

Genome scanning and biodiversity

UDC 636.2:575.174:575.2:51-76

doi: 10.15389/agrobiology.2016.6.788rus

doi: 10.15389/agrobiology.2016.6.788eng

STUDY OF GENETIC DIVERSITY AND POPULATION STRUCTURE OF FIVE RUSSIAN CATTLE BREEDS USING WHOLE- GENOME SNP ANALYSIS

N.A. ZINOVIEVA¹, A.V. DOTSEV¹, A.A. SERMYAGIN¹, K. WIMMERS²,
H. REYER², J. SÖLKNER³, T.E. DENISKOVA¹, G. BREM^{1, 4}

¹L.K. Ernst All-Russian Research Institute of Animal Husbandry, Federal Agency of Scientific Organizations, 60, pos. Dubrovitsy, Podolsk District, Moscow Province, 142132 Russia, e-mail n_zinovieva@mail.ru, asnd@mail.ru, alex_sermyagin85@mail.ru;

²Institute of Genome Biology, Leibniz Institute for Farm Animal Biology (FBN), Mecklenburg-Vorpommern, 18196 Dummerstorf, Germany, e-mail wimmers@fbn-dummerstorf.de, reyer@fbn-dummerstorf.de;

³Division of Livestock Sciences, University of Natural Resources and Life Sciences, Augasse 2-6, A-1090, Vienna, Austria, e-mail johann.soelkner@boku.ac.at;

⁴Institut für Tierzucht und Genetik, University of Veterinary Medicine (VMU), Veterinärplatz, A-1210, Vienna, Austria, e-mail gottfried.brem@agrobiogen.de

Acknowledgements:

We thank Department of Animal Husbandry and Breeding (Ministry of Agriculture of the Russian Federation), Dr V.S. Matyukov (Komi Republic Research Institute of Agriculture) and Dr D.K. Nekrasov (Ivanovo State Academy of Agriculture) for assistance in collecting samples.

Supported by Russian Science Foundation, project № 14-36-00039

Received September 26, 2016

Abstract

With the publication of the complete sequence of a cattle genome, it became possible to trace the history of breed origins and to evaluate genetic relationships between modern breeds, based on the results of genome-wide SNP screening. Whilst numerous studies have been undertaken to characterize the commercial breeds and some local cattle breeds of Europe, North America, Asia and Africa at whole-genome level, little is known about genetic differences, relationships and population genetic structure of the Russian native cattle breeds. The aim of our work was to study the genetic diversity and population structure of five locally-developed Russian cattle breeds, based on genome-wide single nucleotide polymorphisms (SNPs) generated using Illumina Bovine SNP50 BeadChips (Illumina, San Diego, CA, USA). In total, 116 samples (sperm or tissue) collected from five breeds were analyzed, including Bestuzhev (BEST, $n = 27$), Kholmgor (KHLM, $n = 25$), Kostromsky (KSTR, $n = 20$), Red Gorbатов (RGBT, $n = 23$) and Yaroslavl breeds (YRSL, $n = 21$). Samples of Holstein cattle (HLST, $n = 29$) were used for comparison. Quality filtering of genetic markers was performed in PLINK v 1.07. Data processing was performed using software PLINK 1.07, HP-Rare 1.1, STRUCTURE, ver. 2.3.4, Phylip, ver. 3.695, FigTree 1.4.2, Arlequin suite, ver. 3.5.2.2 and R pocket. The final set of markers passed through the quality control and selected for further analysis included 35874 SNPs. Average heterozygosity within breeds ranged from 0.378 in BEST to 0.390 in KHLM and was higher comparing to HLST (0.377). Allelic richness was ranging from 1.914 ± 0.001 in KSTR to 1.955 ± 0.001 in BEST. A slight heterozygote excess was detected in all breeds studied (F_{IS} from -0.015 in BEST to -0.054 in KHLM). The multidimensional scaling (MDS) showed the presence of non-overlapping breed specific clusters, whereas the first principal component (PC1) accounted for 5.46 % and the second principal component (PC2) was responsible for 5.05 % of the genotypic variance. Phylogenetic analysis based on parsimony method grouped individuals into six clusters according to their breeds. The STRUCTURE analyses supported the assumption that the ancestry of the locally developed Russian cattle breeds is distinct from Holsteins and Holstein-related breeds. The highest ΔK showing the assumed number of populations was observed for $k = 6$. At $k = 6$, the genetic structure is in agreement with breed origin of individuals: $Q_{1/6} = 0.855 \pm 0.018$ for BEST, $Q_{2/6} = 0.818 \pm 0.029$ for KHLM, $Q_{3/6} = 0.923 \pm 0.015$ for KSTR, $Q_{4/6} = 0.816 \pm 0.027$ for RGBT, $Q_{5/6} = 0.873 \pm 0.031$ for YRSL and $Q_{6/6} = 0.935 \pm 0.014$ for HLST. Analysis of molecular variance (AMOVA) showed highly significant results for genetic differentiation ($p < 0.001$) in studied breeds. AMOVA revealed that most of the genetic variation in cattle breeds was found within populations (91.2 %), and less among populations (8.8 %). The emerging structure of the phylogenetic tree constructed on the Nei genetic distances, is in full concordance with the his-

torical origin of breeds and confirms the MDS and STRUCTURE results. Thus, using the method of genome-wide SNP studies we were able for the first time to study the population structure and genealogical relationships among the five Russian cattle breeds. The received information is the first step towards the evaluation of the value of these breeds regarding their conservation and usage in the agricultural production of the future.

