Intra- and interbreed genetic heterogeneity and divergence in four commercial pig breeds based on microsatellite markers

E. A. Snegin*, A. S. Kramarenko**, O. Y. Artemchuk*, S. S. Kramarenko**

*Belgorod National Research University, Belgorod, Russia
**Mykolaiv National Agrarian University, Mykolaiv, Ukraine

In recent years, there has been an increasing amount of attention paid to the genetic health of domesticated animals and its relationship with the level of inbreeding and genetic diversity. At the same time, insufficient attention is still paid to the study of intra- and interbreed genetic diversity and intra- and interbreed stratification. The main goal of our work was to analyze the intra- and interbreed genetic diversity of commercial pig breeds on the basis of DNA microsatellite (MS-DNA) polymorphism. In total, the work used data for 3,308 pigs, which represented 11 herds. The animals belonged to four commercial pig breeds – Duroc (DR), Yorkshire (YR), Landrace (LN) and Large White (LW). 12 microsatellite loci recommended by ISAG-FAO and arranged in one multiplex panel (S0101, S0155, S0228, S0355, S0386, SW24, SW240, SW72, SW857, SW911, SW936, SW951) were used as DNA markers. When analyzing the intra- and interbreed variability of 11 herds, we found that all studied breeds significantly differed in terms of the proportion of both rare and common alleles. At the same time, the noted differences were determined, first of all, by the variability between individual herds within their breed. The location of herd centroids is random and is not consistent with their breed affiliation at all. When individuals belonging to the same breed are combined, the centroids of pig breeds in the space of first two axes from a Principal Coordinate Analysis form two clusters. The first one contains the only red pig breed (DR) used in the analysis, while the second one contains white pig breeds. In six pig herds the Ne estimates were below 50 inds., in two herds they were in the range of 50–100 inds., and finally in three herds the Ne estimates exceeded 100 inds. The analysis of the genetic variability of pigs of four commercial breeds showed that the high level of interbreed differences is caused, first of all, by the high variability among pig herds within each studied breed. Such intra- and interbreed stratification can be formed due to the manifestation of many causes: different genetic basis of the founders of intra- and interbreed genealogical groups, geographical isolation, different directions of selection within individual herds, exchange of animals between separate herds, the use of inbreeding in the practice of selection together with isolation, etc. Important consequences of intra- and interbreed stratification are an increase in the level of interherd diversity (which is not lower than the level of interbreed diversity) against the background of a decrease in variability within individual herds, as well as a significant deficit of heterozygotes and an increase in the role of negative genetic and demographic processes. Thus, the existence of genetic heterogeneity within commercial pig breeds should be considered as an essential element in the history of their formation and breeding.

Keywords: microsatellite DNA loci; intra- and interbreed stratification; commercial pig breeds.

Introduction

In recent years, an increasing amount of attention has been paid to the genetic health of domesticated animals and its relationship with the level of inbreeding and genetic diversity. This problem is especially manifested in those species of farm animals for which there are negative consequences of the founder effect in the creation of breeds, extensive use of sires and intensive selection, which increase the possibility of genetic abnormalities. An important condition for preventing these negative consequences should be constant genetic monitoring of genetic diversity, intra- and interbreed structure and the degree of inbreeding in the breeds of the main types of farm animals, including pigs (Wiener et al., 2017). In addition, obtaining an adequate evaluation of the genetic structure and diversity of breeds (especially endangered and/or local breeds) is a background for developing a strategy for their conservation (Wilkinson et al., 2011).

Obtaining as complete and reliable information as possible about the level of intra- and interbreed genetic diversity is a key element in the selection of donors when developing a cryopreservation program of germ cells from the most promising sires and dams. However, complete information is not always available to breeders (for example, due to errors in pedigrees), which can lead to inaccurate identification of the intra- and interbreed genetic diversity. Therefore, it is necessary to use other laboratory diagnostic methods, that are more accurate and less dependent on registration errors, for example, the use of highly polymorphic genetic markers – DNA microsatellites (MS-DNA) (Durnasy et al., 2012). Thus, it is necessary to develop optimal management and breeding programs that use genetic data to minimize inbreeding, as programs to maintain overall genetic and allelic diversity and breed identity, and take into account the intra- and interbreed genetic structure (Martinez et al., 2015).

The formation of a complex intra- and interbreed population structure, therefore, cannot be regarded as an exceptional phenomenon for some breeds solely, and such stratification was most often explained by differences in the geographical origin of individual intra- and interbreed groups, different criteria for breeding work, or their joint effect (Chang et al., 2009). To date, there are already many studies devoted to the analysis of intra- and interbreed genetic diversity and the assessment of interbreed differences in the genetic structure of farm animals, incl. pigs (SanCristobal et al., 2006; Sollero et al., 2009). At the same time, insufficient attention is still paid to the study of...
intrabreed genetic diversity and intrabreed stratification. Thus, the main goal of our work was to analyze the intra- and interbreed genetic diversity of commercial pig breeds on the basis of DNA microsatellite polymorphism.

Materials and methods

In total, the work used data for 3,308 pigs, which represented 11 herds. The animals belonged to the four commercial pig breeds – Duroc (DR), Yorkshire (YR), Landrace (LN) and Large White (LW). Information on the breeds and herds of pigs included in the analysis is presented in Table 1.

