

ОСОБЕННОСТИ КЛАСТЕРНОГО АНАЛИЗА НА ОСНОВЕ ЧАСТОТ ЭРИТРОЦИТАРНЫХ АНТИГЕНОВ ГРУПП КРОВИ СВИНЕЙ

¹К.С. Шатохин, ²С.В. Никитин, ¹Н.Н. Кочнев

¹Сибирский государственный университет инженерии и биотехнологий, Новосибирск, Россия

²Институт цитологии и генетики СО РАН, Новосибирск, Россия

E-mail: true_genetic@mail.ru

Для цитирования: Шатохин К.С., Никитин С.В., Кочнев Н.Н. Особенности кластерного анализа на основе частот эритроцитарных антигенов групп крови свиней // Вестник Университета биотехнологий. – 2026. – Т. 21, № 1. – С. 193–206. – DOI: 10.31677/3033-8433-2026-21-1-193-206.

Реферат. Работа посвящена оценке эффективности различных методов расчета дистанций на основе частот аллелей эритроцитарных антигенов групп крови для филогенетического кластерного анализа популяций свиней. Была поставлена цель сравнить точность и адекватность различных метрик расстояний при восстановлении известных заранее родственных связей между группами домашних и диких свиней. В анализе задействованы дистанции Эвклида, Манхэттена, Канберра, Чебышева, а также специализированные генетические дистанции Нея, Эдвардса, Роджерса, Превости и Рейнольдса. В качестве объектов исследования использовались 12 популяций, включая мини-свиней Института цитологии и генетики СО РАН, крупную белую породу, ландрас, кемеровскую породу, а также различные подвиды дикого кабана. Для валидации результатов использовалась внешняя оценка на основе известной генеалогической истории популяций. Точность методов оценивалась с помощью корреляции Спирмена между дистанциями и рангами родства. Результаты показали, что большинство протестированных дистанций демонстрируют слабую или умеренную корреляцию с реальными филогенетическими связями. Дендрограммы, построенные на основе различных метрик, часто группировали неродственные популяции вместе и не отражали ожидаемой кластеризации по происхождению. Была выдвинута гипотеза, что это связано с относительно низким уровнем полиморфизма эритроцитарных антигенов групп крови по сравнению с такими современными маркерами, как микросателлиты или SNP. Сделан вывод о непригодности частот аллелей эритроцитарных антигенов групп крови для точного восстановления филогенетических отношений между популяциями свиней.

Ключевые слова: мини-свиньи, филогенетический анализ, эритроцитарные антигены, группы крови, генетические дистанции, дендрограммы, сельское хозяйство.

FEATURES OF CLUSTER ANALYSIS BASED ON THE FREQUENCIES OF ERYTHROCYTE ANTIGENS OF PIG BLOOD GROUPS

¹K.S. Shatokhin, ²S.V. Nikitin, ¹N.N. Kochnev

¹Siberian State University of Engineering and Biotechnology, Novosibirsk, Russia

²Institute of Cytology and Genetics, Novosibirsk, Russia

E-mail: true_genetic@mail.ru

Abstract. This work is devoted to evaluating the effectiveness of various methods for calculating distances based on the frequencies of alleles of erythrocyte antigens of blood groups for phylogenetic cluster analysis of pig populations. The aim was to compare the accuracy and adequacy of various distance metrics when restoring previously known kinship relationships between groups of domestic and wild pigs. The analysis involved the distances of Euclid, Manhattan, Canberra, Chebyshev, as well as the specialized genetic distances of Ney, Edwards, Rogers, Prevost, and Reynolds. 12 populations were used as research objects, including mini pigs of ICiG SB RAS, large white breed, Landrace, Kemerovo breed, as well as various subspecies of wild boar. An external assessment based on the known genealogical history of the populations was used to validate the results. The accuracy of the methods was assessed using Spearman's correlation between distance and kinship ranks. The results showed that most of the tested distances show weak or moderate correlation with real phylogenetic relationships. Dendrograms based on various metrics often grouped unrelated populations together and did not reflect the expected clustering by origin. It has been hypothesized that this is due to the relatively low level of polymorphism of erythrocyte antigens of blood groups compared with such modern markers as microsatellites or SNPs. It is concluded that the frequencies of alleles of erythrocyte antigens of blood groups are unsuitable for accurate reconstruction of phylogenetic relationships between pig populations.

Keywords: mini-pigs, phylogenetic analysis, erythrocyte antigens, blood groups, genetic distances, dendrograms, agriculture.

Филогенетический анализ является достаточно обширным разделом, объединяющим в себе методы биологии и статистики. Его сущность заключается в реконструкции эволюционной истории биологических видов, их популяций или отдельных генов путем их объединения в иерархические кластеры – филогенетические деревья, отражающие схему их родственных связей и последовательность дивергенции от общих предков [1, 2]. Классический подход в рамках данного анализа предполагает количественную оценку степени сходства или различия между изучаемыми популяциями. Данные о сходстве трансформируются в меру эволюционного расхождения, так называемые генетические дистанции, которые количественно отражают степень накопленных различий [3, 4]. Серьезной проблемой, способной существенно исказить достоверность филогенетических реконструкций, являются естественные популяционно-генетические процессы. К ним относятся эффект основателя, генетический дрейф, давление отбора или миграция особей. Эти процессы могут независимо исказить распределение частот аллелей и генотипов в популяциях, создавая картину эволюционного родства там, где его нет (явление гомоплазии), или, наоборот, маскируя истинное родство [5]. Для минимизации подобных искажений и повышения точности анализа рекомендуется комплексный подход, в частности, учет возможности интродукций [6].

Первичным материалом для филогенетического кластерного анализа обычно выступают частоты гаплотипов Y-хромосомы [7], митохондриальной ДНК [8], SNP либо SRT [9–11]. В начале XXI в. для оценки родства между популяциями пользовались частотами эндогенных ретровирусов. Во второй половине прошлого века популярным был анализ генетических дистанций, базирующихся на частотах аллотипов сывороточных белков, а также аллелей и антигенов групп крови [12]. Несмотря на архаичность, группы крови до недавнего времени использовались в качестве первичных данных для кластерного анализа [13].

При выборе метода кластеризации для филогенетического анализа существует одна проблема, состоящая в огромном количестве методов кластерного анализа [14, 15]. Среди них выделяют как специализированные методы оценки родства между популяциями, основанные на аллельных частотах [16], так и общие методы расчета дистанций, такие как Евклидово, расстояние Канберра, расстояние Минковского, расстояние Чебышёва [13]. В связи с этим возникает вопрос о том, каким именно методом лучше всего оценивать генетические дистанции при необходимости проведения филогенетического анализа.

Цель исследования состояла в сравнении точности различных типов дистанций на примере нескольких групп домашних и диких свиней, чье родство было заранее известно по результатам скрещиваний.

