

RESEARCH ARTICLE

Genetic diversity and population structure of native and translocated Bashkir horse (*Equus ferus caballus*) populations based on microsatellite markers



Tatiana Alexandrovna Sedykh^{1,2} , Aleksey Victorovitch Vishniakov³ , Farit Ravilovitch Valitov⁴ , Ruslan Rinatovitch Gizatullin⁵ , Zaitsev Aleksandr Mikhailovitch⁶ , Diana Ildarovna Gareeva¹ , Elina Nailevna Latypova¹ , Irina Sergeevna Gavriliicheva⁶ , Kamil Dinarovitch Farkhutdinov^{4,5} , Rustam Sultangareevich Kiraev¹ and Elena Aleksandrovna Gladyr⁷

1. Bashkir Scientific Research Institute of Agriculture of the Ufa Federal Research Center of the Russian Academy of Sciences, Ufa, Republic of Bashkortostan, Russia.
2. Bashkir State Pedagogical University named after M. Akmulla, Ufa, Republic of Bashkortostan, Russia.
3. SNAIP Limited Liability Company, Tver Region, Russia.
4. Bashkir State Agrarian University, Ufa, Republic of Bashkortostan, Russia.
5. Research Institute for the Study of Bashkir Horses "Bashkort Aty (Bashkir Horse)", Ufa, Republic of Bashkortostan, Russia.
6. All-Russian Research Institute of Horse Breeding named after Academician V.V. Kalashnikov, Ryazan Region, Russia.
7. Federal Research Center for Animal Husbandry – All-Russian Institute of Animal Husbandry named after Academician L.K. Ernst, Moscow Region, Russia.

ABSTRACT

Background and Aim: The Bashkir horse (*Equus ferus caballus*) is an indigenous Russian breed with a unique adaptive gene pool that is important for biodiversity conservation and sustainable breeding. However, information on the genetic differentiation between native and recently translocated populations remains limited. This study evaluated the genetic diversity, population structure, and within-breed differentiation of native Bashkir horses from the Republic of Bashkortostan and a translocated population established in the Tver Region using microsatellite markers.

Materials and Methods: A total of 240 Bashkir horses were included, comprising a native population from the Republic of Bashkortostan (n = 190) and a translocated population from the Tver Region (n = 50). Genomic DNA extracted from hair follicles was genotyped at 17 International Society for Animal Genetics (ISAG)-recommended microsatellite loci using the COrDIS Horse kit. Genetic diversity indices, including the number of alleles, effective number of alleles, Shannon index, observed and expected heterozygosity, polymorphic information content (PIC), and Wright's fixation indices, were calculated. Population differentiation was evaluated using principal coordinates analysis (PCoA), pairwise fixation index among subpopulations (FST), permutational multivariate analysis of variance (PERMANOVA), and Bayesian clustering with STRUCTURE and the Evanno method.

Results: Analysis of the 17 microsatellite loci identified 160 alleles in the native population and 137 alleles in the translocated population, with mean allele numbers of 9.41 and 8.06 per locus, respectively. The loci ASB17, ASB2, and HTG10 contributed most to genetic diversity in both populations. Mean effective allele numbers were 4.783 and 4.713, whereas the Shannon index averaged 1.741 and 1.676 for the native and translocated populations, respectively. High PIC values confirmed the excellent discriminatory power of the marker panel. Low mean inbreeding coefficient within subpopulations (0.012) and overall inbreeding coefficient (0.022) values indicated minimal inbreeding. PERMANOVA detected statistically significant but extremely weak population differentiation ($F = 3.28$, $R^2 = 0.0136$, $p = 0.001$, whereas bootstrapped F_{ST} (0.0072 ± 0.0012 ; 95% confidence interval: 0.0045–0.0093) demonstrated very low genetic divergence. Bayesian clustering identified $K = 3$ as the optimal model, revealing subdivision within the native Bashkir population while preserving the overall genetic integrity of the breed.

Conclusion: Native and translocated Bashkir horse populations have retained high genetic diversity and exhibit only minimal genetic differentiation despite approximately two decades of geographic separation. The combined use of microsatellite markers with complementary population genetic analyses provides robust evidence supporting the conservation of the Bashkir horse gene pool while identifying subtle within-breed substructure. These findings provide a scientific foundation for genetic monitoring, conservation planning, pedigree management, and sustainable breeding programs for this valuable indigenous horse breed.

Corresponding Author: Tatiana Alexandrovna Sedykh

E-mail: s_ta@inbox.ru

Received: 28-04-2026, **Accepted:** 27-08-2026, **Published online:** 17-10-2026

Co-authors: AVV: mcherry.89@mail.ru; FRV: fvalitov@mail.ru; RRG: g.ruslan.07@mail.ru; ZMA: amzaitceff@mail.ru; DIG: diana.razgar@mail.ru; ENL: elinalatypova02@mail.ru; ISG: labgenetics79@gmail.com; KDF: anobashkortati@mail.ru; RSK: bniish@rambler.ru; EAG: elenagladyr@mail.ru

How to cite: Sedykh TA, Vishniakov AV, Valitov FR, Gizatullin RR, Zaitsev AM, Gareeva DI, et al. Genetic diversity and population structure of native and translocated Bashkir horse (*Equus ferus caballus*) populations based on microsatellite markers. Vet World. 2026;19(10):4589-4608.

Copyright: Sedykh et al. This article is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>).



Keywords: Bashkir horse, conservation genetics, genetic diversity, indigenous horse breed, microsatellite markers, population genetics, population structure, short tandem repeats polymorphism.

INTRODUCTION

One of the primary objectives of animal husbandry, particularly horse breeding, is the conservation and sustainable utilization of valuable genetic resources, with local horse breeds serving as one of the principal reservoirs of such diversity. The Bashkir horse (*Equus ferus caballus*) is a well-recognized indigenous breed that represents an important component of the historical and cultural heritage of the Republic of Bashkortostan. Through long-term selective breeding, this breed has developed distinctive characteristics, including hardiness, low feed requirements, the ability to graze beneath snow cover (tebenevka), and the capacity to maintain body condition under harsh climatic conditions. It is also characterized by rapid physiological recovery, stress resistance, well-developed herd behavior, psychological stability, manageability, curiosity, and pronounced socio-emotional responsiveness to humans [1–3]. These traits, developed through long-term selection, contribute to high milk and meat productivity while supporting the use of Bashkir horses in agricultural work, sports, tourism, and therapeutic applications.

Numerous studies have highlighted the importance of investigating indigenous horse breeds using molecular genetic approaches. Such studies are essential for evaluating genetic diversity, inbreeding levels, and other genetic parameters that reflect population status and potential conservation risks [4–10]. Microsatellite markers (short tandem repeats [STRs]) are widely used for breed conservation, studies of breed origin, and the assessment of genetic diversity within populations [11–15]. Together with single nucleotide polymorphism (SNP) genotyping, microsatellite analysis has become a principal molecular genetic tool in animal breeding for parentage verification, individual identification, and population genetic studies [16–21]. Advances in DNA-based technologies have enabled continuous genetic monitoring at both individual and population levels, thereby helping to prevent inbreeding depression, optimize breeding strategies, and accelerate genetic improvement programs [22–27].

Breeding Bashkir horses under year-round pasture conditions in two geographically distant regions, the core population in the Republic of Bashkortostan (Southern Urals) and the translocated population in the Tver Region (Upper Volga), using different breeding objectives may have resulted in microevolutionary changes within these populations. These contrasting breeding objectives, together with the use of selected breeding stallions in the translocated population, may have influenced allele frequencies at the neutral microsatellite loci recommended by the International Society for Animal Genetics (ISAG). Nevertheless, because both populations share a common origin and have not experienced long-term reproductive isolation, directional selection is unlikely to have substantially altered their overall genetic structure. Previous studies have demonstrated that the Bashkir horse, as an indigenous breed with a relatively large population size, exhibits a high degree of genetic consolidation and maintains a stable gene pool [2, 3, 12]. Therefore, although some genetic differences between populations may be expected, overall genetic diversity is likely driven primarily by within-breed variation, while interpopulation differentiation is expected to reflect only localized genetic variation without compromising overall breed integrity. Furthermore, because horses were translocated to the Tver Region relatively recently, genetic differentiation between populations may be limited. Few published studies have evaluated genetic differences between core and recently translocated animal populations using complementary statistical approaches. Therefore, to assess within-breed differentiation and characterize genetic variation between the core and translocated Bashkir horse populations, several analytical approaches were employed, including principal coordinates analysis (PCoA) for visualization of genetic relationships, permutational multivariate analysis of variance (PERMANOVA) for testing group differentiation, and Bayesian clustering using STRUCTURE with the Evanno method to determine the optimal number of genetic clusters.

Although microsatellite markers have been used to investigate the genetic diversity of indigenous horse breeds, information on genetic differentiation between core and recently translocated populations of the Bashkir horse remains limited. Moreover, few studies have combined multiple complementary analytical approaches, including PCoA, PERMANOVA, and Bayesian clustering, to comprehensively evaluate within-breed genetic subdivision while accounting for the relatively recent establishment of translocated populations. Consequently, the extent to which translocation and region-specific breeding strategies have influenced the genetic structure of the Bashkir horse population has not been fully elucidated.

This study aimed to evaluate the polymorphism of microsatellite markers in Bashkir horses bred in the Republic of Bashkortostan and the Tver Region. Specifically, the study sought to establish representative

experimental groups, collect biological samples, genotype horses using 17 microsatellite loci, identify individual genotypes, and perform population genetic analyses to determine genetic diversity indices and assess the degree of within-breed genetic differentiation between the core and translocated Bashkir horse populations.

MATERIALS AND METHODS

Ethical approval

The study protocol was reviewed and approved by the Scientific Ethics Committee of the Bashkir Research Institute of Agriculture, Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russian Federation (Approval No. 0003/BNIISH UFRC RAS/2026). The study involved 240 Bashkir horses maintained under routine herd-based production and breeding conditions in the Republic of Bashkortostan and the Tver Region. Biological sampling was limited to the collection of mane hairs with follicles for genomic DNA extraction and microsatellite genotyping.

All procedures involving the animals were conducted in accordance with the approved study protocol and applicable institutional standards for animal welfare and ethical research. Animal handling and sample collection were performed with due consideration for animal welfare and were designed to minimize unnecessary restraint, discomfort, and stress. The study did not involve experimental treatment, invasive surgical procedures, or euthanasia of animals for research purposes. The conduct and reporting of the study followed the relevant principles of the ARRIVE 2.0 guidelines and recommendations of the World Organization for Animal Health.

Study period and location

This study was conducted between 2024 and 2025 using biological samples collected from Bashkir horses representing the core population in the Pre-Ural and Trans-Ural regions of the Republic of Bashkortostan and a translocated population in the Tver Region, Russian Federation. The collection of biological material for the study was conducted from April to November 2024, and molecular genetic studies were carried out in laboratories as the horse biological material arrived, up until February 2025. The core population comprised horses from JSC "Ufa Stud Farm No. 119" (Ufimsky District; n = 97), LLC "Bashagrogen" (Kushnarenkovsky District; n = 13), the peasant farm "Idiyatullin Kh.S." (Iglinsky District; n = 8), the State Budgetary Institution of the Republic of Bashkortostan "Ufa Breeding Center" (n = 35), LLC "Yuldash" (Uchalinsky District; n = 20), and LLC "Tolpar" (Baymasky District; n = 17). The translocated population consisted of 50 horses maintained at LLC "SNAIP" in the Tver Region.

