – мерой расстояния или дистанции.  $\theta$  характеризует не только степень дифференциации, но и представляет «коэффициент взаимного родства» (co-ancestry) индивидов [18]. Если  $\theta$  используют в качестве описательной статистики для характеристики распределения генетического разнообразия среди субпопуляций (дифференциации), то нет необходимости в предположениях о дискретности анализируемых субпопуляций и об эволюционных процессах (мутации, миграции, отборе), которые могли бы привести к различиям между ними.

Excoffier с соавторами [19] трансформировали ANOVA Кокерхэма и Вейра в анализ молекулярной дисперсии (Analysis of Molecular Variance, AMOVA) для исследования изменчивости гаплотипов. Средние квадраты рассчитывают из матрицы квадратов евклидова расстояния между парами индивидов по гаплотипам ДНК. На выходе AMOVA даёт оценки варiances компонентов и статистику  $F_{PT}$  – аналог  $F_{ST}$  Райта. Если используют бинарные расстояния между гаплотипами, 1 – для идентичных и 0 – для разных гаплотипов, то оценки  $F_{PT}$  совпадают с  $\theta$ . Метод достаточно гибок, чтобы приспособить несколько альтернативных входных матриц дистанций, соответствующих различным типам молекулярных данных, включая микросателлиты [20].

Slatkin [21] предложил  $R_{ST}$ -статистику (также подобную  $F_{ST}$  Райта), которая особенно подходит для маркеров с пошаговой моделью мутации (Stepwise Mutation Model, SMM). По SMM каждая мутация создаёт новый аллель, добавляя или удаляя повторный мотив с равной вероятностью. Иногда допускают, что нет никаких ограничений на число повторов, возможных в локусе. Аллели разного размера будут менее связаны, чем аллели аналогичного размера. Считают, что SMM имеет «память» размера аллели и более точно отражает

процесс мутации микросателлитов. Эту дополнительную «филогенетическую» информацию используют в ANOVA для оценки компонентов дисперсии размера аллелей или числа повторов пар оснований в каждом локусе с последующим расчётом  $R_{ST}$ -статистики:  $R_{ST} = (S - S_w)/S$ , где  $S$  – средняя сумма квадратов различий в размере аллелей между всеми парами аллелей;  $S_w$  – то же в пределах каждой субпопуляции.

Нами в предыдущей работе [22] были представлены результаты сравнительного анализа изменчивости аллелей 11 микросателлитных локусов семи породных выборок быков-производителей. Генетическую дифференциацию выборок оценивали статистиками  $F_{ST}$ ,  $G_{ST}$ ,  $G'_{ST}$ ,  $G''_{ST}$  и  $D_{EST}$ . Также были использованы статистики GDN и uGDN (генетическая дистанция Нея и её несмещённая оценка [23, 24]).

**Цель аналитического обзора** – сравнить разные подходы к использованию анализа молекулярной дисперсии (AMOVA) для оценки генетической дифференциации популяций; на микросателлитных данных по семи породным выборкам рассчитать сводные  $\theta$ ,  $\Phi_{PT}$  и  $R_{ST}$  статистики, сопоставить полученные оценки, сравнить с результатами предыдущего исследования [22], рассчитать матрицы парных генетических дистанций и соотнести их 2-мерные проекции, двухфакторным AMOVA оценить «региональную» компоненту и рассмотреть последствия объединения породных выборок с малыми генетическими дистанциями (сходными аллельными профилями) на внутрипопуляционное разнообразие набора «новых» пород и их генетическую дифференциацию.

**Материал и методы.** Использованы те же данные, что и в предыдущем исследовании [22]. В частности, 84 быка ( $= N$ ), каждый генотипирован по 11 STR-локусам (Simple Tandem Repeats – простые тандемные последовательности, также Short Tandem Repeats – короткие тандемные повторы)\*, именно: 10 быков джерсейской породы (JER), 10 быков айрширской породы (AYR), 10 – красной датской (RDAT), 9 – красной шведской (RSH) и 45 быков голштинской породы трёх «экотипов»\*\* (отродий): 13 быков из Германии (H-DEU), 17 – из Нидерландов (H-NLD), 15 – из США (H-USA).

\* База генетических данных по быкам на сайте ВНИИплем. [Электронный ресурс]. URL: <http://www.vniipleм.ru/rus/files/Database/DNK/mikrosatellity.pdf> (дата обращения: 09.12.2018).

\*\* Экотип – субпопуляция внутри породы, адаптированная к определённой среде обитания.

\*\*\* Используют для сравнения с оценками дифференциации по иным, чем STR, типам маркеров.

Основные расчёты проводили с помощью компьютерной программы GenAlEx 6.502 [25, 26]. В модуле AMOVA для обработки кодоминантных данных использовали опции: **Codominant-Allelic** – генерирует матрицу парных дистанций между аллелями размером  $2N \times 2N$  с последующим расчётом  $F_{ST}$ -статистики ( $= \theta$ ) – в данной работе обозначается  $F_{ST(W\&C)}$  ( $F_{ST}$  по методу Weir & Cockerham), чтобы не путать с  $F_{ST}$  Райта и Нея (вариант AMOVA1); **Codominant-Genotypic** – генерирует матрицу парных дистанций между генотипами размером  $N \times N$  с последующим расчётом  $\Phi_{PT}$ -статистики с подавлением внутрииндивидуальной изменчивости\*\*\* (вариант AMOVA2); **Codominant-Microsat** – генерирует матрицу дистанций, как и при опции **Codominant-Allelic**, но на основе *размеров* аллелей с последующим расчётом  $R_{ST}$ -статистики (вариант AMOVA3).

Все три статистики оценивают долю STR-изменчивости между субпопуляциями относительно общей изменчивости во всей популяции:

$$F_{ST(W\&C)} = \frac{\sigma_{AP}^2}{\sigma_{AP}^2 + \sigma_{AI}^2 + \sigma_{WI}^2},$$

$$\Phi_{PT} = \frac{\sigma_{AP}^2}{\sigma_{AP}^2 + \sigma_{WP}^2}, \quad R_{ST} = \frac{\sigma_{AP}^2}{\sigma_{AP}^2 + \sigma_{AI}^2 + \sigma_{WI}^2},$$

где  $\sigma^2$  – дисперсия (дисперсия); субиндексы:

AP – между субпопуляциями,

AI – между индивидами,

WI – внутри индивидов,

WP – внутри популяции.

Статистическая значимость отклонения оценок от нуля в программе GenAlEx проверяется пермутационным тестом: нулевая гипотеза ( $H_0$ ) – нет генетических различий между субпопуляциями ( $F_{ST(W\&C)} = 0$ , или  $\Phi_{PT} = 0$ , или  $R_{ST} = 0$ ); альтернатива ( $H_1$ ) – есть генетические различия между субпопуляциями ( $F_{ST(W\&C)} > 0$  или  $\Phi_{PT} > 0$ , или  $R_{ST} > 0$ ). В общем,  $H_0$  в AMOVA заключается в том, что анализируемые субпопуляции допускаются частями одной большой случайно спариваемой генетической популяции. Если это истинно, то любая подразделённость случайна, и субпопуляции являются лишь выборками из единого генофонда.

Было показано [12, 21], что с увеличением уровня генетического разнообразия в пределах субпопуляций статистики дифференциации изменяются не линейно. Другими словами, чем больше внутрисубпопуляционная гетерозиготность, тем меньше их приращение. Программа GenAlEx выполняет линеаризацию оценок, например  $F_{ST(W\&C)}$ -оценки:

$$\ln F_{ST(W\&C)} = F_{ST(W\&C)} / (1 - F_{ST(W\&C)}).$$

Уровень гетерозиготности в (суб-)популяции не полностью зависит от числа аллелей (по нашим данным  $r = 0,68$ ,  $r^2 = 0,46$ ), т. к. верхний предел, равный 1, одинаков для любого числа аллелей. Для STR-маркеров, внутрисубпопуляционная дисперсия часто почти столь же большая, как общая дисперсия. Это приводит к низким оценкам дифференциации, даже если анализируемые субпопуляции общих аллелей не имели [10, 27, 22]. В GenAlEx эта проблема решается стандартизацией статистик дифференциации посредством деления первоначальной оценки на её *максимально возможное* значение при фактической внутрисубпопуляционной гетерозиготности [27]. Например, стандартизированная  $F_{ST(W\&C)}$  есть

$$F'_{ST(W\&C)} = F_{ST(W\&C)} / F_{ST(W\&C)max}.$$

Результаты AMOVA по GenAlEx сравнивались с таковыми по программам Arlequin v.3.5 [28], GenePop v.4.7.3 [29, 30], RST22 [31], SpaideR [32]. Для регрессионно-корреляционного анализа использовалась программа STATGRAPHICS® Centurion XVI [33].

Для сравнения подходов к использованию анализа молекулярной дисперсии (AMOVA) проведен отбор научных статей на английском языке в библиографических базах «Google Scholar» и «Scopus», статей на русском языке – в библиографической базе «Google Scholar» и «Научной электронной библиотеке eLIBRARY.RU». Также были изучены пристатейные списки литературы отобранных публикаций для выявления дополнительных релевантных статей. Поиск публикаций осуществляли по следующим ключевым словам: molecular markers, population genetics, population structure, biodiversity, diversity indices, analysis of molecular variance (AMOVA), genetic resources, conservation genetics, dairy cow, data analysis, multivariate statistical analyses. Также в комбинации с предыдущими терминами для поиска были использованы следующие ключевые слова: microsatellite, heterozygosity, Wright's F-statistics, Nei's

G-statistics, Jost's D-statistic, genetic differentiation among populations, distance matrix, ordination, PCoA, Mantel-test.

**Результаты и обсуждения.** Декомпозиция общей STR-дисперсии. В таблице 1 представлены результаты разложения STR-дисперсии. В рамках фиксированной статистической модели (до оценки варiances) наибольший интерес представляет отношение  $\eta^2 = SS_{AP} / SS_{Tot}$  ( $\eta$  – эта), которое в определённой степени характеризует *размер эффекта* – относительную долю (или процент) генетических различий между породными выборками. По AMOVA1 это отношение составило 12,4 %, в AMOVA2 – 16,6 % и 12,4 % по AMOVA3. Нулевая гипотеза,  $\eta^2 = 0$ , проверялась F-критерием Фишера через отношение  $MS_{AP}$  к  $MS_{WI}$  (для AMOVA2 –  $F = MS_{AP} / MS_{WP}$ ). F-критерии Фишера составили 3,45, 2,6 и 3,45 для AMOVA1, AMOVA2 и AMOVA3, что, учитывая соответствующие степени свободы (df), говорит о статистически значимых на уровне  $\alpha \leq 0,05$  межпородных различиях. Следует, однако, отметить, что этот тест в строгом смысле не является корректным для данных анализов. Распределение индикаторных переменных (аллелей),  $x_{ij}$ , не является нормальным. Поэтому отношение  $MS_{AP} / MS_{WP}$  не подчиняется F-распределению и не может служить статистикой критерия [18], но  $\eta^2$  может быть предварительной оценкой и/или дополнением к анализу результатов по модели случайного типа.

Когда используется статистическая модель с рандомизированными (случайными) эффектами, то конечным продуктом процедуры AMOVA являются варiances компоненты – Est.Var. или  $\sigma^2$ . Их оценки представлены в последнем столбце таблицы 1.

Для трёх вариантов AMOVA матрицы дистанций имели разные основы. Поэтому абсолютные значения варiances для одних и тех же уровней структуры данных значительно различались. Чтобы сопоставить величины варiances из разных анализов, использовали соотношение  $\sigma_{WI}^2 / \sigma_{AP}^2$ , которое для AMOVA1 и AMOVA3 было равно 9. В AMOVA2 изменчивость между аллелями внутри индивидов игнорируется, но рассчитывается внутрисубпопуляционная варiances. Её отношение к межсубпопуляционной варiances составило около 8 (достаточно близко к 9).