Table 1: Information on breeds and herds of pigs included in the analysis

| Breed      | Herd abbreviation | Origin of the samples | Sample size (n) |
|------------|-------------------|-----------------------|-----------------|
| Duroc      | DR1               | Russia, Belgorod region | 520             |
| Duroc      | DR2               | Russia, Kursk region   | 44              |
| Yorkshire  | YR1               | Russia, Belgorod region | 420             |
| Yorkshire  | YR2               | Russia, Voronezh region | 255             |
| Landrace   | LN1               | Russia, Belgorod region | 420             |
| Landrace   | LN2               | Russia, Voronezh region | 220             |
| Landrace   | LN3               | Russia, Kursk region   | 11              |
| Large White| LW1               | Russia, Belgorod region | 420             |
| Large White| LW2               | Russia, Lipetsk region | 192             |
| Large White| LW3               | Russia, Voronezh region | 354             |
| Large White| LW4               | Russia, Kursk region   | 452             |

12 microsatellite loci recommended by ISAG-FAO and arranged in one multiplex panel (S0101, S0155, S0228, S0355, S0386, SW24, SW240, SW72, SW857, SW911, SW936, SW951) were used as DNA markers. Primers for PCR were selected with consideration of the amplification of all 12 loci in one test-tube. The size of all amplified PCR products, taking into consideration all known alleles, was <300 base pairs.

The PCR reaction was carried out on a Verity amplifier (Applied Biosystems, USA) in 20 μL of a mixture containing 20 ng of genomic DNA, PCR buffer (10 mmol Tris-HCl (pH 8.3), 50 mmol KCl, 2 mmol MgCl₂), 0.25 mmol dNTP, 0.5 μmol primer, 1 unit of Taq DNA polymerase (inhibited for hot start). PCR parameters: 94 °C – 3 min; (98 °C – 30 s, 59 °C – 120 s, 72 °C – 75 s) – 6 cycles; (90 °C – 30 s, 59 °C – 120 s, 72 °C – 75 s) – 20 cycles; 68 °C – 30 min. In this case, the heating rate from 59 °C to 72 °C was no more than 0.3°C/s. Fragment analysis of PCR products was performed on an ABI PRISM 3500 automatic capillary DNA sequencer (Applied Biosystems, USA), using 50 cm capillaries and a POP-7TM polymer matrix. The primers were labeled with four fluorescent dyes detectable in the Blue (FAM), Green (ROX), Yellow (TAMRA), Red (ROX) channels (Table 2).

The length standard SD 450 (Synthol, Russia) was labeled with a ROX dye and was detected in a separate Orange channel simultaneously with PCR products. After amplification, 9 μL of Hi-DTMM formamide and 0.5 μL of a SD 450 dimensional standard solution were added to 1 μL of PCR product. Samples prepared in this way were analyzed on an ABI PRISM 3500 DNA sequencer (Applied Biosystems, USA). Fragment size analysis was performed using GeneMapper R Software v.4.1 (Applied Biosystems).

Table 2: Characteristics of microsatellite loci recommended by ISAG for determining the reliability of the origin of pigs

| Locus   | Allele length, bp | Dye   |
|---------|-------------------|-------|
| S0101   | 193–221           | R6G   |
| S0155   | 142–166           | TAMRA |
| S0228   | 218–270           | TAMRA |
| S0355   | 225–277           | FAM   |
| S0386   | 164–182           | FAM   |
| SW24    | 93–125            | ROX   |
| SW240   | 93–125            | R6G   |
| SW72    | 97–125            | TAMRA |
| SW857   | 137–161           | R6G   |
| SW911   | 149–177           | ROX   |
| SW936   | 81–117            | FAM   |
| SW951   | 124–134           | FAM   |

For each sample of pigs, estimates of the frequencies of genotypes and alleles, the number of alleles (Ae), observed (Ho) and expected (He) heterozygosity, as well as the inbreeding coefficient (Fis) for individual MS-DNA loci were calculated using the GenAIEx v.6.5 program (Peakall & Smouse, 2012). In addition, the M-ratio (Garza & Williamion, 2001) estimates were calculated for each herd and MS-DNA locus.

The hypothesis of the absence of significant differences between the studied herds and breeds of pigs in terms of the frequencies of rare and most common alleles was tested using the Pearson Chi-square test in the PAST v. 2.14 software (Hummer et al., 2001). To check the adequacy of genotypes distribution in each MS-DNA locus in each pig herd to the Hardy-Weinberg equilibrium (HWE) based on the likelihood ratio G-test, the PopGen v.1.31 software (Yeh et al., 1999) was used.

To test the hypothesis that there were no significant differences in the indicators of genetic diversity used for individual herds, a non-parametric Friedman two-way ANOVA by ranks using the PAST v. 2.14 software was carried out (Hummer et al., 2001).

Nested two-way ANOVA was used to assess differences within and between individual pig breeds in terms of genetic diversity indicators and M-ratio estimates, considering different MS-DNA loci as independent implementations using the PAST v. 2.14 software (Hummer et al., 2001). Wright’s F-statistic estimates (Fis, Fit and Fst) for each MS-DNA locus and each pig herd were obtained using the GenAIEx v.6.5 program (Peakall & Smouse, 2012). The significance level of the deviation of the obtained estimates from zero was calculated using the permutation test with 999 permutations. In addition, estimates of the coefficient of genetic differentiation (Fst) were calculated for individual herds within each of the four pig breeds used.

Based on the hierarchical analysis of molecular variance (AMOVA) algorithm for each MS-DNA locus, estimates of ΦST (differentiation between breeds), ΦWI (differentiation between herds within their own breed) and ΦST (differentiation between herds) were calculated using the GenAIEx v.6.5 program (Peakall & Smouse, 2012).