ОБЪЕКТЫ И МЕТОДЫ ИССЛЕДОВАНИЙ

В качестве объекта исследований был выбран ряд популяций, отвечающих нескольким требованиям. Во-первых, популяции должны быть генеалогически преемственными, а их родственные связи должны быть заранее известны. Во-вторых, каждая популяция должна была быть протестирована на полиморфизм каких-либо маркеров. Такими популяциями оказались: мини-свиньи Института цитологии и генетики (ИЦИГ СО РАН), светлогорские мини-свиньи, минисибс, крупная белая порода (1992 г.), ландрас, а также средн азиатский и европейский подвиды дикого кабана. Для контраста в исследование включили три группы свиней, чьи родственные связи с указанными не задокументированы. Это выборки заводских пород, кемеровской и крупной белой 2016 г., а также сибирский подвид дикого кабана. Каждая из представленных совокупностей тестировалась на полиморфизм аллелей эритроцитарных антигенов групп крови (табл. 1).

Таблица 1

Перечень используемых в анализе выборок
List of samples used in the analysis

Популяция	Сокращенное название	Место взятия образцов биологического материала	Источник	n
Мини-свиньи ИЦиГ СО РАН, 2013	MSI_2013	Питомник мини-свиней ИЦиГ СО РАН, Новосибирская обл.	[17]	120
Мини-свиньи ИЦиГ СО РАН, 2001	MSI_2001		[18]	62
Минисибс	MSB	Экспериментальное хозяйство СО РАН, Новосибирская обл.	[19]	167
Светлогорские мини-свиньи	SG	Научный центр биомедицинских технологий РАМН, пос. Светлые Горы, Московская обл.	[20]	62
Крупная белая, 1992	LW_1992	Новосибирская обл.	[21]	383
Крупная белая, 2016	LW_2016	Новосибирская обл.	[22]	65
Ландрас	LND	Экспериментальное хозяйство СО РАН, Новосибирская обл.	[19]	905
Кемеровская	KEM	Кемеровская обл.	[12]	403
Сибирский кабан	SWB	Респ. Тыва	[23]	33
Среднеазиатский кабан	MAWB	Кыргызская Республика	[12]	83
Европейский кабан	EWB	Респ. Беларусь, Воронежская обл.	[12]	370
Вьетнамская масковая Й	VY	Популяция Рижского зоопарка	[19]	60

Все перечисленные популяции свиней (табл. 2) были тестированы иммунно-специфическими антисыворотками, прошедшими международные испытания. Большинство из них было получено ИЦиГ СО РАН согласно описанным методам [24]. Непосредственную паспортизацию животных по эритроцитарным антигенам групп крови проводили согласно рекомендациям [24].

Исследование проводили по следующей схеме. Вначале на основании частот аллелей групп крови (см. табл. 2) были построены дистанции, которые визуализировали в виде дендрограмм. В ходе исследования использовались дистанции, рассчитываемые базовой функцией *dist*: Манхэттена, Минковского, Евклида, Канберра, Чебышева («базовые дистанции»). Также в работе использовались генетические дистанции, вычисляемые популярной библиотекой *poppr*: Нея, Превости, Эдвардса, Роджерса и Рейнольдса [16]. Кластеризация осуществлялась при помощи функции *pvclust* с 1000 бутстрэп-повторами, что минимизировало ошибки, вызванные малочисленностью выборок и небольшого количества локусов [25, 26]. В качестве способа кластеризации был выбран метод Уорда. Статистика Хопкинса (0.5197) показывает пригодность исходных данных (см. табл. 2) для выполнения кластерного анализа [27]. Оценку правдоподобности кластерного анализа проводили методом внешней валидации [28]. В качестве априори известных расстояний использовались ранги, построенные на основании генеалогического родства (рис. 1). Точность дистанций оценивали при помощи корреляции с рангами родословной по методу Спирмена. Исходя из числа исследуемых популяций, количество вариантов в корреляционном анализе ($n = 11$) может вызвать некоторые вопросы о допуске по минимальному объему выборки, требования к которому существенно разнятся [29, 30]. Поэтому для нивелирования малочисленности было использовано расширение выборки с помощью бутстрэпа (1000), для чего использовали функцию *resample* из пакета *mosaic* [31]. Анализ данных осуществляли в среде программирования R.

Таблица 2

Исходные частоты аллелей групп крови
Initial frequencies of blood group alleles

Популяция	Acp	A-	Ba	Bb	Da	Db	Ebdg	Eaeg	Eedg	Eedf	Fa	Fb	Ga	Gb
Мини-свины ИЦиГ СО РАН, 2013	0,933	0,067	0,883	0,118	0,363	0,638	0,083	0,354	0,550	0,013	0,304	0,696	0,475	0,525
Мини-свины ИЦиГ СО РАН, 2001	0,905	0,095	1,000	0,000	0,295	0,705	0,068	0,416	0,516	0,000	0,458	0,542	0,332	0,668
Минисибс	0,500	0,500	1,000	0,000	0,696	0,304	0,134	0,560	0,242	0,061	0,045	0,955	0,089	0,911
Светлогорские	0,380	0,620	1,000	0,000	0,020	0,980	0,100	0,170	0,560	0,170	0,770	0,230	0,500	0,500
Крупная белая, 1992	0,172	0,828	1,000	0,000	0,038	0,962	0,329	0,307	0,230	0,133	0,000	1,000	0,398	0,602
Крупная белая, 2016	0,510	0,490	1,000	0,000	0,262	0,738	0,030	0,331	0,431	0,208	0,007	0,993	0,285	0,715
Ландрас	0,295	0,705	1,000	0,000	0,000	1,000	0,328	0,129	0,322	0,218	0,156	0,844	0,290	0,710
Кемеровская	0,120	0,880	0,990	0,010	0,000	1,000	0,280	0,180	0,430	0,112	0,403	0,597	0,715	0,285
Сибирский кабан	0,740	0,260	1,000	0,000	0,923	0,077	0,000	0,154	0,846	0,000	0,000	1,000	0,000	1,000
Среднеазиатский кабан	0,460	0,540	1,000	0,000	0,200	0,800	0,000	0,000	1,000	0,000	0,000	1,000	0,301	0,699
Европейский кабан	0,849	0,151	1,000	0,000	0,011	0,989	0,000	0,299	0,729	0,000	0,000	1,000	0,981	0,019
Вьетнамская масковая Й	0,143	0,857	1,000	0,000	0,725	0,276	0,031	0,633	0,337	0,000	0,770	0,230	0,174	0,827

