Original Scientific Paper
DOI: 10.7251/AGRENG1603110K
UDC 575:636.2(437.6)

GENOMIC DETERMINATION OF THE MOST IMPORTANT FATHER LINES OF SLOVAK PINZGAU COWS

Veronika KUKUČKOVÁ*, Nina MORAVČÍKOVÁ, Radovan KASARDA

Department of Animal Genetics and Breeding Biology, Faculty of Agrobiology and Food Resources, Slovak University of Agriculture in Nitra, Slovakia
*Corresponding author: veron.sidlova@gmail.com

ABSTRACT

The aim of this study was to assess genetic structure of Slovak Pinzgau population based on polymorphism at molecular markers using statistical methods. Female offspring of 12 most frequently used bulls in Slovak Pinzgau breeding programme were investigated. Pinzgau cattle were found to have a high level of diversity, supported by the number of alleles observed across loci (average 5.31, range 2-11) and by the high within-breed expected heterozygosity (average 0.66, range 0.64-0.73). The state of genetic diversity is satisfying and standard for local populations. Detection of 12 possible subpopulation structures provided us with detailed information of the genetic structure. The Bayesian approach was applied, detecting three, as the most probable number of clusters. The similarity of each subpopulation using microsatellites was confirmed also by high-throughput molecular data. The observed inbreeding (F_{ROH}=2.3%) was higher than that expected based on pedigree data (F_{PED}=0.4%) due to the limited number of available generations in pedigree data. One of the most important steps in development of efficient autochthonous breed protection programs is characterization of genetic variability and assessment of the population structure. The chosen set of microsatellites confirmed the suitability in determination of the subpopulations of Pinzgau cattle in Slovakia. The state of genetic diversity at more detailed level was successfully performed using bovineSNP50 BeadChip.

Keywords: *genetic differentiation, microsatellites, Pinzgau cattle, SNP chip, structure.*

INTRODUCTION

Slovak Pinzgau cattle belong to the traditional livestock breeds, mainly in upland regions in Slovakia. Nowadays, this breed belongs to the endangered populations (Kadlečík et al., 2004) due to drastic decreasing of the animal counts. Currently loss of genetic resources concerns not only the extinction of traditional breeds, but also the loss of genetic diversity within breeds. Most of the endangered breeds are specialized in a particular habitat or production system and represent, in both developed and developing countries, a unique resource to meet present and future

breeding objectives. Therefore, thorough information on diversity and population structure in cattle is urgently needed to serve as a rational basis for the conservation and possible use of indigenous cattle breeds as genetic resources to meet potential future demands (Taberlet et al., 2008).

Markers are used by population geneticists to investigate the origin, genetic diversity and population structure of alleles, by evolutionists to describe genetic relationship among species or populations and by geneticists to study linkage disequilibrium within or between genes (Liu and Muse, 2005). Molecular markers based on DNA have a very high polymorphism level, and they have been successfully used for evaluation of genetic diversity and variation in breeding programs with an impact on the level of genetic conservation schemes (Židek and Kasarda, 2010).

The inbreeding coefficient is defined as the probability that a pair of alleles is identical by descent (IBD). Historically, geneticists have estimated this probability using pedigree data though genomic information should lead to a more accurate depiction (Bjelland et al., 2013). Increased levels of inbreeding would appear genomically as an increase in the frequency of homozygous alleles. A problem with this method is that alleles that are IBD and identical by state (IBS) cannot be distinguished and are both included in this measure of inbreeding. An alternative method involving genomic runs of homozygosity (ROH) attempts to distinguish these differences and has been used in human (Kirin et al., 2010) as well as cattle genomics (Bjelland et al., 2013; Ferenčaković et al., 2013), examining population history. The ROH are consequence of inbreeding and relatively close relationships between parent pairs, especially in small endangered populations (Mészáros et al., 2015).

The aim of this study was to evaluate the genetic diversity and population structure of Slovak Pinzgau cattle based on polymorphism in genotyping data using statistical programs.

MATERIALS AND METHODS

Selected cows of Pinzgau cattle originated from Slovakia were analysed. DNA of 140 animals was isolated from hair roots and amplified in one multiplex PCR with 10 microsatellites. To determine the polymorphism of microsatellite DNA sequences fluorescent fragmentation analysis by capillary electrophoresis (ABI PRISM 310 Genetic Analyser) was used and the alleles' sizes were evaluated using software Gene Mapper 4.0. Average number of alleles per subpopulation of fathers, Shannon information index, observed heterozygosity, gene diversity (expected heterozygosity) and inbreeding coefficient (F_{MST}) were calculated by GenAlex 6.5 (Peakall and Smouse, 2012).