The Tver population was established in 2005 using stallions and mares with high milk productivity obtained from several farms in the Republic of Bashkortostan, including JSC "Ufa Stud Farm No. 119" (Ufimsky District), the peasant farm "Idiyatullin Kh.S." (Iglinsky District), LLC Agrofirma "Irandyk," the agricultural production cooperative "Sibaysky," the state agricultural enterprise "Zilairsky," and the experimental production farm "Baymakskiye" (Baymasky District). Approximately 1,200 horses were transferred to establish a koumiss-producing farm.

Genotyping of horses from the Bashkir subpopulation and analysis of short tandem repeat (STR) loci were performed at the Laboratory of Molecular Genetics and Animal Breeding, Bashkir Research Institute of Agriculture (Ufa), the Laboratory of Molecular Genetics of Farm Animals, L.K. Ernst Federal Science Center for Animal Husbandry (Dubrovitsy, Moscow Region), and the Laboratory of Genetics, All-Russian Research Institute of Horse Breeding (Divovo, Ryazan Region). Genotyping of horses from the Tver subpopulation and STR analysis were conducted at the Laboratory of Molecular Genetic Expertise, My Gen LLC (Moscow).

Study design

The study design was determined by the scientific hypothesis in which we suggested that the native and relatively recently relocated horse populations may have minor genetic differences that can only be identified using highly sensitive statistical methods, for which a cross-sectional observational population-genetic study was applied. No experimental intervention was conducted.

Study animals and group formation

The study included Bashkir horses maintained in two regions of the Russian Federation. Group 1 consisted of horses from the Republic of Bashkortostan (n = 190), including 84 stallions and 106 mares, whereas Group 2 comprised horses from LLC "SNAIP" in the Tver Region (n = 50), including 19 stallions and 31 mares. All horses were managed under herd-based production systems. Experimental groups were established using representative animals of the breed that were not closely related. Before obtaining the samples, zootechnical records from all horse breeding farms were studied to select non-related typical Bashkir horses. Subsequently, based on the obtained microsatellite profiles, the zootechnical records were confirmed - no erroneous entries in the pedigrees were found.

During the breeding season, horses were maintained in harem groups, with one stallion assigned to 15–30 mares. Animals grazed on natural pastures throughout the year, including winter grazing beneath snow cover (tebenevka), and had access to shelters during snowstorms and blizzards.

DNA extraction and microsatellite genotyping

For genetic analysis, only hairs with visible hair follicles (root sheaths), dark or light thickenings at the end, were selected; hairs without follicles (dead, shed after molting, cut) are unsuitable for analysis. The selection was carried out as follows: a section on the neck, withers, or mane of the animal was chosen, a small strand of hair was grasped at the very base (close to the skin), and a sharp motion was used to pull out the tuft without cutting it with scissors; 20–40 hairs were collected from each animal (up to 50–60 if the bulbs were poorly developed). The pulled tuft was carefully placed in a paper envelope, ensuring that the bulbs were not torn off, and a unique animal identifier (inventory number) was clearly and legibly written on the envelope or vial with a waterproof pen or marker. Samples were stored and transported in a dry, dark place at room temperature (20–22°C). DNA extraction was optimally performed within 1–2 weeks after collection, with storage allowed for up to 1–3 months in dry conditions.

Hair follicles collected from the mane were used for DNA extraction. Genomic DNA was extracted using the COrDIS Extract kit (Gordiz, Russia).

DNA concentration and purity were determined using a NanoDrop ND-500 spectrophotometer (Thermo Fisher Scientific, USA) by measuring absorbance at 260, 280, and 230 nm. Samples with A260/A280 ratios of 1.8–2.0 and A260/A230 ratios >1.8 were considered suitable for subsequent analyses.

Genotyping was performed using the COrDIS Horse kit (Gordiz), which includes 17 ISAG-recommended microsatellite loci. Multiplex amplification of the 17 loci (AHT4, AHT5, ASB2, ASB17, ASB23, CA425, HMS1, HMS2, HMS3, HMS6, HMS7, HTG4, HTG7, HTG6, HTG10, VHL20, and LEX3) was carried out according to the manufacturer's protocol to ensure reproducibility and comparability with international studies. Polymerase chain reaction was performed in a 25 µL reaction volume containing 1–5 ng DNA under the following cycling conditions: initial denaturation at 94°C for 3 min; 32 cycles of 94°C for 30 s, 59°C for 120 s, and 72°C for 90 s (ramp rate ≤0.3°C/s); followed by a final extension at 68°C for 10 min. A positive control (HRS-12 DNA) and a negative control (double-distilled water) were included in each amplification run.

Detection of amplification products

Amplified DNA fragments were separated and detected by capillary electrophoresis using a Nanofor 05 genetic analyzer (LLC SPF, Sintol, Russia). Fragment analysis was performed using GeneMapper v4.0 (Applied Biosystems (Thermo Fisher Scientific CLJA, Russia) with the control DNA profile supplied in the COrDIS Horse kit (Gordiz). An S550 internal size standard was used for allele sizing, and stutter peaks representing <20% of the corresponding major peak were excluded from the analysis.

Genetic and population analyses

Potential microsatellite genotyping errors were evaluated using Micro-Checker software [28], which detects null alleles, large allele dropout, stutter peaks, and potential scoring errors. The analysis indicated that although 50% of the alleles at loci HTG4 and HTG6 belonged to the same size class, they represented valid alleles and therefore required no correction. Null alleles were detected at locus LEX3. Although this locus could have been excluded from subsequent analyses, it was retained because the generated data were considered scientifically informative.

Genetic and population analyses were performed using GenAlEx v6.41 [29, 30]. The following parameters were calculated: number of alleles (Na), effective number of alleles (Ne), Shannon index (I), observed heterozygosity (Ho), expected heterozygosity (He), and Wright's fixation index (F).

Genetic relationships between the two subpopulations were visualized using PCoA based on a matrix of genetic distances derived from multilocus genotypes. Genetic differentiation between the subpopulations was assessed using pairwise fixation index among subpopulations (FST) values.

Because of the unequal sample sizes (190 vs. 50 animals), Bonferroni correction was applied when testing for Hardy–Weinberg equilibrium by adjusting the significance level to $\alpha = 0.05/17$. Allelic richness (AR) was estimated using rarefaction to account for differences in sample size [31]. Rarefaction was performed to 50 individuals for 16 loci and to 31 individuals for locus LEX3, allowing unbiased comparisons of allelic diversity between the subpopulations.

The polymorphic information content (PIC) was calculated for each locus using Botstein's formula [32]

implemented in R v4.2.1 (R Core Team, 2022, Vienna, Austria).

Hardy–Weinberg equilibrium (HWE) was evaluated for each locus in each subpopulation using the Markov chain exact test implemented in the Pegas package in R [33]. To control the family-wise error rate resulting from multiple hypothesis testing, Bonferroni correction was applied, and the corrected significance level was calculated as $\alpha = 0.05/N$, where N represents the total number of statistical tests performed.

Genetic differentiation between the studied subpopulations was further evaluated using the nonparametric PERMANOVA. To assess the significance of differences between the studied subpopulations, a non-parametric PERMANOVA (permutational multivariate analysis of variance) was applied, implemented in the vegan package (version 2.7-5) in the R environment [34]. This method allows for the comparison of multivariate groups without strict assumptions about the distribution of the data. No differences in dispersion between groups were found ($p = 0.38$), which confirms the validity of using PERMANOVA. The matrix of genetic distances between individuals was calculated based on Euclidean distance from genotypic data. Before interpreting the PERMANOVA results, the homogeneity of multivariate dispersion was checked using the betadisper function (the multivariate analog of Levene's test) with 999 permutations. [35]. Pairwise group comparisons were adjusted using the Bonferroni correction to control the Type I error rate. In addition, confidence intervals for F_{ST} were estimated by bootstrapping with 1,000 iterations and locus resampling to obtain robust estimates of genetic differentiation.

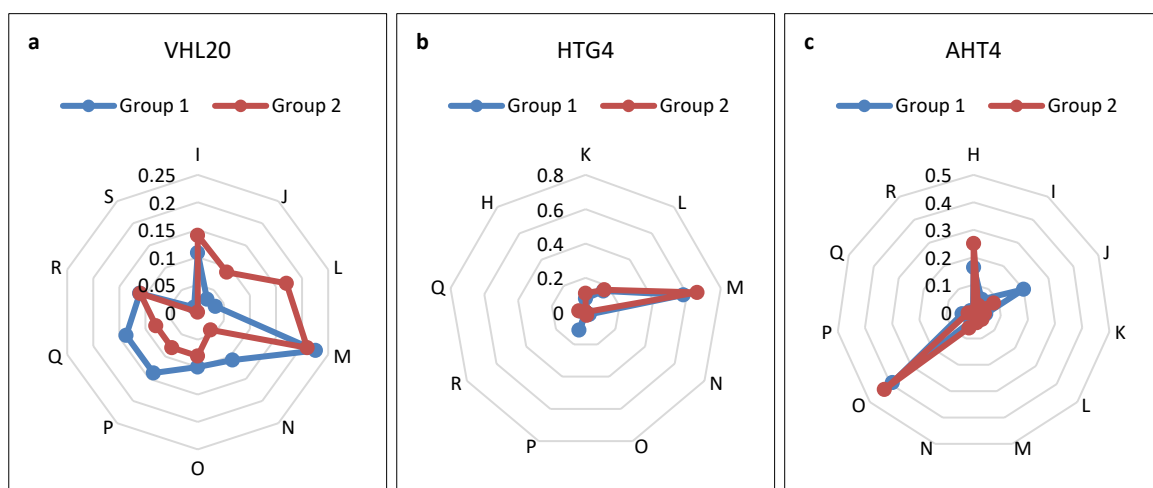
Determination of within-breed genetic subdivision

Within-breed genetic subdivision and the optimal number of genetic clusters (K) were investigated using Bayesian clustering analysis implemented in STRUCTURE v2.3.4 using the admixture model (NOADMIX = 0) and the correlated allele frequencies model (FREQSCORR = 1) [36]. Loci were assumed to be unlinked (LINKAGE = 0). Prior information about the samples' population membership was not used (USEPOPINFO = 0, LOCPRIOR = 0). For each value of K from 2 to 10, five independent MCMC runs were performed with a burn-in of 10,000 iterations and 100,000 data collection iterations. This method identifies genetic clusters and estimates individual membership coefficients (Q) without prior knowledge of sample origin. Individuals showing nonzero membership coefficients for more than one cluster were considered genetically admixed.

Analyses were performed for K values ranging from 2 to 10. For each K value, five independent runs were conducted using a burn-in period of 10,000 iterations followed by 100,000 Markov chain Monte Carlo iterations. The optimal number of genetic clusters was determined using the ΔK method described by Evanno *et al.* [37]. The pophelper package in R was used to calculate ΔK values and generate likelihood plots [38].