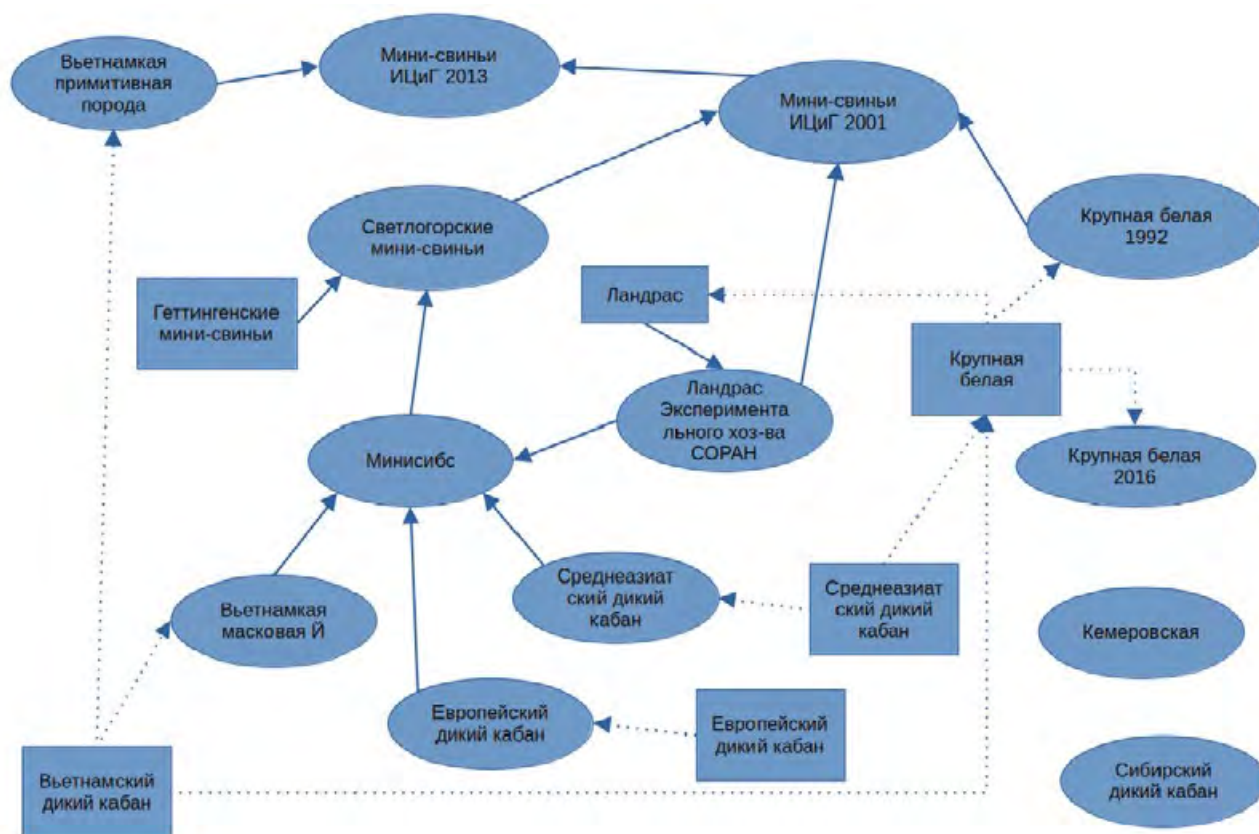


Рис. 1. Схема филогенетических отношений между изученными группами: в прямоугольниках обозначены большие массивы животных, относящиеся целиком к породе/виду; в эллипсах — конкретные выборки животных, задействованные в анализе; сплошными стрелками указаны непосредственные интродукции, пунктиром — опосредованная интродукция, включающая много поколений животных и возможное наличие переходных форм

A diagram of the phylogenetic relationships between the studied groups: rectangles represent large sets of animals belonging entirely to a breed/species; ellipses represent specific samples of animals used in the analysis; solid arrows indicate direct introductions; dotted lines indicate indirect introductions, including many generations of animals and the possible presence of transitional forms.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

По результатам частот аллелей эритроцитарных антигенов групп крови табл. 2 были построены различные дистанции между изученными выборками табл. 3. Далее на основе филогенетических отношений определили ранги удаленности, используемые в качестве дистанций по родословным (см. рис. 1). Ранги удаленности определяли по отношению к популяции мини-свиней ИЦиГ СО РАН 2013 г. Наименее удаленной дистанции присваивали «1», наиболее удаленной — «11». Если две и более выборки были генеалогически равноудалены от мини-свиней ИЦиГ СО РАН 2013, ранги присваивали в соответствии с принципами ранжирования данных [32].

Наиболее близкой популяцией по отношению к мини-свиньям 2013 г. само собой является типированная в 2001 г. селекционная группа мини-свиней ИЦиГ СО РАН. Каждая из используемых дистанций относительно мини-свиней 2013 г. показывает наименьшее расстояние с выборкой 2001 г. Ранг «2» следовало бы присвоить вьетнамской примитивной породе, «прилитие крови» которых было в 2007 и 2010 гг. Их вклад в генофонд стада составил около 20 % [17]. Однако это не было учтено, так как представители породы не имеют результатов иммуногенетической паспортизации. Поэтому ранг «2» был присвоен ландрасам, обогатившим генофонд стада в 1995 г. Третий и четвертый ранги были поделены между родоначальниками мини-свиней ИЦиГ СО РАН: светлогорскими хряками и свиноматками крупной белой породы (1992). Далее ранги распределились между минисибс и их предковыми формами: вьетнамской масковой породой Й, среднеазиатским и европейским подвидами дикого кабана. Предпоследние ранги получили домашние свиньи, напрямую не участвовавшие в создании мини-свиней ИЦиГ СО РАН. Наиболее удаленный ранг получил сибирский подвид дикого кабана (см. табл. 3).

Таблица 3

Матрица дистанций популяций домашних и диких свиней по отношению к мини-свиньям ИЦиГ СО РАН, 2013 г.

Matrix of distances of populations of domestic and wild pigs in relation to minipigs ICG SB RAS 2013

Популяция	По родословным	Евклидова	Чебышева	Манхеттена	Канберра	Минковского	Нея	Эдвардса	Превости	Роджерса	Рейнольдса
Вьетнамская масковая Й	7	1,51	0,79	4,63	7,99	1,51	0,34	0,82	0,29	0,20	0,29
Европейский кабан	7	1,01	0,51	3,06	8,38	1,01	0,11	0,82	0,19	0,14	0,21
Крупная белая, 1992	3,5	1,33	0,76	3,90	7,84	1,33	0,24	0,82	0,24	0,17	0,26
Крупная белая, 2016	9,5	0,84	0,42	2,65	6,17	0,84	0,10	0,81	0,17	0,12	0,17
Кемеровская	9,5	1,35	0,81	3,83	7,15	1,35	0,26	0,82	0,24	0,17	0,27
Ландрас	2	1,19	0,64	3,81	8,16	1,19	0,20	0,82	0,24	0,17	0,24
Мини-свиньи ИЦиГ СО РАН 2001	1	0,36	0,15	1,14	3,70	0,36	0,02	0,79	0,07	0,05	0,07
Минисибс	5	1,09	0,43	3,67	7,63	1,09	0,16	0,81	0,23	0,16	0,22
Светлогорские	3,5	1,17	0,55	3,38	6,51	1,17	0,19	0,81	0,21	0,15	0,23
Сибирский кабан	11	1,23	0,56	3,89	9,24	1,23	0,17	0,82	0,24	0,17	0,25
Среднеазиатский кабан	7	1,05	0,47	3,36	8,48	1,05	0,13	0,82	0,21	0,15	0,21