The most important fathers of cows were genotyped using BovineSNP50 v2 BeadChip (Illumina Inc., San Diego, CA). Only 12 bulls (fathers of 140 cows) with minimum of 5 and maximum of 34 daughters were chosen. SNP markers with more than 10% of missing genotypes, SNPs with less than 0.01 minor allele frequency (MAF) as a threshold to declare a polymorphic SNP and individuals

with low genotyping (< 95%) were excluded. The inbreeding coefficient was calculated first by GenAlex 6.5 software (Peakall and Smouse, 2012; F_{SNP}) and then as ROH-based estimates of autozygosity (F_{ROH_MAF}). Pruning SNPs that show low MAF can affect the results (Albrechtsen et al., 2010) thus quality control setting GeneCall ≤ 0.7 and GeneTrain ≤ 0.4 score was used to evaluate inbreeding coefficient as well ($F_{ROH\_GC\_GT}$). In our analysis autozygosity was defined by ROHs that were > 4 Mb following the study of Ferenčaković et al. (2013).

Subsequently, estimation of subpopulation structure using prior information about fathers was performed. Mixture partition based on pre-defined clustering using Bayesian Analysis of Population Structure (BAPS v. 6.0) software was executed, further described in Cheng et al. (2013). For analysis of relatedness and principal component analysis (PCA) of SNP data a high-performance computing toolset gdsfmt and SNPRelate (R packages for multi-core symmetric multiprocessing computer architectures) were used according to Zheng et al. (2012).

RESULTS AND DISCUSSION

All evaluated cows were divided to 12 groups by fathers and summary statistics for each group were calculated (Table 1). The number of alleles over subpopulations and loci ranged from 2-11 with the mean 5.31±0.15. Regarding the Shannon's information index (I), all groups of fathers presented a value distant from zero with an overall mean of 1.31±0.03. The overall average of observed heterozygosity (H_o=0.77±0.02) has reached higher values than expected (H_e =0.66±0.01) and indicated the presence of high level of heterozygosity in native local cattle breeds. Expected heterozygosity and mean number of alleles calculated here were similar to those obtained in endangered German Pustertaler Sprinzen (0.69 and 5.3), Pinzgauer (0.71 and 6) and Simmental (0.58 and 5.2; Edwards et al., 2000).

The average value of F_{MST} reached a negative number (-0.17±0.02), generally it can be concluded there is no reduction of heterozygosity in daughters of evaluated bulls, whereas the inbreeding in bull Nero was F_{SNP} = 0.003. Positive F values could be derived from inbreeding or from the presence of a substructure within the population. The ROH greater than 4 Mb cover on average 2.3% of genome (F_{ROH_MAF} = 0.0225 and $F_{ROH\_GC\_GT}$ = 0.0234). The observed inbreeding was higher than that expected based on pedigree data (F_{PED}=0.4%). According to pedigree data only 5 animals have arisen by breeding of related animals, whereas based on $F_{ROH>4}$ even 11 animals were inbred. Ferenčaković et al. (2013) showed higher inbreeding level in Pinzgau from Austria (F_{ROH4} = 0.037) compared to Slovak Pinzgau (F_{ROH4} = 0.023) from this study. The Austrian bull Nero had the highest inbreeding $F_{ROH\_GC\_GT}$ = 5.1%. (4.6% calculated using F_{ROH_MAF}), while Carlo with Canadian origin had zero inbreeding. It was noticeable that sires with Austrian origin had overall higher F_{ROH} levels.

Table 1. The number of daughters (N) and alleles (N_A), information index (I), observed (H_o) and expected (H_e) heterozygosity and inbreeding coefficient based on microsatellite data (F_{MST}), based on high-throughput molecular data (F_{SNP}), based on runs of homozygosity (F_{ROH_MAF} and $F_{ROH\_GC\_GT}$) values per subpopulation of fathers