Keywords: Russian cattle breeds, whole-genome SNP screening, biodiversity

The development of intensive livestock farming systems based on a limited number of breeds leads to a decrease of the diversity in livestock species [1, 2]. Identification and maintenance of the unique variability of local breeds in terms of growth of anthropogenic load and climate change can become an indispensable source of new valuable genotypes for future animal breeding. The centuries-old history of the territory of Russia formed cattle populations that are well adapted to the local climatic environment and economic conditions, while foreign cattle breeds were not massively imported to the old Russia in the first quarter of the XVIII century [3]. The chaotic importation of different cattle breeds from the European countries to Russia, with the aim of improving local cattle, began in 1725 with the creation of the Tsarskoe Selo farm. From the eighteenth century to the first quarter or even the half of the nineteenth century, the Holland cattle was the most imported breed. In parallel, with the beginning of the nineteenth century, Tyrolean cattle were imported in large quantities for several decades. From the end of the first half of the XIX century to early XX century (1930s), a large number of Simmental and Brown Swiss cattle were brought to the old Russia [4]. Along with these breeds, which dominated over other imported cattle, a limited number of almost all cattle breeds that were bred in Western Europe were imported to the old Russia [5-7]. The decision on import of a particular breed to improve the local cattle was taken independently by each landowner and was often based on fashion for a particular breed in a certain period of time [3]. Based on this unsystematic use of a variety of breeds in the regions of the old Russia very diverse cattle populations occurred. In cases where the local natural and economic conditions were favorable (as, for example, in Kholmogor district of the Arkhangelsk region), groups of improved and more productive cattle were formed [8]. However, in most areas the importation of foreign cattle had no visible influence and did not affect the development of local livestock farming. In the 20s of the XX century, based on a territorial principle and phenotypic traits, 12 large arrays of cattle arose in the USSR — the so-called Russian spawn (currently defined as breeds), which acquired considerable notoriety, has been of a great economic interest and became the subject of breeding. In addition, there were a few small groups of local cattle [3].

In the 20s to 30s of the XX century, the best part of the population became the basis for the development of breeds. So, such breeds were officially recognized as Yaroslavl (1925), Red Gorbатов (1926), Kholmogor (1927), and Bestuzhev (1928). In the 30s, after the civil war was over and the agriculture was restored, the breed-zoning plan was developed, which defined the target cattle breeds and regions of their breeding due to economic, naturally historical and zootechnical assumptions. In total, 17 breeds, including twelve indigenous and five foreign, were selected as planned breeds. In the European part of Russia the breeding of 11 local breeds has been recommended including Kholmogor, Tagil, Red German, White-headed Colonist, Yaroslavl, Town Caucasian, Bestuzhev, Red Gorbатов, Kalmyk, Kirghiz, and Grey Ukrainian and 5 foreign: Holland, Simmental, Brown Swiss, Shorthorn and Hereford [4]. In subsequent years, crossbreeding of domestic breeds with foreign ones has been practiced, which led to the extinction of many indigenous breeds and the development of new crossbred breeds. For example, based on the local cattle of Kostroma region, which

was previously slightly improved by crossing with Yaroslavl and Algaus cattle, the new Kostromsky breed was created through intensive crossing with Brown Swiss cattle and received the official recognition in 1944 [9]. Starting in the 80s-90s of last century, due to the lack of economic competitiveness, which was mainly based on low yield and unsuitability to industrial scale, the number of domestic cattle began to decline significantly. For example, from 1990 to 2013, the number of Kholmogor cattle decreased from 2137.0 to 285.8 thousand heads, Yaroslavl — 746.2 to 62.4 thousand heads, Bestuzhev — 981.9 to 33.1 thousand heads, Kostromsky — 384.0 to 13.0 thousand heads, Red Gorbatov — 38.0 to 1.5 thousand heads [10, 11]. Taking into account a significant contribution of indigenous breeds and regional cattle populations to the genetic diversity of domesticated species [12, 13], the reduction of their population size and "dilution" of their allele pool may be one of the main reasons for the decline of the biodiversity of domesticated species.

It should be noted that the improvement of local cattle breeds occurred mainly by crossbreeding to obtain crosses with different contributions from improved breeds [9]. Given the significant impact of environmental conditions on the formation of allele pool, it can be assumed that many alleles, which are specific for aboriginal breeds and associated with the adaptation to local climatic conditions, were kept by selection pressure and have survived in modern populations.

Generating new knowledge concerning the structure of animal genome, the improvement of methodical approaches, the development of high-performance technologies of genomic analysis, and the creation of the analytical equipment of the new generation has led to the identification and application of different types of genetic markers in studies of animal diversity [14]. Until recently, mtDNA polymorphisms and microsatellites were most widely used in studies of domestication investigating the origin and demographic history of cattle breeds. Previous studies of the D-loop of mtDNA polymorphisms revealed the strongest evidence for an independent domestication of zebu and taurine cattle [15-17] that was later confirmed by microsatellite analysis [18, 19]. Based on microsatellite diversity, the hybrid origin of the Near East cattle breeds was verified [19, 20] and the different histories of Mediterranean and Northern European cattle populations were proven [21]. Microsatellites were successfully applied to study genetic diversity, genetic structure and diversity of several Russian cattle breeds [22-24]. With the publication of the complete sequence of a cattle genome in 2009 [25, 26], it became possible to trace the history of breed origins and to evaluate genetic relationships between modern breeds, based on the results of genome-wide SNP screening [27-31]. Conducting genome-wide studies of the allele pool of the local breeds will contribute to the identification of new polymorphisms and to the understanding of biological mechanisms that allow these breeds to adapt and survive in different local environmental conditions.

In the present work, we were the first to show the uniqueness of the allele fund of domestic dairy breeds at genomic level, which allows us to consider these breeds as a reserve of variability for animal husbandry.

The aim of our work was to study the genetic diversity and population structure of five locally-developed Russian cattle breeds, including Bestuzhev, Kholmogor, Kostromsky, Red Gorbatov and Yaroslavl, based on genome-wide single nucleotide polymorphisms (SNP) generated using Illumina Bovine SNP50 BeadChips.

Technique. In total, 116 samples (sperm or tissue) collected from five locally developed Russian cattle breeds were analyzed, including Bestuzhev (BEST, $n = 27$), Kholmogor (KHLM, $n = 25$), Kostromsky (KSTR, $n = 20$), Red Gorbatov (RGBT, $n = 23$) and Yaroslavl breeds (YRSL, $n = 21$). The sam-

ples of North American and German Holstein cattle (HLST, $n = 29$) were used for comparison.

Genomic DNA was extracted from biomaterial samples using Nexttec column (Nexttec Biotechnology GmbH, Germany) following the manufacturer's instructions. The DNA concentration was estimated by measuring the absorbance at 260 nm and the DNA quality was checked by separation on agarose gels. Whole-genome SNP screening was performed using Bovine SNP50 BeadChip (Illumina, San Diego, CA, USA).