Таблица 1 – AMOVA-результаты по программе GenAlEx /  
Table 1 – AMOVA-results of the program GenAlEx

Вариант / Variant	Источник / Source	df	SS	MS	Est. Var.
AMOVA1	AP	6	86,9	14,5	0,46
	AI	77	266,7	3,5	0,0*
	WI	84	349,5	4,2	4,16
	Tot	167	703,1	-	4,62
AMOVA2	AP	6	28,1	4,6	0,24
	WP	77	141,7	1,8	1,84
	Tot	83	169,8	-	2,08
AMOVA3	AP	6	9199,9	1533,3	49,40
	AI	77	27604,1	358,5	0,0*
	WI	84	37308,5	444,1	444,15
	Tot	167	74112,5	-	493,56

Примечания: AP – между субпопуляциями, AI – между индивидами, WI – внутри индивидов, WP – внутри популяции, Tot – общая; df – степень свободы; SS – сумма квадратов отклонений; MS – средний квадрат; Est.Var. – оценка вариационного компонента. \* – отрицательное значение вариации приравнено нулю /

Notes: AP – between populations, AI – between individuals, WI – within individuals, WP – within population, Tot – total; df – degree of freedom; SS – sum of squared deviations; MS – mean square; Est.Var. – estimation of the variance component. \* – the negative value of the option is equal to zero.

Оценки коэффициентов инбридинга и дифференциации. По вариантам таблицы 1 получены сводные коэффициенты инбридинга и генетической дифференциации породных выборок (табл. 2).

По AMOVA1 усреднённый внутривыборочный коэффициент инбридинга (точнее – индекс фиксации,  $F_{IS}$ ) составил -9,1 %, что могло бы указывать на некоторый избыток гетерозигот. Но оценка была статистически незначимой ( $p_{perm} > 0,05$ ). С другой стороны, индекс фиксации в обобщенной выборке ( $F_{IT}$ ) был +2,7 %, что могло свидетельствовать

о некотором преобладании родственного разведения, если бы оценка была статистически значимой. В AMOVA3 расчёты основывались не на анализе частот аллелей, как в AMOVA1, а на размере аллелей. Несмотря на это, две оценки, характеризующие систему спаривания,  $R_{IS}$  и  $R_{IT}$ , оказались относительно близкими к таковым по AMOVA1, соответственно, -10,7, +1,5 % и также были статистически незначимыми ( $p_{perm} > 0,05$ ). В общем, по имеющимся данным в породных выборках статистически значимых отклонений от HWE не было установлено.

Таблица 2 – Оценки коэффициентов инбридинга и статистик генетической дифференциации выборок по программе GenAlEx /

Table 2 – Estimates of inbreeding coefficients and statistics of genetic differentiation of samples according to the GenAlEx program

AMOVA1			AMOVA2			AMOVA3		
Stat.	Est.	$p_{perm}$	Stat.	Est.	$p_{perm}$	Stat.	Est.	$p_{perm}$
$F_{IS}$ (f)	-0,091	1,000	-	-	-	$R_{IS}$	-0,107	0,985
$F_{IT}$ (F)	0,027	0,081	-	-	-	$R_{IT}$	0,015	0,375
$F_{ST(W\&C)}$ ( $\theta$ )	0,108	0,001	$\Phi_{PT}$	0,115	0,001	$R_{ST}$	0,110	0,001

Примечания: Stat. – статистика; Est. – оценка;  $p_{perm}$  – достигнутый уровень статистической значимости (пермутационный тест с 999 перестановками);  $F_{IS}$  ( $R_{IS}$ ) – коэффициент инбридинга в пределах выборок ( $= \sigma_{AI}^2 / \sigma_{Tot}^2$ );  $F_{IT}$  ( $R_{IT}$ ) – коэффициент инбридинга в объединённой выборке ( $= (\sigma_{AP}^2 + \sigma_{AI}^2) / \sigma_{Tot}^2$ );  $F_{ST(W\&C)}$  ( $\theta$ ),  $\Phi_{PT}$ ,  $R_{ST}$  – статистики генетической дифференциации породных выборок /

Notes: Stat. – statistic; Est. – estimation;  $p_{perm}$  – achieved level of statistical significance (permutation test with 999 permutations);  $F_{IS}$  ( $R_{IS}$ ) – inbreeding coefficient within the samples ( $= \sigma_{AI}^2 / \sigma_{Tot}^2$ );  $F_{IT}$  ( $R_{IT}$ ) – inbreeding coefficient in the combined sample ( $= (\sigma_{AP}^2 + \sigma_{AI}^2) / \sigma_{Tot}^2$ );  $F_{ST(W\&C)}$  ( $\theta$ ),  $\Phi_{PT}$ ,  $R_{ST}$  – statistics of genetic differentiation of breed samples.

Главный интерес представляли статистики, характеризующие уровень STR-дифференциации породных выборок. Рассчитанная по AMOVA1 статистика  $F_{ST(W\&C)}$  составила 0,108 или 10,8 % при  $p_{perm} < 0,001$ . В варианте AMOVA2, в котором внутрииндивидуальная STR-изменчивость игнорировалась,  $\Phi_{PT}$ -оценка незначительно превышала таковую варианта AMOVA1 и составила 11,5 % при  $p_{perm} < 0,001$ . В AMOVA3, где расчёты базировались на размерах аллелей,  $R_{ST}$ -оценка составила 11 % при  $p_{perm} < 0,001$ , что практически эквивалентно значению по AMOVA1. Полученные величины интерпретируют как «представляющие довольно слабую (rather weak) популяционную структуру» [11]. Отметим, рассчитанные выше оценки размера эффекта ( $\eta^2$ ) были лишь на 1-5 процентных пункта выше.

**Полокусные оценки дифференциации.** Для исследования степени связности разных мер STR-дифференциации имеющийся набор данных был разделён на 11 псевдонаборов каждый с разным локусом. Все псевдонаборы были обработаны тремя вариантами AMOVA для получения соответствующих полокусных оценок дифференциации и внутривыборочной гетерозиготности ( $cH_s$ ). В результате сформировалась выборка с четырьмя переменными (три оценки AMOVA +  $cH_s$ ) и 11 записями (локусы). Взаимосвязь между переменными оценивалась непараметрическим коэффициентом корреляции Спирмена ( $r_s$ ) и тестировалась достигнутым уровнем статистической значимости ( $p_{value}$ ).

Между полокусными оценками  $F_{ST(W\&C)}$  и  $\Phi_{PT}$  корреляция Спирмена была 0,99 при  $p_{value} < 0,0001$ . В то же время между указанными статистиками и  $R_{ST}$  значимой взаимосвязи не было установлено ( $r_s$  для  $F_{ST(W\&C)} \times R_{ST}$  была 0,28, для  $\Phi_{PT} \times R_{ST}$  – 0,32 обе с  $p_{value} = 0,31$ ). При практически равных сводных оценках, отсутствие взаимосвязи между полокусными оценками было неожиданным.

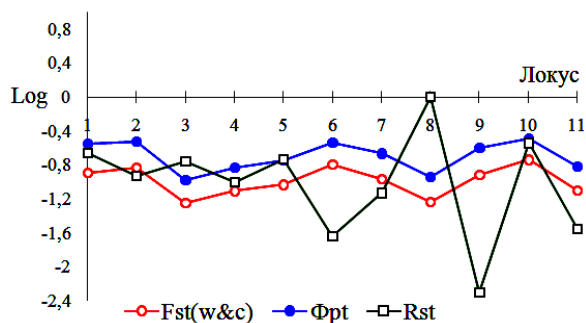


Рис. 1. Логарифмы  $F_{ST(W\&C)}$ -,  $\Phi_{PT}$ - и  $R_{ST}$ -статистик по 11 STR-локусам /

Fig. 1. Logarithms of  $F_{ST(W\&C)}$ -,  $\Phi_{PT}$ - and  $R_{ST}$ -statistics for 11 STR loci

Графики рисунка 1 прояснили причину получения таких результатов, именно: логарифмы ( $\log_{10}$ ) оценок  $R_{ST}$  по шестому, восьмому и девятому локусам очень сильно отклонялись от таковых по  $F_{ST(W\&C)}$  и  $\Phi_{PT}$  мерам, изменение которых по локусам было практически синхронным. Отметим также, коэффициенты вариации полокусных  $F_{ST(W\&C)}$  и  $\Phi_{PT}$  оценок были соответственно 37,8 и 36,6 %, в то время как  $R_{ST}$  оценок – 91 %!  $R_{ST}$ -оценки по локусам имели очень большой разброс, что может характеризовать вариант AMOVA3 не с лучшей стороны (ненадёжный при оценке по ограниченному числу локусов). Кроме того, имеется мнение, что, если даже мутации строго пошаговые,  $R_{ST}$  не всегда является лучшей оценкой, чем  $F_{ST(W\&C)}$  [34]. Это может быть в случае, когда интересующий временной период короткий, а влияние мутаций сравнительно невелико [21]. Также полагают, что микросателлиты почти никогда строго не следуют пошаговой мутационной модели. Поэтому советуют избегать использования  $R_{ST}$ -статистики [35].

**AMOVA-оценки по иным, чем GenAlEx программам.** Была проведена проверка степени подобия статистик дифференциации, полученных по программе GenAlEx, с таковыми при использовании (на тех же самых данных) модуля AMOVA в других компьютерных программах. Результаты представлены в таблице 3.

Оценки STR-дифференциации породных выборок,  $F_{ST(W\&C)}$  ( $R_{ST}$ ), имели высокий уровень статистической значимости ( $p_{perm} < 0,001$ ). Все оценки по программе Arlequin были идентичны таковым по программе GenAlEx. Поэтому, как представляется, 95 % доверительные интервалы (confidence interval, CI) оценок по  $F_{ST}$  и  $R_{ST}$  мерам, рассчитанные в программе Arlequin, можно приложить и к оценкам по программе GenAlEx.

По программе GenePop оценка была идентичной [GenAlEx] только в варианте AMOVA1; по варианту AMOVA3  $R_{ST}$ -оценка была на 22 % ниже, но не выходила за границы 95 % CI, рассчитанного в программе Arlequin.

По программе RST22 были получены две оценки  $R_{ST}$ : усреднённая по локусам, равная 0,115, и средневзвешенная (по средним квадратам отклонений) – 0,122. Последняя была на 6,1 % выше, но 95 % CI двух оценок перекрывались, следовательно различие между ними было статистически незначимым. Оценки по RST22 были на 4,5-10,9 % выше, чем в программе GenAlEx, но в пределах 95 % CI, рассчитанного в программе Arlequin.

Таблица 3 – AMOVA по программам Arlequin v.3.5, GenePop v.4.7.3 и RST22 /  
Table 3 – AMOVA for Arlequin v.3.5, GenePop v.4.7.3 and RST22 programs

AMOVA1				AMOVA3			
Arlequin v.3.5							
Stat.	Est.	95% CI <sub>L</sub>	95% CI <sub>U</sub>	Stat.	Est.	95% CI <sub>L</sub>	95% CI <sub>U</sub>
F <sub>IS</sub> (f)	-0,091	-0,142	-0,046	R <sub>IS</sub>	-0,107	-0,175	0,008
F <sub>IT</sub> (F)	0,027	0,011	0,064	R <sub>IT</sub>	0,015	-0,030	0,129
F <sub>ST</sub> (θ)	0,108	0,086	0,133	R <sub>ST</sub>	0,110	0,036	0,174
GenePop v.4.7.3							
F <sub>IS</sub> (f)	-0,091	-	-	R <sub>IS</sub>	-0,123	-	-
F <sub>IT</sub> (F)	0,026	-	-	R <sub>IT</sub>	-0,025	-	-
F <sub>ST</sub> (θ)	0,108	-	-	R <sub>ST</sub>	0,086	-	-
RST22							
AMOVA3 Est.av.				AMOVA3 Est.w.			
R <sub>ST</sub>	0,115	0,109	0,180	R <sub>ST</sub>	0,122	0,110	0,218

Примечания: Est.av. – усреднённая по локусам; Est.w. – взвешенная по локусам; 95% CI – 95% доверительный интервал; субиндексы L и U – нижняя и верхняя границы CI /

Notes: Est.av. – averaged by loci; Est.w. – weighted by loci; 95 % CI – 95 % confidence interval; subindexes L and U – lower and upper bounds of CI.