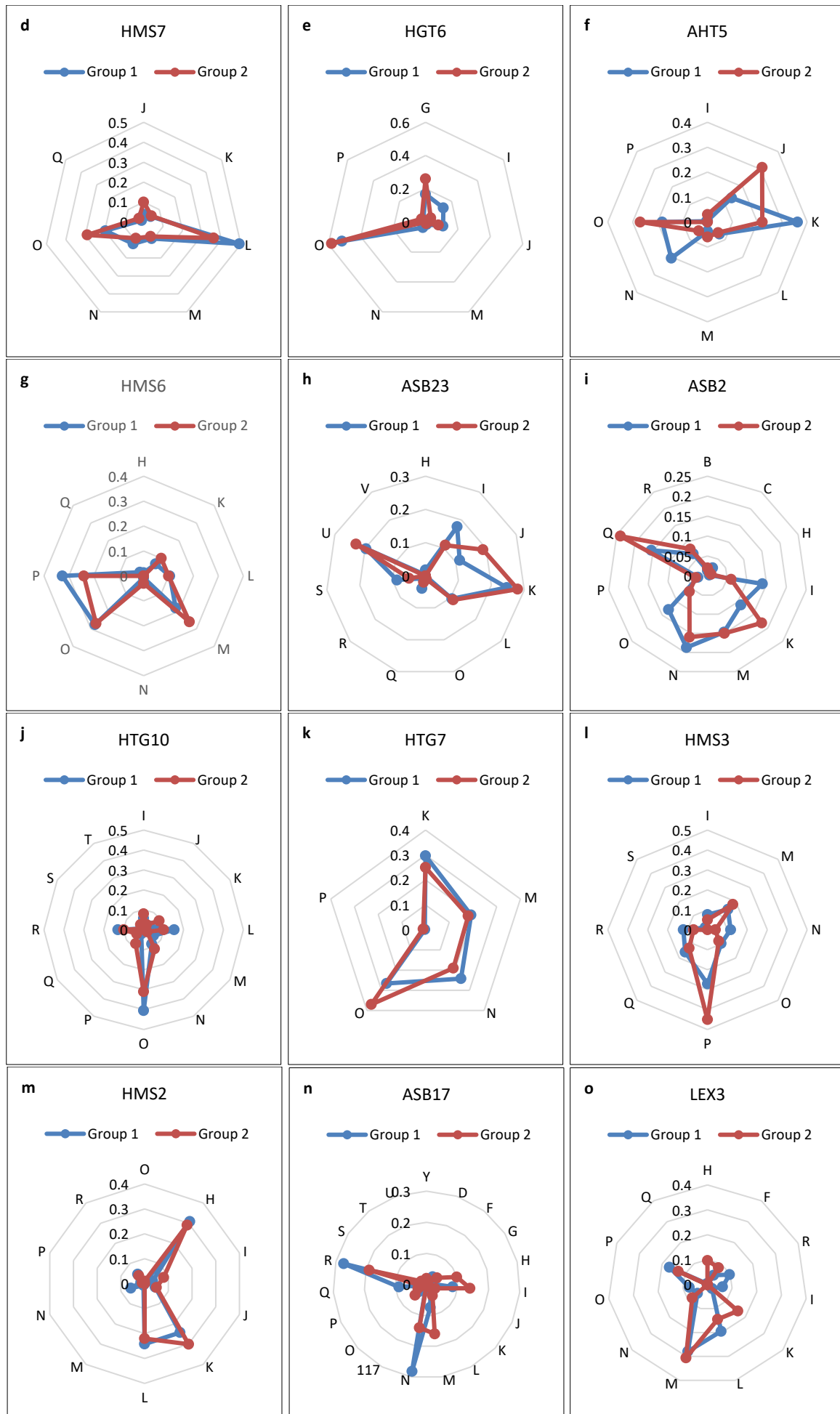
RESULTS

Analysis of 17 STR loci in Bashkir horses revealed genetic differences between the two study groups. The polymorphism of the STR loci is illustrated in Figures 1a–q.



Allelic and genotypic distributions

The allelic composition of the microsatellite loci indicated a high level of polymorphism. A total of 160 alleles were identified in Group 1, comprising horses bred in the Republic of Bashkortostan, whereas 137 alleles were identified in Group 2, comprising horses bred in the Tver Region. The genotype and allele distributions in the core and translocated populations are described below, highlighting similarities and differences in the genetic architecture of the studied loci.



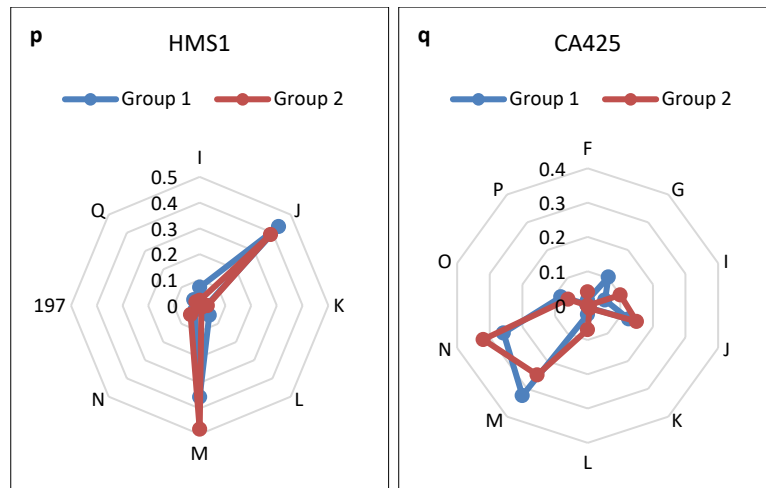


Figure 1: Polymorphism of microsatellite loci in the core population (Group 1, n = 190) and translocated population (Group 2, n = 50) of Bashkir horses: (a) VHL20; (b) HTG4; (c) AHT4; (d) HMS7; (e) HTG6; (f) AHT5; (g) HMS6; (h) ASB23; (i) ASB2; (j) HTG10; (k) HTG7; (l) HMS3; (m) HMS2; (n) ASB17; (o) LEX3; (p) HMS1; and (q) CA425.

VHL20 locus

In Group 1, homozygous individuals accounted for 10.00% of the population, with genotype MM being the most common (5.26%). Heterozygous individuals accounted for 90.00%, with genotypes MP (6.84%) and MR (5.26%) occurring most frequently. In Group 2, homozygous individuals accounted for 22.00%, with seven homozygous genotypes identified: MM (10.00%), II (2.00%), QQ (2.00%), PP (2.00%), LL (2.00%), NN (2.00%), and RR (2.00%). Heterozygous individuals accounted for 78.00%, with genotype IL being the most common (12.00%).

A total of 10 alleles were identified in Group 1, with alleles M (0.226) and R (0.111) showing the highest frequencies. In Group 2, nine alleles were identified, with alleles M (0.210) and I (0.140) showing the highest frequencies. Allele S was rare in Group 1 (0.011). In Group 2, allele N was rare (0.040), whereas allele S was absent (Figure 1a).

HTG4 locus

In Group 1, homozygous individuals accounted for 42.11%, with genotype MM being the most common (37.37%). Heterozygous individuals accounted for 57.89%, with genotype LM being the most frequent (18.42%). In Group 2, homozygous individuals accounted for 50.00%, with genotype MM also being the most common (44.00%). Heterozygous individuals accounted for 50.00%, with genotype LM being the most frequent (24.00%).

Nine alleles were identified in Group 1, with alleles M (0.579) and L (0.161) showing the highest frequencies. Five alleles were detected in Group 2, with alleles M (0.660) and L (0.170) predominating. Alleles R and H were rare in Group 1, each with a frequency of 0.003. In Group 2, allele O occurred at a low frequency (0.020), whereas alleles P, N, R, and H were absent, although they occurred at low frequencies in Group 1 (Figure 1b).

AHT4 locus

In Group 1, homozygous individuals accounted for 27.37%, with genotype OO being the most common (17.89%). Heterozygous individuals accounted for 72.63%, with genotypes HO and JO being the most frequent, each occurring in 11.05% of the population. In Group 2, homozygous individuals accounted for 26.00%, with genotype OO being the most common (18.00%). Heterozygous individuals accounted for 74.00%, with genotype HO being the most frequent (18.00%).

Eleven alleles were identified in Group 1, with allele O showing the highest frequency (0.392). Nine alleles were identified in Group 2, with alleles O (0.430) and H (0.250) predominating. Alleles M (0.016), Q (0.013), and R (0.005) were rare in Group 1. In Group 2, alleles P (0.020), I (0.030), K (0.040), and L (0.040) occurred at low frequencies, whereas alleles Q and R were absent (Figure 1c).

HMS7 locus

In Group 1, homozygous individuals accounted for 30.53%, with genotype LL being the most common (23.68%). Heterozygous individuals accounted for 69.47%, with genotypes LO (20.00%) and LN (12.11%) being the most frequent. In Group 2, homozygous individuals accounted for 26.00%, with genotype LL being the most common (16.00%). Heterozygous individuals accounted for 74.00%, and 13 heterozygous genotypes were identified, with LO (16.00%) and JL (12.00%) occurring most frequently.

Seven alleles were identified in each group. Allele L showed the highest frequency in Group 1 and Group 2, at 0.492 and 0.360, respectively. Allele Q occurred at low frequencies in both groups, at 0.013 and 0.030, respectively (Figure 1d).

HTG6 locus

In Group 1, homozygous individuals accounted for 29.47%, with genotype OO being the most common (23.68%). Heterozygous individuals accounted for 70.53%, with genotypes GO (20.53%) and JO (11.05%) being the most frequent. In Group 2, homozygous individuals accounted for 46.00%, with genotype OO being the most common (38.00%). Heterozygous individuals accounted for 54.00%, with genotype GO being the most frequent (28.00%).

Seven alleles were identified in Group 1, with alleles O (0.516) and G (0.168) showing the highest frequencies. Six alleles were identified in Group 2, with alleles O (0.580) and G (0.260) predominating. Alleles P (0.018) and M (0.016) were rare in Group 1. In Group 2, allele N, which was present in Group 1, was absent, whereas alleles P (0.030) and M (0.010) occurred at low frequencies (Figure 1e).

AHT5 locus

In Group 1, homozygous individuals accounted for 22.11%, with genotype KK being the most common (14.21%). Heterozygous individuals accounted for 77.89%, with genotypes KO (14.74%) and KN (14.21%) being the most frequent. In Group 2, homozygous individuals accounted for 22.00%, with genotype JJ being the most common (10.00%). Heterozygous individuals accounted for 78.00%, with genotype JO being the most frequent (16.00%).

Eight alleles were identified in Group 1, with alleles K (0.361) and N (0.205) showing the highest frequencies. Seven alleles were identified in Group 2, with alleles J (0.310) and O (0.270) showing the highest frequencies. Alleles I and P were rare in Group 1, each with a frequency of 0.001. In Group 2, allele I was rare (0.030), whereas allele P was absent (Figure 1f).

HMS6 locus

In Groups 1 and 2, homozygous individuals accounted for 22.63% and 18.00%, respectively, with genotypes PP (10.53%) and MM (6.00%) being the most common. Heterozygous individuals accounted for 77.37% and 82.00% in Groups 1 and 2, respectively, with genotype OP being the most frequent in Group 1 (20.53%) and genotype MO being the most frequent in Group 2 (20.00%).