Father	N	N_A	I	H_o	H_e	F_{MST}	F_{SNP}	F_{PED}	F_{ROH_MAF}	$F_{ROH\_GC\_GT}$
ATLAS	13	6	1.408	0.754	0.707	-0.105	-0.020	0.0001	0.045	0.045
CARLO	16	6.3	1.448	0.806	0.725	-0.147	-0.111	0	0	0
GOMOL	13	5.8	1.325	0.723	0.672	-0.115	-0.081	0	0.002	0.002
LODRON	5	4	1.171	0.82	0.693	-0.299	-0.021	0.004	0.033	0.035
LOLTEL	6	4.1	1.114	0.667	0.638	-0.167	-0.032	0	0.025	0.027
LUTGO	13	5.2	1.262	0.769	0.668	-0.204	-0.009	0.008	0.043	0.046
LUTLUX	13	5.6	1.366	0.754	0.707	-0.116	-0.033	0.008	0.015	0.014
NERO	34	7.8	1.543	0.824	0.738	-0.131	0.003	0	0.045	0.051
NOBMON	7	4.7	1.227	0.786	0.669	-0.255	-0.057	0.031	0.007	0.012
NOBTELO	7	4.5	1.195	0.771	0.677	-0.203	-0.037	0	0.020	0.020
ROMIL	7	4.9	1.31	0.771	0.721	-0.157	-0.073	0	0.027	0.026
SAMFO-ET	6	4.8	1.344	0.8	0.747	-0.188	-0.082	0	0.007	0.007

Further analysis was performed using prior information about subpopulations from microsatellite markers. Partitioning of Pinzgau cows according father of cows is visible in figure 1. Each individual that was clustered is represented by a vertical bar having the colour corresponding to the cluster where it was placed. From 12 fathers 3 main clusters was created based on Bayesian approach. Red colour marked cluster represents line COS, bull Carlo with Canadian origin. Austrian bull Nero representing line NUS, Atlas (Austrian origin) representing AER line and Slovak bull Loltel from line LOZ belong to the second cluster (green colour).

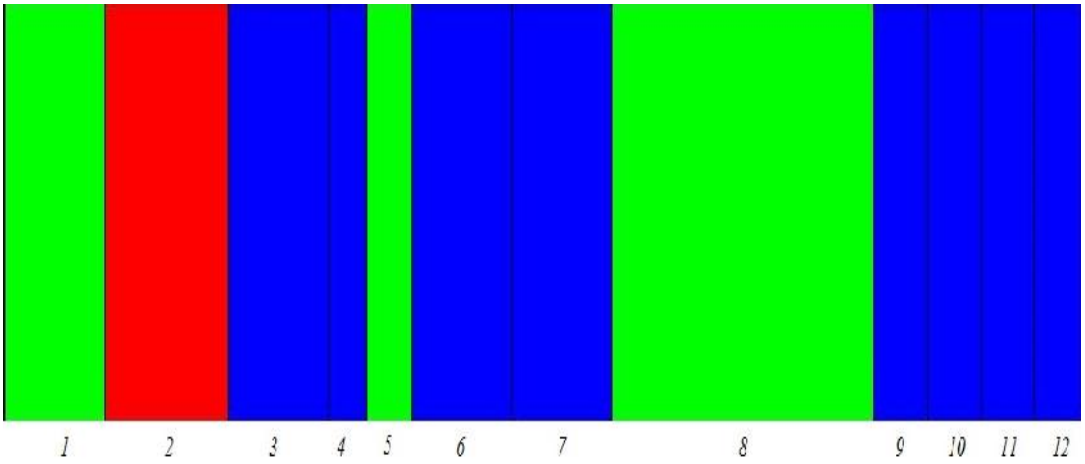

Figure 1. Graphical presentations of the population structure analyses for a sample of 140 Pinzgau cows (using prior information about subpopulations) based on father of cows. Atlas (1), Loltel (5) and Nero (8) in green; Carlo (2) in red; Gomol (3), Lodron (4), Lutgo (6), Lutlux (7), Nobmon (9), Nobtelo (10), Romil (11) and Samfo-et (12) in blue colour.

The most important 12 Pinzgau bulls used in breeding were successfully genotyped using Illumina BovineSNP50 BeadChip with total call rate 99.95%. Genotyping results revealed that 43,068 SNPs (78.87%) were polymorphic (MAF>0.01) with average minor allele frequency ranged from 0.2588 ± 0.1433 on chromosome 2 to 0.2766 ± 0.1403 on chromosome 23. The average values of MAF groups are summarized in table 2.