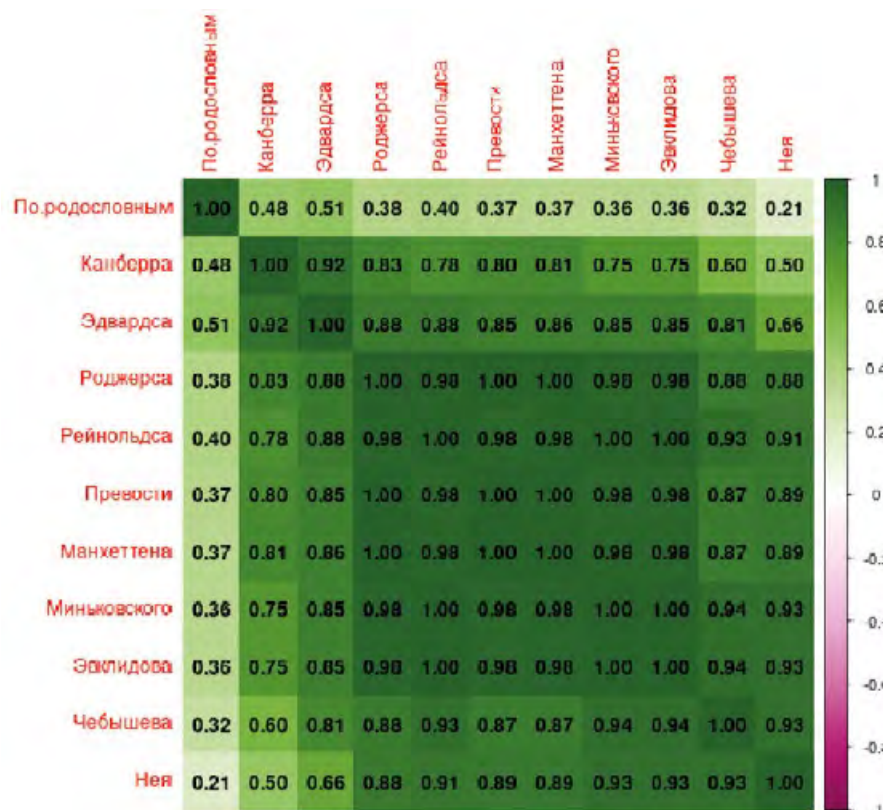


Рис. 2. Коэффициенты корреляции Спирмена между дистанциями и рангами расстояния между популяциями по родословным, $P < 0,001$

Spearman correlation coefficients between distances and distance ranks between populations by pedigrees, $P < 0.001$

Table 1



0

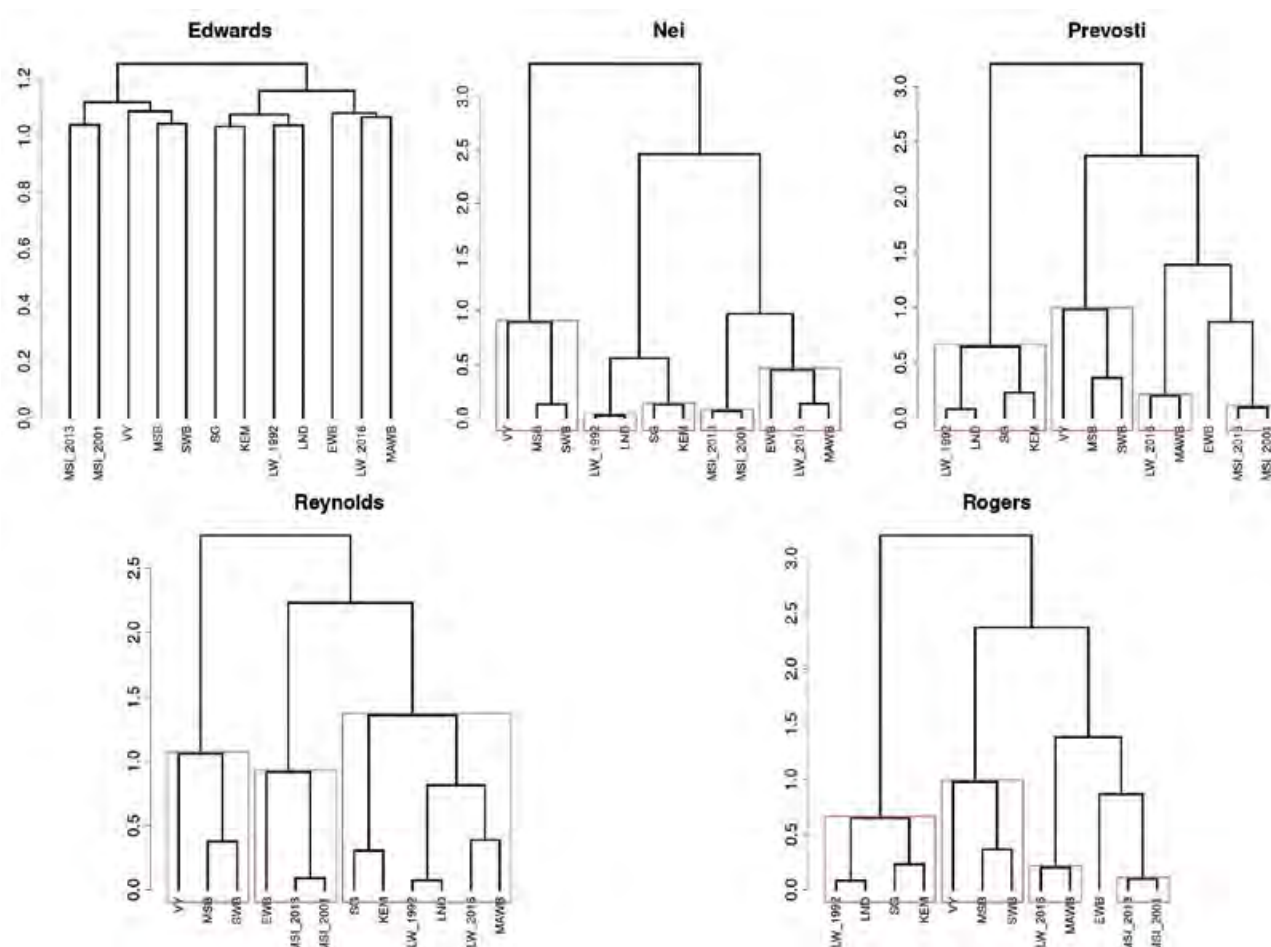


Рис. 4. Дендрограммы на основе генетических дистанций, кластеризованных методом Уорда
Dendrograms based on genetic distances clustered by Ward's method

Выборка крупной белой породы 2016 г. вообще не имела четкого местоположения на дендрограммах. Причем разные дистанции «ставили» ее по соседству с совершенно неродственными группами: минисибс (дистанции Канберра) и среднеазиатским диким кабаном (Евклидовы, Манхэттена, Минковского и все генетические дистанции). Наиболее логичная крупная белая порода 2016-го локализация наблюдалась на дендрограмме, построенной на основе дистанций Чебышева, где она располагалась по соседству с кластером крупной белой породы 1992 г. и ландрасами (см. рис. 3–4). Для удобства состав соответствующих кластеров сгруппирован в табл. 4.