Таким образом, все используемые компьютерные программы обеспечивали получение эквивалентных или близких (в пределах 95 % CI) оценок уровня STR-дифференциации породных выборок.

Сравнение с оценками по Нею (Nei). Ранее [22] на этих же STR-данных в программе GenAlEx были получены оценки генетической дифференциации на основе анализа внутри- и межвыборочной ожидаемой гетерозиготности. В частности, были рассчитаны оценки F<sub>ST</sub> – адаптированного Неем индекса фиксации Райта [23], G<sub>ST</sub> – коэффициента генной дифференциации Нея [8] и G<sub>ST(NEI)</sub> – модифицированной Неем G<sub>ST</sub>-меры, учитывающей число анализируемых выборок [36]. Эти оценки приведены в таблице 4.

Таблица 4 – Сводные оценки статистик дифференциации по Нею [22] /

Table 4 – Summary estimates of the differentiation statistics for Nei's [22]

	F <sub>ST</sub>	G <sub>ST</sub>	G <sub>ST(NEI)</sub>
<b>Est.</b>	<b>0,134</b>	<b>0,103</b>	<b>0,118</b>
Error	0,013	0,014	0,015
95% CI <sub>L</sub>	0,111	0,080	0,093
95% CI <sub>U</sub>	0,159	0,130	0,148
p <sub>perm</sub>	0,001	0,001	0,001

Если исключить F<sub>ST</sub>-оценку, при расчёте которой не учитывается ни число выборок,

ни их размеры, то оценки G-статистик были достаточно близки к AMOVA-оценкам. При расчёте G<sub>ST</sub> учитывались численности выборок через их среднегармонический размер, при расчёте G<sub>ST(NEI)</sub> также число выборок. Поэтому оценка G<sub>ST(NEI)</sub> = 0,118 представляется более корректной. Её величина особенно хорошо соотносится с оценкой Ф<sub>PT</sub> = 0,115, полученной в программе GenAlEx, и с оценками R<sub>ST</sub>, рассчитанными в программе RST22 – 0,115 и 0,122.

В общем можно полагать, что все используемые методы и компьютерные программы производили вполне сопоставимые оценки статистик генетической дифференциаций. Коэффициенты ранговой корреляции между полокусными F<sub>ST</sub>-, G<sub>ST</sub>- и G<sub>ST(NEI)</sub>-оценок с F<sub>ST(w&c)</sub>- и Ф<sub>PT</sub>-оценками были равны 0,96 при p<sub>value</sub> = 0,0023, с R<sub>ST</sub>-оценками – 0,29 при p<sub>value</sub> = 0,3576. Эти соотношения были близки к таковым, полученными выше между оценками по F<sub>ST</sub>-, Ф<sub>PT</sub>- и R<sub>ST</sub>-мерам.

Стандартизация AMOVA-оценок. На гипотетических примерах было показано, что при анализе мультиполиморфных маркеров уровень гетерозиготности в субпопуляциях (сНs) оказывает негативное влияние на статистики дифференциации, занижая их оценки [10, 37]. Это нашло подтверждение при расчёте G-статистик [22]. Полокусные наборы данных были использованы для изучения влияния внутривыборочной гетерозиготности (сНs) на AMOVA-оценки (рис. 2).



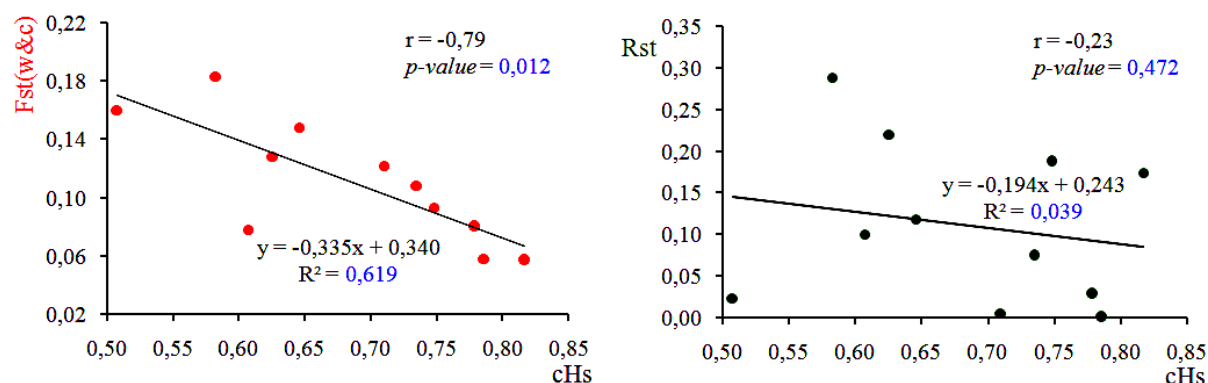


Рис. 2. Зависимость  $F_{ST(W\&C)}$ - и  $R_{ST}$ -оценок от внутривыборочной гетерозиготности (cHs) (оценки по каждому локусу;  $r$  – корреляция по Спирмену,  $R^2$  – коэффициент детерминации линейной модели) /

Fig. 2. Dependence of  $F_{ST(W\&C)}$ - and  $R_{ST}$ -estimates on intrasample heterozygosity (cHs) (estimates for each locus;  $r$  – Spearman correlation,  $R^2$  – coefficient of determination of the linear model)

С повышением cHs оценки  $F_{ST(W\&C)}$  снижались, их диапазон сужался. Подобная закономерность имела место для полокусных  $\Phi_{PT}$ -оценок ( $y = -0.602x + 0.628$ ,  $R^2 = 0.565$ ; на рисунке нет), но отсутствовала для  $R_{ST}$ -оценок (рисунок 2 справа). Негативная зависимость  $F_{ST(W\&C)}$  и  $\Phi_{PT}$  оценок от внутривыборочной гетерозиготности была достаточно высокой и статистически значимой. В случае с  $R_{ST}$  такой зависимости не наблюдалось.

Для устранения занижения оценок дифференциации при высокой внутрисубпопуляционной гетерозиготности Hedrick ввёл стандартизацию  $G_{ST}$ -статистики на её возможное максимальное значение –  $G_{ST(max)}$  [10]. Следуя за подходом Hedrick, Meirmans [27] предложил вычислять стандартизованную  $\Phi_{PT}$ -статистику:  $\Phi'_{PT} = \Phi_{PT}/\Phi_{PT(max)}$ . Оценку  $\Phi_{PT(max)}$  получают путём максимизации межпопуляционной дисперсии,  $\sigma^2_{AP}$ , через межпопуляционную сумму квадратов отклонений  $SS(AP)$  при фактической внутрисубпопуляционной дисперсии,  $\sigma^2_{WP}$ :

$$\Phi_{PT(max)} = \sigma^2_{AP(max)} / (\sigma^2_{AP(max)} + \sigma^2_{WP}).$$

Этот подход заложен в программе GenAlEx. В нашем случае оценка  $\Phi_{PT(max)}$  составила 0,308, а стандартизованное значение  $\Phi_{PT}$ :

$$\Phi'_{PT} = 0,115/0,308 = 0,373 \text{ или } \approx 37 \ \%.$$

Стандартизованная оценка  $F_{ST(W\&C)}$ ,  $F'_{ST(W\&C)}$ , была 0,351 или  $\approx 35 \ %$  от максимального значения. Обе стандартизованные оценки более чем в три раза превышали исходные. Сводная  $F_{ST}$ -оценка по программе Arlequin v.3.5 была 10,8 % (как и  $F_{ST(W\&C)}$ ) с 95 % CI 8,6...13,3 % (табл. 3). Если стандартизовать значения доверительного интервала, то можно полагать, что с вероятностью 95 %

истинная  $F'_{ST(W\&C)}$  находится в диапазоне от 28 до 43 %. Эта интервальная оценка указывает на очень выраженную популяционную структуру анализируемых данных (в противоположность заключению по первичным оценкам).

На рисунке 3 показано распределение полокусных оценок  $F'_{ST(W\&C)}$  в зависимости от внутривыборочной гетерозиготности. Коэффициент детерминации ( $R^2$ ) линейной модели был 0,002, а коэффициент корреляции Спирмена –  $-0,05$  при  $p\text{-value} = 0,875$ . Стандартизация устранила влияние внутривыборочной гетерозиготности на оценки дифференциации.

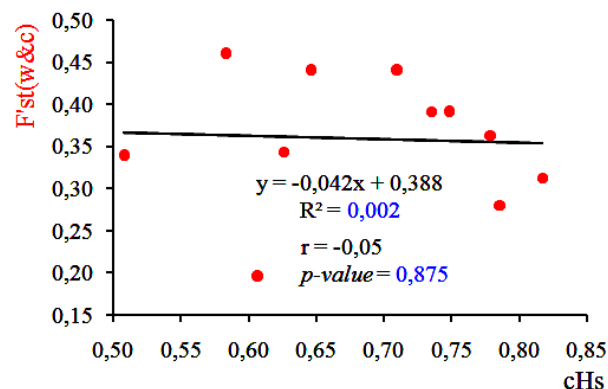


Рис. 3. Зависимость  $F'_{ST(W\&C)}$ -оценок от внутривыборочной гетерозиготностью (cHs) /

Fig. 3. Dependence of  $F'_{ST(W\&C)}$  estimates on intrasample heterozygosity (cHs)

Аналогичные результаты были получены в наших предыдущих исследованиях [22]. Так, сводные оценки генетической дифференциации методами Нея ( $G_{ST}$  и  $G_{ST(NEI)}$ ) были 10,3-11,8 %. После стандартизации по Hedrick [10] они составили 36,4-37,5 %. При использовании метода Джоста (Jost) на базе эффективного числа аллелей [12], результаты которого

не зависят от внутрисубпопуляционной гетерозиготности, сводная оценка ( $D_{EST}$ ) была 29,2 %.

Расчёты в рамках данной работы методом Морисита-Хорна (Morisita-Horn) по программе SpadeR (при «number Hill»  $q = 2$ ; [32]) показали STR-дифференциацию породных выборок на уровне 30,8 % с 95 % CI 27,5...34,2 % (в среднем около 1/3 аллелей одной выборки не являлись общими для других породных выборок). Однако заметим, оценка по этой же программе региональной дифференциации (Regional diff.) была на уровне 5 % с 95 % CI 4...6 %, т.е. 5 % аллелей в совокупной выборке не являлись общими для той или иной породной выборки.

**Генетические дистанции.** Сводные оценки характеризовали обобщённую дифференциацию породных выборок, нивелируя двухсторонние генетические отношения. Попарные

генетические дистанции (Genetic Distance, GD) позволяют выявить наличие или отсутствие пространственной структуры среди породных выборок. Упрощенно, GD между двумя животными (породами) можно представить, как долю или процент аллелей локуса(-ов), которые не являются общими; при GD равной нулю оба животных (породы) имеют одни и те же (общие) аллели с одинаковыми частотами.

В таблице 5 даны треугольные матрицы парных GD, которые были получены по разным вариантам AMOVA. В большей степени оценки GD варьировали в пределах матриц, указывая на неоднородность генетических отношений между породными выборками. Между матрицами оценки GD варьировали в меньшей степени, свидетельствуя о незначительном влиянии вида STR-данных, вводимых в модуль AMOVA программы GenAlEx.