Quality filtering of genetic markers was performed in PLINK 1.07 [32]. During an initial quality check, GenCall (GC) and GenTrain (GT) scores were used to assess the accuracy and efficiency of SNP genotyping. A GC score of 0.5 and GT score of 0.3 cutoff was applied to determine valid genotypes for each SNP [33]. At the next stage, the following filters were used: SNPs with less than 90 % of individuals genotyped (--geno 0.1), MAF less than $< 5\%$ (--maf 0.05), Hardy-Weinberg equilibrium test $p < 10^{-6}$ (--hwe 1e-6) and in linkage disequilibrium (--indep-pairwise 50 5 0.5) were pruned. SNPs located on sex chromosomes as well as SNPs with unknown map positions were also deleted. Individuals were initially quality controlled in PLINK 1.07: ones with less than 90 % SNPs genotyped (--mind 0.1) were removed.

Observed and expected heterozygosity, test for deviation from Hardy-Weinberg equilibrium were calculated in PLINK 1.07 [32]. Allelic richness (Ar) for each population was calculated using the program HP-Rare 1.1 [34]. In order to compare the allelic richness of samples having different sizes the procedure of rarefaction was applied.

F_{IS} was calculated using formula: $F_{IS} = (H_{EXP} - H_{OBS})/H_{EXP}$, where H_{OBS} is the observed heterozygosity and H_{EXP} is the expected heterozygosity.

Multidimensional scaling (MDS) based on pairwise identical-by-state (IBS) distance was performed with PLINK 1.07 (--cluster, --mds-plot 4) and visualized with R package 3.2.3 [35].

Population structure was evaluated using admixture model in STRUCTURE 2.3.4 software [36]. We evaluated k values (the number of assumed populations) from 2 to 9 using a burn-in of 50 000 and 50 000 Markov chain Monte Carlo (MCMC) iterations for each value of K . Ten iterations for each value of K were carried out. Average values of similarity coefficient Q in the i -th cluster for the total number of clusters k (Q_i/k) were calculated for each breed.

An unrooted individual phylogenetic tree was constructed using parsimony method with the program Dnapars from the Phylip 3.695 package [37] and visualized with FigTree 1.4.2 (38).

Pairwise Nei's genetic distances [39] were calculated using R package StAMPP [40]. Software Arlequin suite 3.5.2.2 [41] was used to determine pairwise fixation indices (F_{st}) [42] and to perform an analysis of molecular variance (AMOVA).

A rooted tree was constructed using neighbor-joining analysis using M. Nei distances [39] in the Neighbour program from the Phylip 3.695 package [37] and visualized with FigTree 1.4.2 [38]. Reindeer whole-genome genotyping data ($n = 12$) generated applying the same Bead Array were used as an outgroup for the tree.

R version 3.2.3 was used as an instrument for creating input files [35].

Results. After filtering for GT and GC, 1764 SNPs on sex chromosomes were excluded from the initial set of 54609 SNPs. At the subsequent step of the quality filtering, 4901 SNPs failed missingness test ($GENO > 0.1$), 9343 SNPs failed frequency test ($MAF < 0.05$) and 127 markers were excluded based on HWE test ($p \leq 1e-006$). Afterwards 6228 SNPs in linkage disequilibrium were

deleted resulting in the final set of markers of 35874 SNPs that was used for further analyses. Summary statistics for genetic diversity are shown in Table 1.

1. Summary statistic for the genetic diversity of five Russian indigenous cattle breeds compared to Holsteins

Pops	<i>n</i>	Plmph	H _{OBS} (±0.001)	H _{EXP} (±0.001)	F _{IS}	Ar (±0.001)	MAF (±0.001)
BEST	27	33631	0.378	0.372	-0.015	1.955	0.282
KHLM	25	31518	0.390	0.370	-0.054	1.921	0.280
KSTR	20	32115	0.379	0.364	-0.040	1.914	0.275
RGBT	23	31383	0.379	0.368	-0.030	1.922	0.278
YRSL	21	31015	0.381	0.375	-0.018	1.925	0.284
HLST	29	32381	0.377	0.371	-0.016	1.933	0.281

Note. Studied cattle breeds: BEST — Bestizhev, HLST — Holstein, KHLM — Kholomogor, KSTR — Kostromsky, RGBT — Red Gorbato, YRSL — Yaroslavl. Plmph — number of polymorphic SNPs, which pass through quality control, H_{OBS} — observed heterozygosity, H_{EXP} — expected heterozygosity, F_{IS} — fixation index, Ar — rarified allelic richness, MAF — minor allele frequency.

Average heterozygosity within breeds ranged from 0.378 in BEST to 0.390 in KHLM and was higher compared to HLST. The observed tendency of higher heterozygosity values in Russian local cattle breeds could be a result of the less intensive use of artificial insemination sires. Significant higher value of observed heterozygosity in KHLM could be due to the application of a natural mating with own bulls in some cases. Allelic richness is a measure of genetic diversity in either a sample or a population. In our study, it was ranging from 1.914 ± 0.001 in KSTR to 1.955 ± 0.001 in BEST. A slight heterozygote excess was detected in all breeds studied. BEST and YRSL breeds were characterized by smaller heterozygote excess ($F_{IS} = -0.015$, $F_{IS} = -0.018$, respectively), which was comparable to HLST ($F_{IS} = -0.016$), while a greater heterozygote excess was observed in KHLM ($F_{IS} = -0.054$) and KSTR ($F_{IS} = -0.040$).

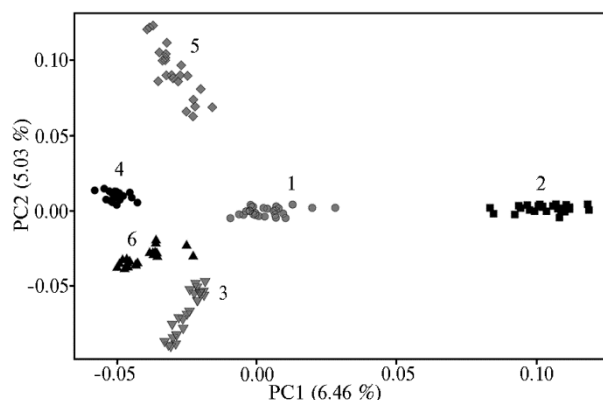


Fig. 1. Genetic variation in 116 individuals of five Russian indigenous cattle breeds and 29 individuals of Holstein breed (comparison group), genotyped at 35874 SNP loci, based on the results of principal component analysis (PCA): 1 — Bestuzhev, 2 — Holstein, 3 — Kholmogor, 4 — Kostromsky, 5 — Red Gorbato, 6 — Yaroslavl (studied cattle breeds).