Таблица 4

Кластеризация популяций свиней
Clustering of pig populations

Кластер	Дистанция					
	Канберра	Евклидовы	Минковского	Превости	Роджерса	Рейнольдса
1	2	3	4	5	6	7
1	—	EWB, MS_2001, MS_2013	EWB, MS_2001, MS_2013	—	—	EWB, MS_2001, MS_2013
2	—	SG, KEM, LW_1992, LND, LW_2016, MAWB	SG, KEM, LW_1992, LND, LW_2016, MAWB	LW_1992, LND, SG, KEM	LW_1992, LND, SG, KEM	SG, KEM, LW_1992, LND, LW_2016, MAWB

1	2	3	4	5	6	7
3	—	VY, MSB,SWB	VY, MSB,SWB	VY, MSB,SWB	VY, MSB, SWB	VY, MSB,SWB
4	LW_1992, LND, MSB, SG, KEM	—	—	—	—	—
5	—	—	—	EWB, MS_2013, MS_2001, MAWB	EWB, MS_2013, MS_2001, MAWB	—
6	EWB, MAWB, SWB	—	—	—	—	—
7	MS_2013, MS_2001, VY	—	—	—	—	—

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Достаточно логично выглядит дендрограмма, построенная по дистанциям Канберра (см. рис. 3). В частности, она оказалась единственным типом дистанций, группирующим диких кабанов отдельно от домашних свиней. Однако и в этом случае наблюдаются некоторые несоответствия реальным филогенетическим отношениям (см. рис. 1). В первом кластере вместе с мини-свиньями ИЦиГ СО РАН оказались также вьетнамская масковая Й, минисибс и выборка крупной белой породы 2016 г. В совместной группировке мини-свиней ИЦиГ СО РАН, минисибс и вьетнамской масковой Й нет ничего необычного, так как эти популяции являются генеалогически преемственными (рис. 1). Присутствие в этом же кластере выборки крупной белой породы 2016 г. нелогично, так как прямых интродукций ни из одной популяции из совместного кластера не было (см. рис. 3). Во втором кластере вместе оказались крупная белая порода 1992, ландрас, светлогорские мини-свиньи и кемеровская порода. Группировка заводских пород в один кластер логична, так как между этими породами существует непосредственная филогенетическая преемственность. Порода ландрас в свое время произошла от скрещивания крупной белой породы и местных датских свиней [34]. Свиньи крупной белой породы участвовали в создании кемеровской породы [35]. Ландрасы являются одной из предковых форм минисибс, которые, в свою очередь, дали начало светлогорским мини-свиньям [19, 20].

При построении кластеров на базе дистанций Рейнольдса, Минковского и Евклида четко выделяются три кластера, которых следовало бы ожидать исходя из распределения частот аллелей эритроцитарных антигенов групп крови. Условно, данная особенность распространяется и на дендрограммы по дистанциям Канберра, Роджерса и Превости. В случае с дистанциями Канберра выделяется два кластера, популяции диких кабанов расположились отдельно, по соседству друг с другом, не образуя единого кластера. На дендрограммах по дистанциям Роджерса и Превости наблюдается два больших кластера, а оставшиеся группы разбиты на два отдельных небольших кластера (см. рис. 3–4).

Несмотря на то, что на дендрограммах на базе дистанций Рейнольдса, Минковского, Евклида, Канберра, Роджерса и Превости количество кластеров было принято равным трем, их состав существенно отличался. Поэтому фактически количество кластеров было равным семи (см. табл. 4). Так как родство в данном случае не было группирующим фактором (см. рис. 2), в качестве объяснения принципа кластеризации, было принято распределение аллелей (см. табл. 2). Первый кластер составили мини-свиньи ИЦиГ СО РАН и европейский подвид дикого кабана (см. табл. 4). Филогенетическая связь между ними достаточно опосредованная (европейский дикий кабан – минисибс – светлогорские мини-свиньи – мини-свиньи ИЦиГ СО РАН), но объединяет их высокая в сравнении с другими группами частота аллеля *Asp*, исходного для гена *EAA* [12].

Согласно дистанциям Превости и Роджерса, второй кластер составляют крупная белая 1992, ландрас, кемеровская и светлогорские мини-свиньи (см. табл. 4). Дистанции Евклида, Минковского и Рейнольдса «добавляют» еще среднеазиатского кабана и крупную белую 2016 г. Объединяющим фак-

тором может быть высокая частота аллеля *Db* при встречаемости аллеля *A* – более 50 % (см. табл. 2). Третий кластер образован минисибс, одной из их предковых форм, вьетнамской масковой породой Й и неродственного им дикого кабана. Наиболее вероятным группирующим фактором является высокая частота «азиатских» аллелей *Da* и *Gb*. Четвертый кластер представляется нелогичным, так как не содержит явного группирующего признака. Образован дистанциями Канберра и включает в себя две родственные группы. Первая – заводские породы, ландрас, кемеровская и крупная белая 1992. Вторая – минисибс и созданные на их основе светлогорские мини-свиньи. Пятый кластер также не имеет явной причины для объединения, включает в себя два подвида дикого кабана (европейский и среднеазиатский) и две выборки мини-свиней ИЦиГ СО РАН. Шестой кластер собран дистанциями Канберра по принципу родства. Здесь оказались дикие кабаны, которых объединяет высокая в сравнении с домашними свиньями частота корневого аллеля *Eedg*. Седьмой кластер также образован дистанциями Канберра. Может показаться, что кластер был собран по причине скрещивания мини-свиней ИЦиГ СО РАН в 2007 и 2010 гг. с хряками примитивной вьетнамской породы [17]. Однако схожих частот аллелей между мини-свиньями ИЦиГ СО РАН и вьетнамской масковой породой Й нет, а родство между Й и примитивной вьетнамской породой остается невыясненным. Более того, у мини-свиней ИЦиГ СО РАН обнаружен не свойственный большинству популяций свиней аллель *Bb*. Так что принцип образования седьмого кластера непонятен.

Функция *pvclust* не выявила кластеров, образованных на основании дистанций Эдвардса (см. рис. 4). Однако визуально выборки делятся на две большие группы, каждая из которых, подразделяется на еще две малые группы. Мини-свиньи ИЦиГ СО РАН образуют первую малую группу, которая соседствует с группой из светлогорских мини-свиней, минисибс и сибирским диким кабаном. Можно предположить, что группирующим фактором для этих групп стали «азиатские» аллели эритроцитарных антигенов групп крови *Da* и *Gb*, что, по-видимому, стало причиной присутствия в названной группе сибирского дикого кабана, который состоит в прямом родстве с остальными выборками. Вторая большая группа (состоящая из третьей и четвертой малых) не менее любопытна. В третью малую группу вошли светлогорские мини-свиньи, выборка крупной белой породы 1992 г., ландрасы и свиньи кемеровской породы. Четвертая группа образована европейскими и среднеазиатскими кабанами с «лишней» выборкой крупной белой породы 2016 г. (см. рис. 4).

Использование дистанций, рассчитанных на основе частот аллелей эритроцитарных антигенов групп крови, позволяет весьма условно сгруппировать по принадлежности к центрам происхождения этих самых аллелей, но не показать более тонкие филогенетические отношения. В случае с дистанциями Канберра кластеры получились довольно логичными: дикие кабаны (кластер № 6), домашние свиньи азиатского корня (кластер № 7) и домашние свиньи европейского корня (кластер № 4). Другие дистанции выделили не менее интересные кластеры: азиатский (кластер № 3), европейский архаичный (кластер № 1) и европейский продвинутый (кластер № 2). Дифференциации на домашних и диких свиней при этом не наблюдается (см. табл. 4). Одной из причин серьезных различий классификации популяции свиней на основе дистанций от реальных филогенетических отношений видится недостаточно высокий полиморфизм локусов эритроцитарных антигенов групп крови. Судя по литературным данным, использование более высокополиморфных маркеров, например множественных SNP и микросателлитов, позволяет избежать подобных ошибок [36–38]. Таким образом, эритроцитарные антигены групп крови являются недостаточно информативным исходным материалом для филогенетического анализа свиней.