**Таблица 5 – Треугольные матрицы парных генетических дистанций (под диагональю – AMOVA-оценки, над диагональю –  $p_{perm}$ ;  $\alpha_{Bonf} = 0,05/21 = 0,0024$ ) /**

**Table 5 – Triangular matrices of paired genetic distances (under the diagonal – AMOVA-estimates, above the diagonal –  $p_{perm}$ ;  $\alpha_{Bonf} = 0.05/21 = 0.0024$ )**

Порода / Breed	AYR	JER	RDAT	RSH	H-DEU	H-NLD	H-USA
AMOVA1 ( $F_{ST(W\&C)}$ )							
AYR	-	0,001	0,429	0,251	0,001	0,001	0,001
JER	0,217	-	0,001	0,001	0,001	0,001	0,001
RDAT	0,001	0,250	-	0,330	0,001	0,001	0,001
RSH	0,008	0,275	0,005	-	0,001	0,001	0,001
H-DEU	0,112	0,181	0,120	0,119	-	0,171	0,241
H-NLD	0,089	0,163	0,100	0,105	0,008	-	0,008
H-USA	0,124	0,216	0,123	0,121	0,006	0,030	-
AMOVA2 ( $\Phi_{PT}$ )							
AYR	-	0,001	0,291	0,248	0,001	0,001	0,001
JER	0,210	-	0,001	0,001	0,001	0,001	0,001
RDAT	0,007	0,246	-	0,281	0,001	0,001	0,001
RSH	0,011	0,266	0,010	-	0,001	0,001	0,001
H-DEU	0,111	0,178	0,123	0,120	-	0,065	0,186
H-NLD	0,105	0,186	0,120	0,124	0,013	-	0,001
H-USA	0,125	0,213	0,128	0,125	0,007	0,038	-
AMOVA3 ( $R_{ST}$ )							
AYR	-	0,001	0,109	0,101	0,001	0,001	0,001
JER	0,175	-	0,001	0,003	0,004	0,001	0,001
RDAT	0,030	0,180	-	0,342	0,002	0,002	0,002
RSH	0,043	0,185	0,006	-	0,011	0,004	0,010
H-DEU	0,203	0,142	0,152	0,113	-	0,376	0,386
H-NLD	0,209	0,161	0,189	0,139	0,000	-	0,372
H-USA	0,169	0,132	0,145	0,090	0,000	0,000	-

Наибольшие GD имели место между JER-выборкой и остальными породными выборками (0,14-0,27). Заметные GD (0,1-0,2) были между группой «красных» пород (AYR, RDAT, RSH) и группой голштинских «экотипов». В то же время в пределах этих двух групп GD между породными выборками были нулевые или близкие к нулю. С вероятностью ошибки  $\alpha_{Bonf} = 0,0024$  (множественный тест Бонферрони) из 21 парных сравнение 15 оценок GD были статистически значимыми. GD между породными выборками AYR с RDAT, AYR с RSH, RDAT с RSH и между голштинскими выборками (кроме H-NLD с H-USA по  $\Phi_{PT}$ ) были статистически незначимыми по всем трём мерам ( $p_{perm} > \alpha_{Bonf} = 0,0024$ ).

В таблице 6 даны четыре треугольные матрицы парных GD: две по AMOVA1 с линеаризованными ( $\text{lin}F_{ST(W\&C)}$ ) и стандартизованными ( $F'_{ST(W\&C)}$ ) элементами, по Нею со стандартизацией Хедрика ( $G''_{ST(HEd)}$ ) и по Джосту ( $D_{EST}$ ); две последние из работы [22]. Диапазон  $\text{lin}F_{ST(W\&C)}$ -дистанций был от 0,001

(AYR с RDAT) до 0,380 (JER с RSH) со средней **0,137**. Диапазон не представленных в таблице 6  $\text{lin}R_{ST}$ -дистанций был от 0 (между всеми голштинскими выборками) до 0,265 (AYR с H-NLD) со средней **0,140**;  $\text{lin}\Phi_{PT}$ -дистанции – от 0,007 до 0,363 со средней **0,142**.

Стандартизованные оценки GD ( $F'_{ST(W\&C)}$ ) были существенно выше первоначальных ( $F_{ST(W\&C)}$ ). Например, если  $F_{ST(W\&C)}$ -дистанция между JER и AYR выборками была 0,217 (табл. 5), а её линеаризованная оценка,  $\text{lin}F_{ST(W\&C)}$ , – 0,277 (табл. 6), то стандартизованная  $F'_{ST(W\&C)}$ -дистанция составила 0,590. Последняя была эквивалентна дистанции по Нею-Хедрику ( $G''_{ST(HEd)} = 0,590$ ), и на 19 % превышала оценку по Джосту ( $D_{EST} = 0,477$ ). Расчёты по программе SpadeR методом Морисита-Хорна (Morisita-Horn), который приравнивается к методу Джоста [32], показали дистанцию, усреднённую по локусам, равную 0,447, а взвешенную на инверсию квадратичной ошибки – 0,516 с 95 % CI 0,441...0,590.

Таблица 6 – Треугольные матрицы парных генетических дистанций по  $\text{lin}F_{ST(W\&C)}$  и  $F'_{ST(W\&C)}$ ,  $G''_{ST(HEd)}$  и  $D_{EST}$  мерам /

Table 6 – Triangular matrices of paired genetic distances by  $\text{lin}F_{ST(W\&C)}$  and  $F'_{ST(W\&C)}$ ,  $G''_{ST(HEd)}$  and  $D_{EST}$  measures

Порода/Breed	AYR	JER	RDAT	RSH	H-DEU	H-NLD	H-USA
Под диагональю $\text{lin}F_{ST(W\&C)}$ , над диагональю $F'_{ST(W\&C)}$ / Under the diagonal $\text{lin}F_{ST(W\&C)}$ , above the diagonal $F'_{ST(W\&C)}$							
AYR	-	0,590	0,005	0,026	0,401	0,337	0,389
JER	0,277	-	0,663	0,729	0,522	0,498	0,563
RDAT	0,001	0,333	-	0,017	0,416	0,371	0,377
RSH	0,008	0,380	0,006	-	0,414	0,388	0,372
H-DEU	0,127	0,222	0,137	0,135	-	0,034	0,020
H-NLD	0,097	0,194	0,111	0,117	0,009	-	0,107
H-USA	0,141	0,276	0,140	0,138	0,006	0,031	-
Под диагональю $G''_{ST(HEd)}$ , над диагональю $D_{EST}$ / Under the diagonal $G''_{ST(HEd)}$ , above the diagonal $D_{EST}$							
AYR	-	0,477	0,003	0,019	0,325	0,270	0,302
JER	0,590	-	0,551	0,626	0,417	0,401	0,443
RDAT	0,005	0,663	-	0,012	0,337	0,300	0,289
RSH	0,026	0,728	0,017	-	0,336	0,316	0,285
H-DEU	0,401	0,524	0,417	0,416	-	0,024	0,014
H-NLD	0,335	0,501	0,371	0,388	0,032	-	0,079
H-USA	0,387	0,565	0,377	0,372	0,019	0,107	-

Примечание: Матрицы с  $G''_{ST(HEd)}$  и  $D_{EST}$  дистанциями из публикации [22] /

Note: Matrices with  $G''_{ST(HEd)}$  and  $D_{EST}$  distances from the publication [22]

Между JER и H-USA выборками  $F_{ST(W\&C)}$ -,  $\text{lin}F_{ST(W\&C)}$ -,  $F'_{ST(W\&C)}$ -,  $G''_{ST(HED)}$ - и  $D_{EST}$ -дистанции были, соответственно, 0,216, 0,276, 0,563, 0,565 и 0,443. По программе SpadeR усреднённая по локусам GD была 0,472, средневзвешенная – 0,373 с 95 % CI 0,304...0,443. Если предпочтение отдать последней оценке (0,373), как, на наш взгляд, наиболее обоснованной, и принять её за 100 %, то оценки  $F_{ST(W\&C)}$  и  $\text{lin}F_{ST(W\&C)}$  были *занижены*, примерно, в 1,5 раза,  $F'_{ST(W\&C)}$  и  $G''_{ST(HED)}$  – *завышены* в 1,5 раза, а оценка по Джосту была ниже на 20 %.

**Тест Мантеля.** Чтобы оценить степень сходства/подобия матриц парных GD, использовали тест Мантеля (корреляцию Мантеля,  $R_M$ ) с рандомизированной пермутационной (random permutation) проверкой статистической значимости. Величина  $R_M = 1$  указывает на то, что увеличение GD между субпопуляциями  $i$  и  $j$  в матрице  $X$  связано с увеличением GD между этими субпопуляциями в матрице  $Y$ , т. е. на сходство матриц. В нашем случае – два типа представления данных для AMOVA обеспечивают получение подобных структур генетических взаимоотношений выборок.  $R_M = -1$  указывает на противоположный результат. Величина  $R_M$ , близкая к нулю, свидетельствует о некоррелированности матриц, т. е. разные типы -данных приводят к независимым оценкам парных GD.

Между  $F_{ST(W\&C)}$  и  $\Phi_{PT}$  матрицами корреляция Мантеля была 0,99 при  $p_{perm} = 0,010$ . Тесты Мантеля этих двух матриц с  $R_{ST}$ -матрицей показали более низкие значения  $R_M$ , соответственно, 0,75 при  $p_{perm} = 0,05$  и 0,80 при  $p_{perm} = 0,02$ .  $R_M$  есть мера линейной зависимо-

сти; квадрат  $R_M$  ( $R_M^2$ ) – это коэффициент детерминации, представляющий долю или процент общей для двух матриц вариации.  $R_M^2$  более точно характеризует сопряжённость матриц, а  $(1-R_M^2)$  – долю «необъяснённой» вариации или *коэффициент неопределённости*. Для рассматриваемых матриц  $R_M^2$  составили, соответственно, 98, 56 и 64 %. Если в первом случае коэффициент неопределённости был всего 2 %, то в двух других – 44 и 36 %. Следовательно, по вариантам AMOVA1 и AMOVA2 были получены очень близкие структуры генетических отношений между породными выборками, тогда как по AMOVA3 породные выборки имели иную структуру генетических взаимоотношений. В частности, по  $F_{ST(W\&C)}$  и  $\Phi_{PT}$  мерам наибольшие парные GD имели место между выборкой джерсейских быков и всеми остальными; по  $R_{ST}$  – между выборкой айрширских быков и прочими породными выборками.

Первоначальные матрицы GD имели очень тесную взаимосвязь с линеаризованными:  $R_M \geq 0,995$  ( $R_M^2 \geq 99$  %) при  $p_{perm} \leq 0,001$ . Это указывало на незначительное воздействие линеаризации на отображение структуры генетических отношений между породными выборками. Также высокой была сопряжённость  $\text{lin}F_{ST(W\&C)}$  и  $\text{lin}\Phi_{PT}$  матриц – 0,994 при  $p_{perm} = 0,001$  (рис. 4). Вместе с тем, тесты Мантеля этих матриц с  $\text{lin}R_{ST}$ -матрицей показали более низкие, чем по первоначальным, оценки  $R_M$ , соответственно, 0,696 и 0,729 при  $p_{perm} = 0,005$ . Заметное *несходство* парных  $\text{lin}F_{ST(W\&C)}$ - и  $\text{lin}R_{ST}$ -дистанций можно видеть по разбросу точек на рисунке 4 (справа) при коэффициенте неопределённости 52 %.