The results of multi-dimensional scaling (MDS) are presented on Figure 1. The principal component 1 (PC1) accounts for 5.46 % of the genotypic variance and splits the Russian domestic breeds from Holsteins. According to G. McVean [43], locations of samples on the principal component analysis plot generated from genome-wide data can be predicted based on the knowledge of the average coalescent time for pairs of samples. The first principal component can be interpreted as the deepest coalescent event in a tree, and the projection of admixed individuals onto this axis can be used to estimate the proportion of mixture between two parental groups [43]. Four out of the five Russian domestic breeds are located at the same region along the first axis indicating that their common ancestry differs from HLST breed. The position of BEST cattle along this axis, between remaining Russian breeds and Holsteins, can be interpreted as the presence of some admixture of Holsteins or Holstein-related cattle in the modern BEST population. The principal component 2 (PC2) is associated with the divergence of KHLM and YRSL (these two breeds have their

origin in the Northern Great Russian cattle) from Red cattle (RGBT), which was formed with great contribution of Tyrolean cattle [44]. The intermediate position of KSTR cattle is probably associated with their crossbred origin and the contribution of both the Northern Great Russian and the Red cattle in their breeding. It is known that the ancestry population of KSTR breed was formed in XIX century by crossing of domestic well adapted to local environmental and forage conditions Northern Great Russian cattle with KHLM bulls and then with Wilstermarsh, Simmental, Ayrshire and Brown Swiss cattle. In the beginning of the XX century, a multiple crossing with Brown Swiss bulls was carried out [9, 44]. In total, the PC2 accounts for 5.03% of the variance in genotypes.

Phylogenetic analysis based on parsimony method grouped individuals into six clusters according to their breeds. All the individuals from the same breeds were consolidated on the neighboring branches of their respective clusters (Fig. 2).

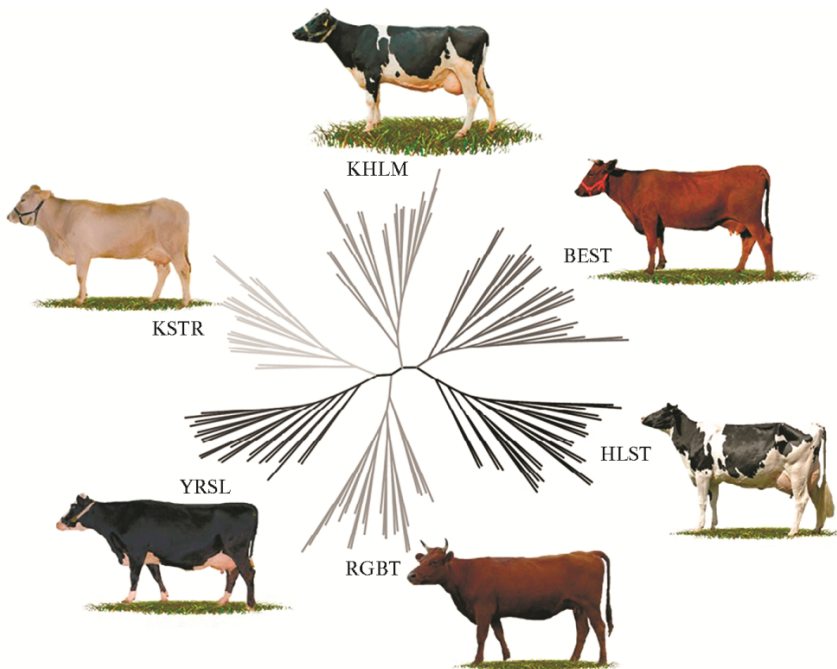


Fig. 2. An unrooted phylogenetic tree of five Russian indigenous cattle breeds (Holstein cattle were included in analysis as comparison group) based on parsimony method: BEST — Bestuzhev, HLST — Holstein, KHLM — Kholomogor, KSTR — Kostromsky, RGBT — Red Gorbatov, YRSL — Yaroslavl (studied cattle breeds).

The STRUCTURE analyses supported the assumption that the ancestry of the locally developed Russian cattle breeds is distinct from Holsteins and Holstein-related breeds (Fig. 3). We found strong split of HLST breed ($Q_{2/2} = 0.956 \pm 0.011$) from the five studied Russian cattle breeds. The HLST contribution in ancestry of four of them is insignificant: $Q_{2/2} = 0.057 \pm 0.008$, 0.003 ± 0.002 , 0.043 ± 0.009 and 0.028 ± 0.008 for KHLM, KSTR, RGBT and YRSL, respectively. The introgression of HLST was observed in BEST ($Q_{2/2} = 0.300 \pm 0.010$), probably, due to using Red Holsteins for the improvement of this breed during last decade. Increasing k values from 3 to 6 showed only slight variation in likelihood scores for HLST from 0.935 ± 0.014 at $k = 6$ to 0.957 ± 0.010 at $k = 4$. At $k = 3$, we observe the KHLM clustering ($Q_{3/3} = 0.862 \pm 0.023$). Some degree of KHLM admixture in YRSL ($Q_{3/3} = 0.389 \pm 0.005$) and BEST ($Q_{3/3} = 0.206 \pm 0.004$) reflects the contribution of KHLM cattle in improvement of Yaroslavl and Bestuzhev breeds

2. Proportion of membership of the studied native Russian cattle breeds in each of the six cluster inferred in STRUCTURE

Pops	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
BEST	0.896±0.015 ^P	0.004±0.002	0.002±0.001	0.005±0.002	0.039±0.006	0.054±0.012
KHLM	0.072±0.012	0.818±0.029 ^P	0.005±0.003	0.013±0.004	0.064±0.009	0.027±0.005
KSTR	0.019±0.007	0.004±0.001	0.923±0.015 ^P	0.009±0.003	0.042±0.006	0.003±0.002
RGBT	0.079±0.013	0.014±0.004	0.018±0.005	0.816±0.027 ^P	0.047±0.006	0.026±0.006
YRSL	0.060±0.014	0.021±0.007	0.017±0.006	0.013±0.005	0.873±0.031 ^P	0.016±0.007
HLST	0.051±0.013	0.002±0.001	0.000±0.000	0.001±0.001	0.011±0.001	0.935±0.014 ^P

Note. BEST — Bestuzhev, HLST — Holstein, KHLM — Kholomogor, KSTR — Kostromsky, RGBT — Red Gorbатов, YRSL — Yaroslavl (studied cattle breeds); the inferred cluster is marked by ^P.