To assess the degree of genetic similarity between herds/breeds of pigs, two approaches were used. First, the Assignment test based on the results of the analysis of microsatellite multilocus genotypes (Paetkau et al., 1995) was carried out both for individual herds of pigs and for individual breeds (all animals of the same breed were combined into one sample) using the GENALEX v.6.5 program (Peakall & Smouse, 2012). Secondly, pairwise Nei’s genetic distance matrix (Nei, 1972) were calculated for individual herds and breeds, which were used to plot the distribution of centroids of herds/breeds in the space of the first two axes from a Principal Coordinate Analysis (PCoA) using the GenAIEx v.6.5 program (Peakall & Smouse, 2012).

Estimates of gene flow (Nm) between herds based on the distribution of allele frequencies of 12 MS-DNA loci were obtained using the divMigrate-online program (https://popgen.shinyapps.io/divMigrate-online) (Keinan et al., 2013).

The hypothesis of a bottleneck effect in pig herds in the past with three models (IAM, SMM and TPM) was tested using the BOTTLENECK v.1.2.03 software (Cornuet & Luikart, 1996). Estimates of the average correlation between alleles (r) and the number of cases of linkage disequilibrium (NL) between individual alleles of 12 MS-DNA loci for pig herds, as well as the results of the Ewens-Watson test for neutrality were obtained using the PopGen v.1.31 software (Yeh et al., 1999).

Estimates of effective population size in individual pig herds (Ne/Neb) were calculated using the NeEstimator v. 2.0 software (Do et al., 2014). The nonparametric Kruskal-Wallis H test with the PAST v. 2.14 software (Hummer et al., 2001) was used to test the hypothesis that there were no significant differences in Ne and Neb estimates between pig breeds.

Results

In total, when analyzing 12 MS-DNA loci in 11 herds of pigs belonging to four breeds, 188 alleles were noted. The highest allelic diversity was noted among pigs of the LN1, YR1 and LW2 herds (167, 165 and 163 alleles, respectively), and the lowest (63 and 62 alleles) – among pigs...
of the LN3 and DR2 herds, respectively (Table 3). The number of alleles was not associated with the size of sample (Spearman’s rank correlation coefficient: \(R_s = 0.468; n = 11; P = 0.147\)).

More than half of the alleles identified in the studied pig herds (with the exception of LN3 and DR2) were represented by very rare alleles (with a frequency of \(\geq 0.200\) alleles \((R_s = -0.855; n = 11; P < 0.001)\). Thus, the increasing in the total number of alleles in the studied pig herds occurred due to the increase in the number of rare alleles (Table 3).

All studied pig herds were significantly different from each other in terms of the proportion of both rare and most common alleles (Pearson Chi-square test: in both cases, \(P < 0.001\)). However, to a greater extent, these differences were determined by intrabreed variability between individual herds within their breed (in both cases: \(P < 0.001\)) rather than by differences between breeds, which were noted only in relation to the proportion of the most common alleles \((\chi^2 = 10.18; df = 3; P = 0.017)\) (Table 3).

Estimates of indicators of genetic diversity and M-ratio (on average per one locus) in the studied pig herds are given in Table 4. For all these parameters, there are significant associations between the herd and the MS-DNA locus used in the analysis (Friedman two-way ANOVA by ranks: \(P < 0.001–0.02\)), i.e., the patterns of variation for individual loci differed significantly in different pig herds.

| Table 4 | Indicators of genetic diversity and estimates of M-ratio (x ± SE) for 12 MS-DNA loci in herds of pigs of different breeds (on average per one locus) |
| --- | --- |
| Herd | Na | Ae | Ho | He | Fis | M-ratio |
| --- | --- | --- | --- | --- | --- | --- |
| DR1 | 11.9 ± 1.0 | 4.4 ± 0.6 | 0.575 ± 0.073 | 0.702 ± 0.052 | 0.213 ± 0.070 | 0.891 ± 0.042 |
| DR2 | 5.2 ± 0.5 | 2.7 ± 0.3 | 0.566 ± 0.088 | 0.561 ± 0.062 | -0.011 ± 0.036 | 0.679 ± 0.069 |
| YR1 | 13.8 ± 1.1 | 7.2 ± 0.8 | 0.689 ± 0.052 | 0.838 ± 0.101 | 0.188 ± 0.059 | 0.939 ± 0.040 |
| YR2 | 10.7 ± 1.3 | 4.3 ± 0.6 | 0.583 ± 0.065 | 0.711 ± 0.042 | 0.182 ± 0.083 | 0.933 ± 0.066 |
| LN1 | 13.9 ± 1.3 | 6.7 ± 0.8 | 0.675 ± 0.025 | 0.824 ± 0.023 | 0.193 ± 0.049 | 0.944 ± 0.033 |
| LN2 | 11.4 ± 1.6 | 4.9 ± 0.9 | 0.665 ± 0.065 | 0.698 ± 0.059 | 0.160 ± 0.094 | 0.880 ± 0.051 |
| LN3 | 5.3 ± 0.5 | 3.4 ± 0.3 | 0.750 ± 0.073 | 0.664 ± 0.049 | -0.118 ± 0.036 | 0.702 ± 0.061 |
| LW1 | 12.3 ± 1.0 | 5.1 ± 0.4 | 0.672 ± 0.049 | 0.786 ± 0.022 | 0.150 ± 0.051 | 0.902 ± 0.034 |
| LW2 | 13.6 ± 1.0 | 4.9 ± 0.5 | 0.695 ± 0.051 | 0.773 ± 0.023 | 0.106 ± 0.054 | 0.939 ± 0.025 |
| LW3 | 7.3 ± 0.8 | 2.8 ± 0.2 | 0.604 ± 0.094 | 0.608 ± 0.040 | 0.034 ± 0.138 | 0.716 ± 0.055 |
| LW4 | 8.0 ± 0.9 | 3.4 ± 0.4 | 0.617 ± 0.076 | 0.664 ± 0.035 | 0.105 ± 0.005 | 0.789 ± 0.064 |