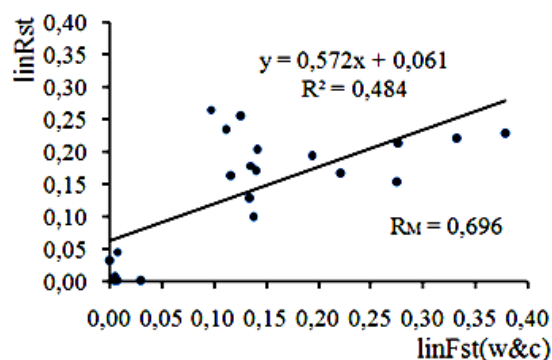
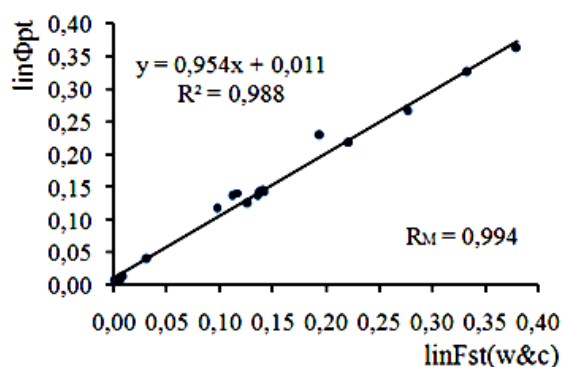


Рис. 4. Тест Мантеля ( $R_M$ ) для GD-матриц  $\text{lin}F_{ST(W\&C)}$ ,  $\text{lin}\Phi_{PT}$  и  $\text{lin}R_{ST}$  /

Fig. 4. Mantel test ( $R_M$ ) for GD-matrices  $\text{lin}F_{ST(W\&C)}$ ,  $\text{lin}\Phi_{PT}$ , and  $\text{lin}R_{ST}$

Корреляция Мантеля матрицы стандартизованных  $F'_{ST(W\&C)}$ -дистанций с матрицей первоначальной  $F_{ST(W\&C)}$ -дистанций была 0,981

при  $p_{perm} = 0,001$ . Стандартизация увеличила в разы оценки парных GD, но незначительно повлияла на структуру генетических взаимо-

отношений породных выборок ( $R^2_M = 96,2\%$ ;  $1 - R^2_M = 3,8\%$ ). С матрицей линейаризованных дистанций коэффициент неопределённости был в 2 раза выше –  $7,6\%$ .

Особый интерес представляла сопряжённость  $F'_{ST(W\&C)}$ -матрицы с матрицами GD из работы [22].  $R_M$  с матрицей по Райту-Нею ( $F_{ST}$ ) была  $0,966$ , с  $G_{ST(NEI)}$ -матрицей –  $0,982$ , с матрицей несмещённых оценок генетических дистанций Нея ( $u_{GDN}$ ) –  $0,977$  (все с  $p_{perm} = 0,001$ ). Соответствующие  $R^2_M$  были  $93,3$ ,  $96,4$ ,  $95,5\%$ . Предельно большие  $R_M$  были получены с матрицей коэффициентов «генной дифференциации» Нея-Хедрика ( $G''_{ST(HEI)}$ ) –  $1,0$  и с матрицей коэффициентов «истинной аллельной дифференциации» Джоста ( $D_{EST}$ ) –  $0,999$  (обе с  $p_{perm} < 0,001$ ). Три меры с разными концепциями и алгоритмами расчётов: по гетерозиготности ( $G''_{ST(HEI)}$ ), по эффективному числу аллелей ( $D_{EST}$ ) и по AMOVA ( $F'_{ST(W\&C)}$ ) выявили структуры генетических отношений между породными выборками со  $100\%$  схожестью (заметим, корреляция Мантеля между  $G''_{ST(HEI)}$  и  $D_{EST}$  матрицами была также очень высокой –  $0,998$  при  $p_{perm} = 0,002$ ). В целом, тесты Мантеля показали достаточно высокую сопряжённость матриц парных GD, полученных разными методами и вариантами AMOVA (кроме  $R_{ST}$ ).

В работе [38] исследовали GD среди 9 линий балканского низинного скота (Balkan Buša / Busha) и 12 европейских коммерческих пород. Были использованы 105 STR-локусов и две меры дифференциации:  $F_{ST}$  и  $D_{EST}$ . Усреднённые по всем популяциям оценки составили  $0,06$  и  $0,093$  соответственно, по балканским Буша –  $0,016$  и  $0,021$ , по северо-западным европейским породам –  $0,085$  и  $0,124$ .  $D_{EST}$ -оценки были выше  $F_{ST}$ -оценок в  $1,3$ - $1,6$  раза. Между представленными в статье матрицами попарных GD корреляция Мантеля (рассчитанная нами) составила  $0,984$  при  $p_{perm} = 0,01$ .

Иные результаты были получены в работе [39]. Авторы изучали 5 исторических и 3 современных российских популяций крупного рогатого скота по девяти локусам STR-маркеров с оценкой парных GD мерами  $F_{ST}$  и  $D_{EST}$ . Было сделано заключение, что «...both pairwise

Jost's D and  $F_{ST}$  values in most cases were similar (Table 3)». Но данные в статье матрицы парных GD свидетельствовали об обратном. В частности, из 28 сравнений в 15 случаях оценки  $F_{ST}$  превышали  $D_{EST}$ ; из них в 10 случаях – более чем в 2 раза (имелись различия в 8, 11, 34 и более раз). Рассчитанная нами корреляция Мантеля составила  $0,508$ ,  $R^2_M = 0,26$ , а коэффициент неопределённости –  $0,74$ , что свидетельствовало о малом соответствии  $F_{ST}$ - и  $D_{EST}$ -матриц (следовательно, и структур генетических взаимоотношений между породами). Особенно настораживает то, что  $F_{ST}$ -оценки превышали  $D_{EST}$ -оценки. По данным литературы в большинстве случаев ситуация, как правило, обратная (см., например, [38]). При небольших значениях внутривидовой гетерозиготности ( $<0,5$ ) иногда  $F_{ST}$  может быть больше  $D_{EST}$  [11]. Но анализируемые «популяции» имели несмещённые оценки гетерозиготности в диапазоне  $0,644...0,852$ . Также озадачивают фактически нулевые GD между великорусским\* и голландским скотом\*\* ( $F_{ST} = 0,006$ ,  $D_{EST} = -0,008$ ), историческим и современным холмогорским скотом ( $F_{ST} = 0,018$ ,  $D_{EST} = 0,016$ ). Между великорусским и современным голштинским скотом GD были меньше ( $F_{ST} = 0,079$ ,  $D_{EST} = 0,016$ ), чем между холмогорским и ярославским ( $F_{ST} = 0,174$ ,  $D_{EST} = 0,071$ ). Отметим, статистики разнообразия характеризуют не только уровень генетических различий [двух] популяций, но и степень генетического сходства животных этих популяций.

В другой российской работе [40] были даны матрицы парных GD между 11 породами свиней по 10 STR-локусам. В большинстве случаев (52 из 55)  $D_{EST}$ -дистанции превышали  $F_{ST}$  (оценки  $D_{EST}$  были выше в  $1,3$  раза). Рассчитанная нами корреляция Мантеля составила  $0,83$  ( $R^2_M = 0,69$ ). Коэффициент неопределённости, равный  $0,31$ , указывал на наличие заметных отличий в структурах генетических взаимоотношений между породами при использовании двух мер. Это можно видеть и при сравнении приводимых в [40] филогенетических дендрограмм (авторы, к сожалению, не указали «правильную» дендрограмму).

\* В отчете академика А. Ф. Миддендорфа за 1884 год коровы великорусской породы характеризовались как «тасканки» и «горемычки» – мелкие, узкие, беднокостные, размножаемые, в основном, для производства навоза (Исследование современного состояния скотоводства в России. Рогатый скот. В. Отчет А. Миддендорфа. М., 1884. Вып. 1. С. 1-50).

\*\* Корицей российской зоотехнической науки проф. П. Н. Кулешов относил голландскую породу к «древним породам, послужившим к улучшению животноводства до 1700-1750 гг.». В частности, он писал: «В отношении крупного рогатого скота можно сказать, что существовала одна единственная голландская порода, выделившаяся из общей массы скота и служившая к улучшению скотоводства во всей Европе; от ней же произошла и наша холмогорская порода» (Кулешов П. Н. Породы домашних животных в исторической последовательности их развития (1926 г.). В кн.: Теоретические работы по племенному животноводству. М.: ОГИЗ «СЕЛЬХОЗГИЗ», 1947. С. 209-221).



*Ординация быков на 2D диаграмме.* Ординация – это упорядочение объектов вдоль некоторых осей методами многомерной статистики (РСА, РСоА и т.п.) с целью установления связи, классификации и проекции взаиморасположения на плоскости или в пространстве. Диаграмма на рисунке 5 представляет собой результат применения *анализа главных координат* (Principal Coordinate Analysis, РСоА) к STR-данным. В процессе РСоА исходная матрица размером 84×84 трансформировалась в более простую матрицу 84×2 (где 2 – число размерностей), содержащую координаты точек для изображения (смысл каждой из координат не имеет значения; главным является взаимное расположение точек-быков). Диаграмма визуализирует STR-вариацию среди быков в общей выборке и существующие в ней структуры. Взаимоотношения 84 быков отображены таким образом,

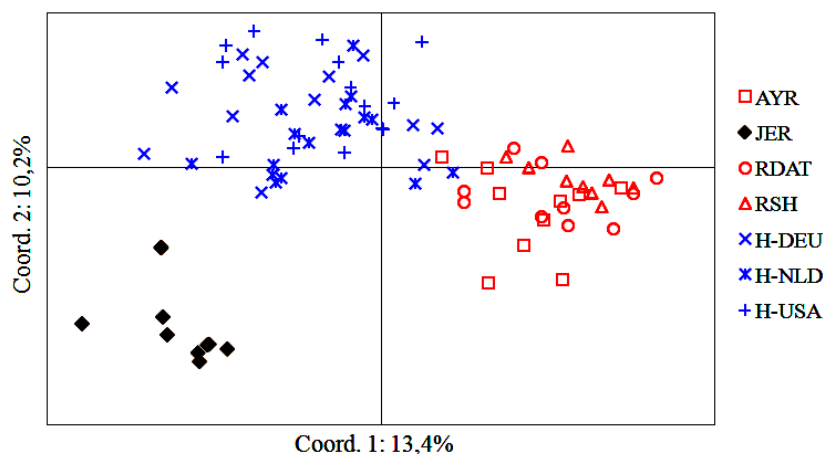


Рис. 5. Ординация быков породных выборок на плоскости первой и второй главных координат /  
Fig. 5. Ordination of the bulls of the breed samples on the plane of the first and second principal coordinates

Несмотря на 100-23,6≈76 % потерю информации, в исследуемых данных выделились три относительно компактные структуры (кластера), что указывает на наличие генетически отличающихся *групп* животных. Координата 1 отделила голштинские «экотипы» от выборок «красных» пород; координата 2 – джерсейских быков от прочих. В верхнем левом углу расположились производители голштинских «экотипов», справа – быки «красных» пород, в нижнем левом углу – более сегрегированная (отличающаяся по генотипам) группа животных джерсейской породы.

Первые две группы (кластера) перекрывались незначительно. Однако внутри этих групп имело место высокая STR изменчивость и *смешение* животных разных породных выборок. Это характеризует слабую генетическую дифференциацию (большое генетическое сходство), которая, возможно, есть следствие

что чем больше GD между животными в исходной матрице, тем дальше они находятся друг от друга на диаграмме. Числа при Coord.1 и Coord.2 указывают на проценты общей дисперсии, которые приходятся на каждую координату. Так, первая главная координата описывала 13,4 % общей дисперсии в структуре данных. Вторая – 10,2 % (третья координата – 6 %, не рассматривается). Суммарный процент дисперсии – это показатель того, насколько полно главные координаты представляют генетическую изменчивость среди быков-производителей (*информативность* РСоА). Отметим, РСоА считается успешным, если первые 2-3 координаты выделяют (учитывают, объясняют) не менее 80 % общей дисперсии. В нашем случае на две первые координаты суммарно приходилось 23,6 % общей дисперсии (около 30 % на три координаты).