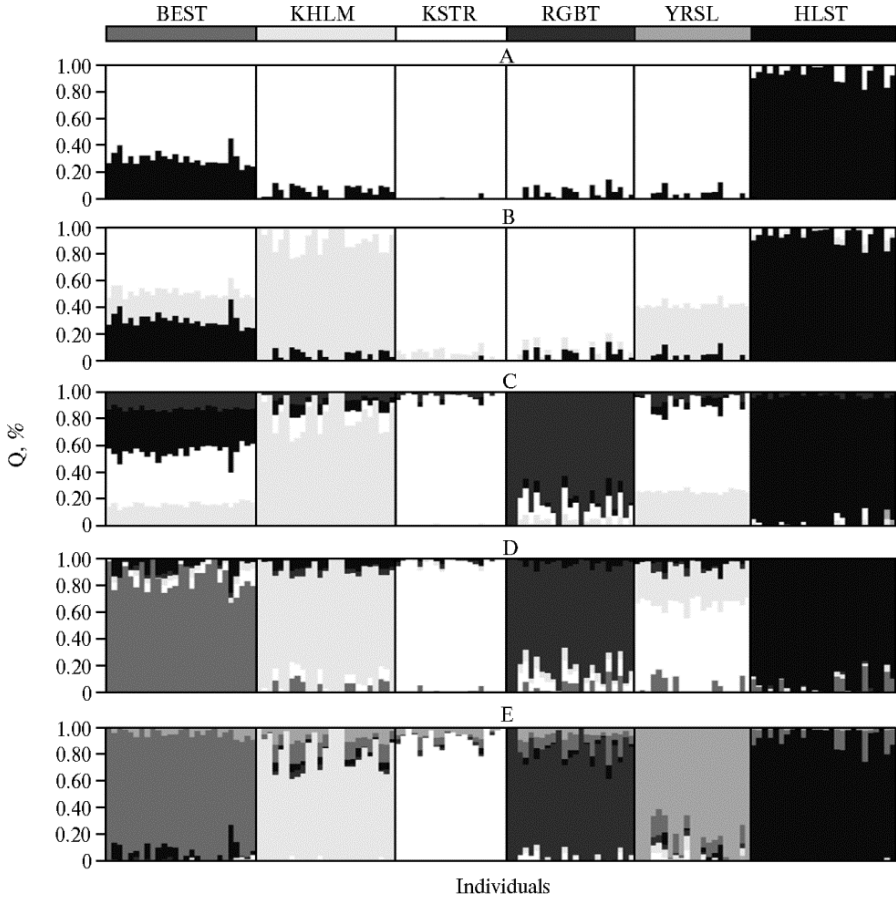


Fig. 3. Population assignment for 145 individuals based on the 35874 markers using STRUCTURE [36] for $k = 2$ (A), $k = 3$ (B), $k = 4$ (C), $k = 5$ (D), $k = 6$ (E): BEST — Bestuzhev, HLST — Holstein, KHLM — Kholomogor, KSTR — Kostromsky, RGBT — Red Gorbатов, YRSL — Yaroslavl (studied cattle breeds); individuals are represented as thin vertical lines, with the proportion of different shades of gray representing their estimated ancestry deriving from different populations. Breeds are labeled by abbreviation at the top of figure.

in XIX-XX centuries [9]. At $k = 4$, the STRUCTURE shows the subdivision of RGBT ($Q_{4/4} = 0.829 \pm 0.026$). The presence of slight RGBT admixture in BEST ($Q_{4/4} = 0.121 \pm 0.003$), probably, reflects the contribution of Tyrolean cattle in improvement of ancestor populations in the XIX century [9, 44]. The highest ΔK showing the assumed number of populations was observed for $k = 6$. At $k = 6$, the genetic structure is in agreement with breed origin of individuals: $Q_{1/6} = 0.855 \pm 0.018$ for BEST, $Q_{2/6} = 0.818 \pm 0.029$ for KHLM, $Q_{3/6} = 0.923 \pm 0.015$ for KSTR, $Q_{4/6} = 0.816 \pm 0.027$ for RGBT, $Q_{5/6} = 0.873 \pm 0.031$ for YRSL and $Q_{6/6} = 0.935 \pm 0.014$ for HLST (Table 2). In total, across a range of k values

from 3 to 6, all of the five studied Russian local cattle breeds consistently showed complex ancestry (Fig. 3). At higher values of k , observed at $k = 6$ among-breed genetic structure predominated.

Analysis of molecular variance (AMOVA) (Table 2) showed highly significant results for genetic differentiation ($p < 0.001$) in studied breeds. AMOVA revealed that most of the genetic variation in cattle breeds was found within populations (91.2 %), and less among populations (8.8 %).

Calculation of the pairwise F_{ST} and Nei's distances (D_N) (Table 3) confirmed the most closeness of BEST breed to HLST ($F_{ST} = 0.0661$, $D_N = 0.0496$), whereas the remaining Russian breeds were approximately equally distant from HLST ($F_{ST} = 0.0963-0.1103$, $D_N = 0.0675-0.0783$).

3. Genetic distances between studied cattle breeds

Pops	BEST	KHLM	KSTR	RGBT	YRSL	HLST
BEST		0.0504	0.0551	0.0542	0.0508	0.0496
KHLM	0.0684		0.0700	0.0670	0.0603	0.0675
KSTR	0.0740	0.1004		0.0715	0.0661	0.0783
RGBT	0.0743	0.0972	0.1027		0.0684	0.0718
YRSL	0.0671	0.0854	0.0931	0.0980		0.0721
HLST	0.0661	0.0963	0.1103	0.1025	0.1014	

Note. BEST — Bestuzhev, HLST — Holstein, KHLM — Kholmogor, KSTR — Kostromsky, RGBT — Red Gorbatov, YRSL — Yaroslavl (studied cattle breeds); pairwise F_{ST} values are shown below diagonal ($p \leq 0,001$); Nei's unbiased distances [39] are shown above diagonal. P-values are given only for F_{ST} values.