Friedman two-way ANOVA by ranks \((\chi^2; P < 0.001)\) 
- Na: \((\chi^2; P = 0.020)\) 
- Ae: \((\chi^2; P = 0.001)\) 
- Ho: \((\chi^2; P = 0.001)\) 
- He: \((\chi^2; P = 0.001)\) 
- Fis: \((\chi^2; P = 0.001)\)

| Table 5 | Results of nested two-way ANOVA of genetic diversity indicators and M-ratio (F; P) estimates of 12 MS-DNA loci in herds of pigs of different breeds |
| --- | --- |
| Source | Indicator | Na | A | Ho | He | Fis | M-ratio |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Between breeds \((df = 3; \chi^2 = 121)\) | 3.64 \((0.015)\) | 6.64 \((0.001)\) | 0.79 \((ms)\) | 4.17 \((0.008)\) | 0.78 \((ms)\) | 1.35 \((ms)\) |
| Between herds within the breed \((df = 7; \chi^2 = 121)\) | 11.15 \((< 0.001)\) | 6.39 \((< 0.001)\) | 0.88 \((ms)\) | 4.51 \((< 0.001)\) | 1.99 \((ms)\) | 5.12 \((< 0.001)\) |
| Between herds \((df = 10; \chi^2 = 121)\) | 8.90 \((< 0.001)\) | 6.40 \((< 0.001)\) | 0.86 \((ms)\) | 4.41 \((< 0.001)\) | 1.63 \((ms)\) | 3.99 \((< 0.001)\) |

The average number of alleles (Na) per locus, as expected, was the highest in LN1, YR1, and LW2 herds (13.9, 13.8 and 13.2 alleles per locus, respectively), and the lowest in LN3 and DR2 herds (5.3 and 5.2 alleles per locus, respectively). The average number of effective alleles (Ae) per locus did not completely repeat the results obtained above—the maximum value was noted in LN1 and YR1 herds (7.2 and 6.7 alleles per locus, respectively), and the lowest—in LW3 and DR2 herds (2.8 and 2.7 alleles per locus, respectively, Table 4).

Average estimates of observed heterozygosity (Ho) varied from 0.563 (LN2 herd) to 0.750 (LN3 herd), while estimates of expected heterozygosity (He) varied over a wider range, from 0.561 (DR2 herd) to 0.838 (YR1 herd). In two cases (for DR2 and LN3 herds) we noted an excess of observed heterozygosity over expected, which led to negative estimates of the inbreeding coefficient (Fis = -0.011 and Fis = -0.118, respectively), while in most cases the deficit of heterozygosity was accompanied by positive estimates of the inbreeding coefficient on average per locus, which varied from 0.034 (LN3 herd) to 0.213 (DR1 herd) (Table 4).

The decrease in allele diversity, which was most noted for pigs of LN3 and DR2 herds, manifests itself against the background of maintaining the same wide interval of alleles as for groups of animals with the maximum number of identified alleles, which leads to a sharp decrease in M-ratio estimates. Thus, only for LN3 and DR2 herds, the obtained estimates of the M-ratio (on average per locus) did not significantly deviate from the critical value of 0.600 (Garza & Williamson, 2001), while in the other studied groups these estimates reliably exceeded it (Table 4).

The results of the nested two-way ANOVA of genetic diversity indicators and M-ratio estimates, taking into account different loci of MS-DNA, as independent realizations, show high homogeneity for different breeds/herds of pigs with respect to observed heterozygosity (Ho) and inbreeding coefficient (Fis) (Table 5). Whereas, for the rest of the indicators used, a high level of interherd diversity was noted, which was determined, to a greater extent, by intrabreed differences between individual herds within their breed (in all cases: \(P < 0.001\)), and not by differences between breeds (\(P < 0.001–0.015\)). In relation to the M-ratio estimates, there was no reliable difference between the breeds (Table 5).

Significant differences were noted with respect to the distribution of allele frequencies of MS-DNA loci among individual herds for all four studied pig breeds. The Fst estimates varied from 0.067 (for two herds of the YR breed) to 0.168 (for two herds of the DR breed), but in all cases Fst
estimates significantly differed from zero (P < 0.001, Table 6). When uniting individuals belonging to the same breed, the estimate of interbreed genetic differentiation was even lower than when comparing herds for individual breeds (Fst = 0.005), although it also significantly differed from zero (P < 0.001).

The results of testing of Hardy-Weinberg equilibrium demonstrate significant interherd differences (Table 7). For DR1, YR1, YR2, LN1, LW1, LW4 herds, all MS-DNA loci used in the analysis demonstrated a significant deviation from the HWE, while for DR2 and LN3 herds, on the contrary, the overwhelming majority of loci were in the HWE. If analyzed in relation to individual MS-DNA loci, then for loci SW936, SW951, for 10 out of 11 studied pig herds, a significant deviation from the HWE was noted, while for other herds the number of such loci was 7–9 (Table 7).