Eight alleles were identified in Group 1, with alleles P (0.326) and O (0.276) occurring at high frequencies. Six alleles were identified in Group 2, with alleles O (0.270) and M (0.260) occurring at high frequencies. In Group 2, allele N was rare (0.030), whereas alleles Q and H were absent, although they occurred at low frequencies in Group 1 (Figure 1g).

ASB23 locus

In Groups 1 and 2, homozygous individuals accounted for 14.21% and 28.00%, respectively, with genotype UU being the most common in both groups (5.26% and 12.00%, respectively). Heterozygous individuals accounted for 85.79% and 72.00% in Groups 1 and 2, respectively, with genotype IK being the most frequent in Group 1 (10.00%) and genotype JK being the most frequent in Group 2 (18.00%).

Ten alleles were identified in Group 1 and eight in Group 2. In both groups, alleles K (0.247 and 0.280, respectively) and U (0.197 and 0.230, respectively) occurred at high frequencies. In Group 1, allele O was absent, whereas alleles V (0.005), R (0.011), and H (0.018) were rare. In Group 2, alleles O (0.010) and Q (0.020) were rare, whereas alleles R, V, and H were absent (Figure 1h).

ASB2 locus

In Groups 1 and 2, homozygous individuals accounted for 12.63% and 18.00%, respectively. Genotype NN was the most common in Group 1 (3.16%), whereas genotypes KK, NN, and QQ each accounted for 6.00% in Group 2. Heterozygous individuals accounted for 87.37% and 82.00% in Groups 1 and 2, respectively, with genotype NQ being the most frequent in both groups (6.32% and 12.00%, respectively).

Eleven alleles were identified in each group. In Group 1, alleles Q (0.155), M (0.147), and I (0.139) occurred at high frequencies. In Group 2, alleles Q (0.240) and K (0.180) occurred at high frequencies. In Group 1, alleles P and C each had a frequency of 0.024, whereas alleles B (0.013) and H (0.005) were rare. In Group 2, alleles H and C each had a frequency of 0.010, whereas alleles B (0.020) and P (0.030) were rare (Figure 1i).

HTG10 locus

In Group 1, homozygous individuals accounted for 25.79%, with genotype OO being the most common (17.37%). Heterozygous individuals accounted for 74.21%, with genotype LO being the most frequent (11.58%). In Group 2, homozygous individuals accounted for 12.00%, with three homozygous genotypes identified: OO (8.00%), RR (2.00%), and II (2.00%). Heterozygous individuals accounted for 88.00%, with genotypes NO (12.00%) and KO (10.00%) being the most frequent.

Eleven alleles were identified in Group 1 and 12 in Group 2. Allele O occurred at a high frequency in both groups, at 0.405 and 0.310, respectively. In Group 1, allele J was absent, whereas alleles P (0.024), S (0.013), and Q (0.016) were rare. In Group 2, alleles S (0.010), M (0.020), J (0.030), and T (0.030) were rare (Figure 1j).

HTG7 locus

In Group 1, homozygous individuals accounted for 26.32%, with genotypes KK (9.47%) and NN (7.37%) being the most common. Heterozygous individuals accounted for 73.68%, with genotypes KO (14.74%) and KN (14.21%) being the most frequent. In Group 2, homozygous individuals accounted for 28.00%, with genotype OO being the most common (16.00%). Heterozygous individuals accounted for 72.00%, with genotype KN being the most frequent (14.00%).

Five alleles were identified in each group. Alleles K and O occurred at high frequencies in Group 1 (0.297 and 0.266, respectively) and Group 2 (0.250 and 0.370, respectively). Allele P was rare in both groups, with a frequency of 0.010 (Figure 1k).

HMS3 locus

In Group 1, homozygous individuals accounted for 14.74%, with genotype PP being the most common (5.26%). Heterozygous individuals accounted for 85.26%, with genotypes PQ (10.53%) and MP (8.95%) being the most frequent. In Group 2, homozygous individuals accounted for 30.00%, with genotype PP being the most common (22.00%). Heterozygous individuals accounted for 70.00%, with genotypes PQ and MP each occurring in 14.00% of the population.

Eight alleles were identified in Group 1, with allele P occurring at a high frequency (0.271). Seven alleles were identified in Group 2, with allele P also occurring at a high frequency (0.450). In Group 1, allele S was rare (0.016). In Group 2, allele N was rare (0.040), whereas allele S was absent (Figure 1l).

HMS2 locus

In Group 1, homozygous individuals accounted for 25.26%, with genotypes HH and KK each occurring in 8.95% of the population. Heterozygous individuals accounted for 74.74%, with genotype HK being the most frequent (17.89%). In Group 2, homozygous individuals accounted for 18.00%, with genotype KK being the most common (8.00%). Heterozygous individuals accounted for 82.00%, with genotype HK being the most frequent (20.00%).

Nine and eight alleles were identified in Groups 1 and 2, respectively. In Group 1, allele O was absent, whereas alleles H (0.308), K (0.242), and L (0.242) occurred at high frequencies. In Group 2, alleles H (0.290), K (0.300), and L (0.220) occurred at high frequencies, whereas alleles N and M were absent. Alleles P and O were rare in Group 2, each with a frequency of 0.010 (Figure 1m).

ASB17 locus

In Group 1, homozygous individuals accounted for 17.89%, with genotype NN being the most common (7.89%). Heterozygous individuals accounted for 82.11%, with genotype NR being the most frequent (15.79%). In Group 2, homozygous individuals accounted for 4.00%, with two homozygous genotypes, MM and RR, each occurring in 2.00% of the population. Heterozygous individuals accounted for 96.00%, with genotype NR being the most frequent (12.00%).

Eighteen alleles were identified in Group 1, with alleles N (0.282) and R (0.274) occurring at high frequencies. Allele Y was absent. Alleles D, H, J, and K each had a frequency of 0.013, whereas alleles P (0.005), S (0.030), and allele 117 (0.003), which had no letter designation, were rare. In Group 2, 15 alleles were identified, with alleles R (0.190) and M (0.160) occurring most frequently. Alleles D, F, T, and 117 were absent. Alleles Y, S, and Q each had a frequency of 0.020; J and P each had a frequency of 0.030; and U and K each had a frequency of 0.010 (Figure 1n).

LEX3 locus

In Groups 1 and 2, homozygous individuals accounted for 51.05%. In Group 1, genotype MM was the most

common (16.84%). Heterozygous individuals accounted for 48.95% in both groups, with genotype LM being the most frequent in Group 1 (5.79%).

Ten alleles were identified in Group 1, with alleles M (0.279) and L (0.195) occurring at high frequencies. Allele H was absent, whereas alleles Q (0.005) and K (0.024) were rare. Seven alleles were identified in Group 2, with alleles M (0.306) and K (0.161) occurring most frequently. Alleles R, I, O, and Q were absent (Figure 1o).

HMS1 locus

In Group 1, homozygous individuals accounted for 28.95%, with genotype JJ being the most common (16.32%). Heterozygous individuals accounted for 71.05%, with genotype JM being the most frequent (32.63%). In Group 2, homozygous individuals accounted for 38.00%, with only two homozygous genotypes identified: MM (22.00%) and JJ (16.00%). Heterozygous individuals accounted for 62.00%, with genotype JM being the most frequent (40.00%).

Eight and seven alleles were identified in Groups 1 and 2, respectively. Alleles J and M occurred at high frequencies in Group 1 (0.434 and 0.355, respectively) and Group 2 (0.390 and 0.480, respectively). In Group 1, allele K (0.026) and allele 197 (0.003), which had no letter designation, were rare. In Group 2, alleles I and Q each had a frequency of 0.020, whereas alleles K (0.030) and L (0.010) were rare (Figure 1p).

CA425 locus

In Group 1, homozygous individuals accounted for 18.42%, with genotypes MM (8.95%) and NN (5.79%) being the most common. Heterozygous individuals accounted for 81.58%, with genotype MN being the most frequent (17.37%). In Group 2, homozygous individuals accounted for 20.00%, with genotype NN being the most common (12.00%). Heterozygous individuals accounted for 80.00%, with genotype MN being the most frequent (16.00%).

Ten alleles were identified in Group 1, with alleles M (0.324) and N (0.258) occurring at high frequencies. Alleles F (0.016), I (0.026), and P (0.005) were rare. Eight alleles were identified in Group 2, with alleles N (0.320) and M (0.250) occurring most frequently. Allele K was rare (0.010), whereas alleles G and P were absent (Figure 1q).

The genetic diversity parameters for the 17 microsatellite loci in the core and translocated Bashkir horse populations are presented in Table 1.

Table 1: Genetic diversity parameters and allelic richness for 17 microsatellite loci in the core and translocated populations of Bashkir horses

Locus	N	Na	AR	Ne	I	Ho	He	F	PIC
Group 1 (core population, n = 190)									
VHL20	190	10.000	9.060	7.337	2.101	0.900	0.864	-0.042	0.852
HTG4	190	9.000	6.235	2.624	1.331	0.579	0.619	0.065	0.483
AHT4	190	11.000	8.785	4.348	1.798	0.726	0.770	0.057	0.702
HMS7	190	7.000	6.300	3.247	1.468	0.695	0.692	-0.004	0.724
HTG6	190	7.000	6.068	3.061	1.415	0.705	0.673	-0.047	0.534
AHT5	190	8.000	6.336	4.350	1.635	0.779	0.770	-0.011	0.736
HMS6	190	8.000	6.509	4.306	1.628	0.774	0.768	-0.008	0.746
ASB23	190	10.000	8.182	6.072	1.944	0.858	0.835	-0.027	0.778
ASB2	190	11.000	9.299	7.443	2.119	0.874	0.866	-0.009	0.828
HTG10	190	11.000	9.339	4.534	1.880	0.742	0.779	0.048	0.833
HTG7	190	5.000	4.132	3.928	1.389	0.737	0.745	0.011	0.685
HMS3	190	8.000	7.553	6.136	1.919	0.853	0.837	-0.019	0.703
HMS2	190	9.000	7.564	4.523	1.717	0.747	0.779	0.040	0.730
ASB17	190	18.000	11.293	5.578	2.101	0.821	0.821	0.000	0.868
LEX3	190	10.000	8.037	5.951	1.974	0.489	0.832	0.412	0.846
HMS1	190	8.000	6.300	3.076	1.380	0.711	0.675	-0.053	0.539
CA425	190	10.000	7.857	4.802	1.796	0.816	0.792	-0.030	0.764
Mean ± SE	190 ± 0.00	9.412 ± 0.692	7.520 ± 0.388	4.783 ± 0.351	1.741 ± 0.066	0.753 ± 0.025	0.772 ± 0.017	0.022 ± 0.026	0.741 ± 0.020
Group 2 (translocated population, n = 50)									
VHL20	50	9.000	8.931	7.485	2.099	0.780	0.866	0.100	0.849
HTG4	50	5.000	4.694	2.089	1.025	0.500	0.521	0.041	0.587
AHT4	50	9.000	8.440	3.771	1.665	0.740	0.735	-0.007	0.743
HMS7	50	7.000	6.846	4.139	1.631	0.740	0.758	0.024	0.657
HTG6	50	6.000	5.317	2.421	1.148	0.540	0.587	0.080	0.638
AHT5	50	7.000	6.824	4.386	1.642	0.780	0.772	-0.010	0.737