достаточно сильного потока генов между породами.

Если маркировать двух джерсейских быков как «выбросы», то разброс остальных животных этой группы указывал на их большую однородность. Гомогенность и отдалённость от двух кластеров, по всей вероятности, есть следствие продолжительной генетической изоляции породы на острове Джерси (с 1763 по 2008 гг., но с возможностью экспорта животных в разные страны [41]).

*Ординация выборок на 2D диаграмме.* На рисунке 6 представлены проекции генетических отношений *семи породных выборок* на 2D плоскости, полученных по разным матрицам парных GD. Первые три проекции – на основе матриц  $F_{ST(W\&C)}$ ,  $\ln F_{ST(W\&C)}$  и  $F'_{ST(W\&C)}$ -дистанций из AMOVA1, четвертая – по  $R_{ST}$ -матрице из AMOVA3 (результаты по AMOVA2 были сходны с таковыми по AMOVA1).

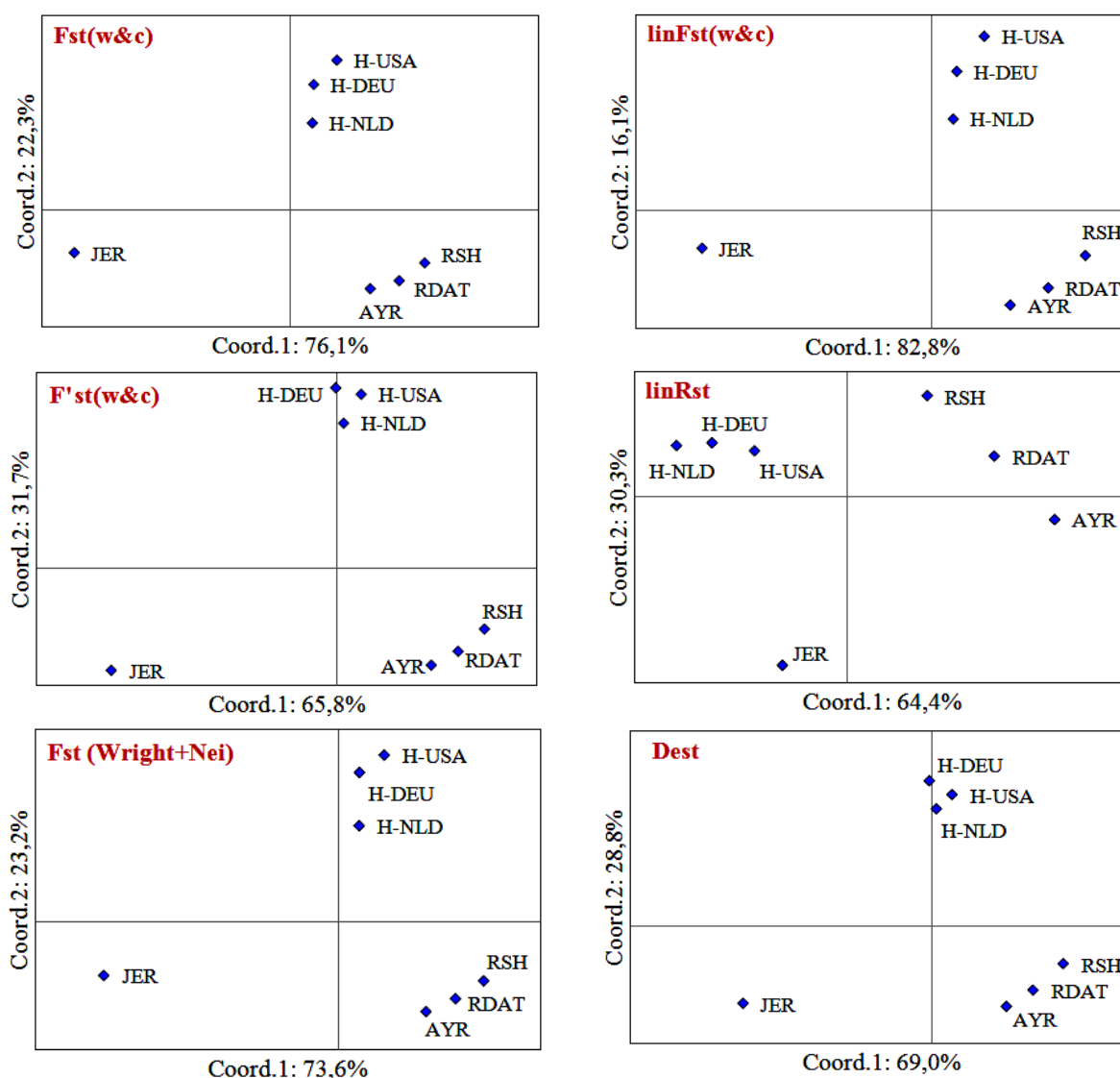


Рис. 6. PCoA по разным матрицам парных генетических дистанций /  
Fig. 6. PCoA for different matrices of paired genetic distances

На первых трёх проекциях местоположение породных выборок относительно друг друга было очень схожим. Чётко выделялись кластер «красных» пород, кластер голштинских «экотипов» и ветка джерсейской породы. Две первые координаты суммарно объясняли 97,5-98,9 % общей дисперсии (максимальная потеря информации 2,5 %).

РсoA-проекция linR<sub>ST</sub>-матрицы (четвёртая) отличалась от предыдущих, но сохранились как кластеры, так и их локализация относительно друг друга. Корреляция Прокруста\* с проекцией linF<sub>ST(w&c)</sub>-матрицы составила 0,846 – даже несколько выше, чем корреляция Мантеля соответствующих матриц ( $R_M = 0,696$ ).

\* Использован *Прокрустов анализ* (Procrustes analysis) – статистический метод, который сравнивает наборы [многомерных] форм, пытаясь преобразовать их в состояние суперналожения. В программе PROTEST это достигается путем минимизации сумм квадратов расстояний между соответствующими точками в каждой форме посредством перемещения, отражения, вращения и масштабирования их координатных матриц [42, 43].

Статистическая значимость корреляции Прокруста была  $p_{perm} = 0,01$ : две проекции показали относительно высокую неслучайную *конкордантность* (concordance), т.е. согласованность, соответствие. Суммарная по двум координатам дисперсия также была высокой – 94,7 % (потеря информации 5,3 %).

На рисунке 6 для сравнения представлены также две диаграммы из работы [22], полученных на том же самом материале, но при использовании иных статистик. В частности, пятая диаграмма – это проекция матрицы парных GD по Райту-Нею ( $F_{ST(Wright-Nei)} = G_{ST}$ ; на базе ожидаемой гетерозиготности).

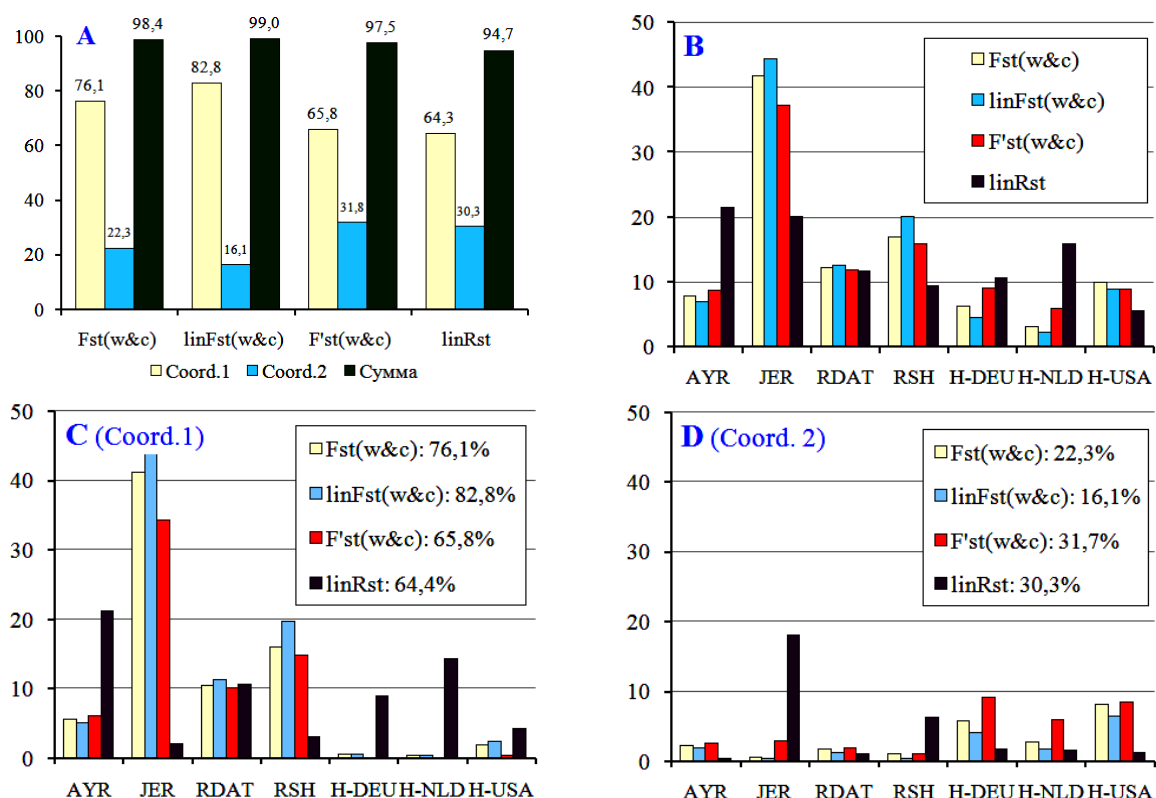


Рис. 7. PCoA-декомпозиция общей дисперсии по разным GD-матрицам (%): А – дисперсии по 1, 2 и (1+2) координатам; В – дисперсия по породам (общая); С – дисперсия по породам для 1-ой координаты; D – дисперсия по породам для 2-ой координаты. Цифры в легенде – % учтённой дисперсии в зависимости от метода /

Fig. 7. PCoA-decomposition of the total variance for different GD matrices (%): A – variance for 1, 2 and (1+2) coordinates; B – variance for breeds (total); C – variance for breeds for the 1st coordinate; D – variance for breeds for the 2nd coordinate. Numbers in the legend – % of the calculated variance, depending on the method

Шестая диаграмма – это проекция матрицы GD по Джосту ( $D_{EST}$ ), в основе расчёта которой вариация эффективного числа аллелей. Как и в первых четырёх диаграммах, информативность (состоятельность) PCoA была высокой (96,8 и 97,7 %). Также имело место хорошее соответствие ординаций с таковыми в трёх первых диаграммах.

Рисунок 7 иллюстрирует информативность PCoA и вклады породных выборок в дисперсию 1 и 2 координат при использовании разных мер GD. Coord.1 объясняла 64-83 % общей дисперсии данных (с max по метрике  $linF_{ST(W\&C)}$ ; рис. 7A), Coord.2 – 16-32 % (с max по метрике  $F'_{ST(W\&C)}$ ). Наибольший вклад в общую дисперсию F-матриц GD вносила JER-выборка – 37-44 % (рис. 7B). Затем следовали выборки RSH – 16-20 % и RDAT – 12-13 %. Вклад AYR-выборки и голштинских «экотипов» был на уровне 2-10 %. По  $linR_{ST}$ -матрице вклады породных выборок были более выравненными: AYR и JER – 21 и 20 %, H-NLD, RDAT и H-DEU – 11-16 %, RSH и H-USA – 6 и 9 %. Вклады породных выборок в дисперсию Coord.1 были, примерно, пропорциональны предыду-

щим (рис. 7C). В дисперсию Coord.2 (рис. 7D) по F-матрицам наибольшие вклады вносили голштинские «экотипы» (2-9 %), в то время как по  $linR_{ST}$ -матрице – JER-выборка (18 %).