Table 6

| Breed | Number of herds | Fst | P |
|-------|-----------------|-----|---|
| DR    | 2               | 0.168 | < 0.001 |
| YR    | 2               | 0.067 | < 0.001 |
| LN    | 3               | 0.108 | < 0.001 |
| LW    | 4               | 0.121 | < 0.001 |

The results of hierarchical analysis of molecular variance (AMOVA) showed that there are significant differences in the distribution of allele frequencies of 12 MS-DNA loci between the studied pig herds (Table 8). The average estimate of interbreed genetic differentiation (FST) is 0.237 ± 0.018 (P < 0.001), with a range for the different loci of MS-DNA used from 0.140 (locus SW240) to 0.323 (locus SW72). We did not identify significant differences between the pig breeds included in the analysis (FST = 0.005 ± 0.014; P > 0.05) and most of the interherd variability is due to intrabreed differences between pig herds within their breed (FST = 0.233 ± 0.019; P < 0.001) (Table 8).

Table 8

| Locus   | DR1  | DR2  | YR1  | YR2  | LN1  | LN2  | LN3  | LW1  | LW2  | LW3  | LW4  |
|---------|------|------|------|------|------|------|------|------|------|------|------|
| S0101   | NS   | <0.01| NS   | <0.01| NS   | <0.01| NS   | <0.01| NS   | <0.01| NS   |
| S0155   | 0.001| NS   | <0.01| <0.01| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |
| S0228   | 0.001| 0.018| 0.018| 0.018| 0.018| 0.018| NS   | <0.01| NS   | <0.01| NS   |
| S0355   | <0.01| NS   | <0.01| <0.01| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |
| S0366   | <0.01| 0.003| <0.01| <0.01| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |
| S0244   | <0.01| NS   | <0.01| <0.01| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |
| SW240   | <0.01| NS   | <0.01| <0.01| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |
| SW72    | <0.01| NS   | <0.01| 0.017| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |
| SW357   | <0.01| NS   | <0.01| <0.01| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |
| SW11    | <0.01| NS   | <0.01| <0.01| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |
| SW951   | <0.01| <0.01| <0.01| <0.01| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |

The high originality of the genetic structure in individual pig herds was also confirmed by the results of the Assignment test based on the distribution of MS-DNA multilocus genotypes (Table 9). The genetic uniqueness of pig herds varied from 72.4% (LN1 herd) to 100% (DR2 herd) and the smallest genetic distance was noted for herds representing Duroc (DR1 and DR2) and Landrace (LN3) breeds (Fig. 1a).

Table 9

| Breed   | DR1  | DR2  | YR1  | YR2  | LN1  | LN2  | LN3  | LW1  | LW2  | LW3  | LW4  |
|---------|------|------|------|------|------|------|------|------|------|------|------|
| Assignment accuracy,% | 82.0 | 81.8 | 80.0 | 80.8 | 82.3 | 81.8 | 82.0 | 81.8 | 82.0 | 82.0 | 82.0 |

The results of testing of Hardy-Weinberg equilibrium for 12 loci of MS-DNA in herds of pigs of different breeds based on the likelihood ratio G-test (Table 7).

Table 7

| Locus   | DR1  | DR2  | YR1  | YR2  | LN1  | LN2  | LN3  | LW1  | LW2  | LW3  | LW4  |
|---------|------|------|------|------|------|------|------|------|------|------|------|
| SW936   | -0.001| NS   | <0.01| <0.01| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |
| SW951   | -0.001| NS   | <0.01| <0.01| <0.01| <0.01| NS   | <0.01| NS   | <0.01| NS   |

The degree of similarity of the genetic structure of pigs was assessed using the distribution of centroids of herds/breeds in the space of first two axes from a Principal Coordinate Analysis (PCoA) based on the pairwise Nei’s genetic distance matrix (Fig. 1). The location of herd centroids is random and is not consistent with their breed affiliation at all. The two herds of Duroc (DR1 and DR2) were the most distant from each other, and the smallest genetic distance was noted for herds representing Duroc (DR2) and Landrace (LN3) breeds (Fig. 1a).

Table 10

| Breed   | DR1  | DR2  | YR1  | YR2  | LN1  | LN2  | LN3  | LW1  | LW2  | LW3  | LW4  |
|---------|------|------|------|------|------|------|------|------|------|------|------|
| Assignment accuracy,% | 82.0 | 81.8 | 80.0 | 80.8 | 82.3 | 81.8 | 82.0 | 81.8 | 82.0 | 82.0 | 82.0 |
When individuals belonging to the same breed are combined, the centroids of pig breeds in the space of first two axes from a Principal Coordinate Analysis form two clusters (Fig. 1b). The first one contains the only red pig breed (DR) used in the analysis, while the second one contains white pig breeds (YR, LN and LW). Although along the second Principal Coordinate (axis) there is a significant differentiation between the YR and LN breeds, on the one hand, and the LW breeds, on the other (Fig. 1b).

Gene flow estimates (Nm) showed the presence of an intensive exchange of genetic information between LW herds from different regions of the Russian Federation, as well as between animals kept in different herds, represented by pigs of all breeds included in the analysis (DR1, YR1, LN1, LN2 and LW1) (Table 11).