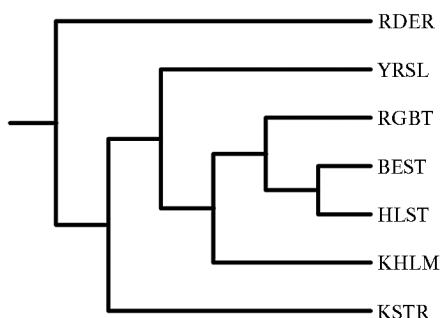


Fig. 4. A rooted phylogenetic tree constructed by neighbor-joining analysis: BEST — Bestuzhev, HLST — Holstein, KHLM — Kholmogor, KSTR — Kostromsky, RGBT — Red Gorbatov, YRSL — Yaroslavl (studied cattle breeds); RDER — reindeer (an outgroup).

The emerging structure of the phylogenetic tree constructed on the Nei genetic distances (Fig. 4), is in full concordance with the historical origin of breeds and confirms the MDS and STRUC-TURE results.

The development of the high throughput genotyping technologies has opened new prospective in the evaluation of genetic diversity in livestock species. The application of genome-wide sets of DNA markers may expand our understanding of population genetic structure of livestock, breed origins and genetic relationship between populations. Whilst numerous studies have been undertaken to

characterize the commercial breeds and some local cattle breeds of Europe, North America, Asia and Africa at whole-genome level [29, 31, 45, 46], little is known about genetic differences, relationships and population genetic structure of the Russian native cattle breeds. The Russian dairy cattle population consists of 8.9 million cows and 949 thousands of them are breeding cows [47, 11]. The most widely used breeds of dairy cattle in Russia are Black-and-White with high degree of Holstein blood and Holsteins. They account for 66 % of the total number of dairy cattle. Russian native cattle breeds account only 15 % of the total dairy cattle population, including 8 % for Kholmogor breed, 6 % for Yaroslavl breed and less than 1 % for all other breeds [47]. Locally developed breeds could however carry the gene pool, which is important for adaptive traits and may become an indispensable source of genetic variability for the future geographically targeted system of milk production. Furthermore, information about genetic diversity and population structure is essential for utilization and conservation of cattle breeds [48].

We evaluated the genetic diversity and population structure of five native

Russian cattle breeds using a set of 35874 polymorphic SNPs generated by the Bovine SNP50K BeadChip (Illumina Inc, San-Diego, USA). We clearly distinguished Russian breeds from each other and from Holstein breed, which was used as comparison. All the investigated cattle breeds revealed a complex origin. Some signals of admixture were observed between several Russian breeds and Holsteins, and between different Russian breeds. The highest contribution of Holsteins or Holstein-related breeds was observed in Bestuzhev breed. Based on PC1 of the MDS analysis the Bestuzhev breed was positioned between other Russian breeds and the Holstein breed showing some degree of membership in the Holstein specific cluster as revealed by STRUCTURE analysis. These two breeds were characterized by lowest pairwise F_{ST} values and formed a common branch at the Nei's phylogenetic tree. The closeness of Bestuzhev breed to Holsteins is probably due to intensive use of Holstein bulls for improvement of this native cattle breed [49]. We observed some admixture of Holsteins or Holstein-related cattle in Kholmogor and Red Gorbatov cattle. In the Kholmogor breed, it is probably associated with the contribution of Dutch cattle in the development of the ancestral populations of this breed. It is known that Dutch cows were distributed among the inhabitants of fertile meads in Kholmogor in the 60s of the XVIII century caused by Empress Catherine II [8]. Geographically close regions of origin of the ancestral populations of Bestuzhev and Red Gorbatov cattle and the similar climatic conditions of their breeding on the banks of the Volga River were depicted by some admixture signals based on the STRUCTURE analysis and the formation of a common branch on the phylogenetic tree. Genetic closeness of Bestuzhev and Red Gorbatov cattle can reflect the contribution of Tyrolean cattle in improvement of both breeds in the XIX century. Most likely, this contribution comes from Tux-Zillertaler cattle, which was widespread in Tyrol due to their unpretentiousness and good productive capacity in poor forage resources. At the agricultural exhibition of 1855 in Tyrol, they were presented as a Tyrolean breed [50]. The export of Tux-Zillertaler cattle to Russia is dated to 1848 [51]. Close localization of Kholmogor and Yaroslavl cattle at the MDS plot, the presence of admixture, observed at values of k from 3 to 5 on the results of the STRUCTURE analysis and relatively low values of F_{ST} and D_N may reflect their common historical origin from the Northern Great Russian cattle [3] and the contribution of Kholmogor cattle in the improvement of Yaroslavl cattle in the XIX and XX centuries [9].

Thus, using the method of genome-wide SNP studies we were able for the first time to study the population structure and genealogical relationships among the five Russian cattle breeds. The received information is the first step towards the evaluation of the value of these breeds regarding their conservation and usage in the agricultural production of the future.

REFERENCES

1. Notter D.R. The importance of genetic diversity in livestock populations of the future. *J. Anim. Sci.*, 1999, 77(1): 61-69 (doi: 10.2527/1999.77161x).
2. Ajmone-Marsan P., GLOBALDIV Consortium. A global view of livestock biodiversity and conservation—GLOBALDIV. *Anim. Genet.*, 2010, 41(s1): 1-5 (doi: 10.1111/j.1365-2052.2010.02036.x).
3. Liskun E.F. *Russkie otrod'ya krupno-rogatogo skota* [Landraces of Russian cattle]. Moscow, 1928 (in Russ.).
4. Diomidov A.M., Zhirkovich E.F. *Razvedenie porod krupnogo rogatogo skota* [Cattle breeding]. Moscow-Leningrad, 1934 (in Russ.).
5. Armfel'd A.A. *Krupnyi rogaty skot v usloviyakh russkikh khozyaistv* [Cattel breeding in Russia]. St. Petersburg, 1904 (in Russ.).
6. Liskun E.F. *Kratkie spravochnye svedeniya o zhivotnykh nekotorykh russkikh khozyaistv* [Brief view of the animals from some farms of Russia]. Petrograd, 1915 (in Russ.).