**Двухфакторный AMOVA.** Ординация породных выборок по всем матрицам GD была в общем сходной: выделились кластер «красных» пород, кластер голштинских «экотипов» и ветка джерсейской породы. Кластеры (определим их как «регионы») включали по три породных выборки. Это позволило провести двухфакторный AMOVA с разложением STR-изменчивости на межрегиональный и межпородный в пределах регионов компоненты. Результаты обобщены в таблице 7.

Суммарные оценки дифференциации были на уровне полученных ранее величин: до стандартизации 0,108 при  $p_{perm} = 0,001$ , после стандартизации – 0,384 (по 7 породным выборкам  $F'_{ST(W\&C)} = 0,351$ ). Доминировало региональное различие, которое составляло около 90 % от общей дифференциации. В пределах регионов дифференциация породных выборок была незначительной и статистически незначимой ( $p_{perm} = 0,063$ ).

Таблица 7 – Результаты 2-х факторного AMOVA1 по 6 породным выборкам (без JER-выборки) /  
Table 7 – Results of 2-factor AMOVA1 for 6 breed samples (without JER-sample)

Статистика / Statistics	Est.	$p_{perm}$
До стандартизации / Prior to the standardization of		
$F_{RT}$	0,101	0,001
$F_{SR}$	0,008	0,063
$F_{ST} = F_{RT} + F_{SR}$	<b>0,108</b>	0,001
После стандартизации / After standardization		
$F'_{RT}$	0,357	-
$F'_{SR}$	0,027	-
$F'_{ST} = F'_{RT} + F'_{SR}$	<b>0,384</b>	-

Примечания:  $F_{RT}$  – дифференциация регионов;  $F_{SR}$  – дифференциация породных выборок внутри регионов /  
Notes:  $F_{RT}$  – differentiation of regions;  $F_{SR}$  – differentiation of breed samples within regions.

Также были проведены однофакторные анализы по каждому «региону». AMOVA1 по RED-региону (AYR, RDAT, RSH) показал статистически незначимую оценку дифференциации породных выборок ( $F_{ST(W\&C)} = 0,005$  при  $p_{perm} = 0,289$ ;  $F'_{ST(W\&C)} = 0,016$ ). При анализе выборок HOL-региона (H-DEU, H-NLD, H-USA) оценка дифференциации была хотя и статистически значимой, но небольшой ( $F_{ST(W\&C)} = 0,016$  при  $p_{perm} = 0,015$ ;  $F'_{ST(W\&C)} = 0,058$ ).

Чем объясняются полученные результаты? На протяжении более 150 лет селекционеры США целенаправленно работали над молочным типом и молочностью голштинго-фризского скота. В Северной Европе большее внимание уделялось жирномолочности коров и молочно-мясному типу скота. Разные цели селекции могли стать причиной «диверсифицированного отбора» – усиления различий в аллельных профилях европейских и голштинской популяциях. Подтверждение находим в полногеномных исследованиях (SNP-маркеры) американских учёных. Так, в работе [41] было показано, что между голштинской породой США (интенсивная система производства) и джерсейской породой острова Джерси (экстенсивная система производства) GD была на уровне 0,15. Это достаточно высокое значение для SNP-маркеров. Например, между голштинской и якутской породами GD оценивалась в 0,19 [44]. В работе [45] одна группа животных голштинской породы 1964 года представляла образцы неотобранных геномов 1950-х годов (контроль). Другая группа состояла из современных голштинских коров, которая представляла

образцы геномов после 40 лет селекции. Анализ главных компонент выявил временные сдвиги «геномного ландшафта» породы. Было показано, что проводимая с 1964 года практическая селекция затронула около 40 % генома. В частности, привела к изменениям частот аллелей по всему геному и связанные с ними изменения гетерозиготности, частоты гаплотипов и локальные изменения частот, которые создавали «пики», «холмы» и «долины» частотных различий между отселектированной и контрольной группами.

С другой стороны, в Европе имели место процессы поглощения и конфлюэнции (слияния) генофондов разных пород. Во-первых, проводимая с 1970-х годов голштинизация немецкого и голландского чёрно-пёстрого скота привела к формированию относительно гомогенных «экотипов», высококровных по голштинской породе. Во-вторых, обмен генетическим материалом между айрширской, красной датской и красно-пёстрой шведской породами способствовал получению потомства различной кровности, смешиванию генофондов и, как следствие, объединению популяций с последующим разведением скота по единой программе селекции VikingRed (<http://www.vikinggenetics.com/>). Известно, что чем ближе в родословных общице предки разных пород, тем животные в большем родстве друг другу. В результате регулярного обмена генетическим материалом различия аллельных профилей между породами (популяциями) стираются и оценки генетической дифференциации (дистанции) приближаются к нулю. В качестве примера можно привести незначительную генетическую дифференциацию (0,006-0,016) популяций гернзейской породы США, Канады, Великобритании и острова Гернси (Guernsey) вследствие международного интенсивного обмена генетическим материалом [46].

Можно полагать, что рассмотренные процессы и явились причиной (или одной из причин) высокого вклада «региональной» компоненты (90 %) в суммарную STR-дифференциацию исследуемых породных выборок быков-производителей.

Моделирование. При нулевой оценке генетической дифференциации субпопуляций или величине, не имеющей практического значения, эти субпопуляции можно считать частями одной популяции. Относительно сохранения генофонда местных пород, то, возможно, целесообразно не поддерживать

много разных малочисленных пород одного вида, а объединить породы со слабой генетической дифференциацией (высоким генетическим сходством). По двухфакторному AMOVA была получена статистически не значимая оценка STR-дифференциации породных выборок в пределах регионов, т.е. региональные выборки представляли генофонды двух популяций. Финская Ассоциация Племенного Скотоводства (Faba) уже в 2010 году *де-факто* объединила AYR, RDAT и RSH породы, присоединив их к международной племенной организации VikingGenetics. Также голштинская порода считается транснациональной, в которой немецкая и нидерландская субпопуляции являются её составными частями. Фиксация *постфактум* генетических взаимоотношений, сформировавшихся при практическом разведении и селекции рассматриваемых пород, – это сопутствующий результат наших исследований. Анализируемые породные выборки представляли интерес не только как объекты для сравнительных исследований методов оценки STR-дифференциации популяций. Они интересны и как *модельные* популяции для исследования разных сценариев сохранения генофонда исчезающих пород.

Допустим, что имеем не семь реальных породных выборок, а семь гипотетических местных малочисленных пород с полученной выше структурой генетических взаимоотношений. Так как генетическое сходство AYR, RDAT и RSH пород очень высокое, то их можно объединить и сформировать новую RED-породу. Аналогично и по H-DEU, H-NLD и H-USA породам – сформировать HOL-породу. Предположим, что мы решились на такой «породообразовательный процесс». Как он отразится на внутри- и межпородном генетическом разнообразии новой породной структуры (JER, RED, HOL)?

Усреднённое по AYR, RDAT и RSH породам число аллелей на локус ( $n_a$ ) было 5,1, эффективных аллелей ( $n_e$ ) – 3,3, ожидаемая гетерозиготность ( $H_e$ ) – 65,8 %. В «новой»

RED-породе эти показатели составили, соответственно, 6,7, 3,5 и 68,3 %. Усреднённые по H-DEU, H-NLD и H-USA породам показатели были:  $n_a = 5,5$ ,  $n_e = 3,8$ ,  $H_e = 70,2$  %. В «новой» HOL-породе они составили, соответственно, 7,0, 4,1 и 72,4 %. Показатели генетического разнообразия RED и HOL пород ожидаемо возрасли: по  $n_a$  в среднем на 29 %,  $n_e$  – на 7 %,  $H_e$  – на 3,5 % (заметим, объединение увеличило и численность «новых» популяций, что немаловажно).

Сводные оценки генетической дифференциации гипотетических местных пород (семи до объединения) были: по  $F_{ST(W\&C)}$  – 0,108,  $F'_{ST(W\&C)}$  – 0,351,  $D_{EST}$  – 0,292 с 95 % CI 0,244...0,339 ( $D_{EST}$  из работы [22]). После объединения (3 породы) оценки составили, соответственно, 0,140, 0,453 и 0,427 с 95 % CI 0,348...0,493. Объединение пород с высоким генетическим сходством увеличило STR-дифференциацию новой породной структуры, именно: оценки  $F_{ST(W\&C)}$  и  $F'_{ST(W\&C)}$  повысились на 29 %,  $D_{EST}$  – на 46 %. Здесь следует отметить, что статистики  $F_{ST(W\&C)}$  и  $F'_{ST(W\&C)}$  относятся к семейству «мер фиксации», т. к. они фактически измеряют близость аллелей к их фиксации в каждой субпопуляции или степень завершения процесса фиксации (аллель достиг частоты 1), а не абсолютную дифференциацию.  $D_{EST}$ -статистика относится ко второму семейству мер – «аллельной дифференциации», когда дифференциация – это отсутствие какой-либо общей аллели(-ей) в субпопуляциях. Эта статистика равна 1, когда аллелофонд каждой субпопуляции состоит полностью из частных аллелей, и равна 0, когда все субпопуляции имеют одни и те же аллели с одинаковыми частотами [4]. Исходя из 95 % CI по  $D_{EST}$  можно полагать, что общих аллелей в аллелофондах трёх «новых» пород 50-65 % (до объединения их было 66-76 %). Парные по породам GD были (в круглых скобках – усреднённые оценки из анализов по семи породным выборкам, в квадратных скобках – % снижения после объединения):

	$F_{ST(W\&C)}$	$F'_{ST(W\&C)}$	$D_{EST}$
JER – RED	0,234 (0,247) [-5,3]	0,652 (0,661) [-1,4]	0,547 (0,551) [-0,7]
JER – HOL	0,171 (0,187) [-8,6]	0,513 (0,528) [-2,8]	0,413 (0,420) [-1,7]
RED – HOL	0,105 (0,113) [-7,1]	0,366 (0,385) [-4,9]	0,292 (0,307) [-4,9]

Наибольшие STR-дистанции были между JER и RED породами, в 2 раза короче (в среднем) – между породой RED и HOL. В противоположность сводным оценкам диф-

ференциации, имело место небольшое сокращение парных GD: в среднем на 7 % по  $F_{ST(W\&C)}$  оценщику, на 3 % – по  $F'_{ST(W\&C)}$  и на 2,4 % – по  $D_{EST}$ . Между JER и RED породами



STR-дистанции сократились в среднем на 2,5 %, между JER и HOL – на 4,4 %, RED и HOL – на 5,6 %.

Изменился и вклад пород в STR-разнообразие. Так, до объединения вклад JER породы (выборки) по PCoA оценивался в 38,1 %, суммарный вклад AYR, RDAT и RSH пород был 37,5%, а H-DEU, H-NLD и H-USA пород – 24,4 %. После объединения вклад JER-породы составил 43,5 % (+5,4 процентных пункта, п.п.), вклад RED-породы – 36,3 % (-1,2 п.п.) и вклад HOL-породы 20,2 % (-4,2 п.п.).