Table 11
Testing results of the hypothesis about the presence of bottleneck effect for pig herds of different breeds in the past, based on estimates of heterozygosity of 12 MS-DNA loci (the theoretically expected/observed number of loci that exhibit excess of heterozygosity are shown)

| Herd       | Model | L95   | U95   |
|------------|-------|-------|-------|
| DR1        | 6.94/6 (ns) | 7.07/5 (P = 0.018) | 6.99/1 (P < 0.001) |
| DR2        | 6.72/6 (ns) | 7.18/6 (P = 0.012) | 6.94/5 (ns) |
| YR1        | 7.24/7 (ns) | 7.09/3 (P = 0.018) | 7.06/1 (P < 0.001) |
| YR2        | 7.06/6 (ns) | 7.10/6 (ns) | 7.02/1 (P < 0.001) |
| LN1        | 7.27/6 (ns) | 7.12/2 (P = 0.003) | 7.05/0 (P < 0.001) |
| LN2        | 7.00/4 (ns) | 7.06/3 (P = 0.019) | 7.06/1 (P < 0.001) |
| LN3        | 6.98/9 (ns) | 7.05/8 (ns) | 7.14/8 (ns) |
| LW1        | 7.16/8 (ns) | 7.06/3 (P = 0.019) | 7.01/1 (P < 0.001) |
| LW2        | 7.25/6 (ns) | 7.90/4 (ns) | 7.04/3 (P = 0.019) |
| LW3        | 6.86/9 (ns) | 6.95/6 (ns) | 7.09/4 (ns) |
| LW4        | 6.94/8 (ns) | 7.04/7 (ns) | 7.11/4 (ns) |

Despite the fact that a-priori MS-DNA are neutral genetic markers, we found that none of 11 studied pig herds had loci for which the null hypothesis of neutrality was rejected based on the results of the Ewens-Watterson test (Table 12). Most of these loci were found in herds YR1 (seven loci) and LN1 (four loci). For only locus S0155, did we never observe a significant deviation from neutrality, while such a deviation was noted for loci SW24, SW240, SW936 (in three pig herds) and SW857 (in five pig herds, Table 12). Most often, MS-DNA loci that significantly deviated from neutrality were found among pigs of YR (9 out of 24 loci, i.e., 37.5%) and LN breeds (7 out of 36 loci, i.e., 19.4%), Table 12.

Table 12
Results of the Ewens-Watterson test for 12 MS-DNA loci for different herds of pigs of different breeds (only loci for which the hypothesis of neutrality is reliably rejected are shown)

| Herd       | Locus | Obs. F | L95–U95   |
|------------|-------|--------|-----------|
| DR1        | SW240 | 0.131  | 0.160–0.651 |
| DR2        | SW936 | 0.254  | 0.306–0.891 |
|            | S0228 | 0.081  | 0.105–0.399 |
|            | S0355 | 0.149  | 0.152–0.620 |
|            | SW24  | 0.108  | 0.143–0.587 |
|            | SW72  | 0.158  | 0.160–0.591 |
|            | SW857 | 0.140  | 0.158–0.666 |
|            | SW936 | 0.097  | 0.140–0.574 |
| YR1        | SW857 | 0.132  | 0.158–0.645 |
| YR2        | SW936 | 0.180  | 0.185–0.698 |
|            | S0228 | 0.081  | 0.100–0.350 |
| LN1        | SW24  | 0.113  | 0.148–0.605 |
| LN2        | SW24  | 0.089  | 0.137–0.517 |
|            | S0101 | 0.123  | 0.130–0.509 |
| LN3        | S0355 | 0.147  | 0.173–0.715 |
| LW1        | SW857 | 0.116  | 0.144–0.519 |
| LW2        | SW936 | 0.284  | 0.347–0.954 |
| LW4        | SW936 | 0.190  | 0.258–0.882 |

Note: Obs. F is the actual sum of the squares of the allele frequencies, L95, U95 – lower and upper values of the 95% confidence interval of the Obs. F estimate calculated based on 1000 simulations.

Estimates of the effective population size (Ne) obtained for the studied pig herds indicate an unfavourable state of the level of their genetic variability, which is apparently caused by a high level of inbreeding. Six...
herds had Ne values below 50 inds. (assuming a 95% confidence interval), two herds had Ne values in the 50–100 inds., and finally, three herds had Ne values greater than 100 inds. (Table 13).

Table 13

| Herd | LD(Ne) | Molecular Coancestry (Neb) |
|------|--------|----------------------------|
| DR1  | 18.1 (16.4–19.8) | 2.3 (1.4–3.4) |
| DR2  | 22.6 (15.0–36.1) | 2.3 (0.8–4.7) |
| YR1  | 29.4 (27.1–31.9) | 10.7 (7.3–14.8) |
| YR2  | 48.6 (40.3–58.5) | 3.5 (2.3–4.8) |
| LN1  | 40.5 (35.7–45.7) | 11.4 (8.3–15.0) |
| LN2  | 110.2 (76.4–168.8) | 6.4 (3.6–9.8) |
| LN3  | 37.4 (16.3–∞) | 13.7 (3.3–31.2) |
| LW1  | 20.4 (18.5–22.4) | 3.8 (1.5–4.5) |
| LW2  | 35.6 (29.6–42.9) | 6.7 (4.0–10.2) |
| LW3  | 85.8 (90.5–142.9) | 172.5 (3.2–865.7) |
| LW4  | 38.2 (25.3–55.7) | 3.6 (2.6–4.9) |

Estimates of the effective number of sires (Neb) obtained using the Molecular Coancestry method (MC) were almost an order of magnitude lower. In five herds, the Neb values did not exceed 5 inds., in another five, the Neb values were in the range of 5–20 inds., and only in one herd (LW3) did the estimate of the effective number of sires exceed one hundred inds. (Table 13).