Locus	N	Na	AR	Ne	I	Ho	He	F	PIC
HMS6	50	6.000	5.878	4.566	1.612	0.820	0.781	-0.050	0.732
ASB23	50	8.000	7.224	5.139	1.770	0.720	0.805	0.106	0.814
ASB2	50	11.000	9.602	6.527	2.044	0.820	0.847	0.032	0.850
HTG10	50	12.000	10.943	6.536	2.151	0.880	0.847	-0.039	0.758
HTG7	50	5.000	4.500	3.731	1.385	0.720	0.732	0.016	0.698
HMS3	50	7.000	6.904	3.743	1.600	0.700	0.733	0.045	0.817
HMS2	50	8.000	6.910	4.288	1.626	0.820	0.767	-0.069	0.745
ASB17	50	15.000	12.869	8.347	2.334	0.960	0.880	-0.091	0.800
LEX3†	31	7.000	8.658	5.555	1.833	0.903	0.820	-0.102	0.812
HMS1	50	7.000	5.856	2.585	1.177	0.620	0.613	-0.011	0.620
CA425	50	8.000	7.421	4.817	1.756	0.800	0.792	-0.010	0.764
Mean ± SE	48.9 ± 1.118	8.059 ± 0.627	7.451 ± 0.513	4.713 ± 0.422	1.676 ± 0.087	0.755 ± 0.029	0.756 ± 0.024	0.003 ± 0.015	0.744 ± 0.019

N = Sample size; Na = number of alleles; AR = allelic richness estimated by rarefaction; Ne = effective number of alleles; I = Shannon index; Ho = observed heterozygosity; He = expected heterozygosity; F = Wright's fixation index; PIC = polymorphic information content; SE = standard error. Values are presented as mean ± SE. †Only 31 animals from Group 2 were evaluated at the LEX3 locus.

Genetic diversity parameters

The number of alleles per locus ranged from 5 to 18 in Group 1 and from 5 to 15 in Group 2, with mean values of 9.41 and 8.06, respectively. In Group 1, the highest numbers of alleles were recorded at ASB17 (18 alleles), AHT4, ASB2, and HTG10 (11 alleles each), and VHL20, ASB23, LEX3, and CA425 (10 alleles each). In Group 2, the highest numbers of alleles were recorded at ASB17 (15 alleles), HTG10 (12 alleles), and ASB2 (11 alleles). The lowest number of alleles was recorded at HTG7 in both groups and at HTG4 in Group 2 (five alleles each). Thus, ASB17, ASB2, and HTG10 were among the principal loci contributing to allelic diversity in both populations. However, not all horses in Group 2 were successfully evaluated at the LEX3 locus.

Despite the difference in sample size between Group 1 and Group 2 (190 vs. 50 animals), the difference in the mean number of alleles was relatively small (9.41 vs. 8.06). To account for unequal sample sizes, AR was estimated using rarefaction. The mean AR values were 7.520 ± 0.388 in Group 1 and 7.451 ± 0.513 in Group 2, indicating comparable allelic diversity between the two subpopulations. Therefore, the difference in the observed number of alleles was likely influenced primarily by the unequal sample sizes rather than by a marked reduction in allelic diversity in the translocated population.

The Ne ranged from 2.62 at HTG4 to 7.44 at ASB2 in Group 1 and from 2.09 at HTG4 to 8.35 at ASB17 in Group 2. The mean Ne values were comparable between Group 1 and Group 2 (4.78 ± 0.35 and 4.71 ± 0.42 , respectively). Values below 3.5 were recorded at HTG4 (2.62 and 2.09), HTG6 (3.06 and 2.42), and HMS1 (3.08 and 2.59) in Groups 1 and 2, respectively, indicating an uneven distribution of allele frequencies at these loci. These loci also showed comparatively low Ho or PIC, indicating lower informativeness within the present dataset.

The Shannon index (I), which reflects overall genetic diversity, averaged 1.74 ± 0.07 in Group 1 and 1.68 ± 0.09 in Group 2, indicating comparable genetic diversity between the two populations. The highest values in Group 1 were recorded at ASB2 (2.12), VHL20 (2.10), and ASB17 (2.10), whereas the highest values in Group 2 were recorded at ASB17 (2.33) and HTG10 (2.15). The lowest values in both groups were recorded at HTG4 (1.33 and 1.03, respectively), consistent with the low Ne at this locus.

Ho ranged from 0.49 to 0.90 in Group 1 and from 0.50 to 0.96 in Group 2, whereas He ranged from 0.62 to 0.87 and from 0.52 to 0.88, respectively. Mean heterozygosity values differed only slightly between the subpopulations (Ho: 0.75 ± 0.03 vs. 0.76 ± 0.03 ; He: 0.77 ± 0.02 vs. 0.76 ± 0.02). High He (He > 0.80) was recorded at VHL20, ASB2, ASB17, LEX3, and ASB23 in both groups. Low He (He < 0.65) was recorded at HTG4 in both groups and at HTG6 and HMS1 in Group 2. In Group 2, several loci, including HMS6, ASB17, and HMS2, showed higher observed than He, as reflected by negative Wright's fixation index (F) values, indicating an excess of heterozygotes at these loci.

Wright's fixation index ranged from -0.053 to 0.412 in Group 1 and from -0.102 to 0.106 in Group 2, with mean values of 0.022 ± 0.026 and 0.003 ± 0.015 , respectively. Most values were close to zero, indicating little evidence of systematic inbreeding in either subpopulation. However, LEX3 in Group 1 showed a high positive F value (0.412), indicating a pronounced heterozygote deficit that was consistent with the detection of null alleles at this locus. Negative F values, indicating heterozygote excess, were recorded at VHL20, HTG6, AHT5, HMS6, ASB23, ASB2, HMS3, HMS1, and CA425 in Group 1 and at AHT4, AHT5, HMS6, HTG10, ASB17, LEX3, HMS2, HMS1, and CA425 in Group 2. Most of these deviations were small.

The PIC values ranged from 0.48 to 0.87 in Group 1 and from 0.59 to 0.85 in Group 2. The mean PIC value

was 0.74 ± 0.02 in both groups, indicating that the microsatellite panel was highly informative for population genetic analysis. In Group 1, PIC values >0.80 were recorded at VHL20, ASB2, HTG10, ASB17, and LEX3. In Group 2, PIC values >0.80 were recorded at VHL20, ASB23, ASB2, HMS3, and LEX3. The lowest PIC value in Group 1 was recorded at HTG4 (0.483), followed by HTG6 (0.534) and HMS1 (0.539). In Group 2, the lowest PIC value was also recorded at HTG4 (0.587). These findings indicate that a high level of genetic variability was retained in the translocated population.

The results of testing genotype distributions for conformity to HWE are presented in Table 2.

Table 2: Testing of genotype distributions for conformity to Hardy–Weinberg equilibrium in Bashkir horse populations

Locus	χ^2	df	p-value	p < 0.05	Bonferroni ($\alpha = 0.00294$)
VHL20	42.10	45	0.5955	ns	ns
HTG4	33.98	36	0.5651	ns	ns
AHT4	47.96	55	0.7381	ns	ns
HMS7	16.34	21	0.7505	ns	ns
HTG6	10.17	21	0.9766	ns	ns
AHT5	35.61	28	0.1529	ns	ns
HMS6	24.90	28	0.6335	ns	ns
ASB23	115.76	55	3.21×10^{-6}	***	***
ASB2	50.67	66	0.9185	ns	ns
HTG10	75.91	66	0.1893	ns	ns
HTG7	9.47	10	0.4879	ns	ns
HMS3	23.11	28	0.7276	ns	ns
HMS2	42.67	45	0.5712	ns	ns
ASB17	325.92	171	9.46×10^{-12}	***	***
LEX3	326.64	66	<0.0001	***	***
HMS1	10.20	21	0.9762	ns	ns
CA425	27.29	45	0.9829	ns	ns

χ^2 = Chi-square statistic; df = degrees of freedom; ns = not significant. Significance levels: p < 0.05 *** p < 0.001, ** p < 0.01, * p < 0.05. The Bonferroni-corrected significance threshold was $\alpha = 0.05/17 = 0.00294$.

The HWE analysis with Bonferroni correction ($\alpha = 0.05/17 = 0.00294$) identified significant deviations from equilibrium at three loci: ASB23 ($\chi^2 = 115.76$, df = 55, $p = 3.21 \times 10^{-6}$), ASB17 ($\chi^2 = 325.92$, df = 171, $p = 9.46 \times 10^{-12}$), and LEX3 ($\chi^2 = 326.64$, df = 66, $p < 0.0001$). The remaining 14 loci did not deviate significantly from HWE after Bonferroni correction.

The deviation observed at LEX3 is consistent with the presence of null alleles detected during quality control. In contrast, the deviations observed at ASB23 and ASB17 may reflect biological processes, such as selection or other population-level factors, although the present data do not allow these mechanisms to be confirmed. Estimates of F-statistics for the analyzed STR loci are presented in Table 3.

The mean inbreeding coefficients; inbreeding coefficient within subpopulations (FIS; 0.012) and overall inbreeding coefficient (FIT; 0.022), were close to zero, indicating a low level of inbreeding and effective breeding management that has limited the accumulation of homozygosity. The low FST value (0.010) indicates a high degree of genetic homogeneity, suggesting that the studied Bashkir horse populations exhibit minimal genetic substructuring and collectively represent a single genetic system.

The migration index, estimated ($Nm = 32.4$), also indicates the maintenance of genetic diversity within the breed. Despite these generally favorable findings, several loci warrant attention. In particular, LEX3 exhibited a comparatively high FIS value (0.157), which is consistent with the presence of null alleles detected during quality control. Negative FIS values at HMS6, HMS2, and ASB17 indicate an excess of heterozygotes at these loci.

PCoA of the studied populations is presented in Figure 2. The translocated population formed a relatively compact cluster within the broader distribution of the core population. Although the centroids of the two groups were separated along the first principal coordinate, the degree of separation was limited. PERMANOVA detected statistically significant differences between the two populations ($F = 2.34$, $p = 0.042$), whereas no significant differences in group dispersion were observed ($p = 0.38$). These findings indicate that, despite minor genetic differentiation, both subpopulations retain substantial genetic similarity consistent with their common breed origin.

The low level of differentiation between the two Bashkir horse populations was further supported by bootstrap analysis of the mean FST values (Figure 3).

Table 3: Estimates of F-statistics for STR loci in the Bashkir horse breed

Locus	FIS	FIT	FST	Nm
VHL20	0.029	0.039	0.010	24.683
HTG4	0.054	0.062	0.009	27.468
AHT4	0.026	0.033	0.008	31.730
HMS7	0.011	0.021	0.011	23.305
HTG6	0.012	0.021	0.010	26.048
AHT5	-0.011	0.016	0.026	9.301
HMS6	-0.029	-0.024	0.005	49.104
ASB23	0.038	0.043	0.004	55.546
ASB2	0.011	0.018	0.007	35.104
HTG10	0.003	0.010	0.007	35.477
HTG7	0.014	0.019	0.005	46.113
HMS3	0.011	0.025	0.014	17.987
HMS2	-0.014	-0.011	0.003	78.168
ASB17	-0.047	-0.031	0.015	15.900
LEX3	0.157	0.170	0.016	15.369
HMS1	-0.033	-0.024	0.009	28.588
CA425	-0.020	-0.012	0.008	30.921
Mean ± SE	0.012 ± 0.011	0.022 ± 0.011	0.010 ± 0.001	32.401 ± 4.131

FIS = Inbreeding coefficient within subpopulations; FIT = overall inbreeding coefficient; FST = fixation index among subpopulations; Nm = estimated number of migrants (gene flow).