ВЫВОДЫ

1. Из всех изученных методов определения дистанций на основе частот аллелей эритроцитарных антигенов групп крови максимальное правдоподобие относительно родословных показали дистанции Канберра и Эдвардса.

2. Использование частот аллелей эритроцитарных антигенов групп крови домашних свиней в качестве первичных данных для кластерного анализа практически не показывает истинных филогенетических отношений между популяциями. Основной причиной представляется недостаточно высокий полиморфизм.

Работа поддержана бюджетным проектом № FWNR-2026-1002.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Genetic Diversity and Population Structure of Six Autochthonous Pig Breeds from Croatia, Serbia, and Slovenia* / M. Zorc, D. Škorput, K. Gvozdanović [et al.] // *Genet Sel Evol.* – 2022. – Vol. 54, N 1. – P. 30. – DOI: 10.1186/s12711-022-00718-6.
2. *Genome-Wide Analysis of Genetic Diversity and Artificial Selection in Large White Pigs in Russia* / S. Bakoev, L. Getmantseva, O. Kostyunina [et al.] // *PeerJ.* – 2021. – Vol. 9. – P. e11595. – DOI: 10.7717/peerj.11595.
3. *Molecular phylogenetic analysis among species of Paridae, Remizidae and Aegithalos based on mtDNA sequences of COI and cyt b* / C. Dai, K. Chen, R. Zhang [et al.] // *Avian Research.* – 2023. – Vol. 1, N 2. – P. 112–123.
4. *McKenzie P.F. Linking Phylogenetic Models to Population Processes, from Species Trees to Genomes.* – Columbia University, 2023.
5. *Kapli P., Flouri T., Telford M.J. Systematic Errors in Phylogenetic Trees* // *Curr Biol.* – 2021. – Vol. 31, N 2. – P. R59–R64. – DOI: 10.1016/j.cub.2020.11.043.
6. *Komarova V.A., Lavrenchenko L.A. Approaches to the Detection of Hybridization Events and Genetic Introgression upon Phylogenetic Incongruence* // *Biology Bulletin Reviews.* – 2022. – Vol. 12, № 3. – P. 240–253. – DOI: 10.1134/S2079086422030045.
7. *Phylogeography and Demographic History of Macaques, Fascicularis Species Group, in East Asia: Inferred from Multiple Genomic Markers* / Y. Zhou, J. Tian, H. Jiang [et al.] // *Mol Phylogenet Evol.* – 2024. – Vol. 194. – P. 108042. – DOI: 10.1016/j.ympev.2024.108042.
8. *Генетическое разнообразие и филогенетические взаимосвязи пород свиней, разводимых в России, на основе анализа полиморфизма d-петли мтДНК* / В.Р. Харзинова, Н.А. Акопян, А.В. Доцев [и др.] // *Генетика.* – 2022. – Т. 58, № 8. – С. 920–932. – DOI: 10.31857/S0016675822080045.
9. *Aitnazarov R.B., Mishakova T.M., Yudin N.S. Assessment of genetic diversity and phylogenetic relationships in Black Pied cattle in the Novosibirsk Region using microsatellite markers* // *Vestnik VOGiS.* – 2022. – Vol. 25, N 8. – P. 831–838. – DOI: 10.18699/VJ21.096.
10. *Population Structure Assessed Using Microsatellite and SNP Data: An Empirical Comparison in West African Cattle* / I. Álvarez, I. Fernández, A. Traoré [et al.] // *Animals (Basel).* – 2021. – Vol. 11, N 1. – P. 151. – DOI: 10.3390/ani11010151.
11. *Cortes O., Cañon J., Gama L.T. Applications of Microsatellites and Single Nucleotide Polymorphisms for the Genetic Characterization of Cattle and Small Ruminants: An Overview* // *Ruminants.* – 2022. – Vol. 2, N 4. – P. 456–470. – DOI: 10.3390/ruminants2040032.
12. *Тихонов В.Н., Жучаев К.В. Микроэволюционная теория и практика пороодообразования свиней.* – Новосибирск, 2008. – 395 с.
13. *Henry O.O., Happiness I.E.N., Makuachukwu J-P.O. Applications of Genetic Distances on Blood-Group Gene Frequencies and Their Statistical Genetic Similarities* // *AJPAS.* – 2022. – Vol. 20, N 7. – P. 1–7. – DOI: 2022/v20i130478.
14. *NetCoMi: Network Construction and Comparison for Microbiome Data in R* / S. Peschel, C.L. Müller, E. von Mutius [et al.] // *Briefings in Bioinformatics.* – 2020. – Vol. 22, N 4. – P. bbaa290. – DOI: 10.1093/bib/bbaa290.
15. *Distance-Based Phylogenetic Inference from Typing Data: A Unifying View* / C. Vaz, M. Nascimento, J.A. Carriço [et al.] // *Briefings in Bioinformatics.* – 2021. – Vol. 22, N 3. – P. bbaa147. – DOI: 10.1093/bib/bbaa147.
16. *Kamvar Z.N., Tabima J.F., Grünwald N.J. Poppr: An R Package for Genetic Analysis of Populations with Clonal, Partially Clonal, and/or Sexual Reproduction* // *PeerJ.* – 2014. – Vol. 2. – P. e281. – DOI: 10.7717/peerj.281.
17. *Генетические особенности миниатюрных свиней ИЦиГ СО РАН* / К.С. Шатохин, В.С. Деева, Г.М. Гончареко [и др.] // *Вестник НГАУ (Новосибирский государственный аграрный университет).* – 2014. – Т. 1(30). – С. 75–80.
18. *Горелов И.Г., Савина М.А., Ермолаев В.И. Сибирские миниатюрные свиньи – новая биомодель для медико-биологических исследований* // *Сибирский вестник сельскохозяйственной науки.* – 2001. – № 3–4. – С. 81–87.
19. *Тихонов В.Н. Лабораторные мини-свиньи. Генетика и медико-биологическое использование.* – Новосибирск, 2010. – 304 с.
20. *Капанадзе Г.Д., Аушев Ж.А. Светлогорская популяция мини-свиней* // *Биомедицина.* – 2007. – № 6. – С. 70–80.
21. *Дмитриева Г.Л., Набродова Н.М., Коломников В.А. Генофонд крупной белой породы свиней Западной Сибири* // *Повышение эффективности селекционно-племенной работы.* – Новосибирск, 1992. – С. 52–57.
22. *Селекционно-генетические особенности помесных свиней, полученных от скрещивания крупной белой и йоркширской пород* / В.А. Бекенев, В.С. Деева, В.И. Фролова [и др.] // *Вестник НГАУ (Новосибирский государственный аграрный университет).* – 2016. – № 4. – С. 112–117.
23. *Горелов И.Г. Биология сибирского кабана.* – Новосибирск, 1994. – 82 с.
24. *Тихонов В.Н. Использование групп крови при селекции животных.* – М., 1967. – 391 с.
25. *Шитиков В.К., Мاستицкий С.Э. Классификация, регрессия и другие алгоритмы Data Mining с использованием R.* – Тольятти; Лондон, 2017. – 351 с.