In general, no significant differences in the Ne and Neb estimates between the pig breeds included in the analysis were identified (Kruskal-Wallis H test: in both cases P > 0.05). Whereas significant intrabreed differences were noted between individual herds of YR, LN and LW breeds, but only in relation to Ne estimates (Table 13).

A high level of interbreeding is also confirmed by a significant number of linkage disequilibrium (NLD) between alleles of 12 MS-DNA loci recorded in most of the studied pig herds, especially in LW1, DR1, YR1 and LN1 (1299, 1247, 1164 and 800 cases, respectively, Table 14). There was no significant difference between the studied breeds in terms of NLD estimates for individual herds of pigs (Kruskal-Wallis H test: H = 1.26; P > 0.05). The estimates of the average correlation between alleles (r) of the MS-DNA loci varied from 0.082 (LW3 herd) to 0.345 (DR1 herd) and in all cases (except for LN3 herd) significant exceeded zero (Table 14). There was no significant difference between the studied breeds in terms of r estimates for individual herds of pigs (Kruskal-Wallis H test: H = 3.89; P > 0.05). As expected, an inverse dependence of the estimates of the average correlation between alleles (r) in different pig herds on the effective population size (Ne) for these herds was noted (Spearman’s rank correlation coefficient: Rs = –0.609; n = 11; P = 0.047, Fig. 3).

Discussion

When analyzing the intra- and interbreed variability of 11 herds, which belonged to four commercial pig breeds, in relation to 12 Ms-DNA loci, we found that all studied breeds significantly differed in terms of the proportion of both rare and the most common alleles. At the same time, the noted differences were determined, first of all, by the variability between individual herds within their breed (Table 3). In addition, for all indicators of genetic diversity used (with the exception of Ho and Fs), a high level of interherd heterogeneity was noted, which was also more determined by intrabreed differences between individual herds within their breed than by differences between breeds (Table 5). Although some authors consider the manifestation of intrabreed stratification to be relatively rare for farm animals (Willkinson et al., 2012), it was previously noted for horses (Głowski-Mullis et al., 2006), cows (Lazurebaya et al., 2020), pigs (Willkinson et al., 2011), goats (Martínez et al., 2015), chickens (Willkinson et al., 2012), dogs (Chang et al., 2009; Wiener et al., 2017) and rabbits (Jochowie et al., 2017). At the same time, the existence of intrabreed genetic heterogeneity should not be viewed exclusively as a negative phenomenon, but as an essential element in the history of their creation (European ..., 2006).

Regul. Mech. Biosyst., 2021, 12(1)
different together with isolation, which manifests itself for populations of different origins and can prevent gene flow and provide genetic differentiation even among individuals of the same breed (Alves et al., 2015). Previously, similar observations were described for dogs (Chang et al., 2009) and sheep (Kijas et al., 2009). In addition, using the example of rabbit breeds, it has been shown (Alves et al., 2015) that continuous and differentiated artificial selection, together with modern breeding that maintains breeds as closed genetic pools, can transform genetic uniformity into a set of highly differentiated genetic interbreed groups.

For most of the studied herds and breeds of pigs, a significant deficit of heterozygotes was noted (Table 4). It has already been noted that, along with inbreeding and genetic drift, the manifestation of which is primarily associated with the small size of populations, the deficit of heterozygotes may also be associated with intrabreed stratification (Martínez et al., 2015). In addition, for rabbit breeds, the revealed intrabreed stratification was accompanied by a significant deviation of the distribution of genotypic frequencies from the HWE, as well as high positive Fis estimates (Jochova et al., 2017). In our study, for DR1, VR1, VR2, LN1, LW1, LW4 herds, all the MS-DNA loci used demonstrated a significant deviation from the HWE (Table 7) against the background of the presence of significant intrabreed stratification. Similar results were obtained earlier when analyzing the intrabreed stratification of the Large White pigs (Lugovoy et al., 2017).

One of the consequences of intrabreed stratification is an increasing in the level of interbreed diversity (which turns out to be not lower than the level of interbreed diversity) against the background of a decreasing in variability within individual herds. We noted this during ordination of pig herds in the space of first two axes from a Principal Coordinate Analysis (Fig. 1a). In this case, the location of the herd centroids is random and is not consistent with their breed affiliation at all. However, when all individuals belonging to the same breed are combined, the breed centroids in the space of first two axes from a Principal Coordinate Analysis form two clusters that clearly reflect interbreed differences (Fig. 1b).

The decrease in the level of genetic diversity within individual pig herds is associated in both the relatively frequent manifestation of the bottleneck effect (Table 11) and the high level of inbreeding, which is confirmed by a significant linkage disequilibrium recorded in most of the studied pig herds (especially in LW1, DR1, YR1 and LN1) and a high estimate of average correlation between alleles (Table 14).

In six pig herds the Ne estimates were below 50 inds., in two herds they were in the range of 50–100 inds., and finally in three herds the Ne estimates exceeded 100 inds. (Table 13). A similar situation was noted for the Large White breed (a widespread breed) and in other countries – Lithuania: Ne = 20–38 inds. (Šveistienė & Rasmaitė, 2013), Czech Republic: Ne = 50 inds. (Krupa et al., 2015) and Brazil: Ne = 40 inds. (Janella et al., 2016). The estimate of the effective population size in the local breed of Ukraine (Ukrainian Meat breed) was only 68 inds. (with a 95% confidence interval – 52–92 inds.) (Lugovoy et al., 2018). It is typical that the estimates of intrabreed differentiation (Fst) obtained for natural populations of wild boar (Sus scrofa L., 1758), on the contrary, indicate their satisfactory condition (for Spain/Portugal: Ne = 180 inds. (Herreero-Medrano et al., 2013); for Australia: Ne = 960–1477 inds. (Covedal et al., 2008).