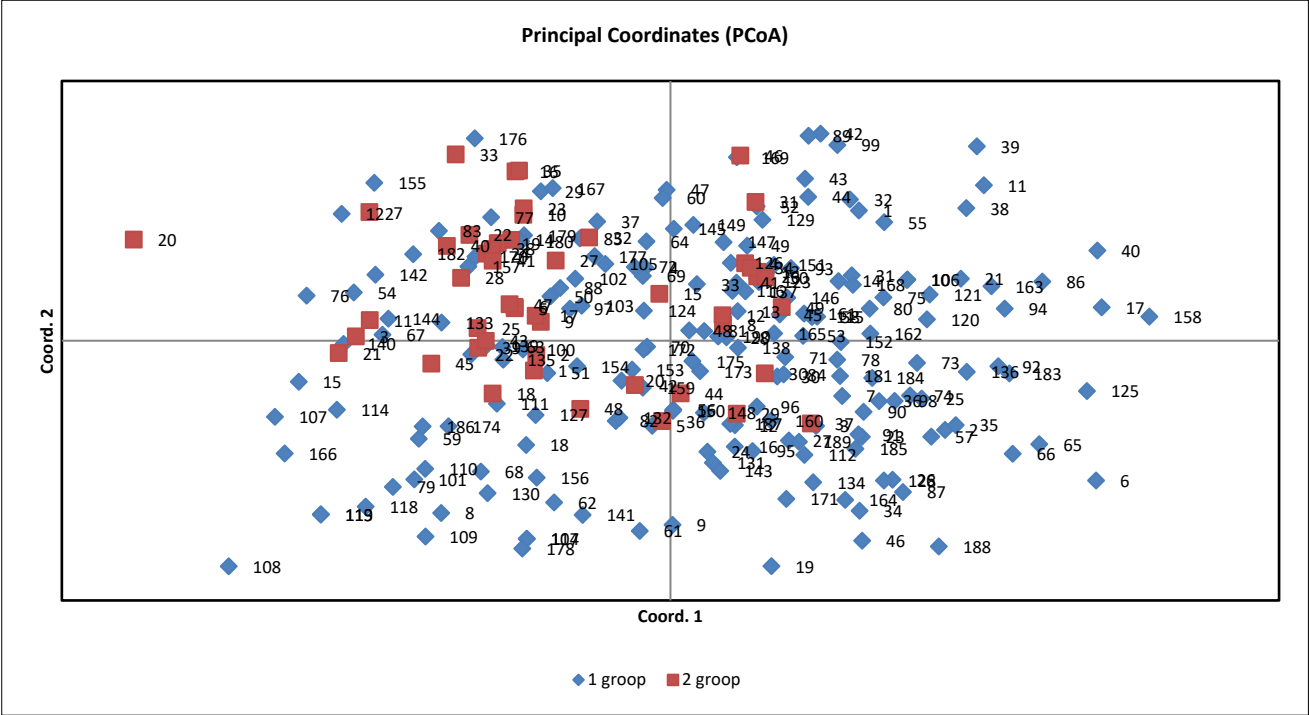


Figure 2: Clustering of the core (Group 1, n = 190) and translocated (Group 2, n = 50) populations of Bashkir horses.

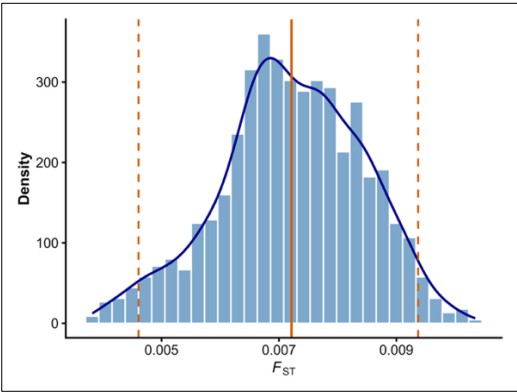


Figure 3: Bootstrap distribution of mean FST based on 1,000 iterations.

Bootstrap analysis (1,000 iterations) confirmed an extremely low level of genetic differentiation within the Bashkir horse breed. The mean F_{ST} value obtained by bootstrapping was 0.0072 ± 0.0012 (95% confidence interval [CI]: 0.0045–0.0093), indicating that only approximately 0.7% of the total genetic variation was attributable to differences between the two populations and indicates a low but statistically supported level of genetic differentiation. The overall point estimate $F_{ST} = 0.010$ is provided for reference; the bootstrap mean was used for the main interpopulation comparison as it is more robust and accompanied by an uncertainty estimate. The narrow 95% CI and the symmetrical distribution of bootstrap estimates demonstrate the robustness of the analysis. Even after accounting for statistical uncertainty, the proportion of between-population genetic variation remained below 1%, confirming the close genetic relationship between the core and translocated populations.

The genetic structure of the Bashkir horse population was further investigated using Bayesian clustering implemented in STRUCTURE. The distribution of individual membership coefficients (Q matrix) is presented in Figure 4.

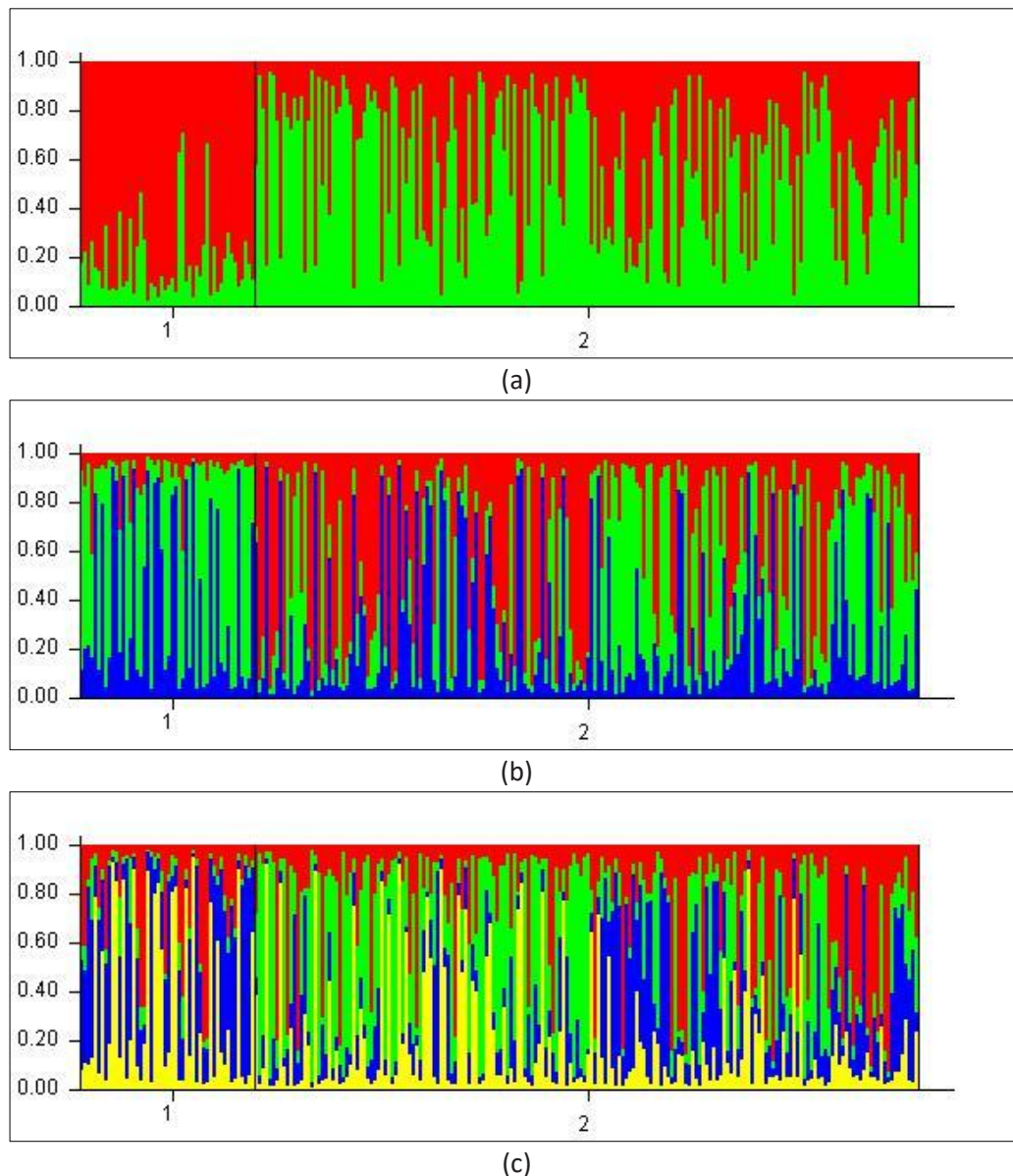


Figure 4: Bayesian clustering of the genetic structure of the Bashkir horse breed: a) division into two genetic clusters; b) division into three genetic clusters; c) division into four genetic clusters.

The $K = 2$ model showed a clear division of individuals into two principal genetic groups. At $K = 3$, an additional cluster emerged that was primarily represented by a subset of individuals within one of the major clusters. Increasing the number of clusters to $K = 4$ did not reveal additional stable genetic groups; instead, it subdivided the existing clusters.

The optimal number of genetic clusters was determined using the method of Evanno *et al.* [37], and the results are presented in Figure 5.

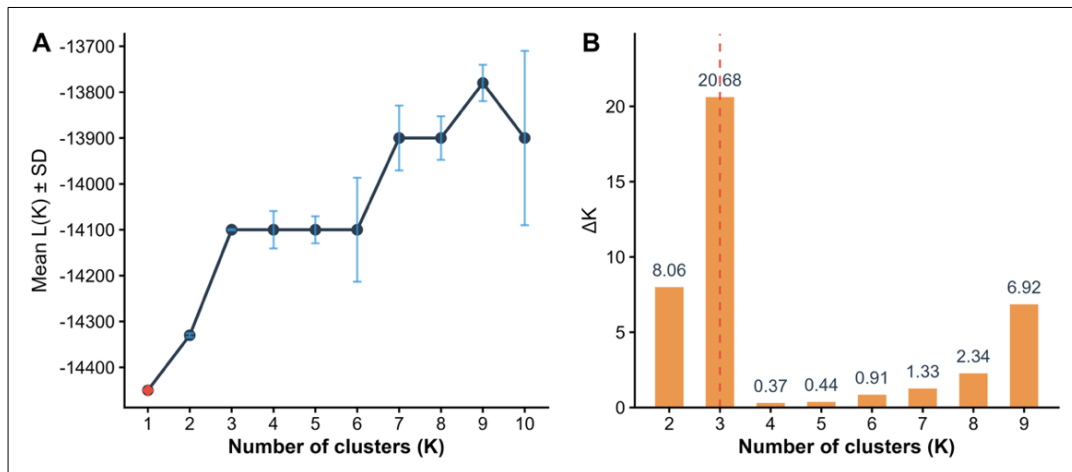


Figure 5: Determination of the optimal number of genetic clusters using the Evanno method. (A) Mean log-likelihood values (Mean $L(K) \pm \text{SD}$) for K values ranging from 1 to 10. Error bars represent SD. (B) ΔK values showing the rate of change in log-likelihood between successive K values normalized by SD.