Table 2. Minor allele frequency (MAF) across autosomes in 12 Slovak Pinzgau bulls with 95% confidence interval (CI) of the mean

MAF	Number of loci	Mean	SD	Min	Max	Lower 95% CI	Upper 95% CI
0.01-0.1	7380	0.0620	0.0208	0.0417	0.0909	0.0615	0.0625
0.1-0.2	7266	0.1459	0.0209	0.125	0.1818	0.1455	0.1464
0.2-0.3	10773	0.2507	0.034	0.2083	0.2917	0.2500	0.2513
0.3-0.4	7664	0.3544	0.0209	0.3182	0.375	0.3539	0.3548
0.4-0.5	9985	0.4505	0.0314	0.4091	0.5	0.4498	0.4511

The PCA and ancestry models were used to cluster animals, to explore the relationships within breed using high-throughput molecular data. First three principal components (PC) are explaining 21.21% of genetic variability (1.PC = 8.44%, 2.PC = 7.38%, 3.PC = 5.39%. Using the three first PC, the mean pairwise distance between the individuals from the Pinzgau population was plotted (Figure 2). First three PC are the most informative plotting on a three-dimensional scatter diagram to allow visual inspection of the relationships among the breed (Dixit et al., 2012). The PCA is used to characterize how different multiple populations are, often using only the two first PC (Albrechtsen et al., 2010). Nobtelo and Nobmon representing line NOB, Lutgo, Lutgo representing line LUZ, Romil and Samfo-et created separate clusters while theirs daughters using microsatellite data were in common cluster also with Gomol and Lodron. Nero, Loltel and Atlas created one cluster by microsatellite analysis and also by SNP Chip data were genetically more similar.

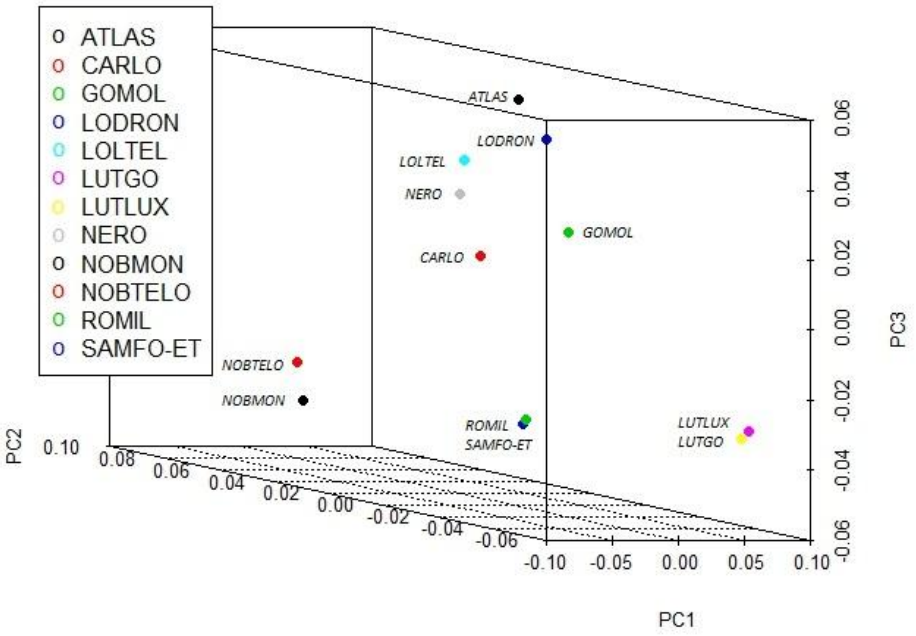

Figure 2. The principal component analysis of 12 Slovak Pinzgau bulls. First 3 principal components (PC) are explaining 21.21% of genetic variability

CONCLUSION

Basic parameters of genetic diversity in traditional Slovak cattle were analysed to determine the level of heterozygosity and inbreeding within population. In spite of significant decrease of population, the state of genetic diversity is satisfying and standard for local populations in comparison to the generally accepted numbers. The proportion of the genome present in ROH provides a good indication of inbreeding levels. The observed inbreeding (F_{ROH}=0.023) was higher than that expected based on pedigree data (F_{PED}=0.004). Genetic structure of Pinzgau cattle has been characterised using set of 10 microsatellites. The similarity of each subpopulation of fathers using microsatellites was confirmed also by high-throughput molecular data. Genomic confirmation of existence of separated breed specific substructures as bull lines allows for more accurate mating strategy and control over inbreeding increase in the breeding programme. Deeper analysis of high-throughput data could provide us with bull line specific regions or SNPs, for which more animals to be sequenced as a basis for preservation of the breed in the original phenotype.