The results of hierarchical analysis of molecular variance (AMOVA) for the studied herds of pigs showed that the average estimate of interherd genetic differentiation (ΦPT) is 0.237 (P < 0.001), and most of this interherd variability was due to intrabreed differences between individual herds within their breed (ΦBPT = 0.233; P < 0.001, Table 8). A significant interherd component was also noted when studying the genetic variability of donkey breeds on different farms (Collí et al., 2013). Thus, the most striking manifestation of intrabreed stratification is the high estimate of the genetic differentiation measure (Fst) obtained in the analysis of intrabreed groups (bred-strain, selection), which turns out to be no lower than when comparing different breeds.

For the studied pigs, the Fst estimates in all cases significantly differed from zero (P < 0.001) and varied from 0.067 (for two herds of the YR breed) to 0.168 (for two herds of the DR breed, Table 6). Whereas when comparing individual breeds, the value of the genetic differentiation index was lower (Fst = 0.065), although it also significantly differed from zero (P < 0.001). When analyzing the intrabreed variability of Large White pigs, represented by 4 separate herds, the estimate of the genetic differentiation index (Fst) was 0.148 ± 0.049 (Lugovoy et al., 2017), i.e., it was close to the value which was noted in this study.

In general, the obtained estimates of intrabreed differentiation were approximately at the same level as the analogous values obtained in the analysis of intrabreed differentiation of pigs: Mexico – Fst = 0.110 (Le- mas-Flóres et al., 2001), the Iberian Peninsula – Fst = 0.130 (Martínez et al., 2000), Brazil – Fst = 0.140 (Sollero et al., 2009), Portugal – Fst = 0.184 (Vicente et al., 2008). When analyzing 48 European breeds of pigs, the value of intrabreed genetic differentiation was higher (Fst = 0.210) (San Cristobal et al., 2006), although it was only slightly higher than the one obtained when analyzing two herds of Duroc pigs in our study (Fst = 0.168). Expectedly higher Fst estimates were obtained when both European and Asian (Chinese and Korean) pig breeds, as well as wild boar, were included in the analysis (Fst = 0.261; Kim et al., 2005).

It is typical that the estimates of intrabreed differentiation (Fst) obtained by us for commercial breeds of pigs are close to the estimates obtained in the study of intrabreed stratification (genetic differentiation between different herds) among sheep (Fst = 0.100–0.170; Dumassy et al., 2012) and goats (Fst = 0.070) (Serrano et al., 2009).

Conclusion

The analysis of the genetic variability of pigs of four commercial breeds showed that the high level of interbreed differences is caused, first of all, by the high variability among pig herds within each studied breed. Such intrabreed stratification can be formed due to the manifestation of many causes: different genetic basis of the founders of intrabreed genealogical groups, geographical isolation, different directions of selection within individual herds, exchange of animals between separate herds, the use of inbreeding in the practice of selection together with isolation, etc. Important consequences of intrabreed stratification are an increase in the level of interbreed diversity (which is not lower than the level of intrabreed diversity) against the background of a decrease in variability within individual herds, as well as a significant deficit of heterozygotes and an increasing in the role of negative genetic and demographic processes.

Thus, the existence of genetic heterogeneity within commercial pig breeds should be considered as an essential element in the history of their formation and breeding.

This work was supported by the RFBR grant No. 20-3480004 / 20, as well as with the financial support of the Ministry of Education and Science of Ukraine (state registration number 0119U001042).

References

Alves, J. M., Caneiro, M., Afonso, S., Lopes, S., Garreau, H., Boucher, S., Allain, D., Queney, G., Esteves, P. J., Bilot, G., & Fernand, N. (2015). Levels and patterns of genetic diversity and population structure in domestic rabbits. PLoS One, 10(12), e0144687.
Bjömerfeldt, S., Hailer, F., Nord, M., & Vila, C. (2008). Assortative mating and fragmentation within dog breeds. BMC Evolutionary Biology, 8(1), 28.
Chang, M. L., Yokoyama, J. S., Branson, N., Dyer, D. J., Hitt, C., Oversea, K. L., & Harrellton, S. P. (2009). Intrabreed stratification related to divergent selection re- sults in purebred dogs may affect the interpretation of genetic association stu- dies. Journal of Heredity, 100(suppl. 1), S28–S36.
Chmarzyńska, A., Maciowska, M., Szydłowski, M., Mel尼克, J., Kamyęczek, M., Eckert, R., Różycki, M., & Świńosik, M. (2004). Polymorphism of intronic microsatellites in the A-FABP and LEPR genes and its association with produc- tive traits in the pig. Journal of Animal and Feed Sciences, 13(6), 615–624.
Collí, L., Perrotta, G., Negrini, R., Bomboi, L., Bigi, D., Zambonelli, P., Verini Suplizi, A., Lio- tta, L., & Ajmone-Marsan, P. (2013). Detecting population struc- ture and recent demographic history in endangered livestock breeds: the case of the Italian autochthonous donkeys. Animal Genetics, 44(1), 69–78.
Cornet, J. M., & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics, 144(4), 2001–2014.
Covedal, B. D., Aldenhoven, J., Odde, I. O., Gurrea, T., Mortan, C., & Lapidge, S. J. (2000). Feral pig population structuring in the rangelands of eastern Australia: Applications for designing adaptive management units. Conservation Genetics, 9(1), 211–224.