The ΔK statistic reached its maximum at $K = 3$ ($\Delta K = 120.68$), substantially exceeding the values obtained for all other K values. This model also exhibited a high mean log-likelihood (Mean $\ln P[D] = -14098.42$) and low variation among replicate runs ($\text{SD} = 1.32$). However, a marked increase in mean log-likelihood was observed between $K = 1$ (-14419.72) and $K = 2$ (-14277.44), indicating the presence of a primary genetic subdivision within the dataset. The $K = 2$ model also showed high stability among replicate analyses ($\text{SD} = 4.56$). Considering the study design, which compared the core and translocated Bashkir horse populations, the $K = 2$ model provides biologically meaningful population separation. Nevertheless, the Evanno method clearly identified $K = 3$ as the optimal clustering solution, suggesting additional genetic subdivision within the core Bashkir horse population. This subdivision may reflect geographic separation between the Pre-Ural and Trans-Ural populations and differences in historical breeding practices. Therefore, although the two-cluster model adequately represents the primary population structure evaluated in this study, the three-cluster model explains a greater proportion of the observed genetic variation within the Bashkir horse breed.

DISCUSSION

Local horse breeds represent an important component of the biological and cultural heritage of many countries. Having evolved over centuries under diverse environmental conditions and traditional breeding systems, these breeds possess unique genetic characteristics, including adaptation to harsh climates, resistance to specific diseases, and relatively low management requirements [4, 6, 8, 12]. Consequently, indigenous horse breeds contribute not only to livestock production but also to food security, biodiversity conservation, cultural heritage, and sustainable agricultural development.

The Bashkir horse is an indigenous breed that originated from local horse populations through the assimilation of horses maintained by Turkic-speaking tribes and has historically been managed under year-round herd-based grazing systems [1, 3]. Consequently, many of its phenotypic and productive characteristics have been shaped by long-term adaptation to local climatic and environmental conditions [1].

Monitoring genetic diversity using microsatellite markers provides valuable information on population structure, genetic variability, and breeding management. Allelic composition reflects the genetic architecture of a population, whereas rare alleles represent low frequency genetic variants that contribute to the maintenance of overall genetic diversity [12, 21, 24, 27].

Genetic diversity of the Bashkir horse populations

The present study provides the first comprehensive comparison of the genetic diversity and population structure of the core Bashkir horse population from the Republic of Bashkortostan and the recently translocated population from the Tver Region using 17 ISAG-recommended microsatellite loci. Unlike previous studies that were limited to descriptive statistics or univariate comparisons, the present investigation combined PCoA, PERMANOVA, and Bayesian clustering using STRUCTURE with the Evanno method to evaluate both the magnitude and statistical significance of genetic differentiation between the two populations. This integrated analytical approach enabled the detection of subtle genetic differentiation that would not have been apparent using

descriptive genetic indices alone.

The study demonstrated a high level of polymorphism across the analyzed loci in both populations. A total of 160 alleles were identified in the core population and 137 in the translocated population, with mean numbers of alleles per locus of 9.41 and 8.06, respectively. The highest levels of polymorphism were consistently observed at ASB17, ASB2, and HTG10, whereas HTG7 exhibited comparatively low allelic diversity in both populations and HTG4 showed low diversity in the translocated population. Similar locus-specific variation has been reported in other indigenous horse breeds, indicating that allele frequency differences largely reflect historical breeding practices, founder effects, and genetic drift rather than recent genetic erosion [37–42]. Comparable studies of Kazakh horse populations identified VHL20 as one of the most polymorphic loci, whereas HTG4, AHT4, AHT5, HTG7, and HMS3 also exhibited high variability across populations [21]. Likewise, previous studies of Bashkir horses reported high frequencies of alleles HTG10O, HTG6O, HTG4M, and HMS7L together with the rare ASB17U allele, findings that are consistent with the present results [12].

Although the observed number of alleles exceeded the N_e at most loci, this pattern is expected in natural populations where a limited number of alleles occur at high frequencies while numerous alleles remain rare. Such allele frequency distributions have previously been associated with founder effects, historical selection, and genetic drift without necessarily indicating a reduction in overall genetic diversity [38].

Despite the smaller sample size of the translocated population, AR remained nearly identical between the two populations after rarefaction, demonstrating that the apparent reduction in the observed number of alleles primarily reflected sample size rather than a true reduction in genetic diversity. Similarly, H_e and PIC remained high in both populations, indicating that the analyzed marker panel was highly informative. The comparable values of the N_e , Shannon index, heterozygosity, and AR indicate that the translocated population has genetic diversity indicators similar to the core population. These findings agree with previous reports describing indigenous horse breeds as maintaining substantial genetic diversity despite differences in breeding history and management practices [42–49].

The values of Wright's F-statistics further support this conclusion. Mean F_{IS} and F_{IT} values remained close to zero, indicating little evidence of inbreeding within either population. Likewise, the low mean F_{ST} value indicates only minimal differentiation between the two populations and is comparable with previous findings reported for indigenous horse breeds in Russia [12, 27].

Comparison with previous investigations suggests that the genetic composition of the Bashkir horse has remained relatively stable over time. Earlier studies involving Bashkir horses reported similar numbers of alleles per locus together with comparable heterozygosity and Wright's F-statistics [12, 46]. Minor increases in heterozygosity and slight reductions in Wright's F-statistics observed in the present study may reflect continued breeding management; however, the overall genetic characteristics remain highly consistent with previous reports. Similar levels of heterozygosity have also been reported for other indigenous horse breeds, including Mugalzhar, Altai, Kyrgyz, and Yakut horses [12].

Genetic differentiation between the core and translocated populations

The combined application of PCoA, PERMANOVA, bootstrap analysis, and Bayesian clustering consistently demonstrated that genetic differentiation between the core and translocated populations remains extremely limited.

PCoA revealed only weak separation between the two populations. The first principal coordinates explained only a small proportion of the total genetic variation, whereas most variation remained distributed within rather than between populations. Such a pattern is characteristic of populations connected by substantial gene flow and shared ancestry.

PERMANOVA detected statistically significant but biologically weak differentiation between the two populations. Only a very small proportion of the total genetic variation was explained by population membership, whereas almost all genetic variability remained within populations. Together with the low F_{ST} value obtained from the bootstrap analysis, these findings demonstrate that the translocated population has retained the genetic characteristics of the original Bashkir population despite approximately two decades of geographic separation.

Bayesian clustering provided additional insight into population structure. The Evanno method identified $K = 3$ as the statistically optimal number of clusters, whereas $K = 2$ corresponded directly to the biological design of the study by separating the core and translocated populations. The additional cluster detected at $K = 3$ most likely reflects subtle subdivision within the geographically widespread core population rather than the emergence of a

genetically distinct lineage. Therefore, the clustering results complement the PCoA, PERMANOVA, and FST analyses by indicating weak within-breed subdivision rather than substantial differentiation between populations.

Taken together, these complementary analytical approaches demonstrate that the translocated Bashkir horse population has successfully retained the genetic diversity and overall genetic structure of the indigenous breed following relocation. The observed differences represent localized variation within a genetically unified breed rather than meaningful population divergence.

Implications for conservation and breeding

Although SNP-based technologies are increasingly used in animal genetics, STR markers remain reliable, highly informative, and internationally accepted tools for parentage verification, genetic identification, and population monitoring in horses. The present findings demonstrate that both the indigenous and translocated Bashkir horse populations retain high levels of genetic diversity with little evidence of inbreeding or substantial genetic subdivision.

Routine genetic monitoring using standardized STR marker panels should therefore continue to support breeding management by verifying pedigrees, minimizing inbreeding during mating decisions, monitoring genetic diversity, maintaining accurate breeding records, and guiding long-term conservation strategies for the Bashkir horse breed.

CONCLUSION

This study demonstrated that both the core Bashkir horse population from the Republic of Bashkortostan and the translocated population from the Tver Region retain a high level of genetic diversity despite approximately two decades of separate breeding. Analysis of 17 STR loci identified 160 alleles in the core population and 137 alleles in the translocated population, while comparable AR, heterozygosity, and PIC confirmed that the translocated population has maintained its genetic variability. The low mean FST value, together with the results of PCoA, PERMANOVA, bootstrap analysis, and Bayesian clustering, indicated only weak genetic differentiation between the two populations, suggesting that the translocated population has preserved the overall genetic structure and identity of the Bashkir horse breed.

These findings have important practical implications for conservation genetics and breeding management. The results indicate that the current breeding strategy has effectively maintained genetic diversity while minimizing inbreeding, supporting the continued use of the translocated population as a valuable genetic resource. Routine genetic monitoring using standardized STR marker panels should be incorporated into breeding programs to verify parentage, monitor genetic diversity, guide mating decisions, prevent inbreeding, and support the long-term conservation of the Bashkir horse gene pool.

A major strength of this study is the comprehensive evaluation of two geographically separated populations using a relatively large sample size and multiple complementary analytical approaches. The integration of descriptive genetic indices with PCoA, PERMANOVA, bootstrap estimation of FST, and Bayesian clustering provided a robust assessment of within-breed genetic diversity and population structure that extends beyond previous descriptive studies of the Bashkir horse.

Several limitations should be considered when interpreting the findings. The study evaluated genetic diversity using 17 ISAG-recommended STR loci, which primarily represent neutral genetic variation and therefore do not directly assess functional genomic variation associated with adaptive or economically important traits. In addition, the translocated population originated relatively recently, which may have limited the extent of detectable genetic divergence. The unequal sample sizes between the two populations, although addressed using rarefaction, also represent a methodological constraint. Furthermore, genomic analyses based on high-density SNP markers or whole-genome sequencing were not performed and could provide higher-resolution insight into population history, selection signatures, and adaptive variation.

Future studies should integrate STR-based monitoring with genome-wide SNP or whole-genome sequencing approaches, expand sampling to additional Bashkir horse populations maintained in other regions, and combine molecular data with pedigree, phenotypic, reproductive, and production records. Longitudinal genetic monitoring will also be valuable for assessing temporal changes in genetic diversity and evaluating the long-term effectiveness of conservation breeding programs.

Overall, the present study demonstrates that the translocated Bashkir horse population has successfully preserved the genetic diversity and genetic integrity of the indigenous breed while exhibiting only limited within-

breed differentiation. These findings provide a scientific basis for evidence-based conservation, sustainable breeding strategies, and long-term genetic management of one of Russia's most valuable indigenous horse breeds.

DATA AVAILABILITY

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

GENERATIVE AI DECLARATION

The authors declare that generative artificial intelligence (AI) tools were used solely to improve the language, grammar, and readability of the manuscript. All scientific concepts, study design, data collection, data analysis, interpretation of the results, and conclusions were developed, verified, and approved by the authors. The authors accept full responsibility for the accuracy, integrity, and originality of the work, and no AI tool was listed as an author.