26. Wang W., Kaffo M. Bootstrap Inference for Instrumental Variable Models with Many Weak Instruments // Journal of Econometrics – 2016. – Vol. 192, N 1. – P. 231–268. – DOI: 10.1016/j.jeconom.2015.12.016.
27. Wright K. Will the Real Hopkins Statistic Please Stand Up? // The R Journal. – 2022. – Vol. 14, N 3. – P. 282–292. – DOI: 10.32614/RJ-2022-055.
28. Kassambara A. Practical Guide To Cluster Analysis in R. Unsupervised Machine Learning. 1st ed. – STHDA, 2017. – URL: <https://xsluulab.github.io/Workshop/2021/week10/r-cluster-book.pdf> (дата обращения: 04.10.2025).
29. Объем выборки для корреляционного анализа / А.М. Гржбовский, М.А. Горбатова, А.Н. Наркевич, К.А. Виноградов // Морская медицина. – 2020. – Т. 6, № 1. – С. 101–106. – DOI: 10.22328/2413-5747-2020-6-1-101-106.
30. Bujang M.A. An Elaboration on Sample Size Determination for Correlations Based on Effect Sizes and Confidence Interval Width: A Guide for Researchers // Restor Dent Endod. – 2024. – Vol. 49, N 2. – P. e21. – DOI: 10.5395/rde.2024.49.e21.
31. Pruim R., Kaplan D.T., Horton N.J. The Mosaic Package: Helping Students to Think with Data Using R // The R Journal. – 2017. – Vol. 9, N 1. – P. 77. – DOI: 10.32614/RJ-2017-024.
32. Ali Abd AlHameed K. Spearmans correlation coefficient in statistical analysis // IJNAA. – 2022. – Vol. 13, N 1. – p. 3249–3255. – DOI: 10.22075/ijnnaa.2022.6079.
33. Wavelet Vegetation Index to Improve the Inversion Accuracy of Leaf V25cmax of Bamboo Forests / K. Guo, X. Li, H. Du [et al.] // Remote Sensing. – 2023. – Vol. 15, N 9. – P. 2362. – DOI: 10.3390/rs15092362.
34. Кабанов В.Д., Терентьева А.С. Породы свиней. – М., 1985. – 336 с.
35. Гудилин И.И., Дементьев В.Н., Тараканов Е.А. Кемеровская порода свиней. – Новосибирск, 2003. – 387 с.
36. Genetic diversity and population structure analysis reveals the unique genetic composition of South African selected macadamia accessions / M. Ranketse, C.A. Hefer, R. Pierneef [et al.] // Tree Genetics & Genomes. – 2022. – Vol. 18, N 2. – P. 15. – DOI: 10.1007/s11295-022-01543-0.
37. Comparative Analysis of Microsatellite and SNP Markers for Genetic Management of Red Deer / J. Pérez-González, J. Carranza, G. Anaya [et al.] // Animals. – 2023. – Vol. 13, N 21. – P. 3374. – DOI: 10.3390/ani13213374.
38. Барсукова М.А., Петров А.Ф., Нароженных К.Н. Оценка генетического разнообразия скота породы герефорд на основе микросателлитных маркеров // Достижения науки и техники АПК. – 2025. – Т. 39, № 2. – С. 52–57. – URL: <https://www.elibrary.ru/item.asp?id=80645325> (дата обращения: 01.09.2025).

REFERENCES

1. Zorc M., Škorput D., Gvozdanović K., Margeta P., Karolyi D., Luković Z., Salajpal K., Savić R., Muñoz M., Bovo S., Djurkin Kušec I., Radović Č., Kušec G., Čandek Potokar M., Dovč P., Genetic Diversity and Population Structure of Six Autochthonous Pig Breeds from Croatia, Serbia, and Slovenia, *Genet Sel Evol*, 2022, Vol. 54, No. 1, pp 30, DOI: 10.1186/s12711-022-00718-6.
2. Bakoev S., Getmantseva L., Kostyunina O., Bakoev N., Prytkov Y., Usatov A., Tatarinova T.V., Genome-Wide Analysis of Genetic Diversity and Artificial Selection in Large White Pigs in Russia, *PeerJ*, 2021, Vol. 9, pp. e11595, DOI: 10.7717/peerj.11595.
3. Dai C., Chen K., Zhang R. [et al.], Molecular phylogenetic analysis among species of Paridae, Remizidae and Aegithalos based on mtDNA sequences of COI and cyt b, *Avian Research*, 2023, Vol. 1, No. 2, pp. 112–123.
4. McKenzie P.F., Linking Phylogenetic Models to Population Processes, from Species Trees to Genomes, Columbia University, 2023, URL: <https://books.google.ru/books?id=2Xwh0AEACAAJ>.
5. Kapli P., Flouri T., Telford M.J., Systematic Errors in Phylogenetic Trees, *Curr Biol*, 2021, Vol. 31, No. 2, pp. R59–R64, DOI: 10.1016/j.cub.2020.11.043.
6. Komarova V.A., Lavrenchenko L.A., Approaches to the Detection of Hybridization Events and Genetic Introgression upon Phylogenetic Incongruence, *Biology Bulletin Reviews*, 2022. T. 12, Vol. 3, pp. 240–253, DOI: 10.1134/S2079086422030045.
7. Zhou Y., Tian J., Jiang H., Han M., Wang Y., Lu J., Phylogeography and Demographic History of Macaques, Fascicularis Species Group, in East Asia: Inferred from Multiple Genomic Markers, *Mol Phylogenet Evol*, 2024, Vol. 194, pp. 108042, DOI: 10.1016/j.ympev.2024.108042.
8. Kharzinova V. R., Akopyan N. A., Akopyan A. V., Deniskova T. E., Sermyagin A. A., Karpushkina, T. V., Solovieva A. D., Brem G., Zinovieva N.A., *Russian Journal of Genetics*, 2022, Vol. 58, No. 8, pp. 944–954, DOI: 10.31857/S0016675822080045. (In Russ.)
9. Aitnazarov R.B., Mishakova T.M., Yudin N.S., Assessment of genetic diversity and phylogenetic relationships in Black Pied cattle in the Novosibirsk Region using microsatellite markers, *Vestn. VOGiS*, 2022, Vol. 25, No. 8, pp. 831–838, DOI: 10.18699/VJ21.096.