ACKNOWLEDGEMENT

This work has been supported by the Slovak Research and Development Agency (Contracts No. APVV-14-0054, SK-AT-2015-0016). Part of the work was done during the stay of the first and second author at BOKU Vienna, supported by the Austrian Agency for International Cooperation in Education and Research (OeAD-GmbH, project SK 07/2016).

REFERENCES

Albrechtsen A., Nielsen F.C., Nielsen R. (2010). Ascertainment biases in SNP chips affect measures of population divergence. Molecular Biology and Evolution, vol. 27, pp. 2534-2547. doi: http://dx.doi.org/10.1093/molbev/msq148

Bjelland D.W., Weigel K.A., Vukasinovic N., Nkrumah J.D. 2013. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. Journal of Dairy Science, vol. 96, pp. 4697-4706. doi: http://dx.doi.org/10.3168/jds.2012-6435

Cheng L., Connor T.R., Sirén J., Aanensen D.M., Corander J. (2013). Hierarchical and spatially explicit clustering of DNA sequences with BAPS software.

Molecular Biology and Evolution, vol. 30, no. 5, pp. 1224-1228. doi: http://dx.doi.org/10.1093/molbev/mst028

Dixit S.P., Verma N.K., Aggarwal R.A.K., Vyas M.K., Rana J., Sharma A. (2012). Genetic diversity and relationship among Indian goat breeds based on microsatellite markers. Small Ruminant Research, vol. 105, no. 1-3, pp. 38-45. doi: http://dx.doi.org/10.1016/j.smallrumres.2011.11.026

Edwards C.J., Dolf G., Looft C., Loftus R.T., Bradley D.G. (2000). Relationship between the endangered Pustertaler-Sprinzen and three related European cattle breeds as analysed with 20 microsatellite loci. Animal Genetics, vol. 31, no. 5, pp. 329-332.

Ferenčaković M., Hamzić E., Gredler B., Solberg T.R., Klemetsdal G., Curik I., Sölkner J. (2013). Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. Journal of Animal Breeding and Genetics, vol. 130, pp. 286-293. doi: http://dx.doi.org/10.1111/jbg.12012

Kadlečík O., Swalve H.H., Lederer J.A., Grosu H. (2004). Development of dual – purpose Pinzgau cattle. SPU, Nitra, Slovak Republic, 128 p. ISBN 80-8069-439-7.

Kirin M., McQuillan R., Franklin C., Campbell H., McKeigue P., Wilson J. (2010). Genomic runs of homozygosity record population history and consanguinity. PLoS One, vol. 5, e13996. doi: http://dx.doi.org/10.1371/journal.pone.0013996

Liu K., Muse S.V. (2005) Integrated analysis environment for genetic marker data. Bioinformatics, vol. 21, no. 9, pp. 2128-2129. doi: http://dx.doi.org/10.1093/bioinformatics/bti282

Mészáros G., Boison S.A., Pérez O´Brian A.M., Ferenčaković M., Curik I., da Silva M.V.B., Utsunoomiya Y.T., Garcia J.F., Solkner J. (2015). Genomic analysis for managing small and endangered populations: A case study in Tyrol Grey cattle. Frontiers in Genetics, vol. 6, pp. 173. doi: http://dx.doi.org/10.3389/fgene.2015.00173

Peakall R. and Smouse P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics, vol. 28, pp. 2537-2539. doi: http://dx.doi.org/10.1093/bioinformatics/bts460

Taberlet P., Valentini A., Rezaei H.R., Naderi S., Pompanon F., Negrini R., Ajmone-Marsan P. (2008). Are cattle, sheep, and goats endangered species? Molecular Ecology, vol. 17, no. 1, pp. 275-284. doi: http://dx.doi.org/10.1111/j.1365-294X.2007.03475.x

Zheng X., Levine D., Shen J., Gogarten S.M., Laurie C., Weir B.S. (2012). A high-performance computing toolset for relatedness and principal component

analysis of SNP data. Bioinformatics Applications Note, vol. 28, no. 24, pp. 3326-28. doi: http://dx.doi.org/10.1093/bioinformatics/bts606

Židek R., Kasarda R. (2010). Distribution of genetic distance within groups with different relationship coefficient. Acta fytotechnica et zootechnica, vol. 13, pp. 73-76.