AUTHORS' CONTRIBUTIONS

STA, VFR, and GRR: Conceptualization, methodology, investigation, supervision, writing—review and editing. GDI, LEN, and GIS: Investigation, methodology, validation, and writing—review and editing. BBB, FCD, and CRS: Investigation, methodology, data curation, and writing—review and editing. ZAM and GEA: Conceptualization, methodology, investigation, and writing—review and editing. All authors have read and approved the final manuscript.

ACKNOWLEDGMENTS

The authors thank the participating horse breeding farms for providing access to biological samples and zootechnical records used in this study. This work was supported by the state assignment entitled “Development of economically beneficial and technological criteria for the improvement of agricultural animals and birds in the conditions of the Republic of Bashkortostan” (Project No. 125020301276-9).

COMPETING INTERESTS

AVV is affiliated with SNAIP Limited Liability Company, and horses from the Tver population were maintained at LLC SNAIP. However, SNAIP had no role in the study at any stage.

PUBLISHER'S NOTE

Veterinary World remains neutral with regard to jurisdictional claims in published institutional affiliations.

REFERENCES

1. Akhatova I. The past, present, and future of the Bashkir horse. *Vatandash*. 2022;9(312):86–92.
2. Sedykh TA, Gizatullin RR, Gareeva DI, Farkhutdinov KD, Gavrilicheva IS, Zaitsev AM, *et al*. Characteristics of genetic polymorphism in Bashkir horses. *Zootech*. 2024;10:14–19.
3. Sedykh TA, Gizatullin RR, Valitov FR, Farkhutdinov KD, Gareeva DI, Latypova EN. Status of the breeding base and prospects for the development of horse breeding in Bashkortostan. *Izv Oren St Agr Univ*. 2024;4(108):294–299.
4. Soloshenkova EA, Voronkova VN, Stolpovskii YuA, Soloshenkov AD. Population structure of indigenous horses in Russia based on microsatellite markers. *Horse Breed Eq Sport*. 2025;1:13–15.
5. Bazon BZ. Genetic structure of herd horses in Transbaikalia. *Horse Breed Eq Sport*. 2021;2:22–25.
6. Voronkova VN, Nikolaeva EA, Piskunov AK, Babayan OV, Takasu M, Tozaki T, *et al*. Assessment of genetic diversity and structure of autochthonous horse breeds in Russia and Mongolia using nuclear and mitochondrial DNA markers. *Genetika*. 2022;8(58):902–919.
7. Gulyaeva AN, Bass SP. Productive traits and biological characteristics of indigenous horse breeds under the conditions of the Udmurt Republic. *Vest Izh St Agric Acad*. 2024;2(78):89–96.
8. Khrabrova LA, Blokhina NV, Belousova NF, Kotran EG. Assessment of the genealogical structure of the Vyatka horse breed (*Equus ferus caballus*) using DNA analysis. *Genetika*. 2022;4(58):457–462.
9. Bass SP, Belousova NF, Atnabaeva NA. Genetic methods in line breeding of Vyatka horses using DNA

- microsatellites. *Izv Oren St Agrar Univ.* 2023;1(99):312–317.
10. Masuda M, Nikadori A, Nikadori E, Uchida N, Takizawa Y, Ishimaru M, et al. Evaluation of genetic diversity using 31 microsatellites in Miyako horses. *J Eq Sci.* 2023;34:1–6.
 11. Avila F, Esdaile E, Bellone RR. Using STR data to investigate the impact of the studbook cap on genetic diversity in the American Standardbred horse from 1998 to 2021. *Genes (Basel).* 2025;16(7):748.
 12. Blokhina NV, Khrabrova LA. Genetic characterization of local horse breeds based on DNA microsatellite loci. *Vavil J Genet Breed.* 2025;1(29):113–121.
 13. Mahrous K, Hassanane M, Mordy M, Shafei H, Hassan N. Genetic variations in horse using microsatellite markers. *J Genet Eng Biotechnol.* 2011;9:103–109.
 14. Rudak AN, German AI, German Yul, Gorbukov MA, Chavlytko VI. Genetic characterization of Trakehner horses based on DNA microsatellite locus polymorphism. *Akt Prob Int Razv Zhivotn.* 2022;25-1:23–30.
 15. Toktosunov BI, Baimukanov DA, Abdurasulov AKh, Mongush SD. STR locus genotype polymorphism in major stallion lines of the indigenous Kyrgyz horse. *Vest Chuv St Agrar Univ.* 2022;3(22):74–82.
 16. Cosenza M, La Rosa V, Rosati R, Chiofalo V. Genetic diversity of the Italian thoroughbred horse population. *Ital J Anim Sci.* 2019;18(1):1–8.
 17. Nolte W, Alkhoder H, Wobbe M, Stock KF, Kalm E, Vosgerau S, et al. Replacement of microsatellite markers by imputed medium-density SNP arrays for parentage control in German warmblood horses. *J Appl Genet.* 2022;63(4):783–792.
 18. Kim D, Lee S, Oyungerel B, Cho G. Evaluation of the effectiveness of single-nucleotide polymorphisms versus microsatellites for parentage verification in horse breeds. *Vet Sci.* 2025;12:890.
 19. Dorofeeva AV, Podobaeva NV, Samandeeva EG, Gavrilicheva IS. Influence of founder breeds on the genetic structure of the Tersk horse breed. *Horse Breed Eq Sport.* 2024;2:37–40.
 20. Vdovina NV, Yuryeva IB. Genetic structure of the herd at the gene pool farm for breeding Mezen horses of SPK RK “Sever”. *Horse Breed Eq Sport.* 2022;6:23–26.
 21. Orazymbetova Z, Ualiyeva D, Dossybayev K, Torekhanov A, Sydykov D, Mussayeva A, et al. Genetic diversity of Kazakhstani *Equus caballus* (Linnaeus, 1758) horse breeds inferred from microsatellite markers. *Vet Sci.* 2023;10(10):598.
 22. Osipov VG, Alferov IV, Ivanov RV, Vinokurov IE, Pak MN. Search for genetic markers for use in the breeding of productive horse breeds in Yakutia (review). *Hippol Vet Med.* 2023;4(50):250–261.
 23. Dubrovin AV, Gavrilicheva IS. Genetic diversity of modern populations of the Novoultai horse breed. *Horse Breed Eq Sport.* 2023;6:23–27.
 24. Baena MM, Diaz S, Moura RS, Meirelles SLC. Genetic characterization of Mangalarga Marchador breed horses based on microsatellite molecular markers. *J Equine Vet Sci.* 2020;95:103231.
 25. Fornal A, Kowalska K, Zabek T, Piestrzynska-Kajtoch A, Musiał AD, Ropka-Molik K. Genetic variability and population structure of Polish Konik horse maternal lines based on microsatellite markers. *Genes (Basel).* 2021;12(4):546.
 26. Kassymbekova ShN, Iskhan K, Rzabaev S, Bimenova Zh, Kabylbekova D, Tursunkulov S. Assessment of genetic diversity using microsatellite markers and milk productivity of Mugalzhar horses. *Vet Sci.* 2024;3(7):29–36.
 27. Kalashnikov V, Khrabrova L, Blohina N, Zaitsev AM, Kalashnikova T. Dynamics of the inbreeding coefficient and homozygosity in Thoroughbred horses in Russia. *Animals.* 2020;10(7):1217.
 28. Van Osterhout C, Hutchinson WF, Wills DP, Shipley P. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes.* 2004;4(3):535–538.
 29. Peakall R, Smouse PE. GenALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics.* 2012;8(19):2537–2539.
 30. Peakall R, Smouse PE. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes.* 2006;6(1):288–295.
 31. Belkhir K, Dawson KJ, Bonhomme F. A comparison of rarefaction and Bayesian methods for predicting the allelic richness of future samples on the basis of currently available samples. *J Hered.* 2006;97(5):483–492.
 32. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet.* 1980;32:314–331.
 33. Paradis E. Pegas: an R package for population genetics with an integrated—modular approach. *Bioinformatics.* 2010;26(3):419–420.
 34. Anderson MJ. Permutational multivariate analysis of variance (PERMANOVA). *Wiley StatsRef: Statistics*

Reference Online. 2014:1–15.

35. Oksanen J, Simpson GL, Blanchet FG, et al. Vegan: community ecology package. R package version 2.7-5. 2024.
36. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945–59.
37. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 2005;14(8):2611–2620.
38. Francis RM. Pophelper: an R package and web app to analyse and visualize population structure. *Mol Ecol Resour*. 2017;17:27–32.
39. Luís C, Cothran EG, Oom MM. Inbreeding and genetic structure in the endangered Sorraia horse breed: implications for its conservation and management. *J Hered*. 2007;98(3):232–237.
40. Kalinkova LV, Osipov VG, Sleptsov ES, Fedorov VI, Vinokurov NV, Pavlova AI. Genetic structure and blood biochemical composition of horses in stallion lines of the Prilenskaya breed and the indigenous type of the Yakut breed. *Hippol Vet Med*. 2020;2(36):34–40.
41. Seo JH, Park KD, Lee HK, Kong HS. Genetic diversity of Halla horses using microsatellite markers. *J Anim Sci Technol*. 2016;58:40.
42. Putnova L, Štohl R, Vrtkova I. Using nuclear microsatellite data to trace the gene flow and population structure in Czech horses. *Czech J Anim Sci*. 2019;64:67–77.
43. An J, Tseveen K, Oyungerel B, Kong HS. Analysis of genetic diversity and structure of Mongolian horse using microsatellite markers. *J Anim Sci Technol*. 2022;64(6):1226–1236.
44. Yoshihara T, Kikuchi M, Akiba Y, Horiguchi M, Takasu M, Tozaki T. Genetic diversity analysis and parentage verification of Taishu horses using 31 microsatellites. *J Equine Sci*. 2022;33(4):63–69.
45. Nakamura M, Tozaki T, Kakoi H, Nakamura K, Rajabi-Toustani R, Ohba Y, et al. Decreased genetic diversity in Kiso horses revealed through annual microsatellite genotyping. *J Vet Med Sci*. 2020;82(4):503–540.
46. Khrabrova L, Blohina N, Suleymanov O, Rozhdestvenskaya G, Pustovoy V. Assessment of line differentiation in the thoroughbred horse breed using DNA microsatellite loci. *Vavil J Genet Breed*. 2019;5(23):569–574.
47. Dolmatova IYu, Niyatshin FI, Urazbakhtin RF. Population genetic characterization of Bashkir horses based on DNA microsatellites. *Horse Breed Eq Sport*. 2017;4:17–19.
48. Liu G, Shafer ABA, Zimmermann W, Hu D, Wang W, Chu H, et al. Evaluating the reintroduction project of Przewalski's horse in China using genetic and pedigree data. *Biol Conserv*. 2014;171:288–298.
49. Takasu M, Hiramatsu N, Tozaki T, Kakoi H, Nakagawa T, Hasegawa T, et al. Genetic characterization of the endangered Kiso horse using 31 microsatellite DNAs. *J Vet Med Sci*. 2012;74(2):161–166.