10. Álvarez I., Fernández I., Traoré A., Menéndez-Arias N.A., Goyache F., Population Structure Assessed Using Microsatellite and SNP Data: An Empirical Comparison in West African Cattle, *Animals (Basel)*, 2021, Vol. 11, No. 1, pp. 151, DOI: 10.3390/ani11010151.
11. Cortes O., Cañon J., Gama L.T., Applications of Microsatellites and Single Nucleotide Polymorphisms for the Genetic Characterization of Cattle and Small Ruminants: An Overview, *Ruminants*, 2022, Vol. 2, No. 4, pp. 456–470, DOI: 10.3390/ruminants2040032.
12. Tihonov V.N., Zhuchaev K.V., *Mikroevoljucionnaja teorija i praktika porodoobrazovanija svinej* (Microevolutionary Theory and Practice of Pig Breeding), Novosibirsk, 2008, 395 pp. (In Russ.).
13. Henry O.O., Happiness I.E.N., Makuachukwu J.P.O., Applications of Genetic Distances on Blood-Group Gene Frequencies and Their Statistical Genetic Similarities, *AJPAS*, 2022, Vol. 20, No. 7, pp. 1–7, DOI: 2022/v20i130478.
14. Peschel S., Müller C.L., von Mutius E., Boulesteix A.-L., Depner M., NetCoMi: Network Construction and Comparison for Microbiome Data in R, *Briefings in Bioinformatics*, 2020, Vol. 22, No. 4, pp. bbaa290, DOI: 10.1093/bib/bbaa290.
15. Vaz C., Nascimento M., Carriço J.A., Rocher T., Francisco A.P., Distance-Based Phylogenetic Inference from Typing Data: A Unifying View, *Briefings in Bioinformatics*, 2021, Vol. 22, No. 3, pp. bbaa147, DOI: 10.1093/bib/bbaa147.
16. Kamvar Z.N., Tabima J.F., Grünwald N.J., Poppr: An R Package for Genetic Analysis of Populations with Clonal, Partially Clonal, and/or Sexual Reproduction, *PeerJ*, 2014, Vol. 2, pp. e281, DOI: 10.7717/peerj.281.
17. Shatohin K.S., Deeva V.S., Gonchareko G.M., Grishina N.B., Goryacheva T.S., Akulich E.G., Kononenko E.V., Ermolaev V.I., Nikitin S.V., *Vestnik NGAU*, 2014, Vol. 30, No. 1, pp. 75–80. (In Russ.)
18. Gorelov I.G., Savina M.A., Ermolaev V.I., *Sibirskij vestnik sel'skohozyajstvennoj nauki*, 2001, Vol. 3–4., pp. 81–87. (In Russ.)
19. Tihonov V.N., *Laboratornye mini-svin'i. Genetika i mediko-biologicheskoe ispol'zovanie* (Laboratory mini-pigs. Genetics and biomedical use), Novosibirsk, 2010, 304 pp. (In Russ.).
20. Kapanadze G.D., Aushev Zh.A., *Biomedicina*, 2007, Vol. 6, pp. 70–80. (in Russ.)
21. Dmitrieva G.L., Nabrodova N.M., Kolomnikov V.A., *Povyshenie effektivnosti selekcionno-plemennoj raboty*, Novosibirsk, 1992, pp. 52–57. (in Russ.)
22. Bekenev V.A., Deeva V.S., Frolova V.I., Botsan I.V., Frolova Yu.V., Orlova K.S., Podvintsev S.I., *Vestnik NGAU*, 2016, Vol. 41, No. 4, pp. 112–117. (in Russ.)
23. Gorelov I.G., *Biologiya sibirskogo kabana* (Biology of the Siberian wild boar), Novosibirsk, 1994. 82 p. (in Russ.)
24. Tikhonov V.N., *Ispol'zovanie grupp krovi pri selekcii zhivotnyh* (The use of blood groups in animal breeding), Moscow, 1967, 391 p. (in Russ.)
25. Shitikov V.K., Mastitsky S.E., *Klassifikaciya, regressiya i drugie algoritmy Data Mining s ispol'zovaniem R* (Classification, regression, and other Data Mining algorithms using R), Tol'yatti; London, 2017, 351 p (in Russ.)
26. Wang W., Kaffo M., Bootstrap Inference for Instrumental Variable Models with Many Weak Instruments, *Journal of Econometrics*, 2016, Vol. 192, No. 1, pp. 231–268, DOI: 10.1016/j.jeconom.2015.12.016.
27. Wright K., Will the Real Hopkins Statistic Please Stand Up? *The R Journal*, 2022, Vol. 14, No. 3, pp. 282–292, DOI: 10.32614/RJ-2022-055.
28. Kassambara A., *Practical Guide To Cluster Analysis in R. Unsupervised Machine Learning.*, 1st ed.; STHDA, 2017, URL: <https://xsluulab.github.io/Workshop/2021/week10/r-cluster-book.pdf>.
29. Grzhbovskij A.M., Gorbatova M.A., Narkevich A.N., Vinogradov K.A., *Morskaya medicina*, 2020, Vol. 6, No. 1, pp. 101–106, DOI: 10.22328/2413-5747-2020-6-1-101-106. (in Russ.)
30. Bujang M.A., An Elaboration on Sample Size Determination for Correlations Based on Effect Sizes and Confidence Interval Width: A Guide for Researchers, *Restor Dent Endod*, 2024, Vol. 49, No. 2, pp. e21, DOI: 10.5395/rde.2024.49.e21.
31. Pruim R., Kaplan D.T., Horton N.J., The Mosaic Package: Helping Students to Think with Data Using R, *The R Journal*, 2017, Vol. 9, No. 1, pp. 77, DOI: 10.32614/RJ-2017-024.
32. Ali Abd AlHameed K., Spearmans correlation coefficient in statistical analysis, *IJNAA*, 2022, T. 13, No 1, pp. 3249–3255, DOI: 10.22075/ijnna.2022.6079.
33. Guo K., Li X., Du H., Mao F., Ni C., Chen Q., Xu Y., Huang Z., Wavelet Vegetation Index to Improve the Inversion Accuracy of Leaf V25cm_{max} of Bamboo Forests, *Remote Sensing*, 2023, Vol. 15, No. 9, pp. 2362, DOI: 10.3390/rs15092362.
34. Kabanov V.D., Terent'eva A.S., *Porody svinej* (The breeds of pigs). Moscow, 1985, 336 p. (in Russ.)
35. Gudilin I.I., Dement'ev V.N., Tarakanov E.A., *Kemerovskaya poroda svinej* (Kemerovo pig breed), Novosibirsk, 2003, 387 p. (in Russ.)
36. Ranketse M., Hefer CA, Pierneef R, Fourie G, Myburg AA., Genetic diversity and population structure analysis reveals the unique genetic composition of South African selected macadamia accessions, *Tree Genetics & Genomes*, 2022, Vol. 18, No. 2, pp. 15, DOI: 10.1007/s11295-022-01543-0.

37. Pérez-González J., Carranza J., Anaya G., Broggin C., Vedel G., de la Peña E., Membrillo A., Comparative Analysis of Microsatellite and SNP Markers for Genetic Management of Red Deer, *Animals*, 2023, Vol. 13, No. 21, pp. 3374, DOI: 10.3390/ani13213374.
38. Barsukova M.A., Petrov A.F., Narozhnyh K.N., *Dostizheniya Nauki I Tekhniki Apk*, 2025, Vol. 39, No 2. pp. 52–57, URL: <https://www.elibrary.ru/item.asp?id=80645325>. (in Russ.)

Сведения об авторах:

К.С. Шатохин, старший научный сотрудник лаборатории прикладной биоинформатики, кандидат биологических наук

С.В. Никитин, старший научный сотрудник лаборатории молекулярной генетики и селекции сельскохозяйственных животных, кандидат биологических наук

Н.Н. Кочнев, доктор биологических наук, профессор

Contribution of the authors:

K.S. Shatokhin, Senior Researcher at the Laboratory of Applied Bioinformatics, Candidate of Biological Sciences

S.V. Nikitin, Senior Researcher at the Laboratory of Molecular Genetics and Breeding of Farm Animals, Candidate of Biological Sciences

N.N. Kochnev, Doctor of Biological Sciences, professor

Вклад авторов:

Все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.