**Заключение.** Три подхода к дисперсионному анализу микросателлитной изменчивости (AMOVA): по матрице дистанций между аллелями (расчёт  $F_{ST(W\&C)}$ -статистики), по матрице дистанций между генотипами ( $F_{PT}$ ) и матрице различий в размерах аллелей ( $R_{ST}$ ), по всей вероятности, приводят к достаточно близким сводным оценкам генетической дифференциации пород. При этом разные компьютерные программы обеспечивают получение схожих AMOVA-оценок. Оценки по AMOVA сопоставимы с G-оценками Нея (Nei) на базе ожидаемой гетерозиготности. Если число локусов для анализа небольшое, то  $R_{ST}$ -оценки

могут быть смещёнными. Статистики  $F_{ST(W\&C)}$  и  $F_{PT}$  подвержены влиянию уровня внутривидовой гетерозиготности и нуждаются в стандартизации, которая может значительно повысить первоначальные оценки. Стандартизированные оценки хорошо согласуются с оценками, рассчитанными методами Джоста (Jost) и Морисита-Хорна (Morisita-Horn), на которые влияния уровня внутривидовой гетерозиготности нет. Сопряжённость матриц парных по породам генетических дистанций, рассчитанных с использованием разных мер, высокая. Их PCoA-проекция на 2D плоскости сопоставимы: разные подходы и методы приводят к сходным структурам пространственных взаимоотношений между породами. Двухфакторный AMOVA позволяет количественно оценить как генетические различия между регионами с разными методами разведения и/или селекции пород, так и генетические различия между породами в пределах регионов. Объединение [вытесняемых] пород с большим генетическим сходством (малыми генетическими дистанциями) может способствовать увеличению аллельного разнообразия «новых» пород и усилению их дифференциации.

### References

1. Wright S. Isolation by distance. *Genetics*. 1943;28(2):114-138. URL: <https://www.genetics.org/content/28/2/114>
2. Wright S. The genetical structure of populations. *Ann. Eugen.* 1951;15(1):323-354. DOI: <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>
3. 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>
4. Kuznetsov V. M. *F-statistiki Rayta: otsenka i interpretatsiya*. [Wright's F-statistics: estimation and interpretation]. *Problemy biologii produktivnykh zhivotnykh*. 2014;(4):80-104. (In Russ.). URL: <https://elibrary.ru/item.asp?id=22833217>
5. Jost L., Archer F., Flanagan S., Gaggiotti O., Hoban S., Latch E. Differentiation measures for conservation genetics. *Evol. Appl.* 2018;11(7):1139-1148. DOI: <https://doi.org/10.1111/eva.12590>
6. Balloux F., Lugon-Moulin N. The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* 2002;11(2):155-165. DOI: <https://doi.org/10.1046/j.0962-1083.2001.01436.x>
7. Wright S. Evolution and the genetics of populations. Vol. 4. Variability within and among natural populations. Univ. Chicago, 1978. 590 p.
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. Кузнецов В. М. Методы Нея для анализа генетических различий между популяциями. Проблемы биологии продуктивных животных. 2020;(1):91-110. Режим доступа: <https://elibrary.ru/item.asp?id=43811467>
10. Kuznetsov V. M. *Metody Neia dlya analiza geneticheskikh razlichiy mezhdu populyatsiyami*. [Nei's methods for analyzing genetic differences between populations]. *Problemy biologii produktivnykh zhivotnykh* = Problems of Productive Animal Biology. 2020;(1):91-110. (In Russ.). URL: <https://elibrary.ru/item.asp?id=43811467>
11. Hedrick P. W. A standardized genetic differentiation measure. *Evolution*. 2005;59(8):1633-1638. URL: <https://www.jstor.org/stable/3449070>
12. Meirmans P. G., Hedrick P. W. Assessing population structure: FST and related measures. *Mol. Ecol. Res.* 2011;11(1):5-18. DOI: <https://doi.org/10.1111/j.1755-0998.2010.02927.x>
13. Jost L.  $G_{ST}$  and its relatives do not measure differentiation. *Mol. Ecol.* 2008;17(18):4015-4026. DOI: <https://doi.org/10.1111/j.1365-294X.2008.03887.x>

13. Whitlock M. C.  $G_{ST}$  and  $D$  do not replace  $F_{ST}$ . *Mol. Ecol.* 2011;20(6):1083-1091.  
DOI: <https://doi.org/10.1111/j.1365-294X.2010.04996.x>
14. Putman A. I., Carbone I. Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecol. Evol.* 2014;4(22):4399-4428. DOI: <https://doi.org/10.1002/ece3.1305>
15. Cockerham C. C. Variance of gene frequencies. *Evolution.* 1969;23(1):72-84.  
DOI: <https://doi.org/10.1111/j.1558-5646.1969.tb03496.x>
16. Cockerham C. C. Analyses of gene frequencies. *Genetics.* 1973;74(4):679-700.  
URL: <https://www.genetics.org/content/74/4/679>
17. Weir B. S., Cockerham C. C. Estimating  $F$ -statistics for the analysis of population structure. *Evolution.* 1984;38(6):1358-1310. DOI: <https://doi.org/10.2307/2408641>
18. Вейр Б. Анализ генетических данных. М.: Мир, 1995. 400 с.  
Very B. *Analiz geneticheskikh dannykh*. [Genetic data analysis]. Moscow: Mir, 1995. 400 p.
19. Excoffier L., Smouse P. E., Quattro J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics.* 1992;131(2):479-491.  
URL: <https://www.genetics.org/content/131/2/479>
20. Michdakis Y., Excoffier L. A Generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics.* 1996;142(3):1061-1064.  
URL: <https://www.genetics.org/content/genetics/142/3/1061.full.pdf>
21. Slatkin M. A. Measure of population subdivision based on microsatellite allele frequencies. *Genetics.* 1995;139(1):457-462. URL: <https://www.genetics.org/content/139/1/457>
22. Кузнецов В. М. Сравнение методов оценки генетической дифференциации популяций по микросателлитным маркерам. *Аграрная наука Евро-Северо-Востока.* 2020;21(2):169-182.  
DOI: <https://doi.org/10.30766/2072-9081.2020.21.2.169-182>
- Kuznetsov V. M. *Sravnenie metodov otsenki geneticheskoy differentsiatsii popu-lyatsiy po mikrosatellitnym markeram*. [Comparison of methods for evaluating genetic differentiation of populations by microsatellite markers]. *Agrarnaya nauka Evro-Severo-Vostoka* = Agricultural Science Euro-North-East. 2020;21(2):169-182. (In Russ.).  
DOI: <https://doi.org/10.30766/2072-9081.2020.21.2.169-182>
23. Nei M. Genetic distance between populations. *Amer. Natur.* 1972;106(949):283-292.  
DOI: <https://doi.org/10.1086/282771>
24. 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>
25. Peakall R., Smouse P. GenAlEx 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>
26. Peakall R., Smouse P. E. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics.* 2012;28(19):2537-2539.  
DOI: <https://doi.org/10.1093/bioinformatics/bts460>
27. 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>
28. Excoffier L., Laval G., Schneider S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online.* 2005;1:47-50.  
URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2658868/>
29. Raymond M., Rousset F. GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. *J. Hered.* 1995;86(3):248-249. DOI: <https://doi.org/10.1093/oxfordjournals.jhered.a111573>
30. Rousset F. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol. Ecol. Res.* 2008;8(1):103-106. DOI: <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
31. Goodman S. J. RST CALC: A collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Mol. Ecol.* 1997;6(9):881-885.  
DOI: <https://doi.org/10.1046/j.1365-294X.1997.00260.x>
32. Chao A., Ma K. H., Hsieh T. C., Chiu C. H. Online program SpadeR (Species-richness Prediction And Diversity Estimation in R). 2016. 88 p. DOI: <https://doi.org/10.13140/RG.2.2.20744.62722>
33. STATGRAPHICS® Centurion XVI User Manual. By StatPoint Technologies, Inc. 2010. 297 p.
34. Balloux F., Goudet J. Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. *Mol. Ecol.* 2002;11(4):771-783. DOI: <https://doi.org/10.1046/j.1365-294x.2002.01474.x>
35. Excoffier L., Hamilton G. Comment on «Genetic Structure of Human Populations». *Science.* 2003;300(5627):1877. DOI: <https://doi.org/10.1126/science.1083411>
36. 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>
37. 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>

38. Medugorac I., Veit-Kensch C. E., Ramljak J., Brka M., Marković B., Stojanović S., Bytyqi H., Kochoski L., Kume K., Grünenfelder H.-P., Bennewitz J., Förster M. Conservation priorities of genetic diversity in domesticated metapopulations: a study in taurine cattle breeds. *Ecol. Evol.* 2011;1(3):408-420. DOI: <https://doi.org/10.1002/ece3.39>
39. Abdelmanova A. S., Kharzinova V. R., Volkova V. V., Mishina A. I., Dotsev A. V., Sermiyagin A. A., Boronetskaya O. I., Petrikeeva L. V., Chinarov R. Yu, Brem G., Zinovieva N. A. Genetic Diversity of Historical and Modern Populations of Russian Cattle Breeds Revealed by Microsatellite Analysis. *Genes.* 2020;11(940):1-15. DOI: <https://doi.org/10.3390/genes11080940>
40. Харзинова В. Р., Зиновьева Н. А. Паттерн генетического разнообразия у локальных и коммерческих пород свиней на основе анализа микросателлитов. *Вавиловский журнал генетики и селекции.* 2020;24(7):747-754. DOI: <https://doi.org/10.18699/VJ20.669>
- Kharzinova V. R., Zinov'eva N. A. *Pattern geneticheskogo raznoobraziya u lokal'nykh i kommercheskikh porod svinей na osnove analiza mikrosatellitov.* [The pattern of genetic diversity of different breeds of pigs based on microsatellite analysis]. *Vavilovskiy zhurnal genetiki i seleksii* = Vavilov Journal of Genetics and Breeding. 2020;24(7):747-754. (In Russ.). DOI: <https://doi.org/10.18699/VJ20.669>
41. Huson H. J., Sonstegard T. S., Godfrey J., Hambrook D., Wolfe C., Wiggans G., Blackburn H., VanTassell C. P. A Genetic Investigation of Island Jersey Cattle, the Foundation of the Jersey Breed: Comparing Population Structure and Selection to Guernsey, Holstein, and United States Jersey Cattle. *Front. Genet.* 2020;11(366):1-17. DOI: <https://doi.org/10.3389/fgene.2020.00366>
42. Peres-Neto P. R., Jackson D. A. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia.* 2001;129(2):169-178. DOI: <https://doi.org/10.1007/s004420100720>
43. Dray S., Chessel D., Thioulouse J. Procrustean co-inertia analysis for the linking of multivariate datasets. *Écoscience.* 2003;10(1):110-119. DOI: <https://doi.org/10.1080/11956860.2003.11682757>
44. Sermiyagin 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., Plemiyashov 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>
45. Ma L., Sonstegard T. S., Cole J. B., Van Tassell C. P., Wiggans G. R., Crooker B. A., Tan C., Praka-penka D., Liu G., Da Y. Genome changes due to artificial selection in U.S. Holstein cattle. *BMC Genomics.* 2019;20(128):1-14. DOI: <https://doi.org/10.1186/s12864-019-5459-x>
46. Cooper T. A., Eaglen S. A. E., Wiggans G. R., Jenko J., Huson H. J., Morrice D. R., Bichard M., de L. Luff W. G., Woolliams J. A. Genomic evaluation, breed identification, and population structure of Guernsey cattle in North America, Great Britain, and the Isle of Guernsey. *J. Dairy Sci.* 2016;99(7):5508-5515. DOI: <http://dx.doi.org/10.3168/jds.2015-10445>

#### 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](mailto:priemnaya@fanc-sv.ru), ORCID: <https://orcid.org/0000-0002-2219-805X>, e-mail: [vm-kuznetsov@mail.ru](mailto: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](mailto:priemnaya@fanc-sv.ru), ORCID: <https://orcid.org/0000-0002-2219-805X>, e-mail: [vm-kuznetsov@mail.ru](mailto:vm-kuznetsov@mail.ru)

✉ – Для контактов / Corresponding author