7. Pridorogin M.I. *Krupnyi rogatyi skot. Vazhneishie porody* [Cattle: the most important breeds]. Moscow, 1919 (in Russ.).
8. Tooke W. View of the Russian empire, during the reign of Catharine the second, and to the close of the eighteenth century. *The Breeding of Sattle*, 1800, 3: 85-129.
9. Dmitriev N.I. *Porody skota po stranam mira* [Cattle breeds — a geography]. Leningrad, 1978 (in Russ.).
10. Ernst L.K., Dmitriev N.G., Paronyan I.A. *Geneticheskie resursy sel'skokhozyaistvennykh zhiivotnykh v Rossii i sopedel'nykh stranakh* [Genetic diversity of farm animals in Russia and neighboring countries]. St. Petersburg, 1994 (in Russ.).
11. *Spravochnik porod i tipov sel'skokhozyaistvennykh zhiivotnykh, razvodimykh v Rossiiskoi Federatsii* /Pod redaktsiei I.M. Dunina, A.G. Dankverta [Breeds and types of farm animals in the Russian Federation: a directory. I.M. Dunin, A.G. Dankvert (eds.)]. Moscow, 2013 (in Russ.).
12. Muir W.M., Wong G.K.-S., Zhang Y., Wang J., Groenen M.A.M., Crooijmans R.P.M.A., Megens H.-J., Zhang H., Okimoto R., Vereijken A., Jungerius A., Albers G.A.A., Lawley C.T., Delany M.E., MacEachern S., Cheng H.H. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *PNAS USA*, 2008, 105(45): 17312-17317 (doi: 10.1073/pnas.0806569105).
13. Dolmatova I.Yu., Zinovieva N.A., Gorelov P.V., Il'yasov A.D., Gladyr E.A., Traspov A.A., Sel'tsov V.I. Characteristics of the allele pool of the Bashkiria population of Simmental cattle using microsatellites. *Agricultural Biology*, 2011, 6: 70-74 (in Engl.).
14. Yang W., Kang X., Yang Q., Lin Y., Fang M. Review on the development of genotyping methods for assessing farm animal diversity. *Journal of Animal Science and Biotechnology*, 2013, 4: 2-6 (doi: 10.1186/2049-1891-4-2).
15. Loftus R.T., MacHugh D.E., Bradley D.G., Sharp P.M., Cunningham E.P. Evidence for two independent domestications of cattle. *PNAS USA*, 1994, 91(7): 2757-2761 (doi: 10.1073/pnas.91.7.2757).
16. Loftus R.T., MacHugh D.E., Ngere L.O., Balain D.S., Badi A.M., Bradley D.G., Cunningham E.P. Mitochondrial genetic variation in European, African and Indian cattle populations. *Anim. Genet.*, 1994, 25: 265-271 (doi: 10.1111/j.1365-2052.1994.tb00203.x).
17. Bradley D.G., MacHugh D.E., Cunningham P., Loftus R.T. Mitochondrial diversity and the origins of African and European cattle. *PNAS USA*, 1996, 93: 5131-5135 (doi: 10.1073/pnas.93.10.5131).
18. MacHugh D.E., Shriver M.D., Loftus R.T., Cunningham P., Bradley D.G. Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics*, 1997, 146: 1071-1086.
19. Loftus R.T., Ertugrul O., Harba A.H., El-Barody M.A.A., MacHugh D.E., Park S.D.E., Bradley D.G. A microsatellite survey of cattle from a centre of origin: the Near East. *Mol. Ecol.*, 1999, 8: 2015-2022 (doi: 10.1046/j.1365-294x.1999.00805.x).
20. Edwards C.J., Baird J.F., MacHugh D.E. Taurine and zebu admixture in Near Eastern cattle: a comparison of mitochondrial, autosomal and Y-chromosomal data. *Anim. Genet.*, 2007, 38: 520-524 (doi: 10.1111/j.1365-2052.2007.01638.x).
21. Cymbron T., Freeman A.R., Isabel Malheiro M., Vigne J.D., Bradley D.G. Microsatellite diversity suggests different histories for Mediterranean and Northern European cattle populations. *Proceedings of the Royal Society B: Biological Sciences*, 2005, 272(1574): 1837-1843 (doi: 10.1098/rspb.2005.3138).
22. Gorelov P.V., Kol'tsov D.N., Zinovieva N.A., Gladyr' E.A. The comparative analysis of blood groups and microsatellites in characteristics of new cattle types of brown Swiss and Sychevskaja breeds. *Agricultural Biology*, 2011, 6: 37-40 (in Engl.).
23. Gladyr' E.A., Zinov'eva N.A., Bagirov V.A., Amirshoev F.S., Volkova V.V., Klenovitskii P.M., Karpov A.P., Ernst L.K. *Dostizheniya nauki i tekhniki APK*, 2012, 8: 58-62 (in Russ.).
24. Kiseleva T.Yu., Kantanen J., Vorobyov N.I., Podobina B.E., Terletsky V.P. Linkage disequilibrium analysis for microsatellite loci in six cattle breeds. *Russian Journal of Genetics*, 2014, 50(4): 406-414 (doi: 10.1134/S1022795414040048).
25. Elsik C.G., Bovine Genome Sequencing and Analysis Consortium. The genome sequence of taurine cattle: A window to ruminant biology and evolution. *Science*, 2009, 324(5926): 522-528 (doi: 10.1126/science.1169588).
26. Zimin A.V., Delcher A.L., Florea L., Kelley D.R., Schatz M.C., Puiu D., Hanrahan F., Pertea G., Van Tassel C.P., Sonstegard T.S., Marçais G., Roberts M., Subramanian P., Yorke J.A., Salzberg S.L. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.*, 2009, 10(4): R42 (doi: 10.1186/gb-2009-10-4-r42).
27. Decker J.E., Pires J.C., Contant G.C., McKay S.D., Heaton M.P., Chen K.,

- Cooper A., Vilkki J., Seabury C.M., Caetano A.R., Johnson G.S., Brenneman R.A., Hanotte O., Eggert L.S., Wiener P., Kim J.-J., Kim K.S., Sonstegard T.S., Van Tassell C.P., Neiberghs H.L., McEwan J.C., Brauning R., Coutinho L.L., Babar M.E., Wilson G.A., McClure M.C., Rolf M.M., Kim J.W., Schnabel R.D., Taylor J.F. Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. *PNAS USA*, 2009, 106(44): 18644-18649 (doi: 10.1073/pnas.0904691106).
28. Kuehn L.A., Keele J.W., Bennett G.L., McDanel D.T.G., Smith T.P., Snelling W.M., Sonstegard T.S., Thallman R.M. Predicting breed composition using breed frequencies of 50,000 markers from the US Meat Animal Research Center 2,000 Bull Project. *J. Anim. Sci.*, 2011, 89(6): 1742-1750 (doi: 10.2527/jas.2010.3530).
 29. McTavish E.J., Decker J.E., Schnabel R.D., Taylor J.F., Hillis D.M. New World cattle show ancestry from multiple independent domestication events. *PNAS USA*, 2013, 110(15): E1398-E1406 (doi: 10.1073/pnas.1303367110).
 30. Decker J.E., McKay S.D., Rolf M.M., Kim J., Molina Alcalá A., Sonstegard T.S., Hanotte O., Götherström A., Seabury C.M., Praharani L., Babar M.E., Regitano L.C.A., Yildiz M.A., Heaton M.P., Wan-Sheng L., Lei C.-Z., Reecy J.M., Saif-Ur-Rehman M., Schnabel R.D., Taylor J.F. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genetics*, 10(3): e1004254 (doi: 10.1371/journal.pgen.1004254).
 31. Decker J.E., Taylor J.F., Kantanen J., Millbrooke A., Schnabel R.D., Alexander L.J., MacNeil M.D. Origins of cattle on Chirikof Island, Alaska, elucidated from genome-wide SNP genotypes. *Heredity*, 2016, 116(6): 502-505 (doi: 10.1038/hdy.2016.7).
 32. Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I., Daly M.J., Sham P.C. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, 2007, 81: 559-575 (doi: 10.1086/519795).
 33. Fan J.B., Oliphant A., Shen R., Kermani B.G., Garcia F., Gunderson K.L., Hansen M., Steemers F., Butler S.L., Deloukas P., Galver L., Hunt S., McBride C., Bibikova M., Rubano T., Chen J., Wickham E., Doucet D., Chang W., Campbell D., Zhang B., Kruglyak S., Bentley D., Haas J., Rigault P., Zhou L., Stuelpnagel J., Chee M.S. Highly parallel SNP genotyping. *Cold Spring Harbor Symposia on Quantitative Biology*, 2003, 68: 69-78.
 34. Kalinowski S.T. HP RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol. Ecol. Notes*, 2005, 5: 187-189 (doi: 10.1111/j.1471-8286.2004.00845.x).
 35. R Development Core Team. *R: A language and environment for statistical computing. 2009. R Foundation for statistical computing.* Vienna, Austria, 2005 (ISBN 3-900051-07-0). Available <http://www.Rproject.org>. No date.
 36. Pritchard J.K., Stephens M., Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*, 2000, 155(2): 945-959.
 37. Felsenstein J. *PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author.* Department of Genetics, University of Washington, Seattle, 1993.
 38. *FigTree*. Available <http://tree.bio.ed.ac.uk/software/figtree/>. No date.
 39. Nei M. Genetic distance between populations. *Am. Nat.*, 1972, 106(949): 283-292.
 40. Pembleton L.W., Cogan N.O.I., Forster J.W. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol. Ecol. Resour.*, 2013, 13(5): 946-952 (doi: 10.1111/1755-0998.12129).
 41. Weir B.S., Cockerham C.C. Estimating F-Statistics for the analysis of population structure. *Evolution*, 1984, 38(6): 1358-1370 (doi: 10.2307/2408641).
 42. Excoffier L., Lischer H.E.L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.*, 2010, 10: 564-567 (doi: 10.1111/j.1755-0998.2010.02847.x).
 43. McVean G. A genealogical interpretation of principal components analysis. *PLoS Genetics*, 2009, 5(10): e1000686 (doi: 10.1371/journal.pgen.1000686).
 44. Liskun E.F. *Otechestvennye porody krupnogo rogatogo skota* [Cattle — Domestic breeds]. Moscow, 1949 (in Russ.).
 45. The Bovine HapMap Consortium. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*, 2009, 324: 528-532 (doi: 10.1126/science.1167936).
 46. Makina S.O., Muchadeyi F.C., Marle-Köster E., Mac Neil M.D., Maiwashe A. Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel. *Frontiers in Genetics*, 2014, 5: 333 (doi: 10.3389/fgene.2014.00333).
 47. Labinov V.V. *Rezultaty i osnovnye napravleniya sovershenstvovaniya selekcionno-plemennoi raboty v molochnom skotovodstve Rossijskoi Federatsii. Ministerstvo sel'skogo khozjajstva RF* [The achievements and improvement of cattle breeding in the Russian Federation. Ministry of Agriculture of the Russian Federation. Available http://www.mcx.ru/documents/file_document/v7_show/33370.133.htm. Ac-

- cessed May 28, 2016.
48. Groeneveld L.F., Lenstra J.A., Eding H., Toro M.A., Scherf B., Pilling D., Negrini R., Finlay E.K., Jianlin H., Groeneveld E., Weigend S., The GLOBALDIV Consortium. Genetic diversity in farm animals — a review. *Anim. Genet.*, 2010, 41: 6-31 (doi: 10.1111/j.1365-2052.2010.02038.x).
 49. Katmakov P.S., Gavrilenko V.P., Bushov A.V., Sten'kin N.I. *Vestnik Ul'yanovskoi GSKHA*, 2014, 1(25): 126-132 (in Russ.).
 50. Moser C., Reiter M. *Die Rinderrasse der Tux-Zillertaler — Ein Stück Tiroler Kultur*. Verlag Edition Tirol, Reith im Alpbachtal, 1996.
 51. Tux Zillertaler. Available http://en.zar.at/Cattle_breeding_in_Austria/Cattle_breeds/Additional_Breeds/Tux_Zillertaler.html. No date.