Genetic Distances and Admixture between Sire Lines of the Old Kladruber Horse

Hana VOSTRÀ-VYDROVÀ [1]
Barbara HOFMANOVÀ [2]
Ivan MAJZLIK [2]
Alexandra NOVOTNÀ [1]
Lubos VOSTRY [1,2] (✉)

Summary

The objective of the study was to determine the genetic distances and admixture in sire lines of Old Kladruber horse (populations in the context of a currently conducted conservation program) base on microsatellite analysis. Basic molecular parameters were estimated. The observed heterozygosity ranged from 0.48 to 0.85. The expected heterozygosity estimated for the sire lines was higher than observed heterozygosity. The minor genetic differences between sire lines of Old Kladruber horse were observed. Differences between the sire lines were identified using discriminant analysis. The first discriminant function separates the individuals of black and grey variants. For the discriminant function, however, the differentiation between sire lines was not so significant. These results showed that the breeding strategy revision is suggested in Old Kladruber horse.

Key words

population structure; genetic diversity, endangered breed

[1] Institute of Animal Science, Přátelství 815, 104 01, Praha - Uhříněves, Czech Republic
✉ e-mail: vostry@af.czu.cz
[2] Czech University of Life Sciences Prague,
Kamýcká 129, 16500, Praha 6 - Suchdol, Czech Republic

Received: May 16, 2017 | Accepted: July 31, 2017

ACKNOWLEDGEMENTS
This work was financially supported through a grant from the Ministry of Agriculture of the Czech Republic (Project no. QJ1510141 and QJ1330189) and "S grant of MSMT CR".

Introduction

Old Kladruber is the only horse breed that can be considered the original Czech breed that was bred especially for ceremonial purposes and aristocracy. The baroque character of the breed has been conserved until now. This breed is an important gene resource with unique characteristics and high cultural and historical value and has been continually kept in the territory of the Czech Republic for more than four hundred years.

The origin of this breed dates back to 1579 on the Kladruby nad Labem stud farm. The main objective was to produce horses for the Imperial Court of the Hapsburgs (Bílek 1957). The Old Spanish horses, which were mainly imported from Italian stud farms, were used for this purpose. A decrease in breeding popularity of the Old Spanish breed at the turn of the 18th and 19th centuries led to successive changes in the breeding program on the Kladruby nad Labem stud farm. The closed herd turnover began to be used, and it has been considered as a separate breed since then. Old Kladruber and Lipizzaner horses were principally bred for the Imperial Court in Vienna on the same basis, i.e., Old Italian and Old Spain blood as the "baroque" horses with different uses: Old Kladruber horses as carriage horses and Lipizzaner horses as riding horses. As a carriage horse, the Old Kladruber differs from the Lipizzaner in size, body conformation and performance capacities.

Affiliation to the paternal line is one of the main criteria of a conservation program for the creation of the next generation parent couples. In a previous paper (Vostrá-Vydrová et al., 2016) we analyzed the genealogical information in order to ascertain the low level of genetic variability within the Old Kladruber horse. All results showed that the small genetic variability exist as the result of a small and closed population.

The objective of this study was to determine the genetic differences between sire lines of the Old Kladruber horse based on microsatellite analysis.

Material and methods

Animal data collection

In this study, 265 individuals from nine sire lines of Old Kladruber horse breed within a 38-year period (1978–2015) were used. The Old Kladruber horse lines, including the year of birth of the line founder, are described in Table 1. In the analysis 17 animals of Generale sire line, 42 animals of Generale-Generalissimus sire line, 30 animals of Favory sire line, 22 animals of Favory-Generalissimus, 82 animals of Sacramoso sire line,18 animals of Rudolfo sire line, 30 animals of Solo sire line, 12 animals of Siglavi Pakra sire line and 30 animals of Romke sire line were included. The total set of 13 microsatellite markers (AHT4, AHT5, ASB2, HMS1, HMS2, HMS3, HMS6, HMS7, HTG4, HTG6, HTG7, HTG10 and VHL20) recommended for parentage testing by the International Society for Animal Genetics (ISAG) and Equine Genetics Standing Committee was used for the analysis.

Genetic diversity

Genetic variability within populations was characterized as allele frequency, mean number of alleles, observed heterozygosity (H_O), genec diversity, which is often called expected heterozygosity (Weir, 1996), testing of the Hardy-Weinberg equilibrium, the values of inbreeding, expressed as Wright's fixation indices – within a population (F_{IS}) and between individuals within subpopulations (F_{ST}) were done using the PEGAS (Paradis, 2010) package. These indices provide more reliable results specifically for microsatellite data. The polymorphism information content (PIC; Botstein et al., 1980) was calculated using PICcalc (Nagy, 2012).

Population structure and genetic relationship

To determine genetic structure and to infer genetic admixtures, a discriminant analysis of principal components (DAPC) implemented in the Adegenet R package (Jombart and Ahmed, 2011) was used. The DAPC approach proposes an optimum distribution of individuals into predefined groups in relation to the discriminant function of principal components. An optimum number of clusters was defined by the K-averaging algorithm that makes use of the Bayesian information criterion. In addition, the DAPC was used to assign individuals and to obtain the membership probability which presents the overall genetic background of an individual. A trade-off between the power of discriminant analysis and overfitting of the given analysis was assessed by the α-score (Jombart and Ahmed, 2011).

Results and discussion

Genetic diversity

Each of the analysed loci appeared as polymorphic, and their alleles were present in or shared by all studied sire lines. The total number of alleles found on 13 microsatellite markers across all sire lines was 92. The average number of alleles per locus was 7.08, ranging from 5 (HMS1) to 9 alleles (VHL20 and HTG10). The acquired PIC values were estimated from 0.289 (HTG7) to 0.78

Table 1. Name, colour variation, year of birth, breed and country of origin of the Old Kladruber horse line founder

Name	Colour	Born	Breed	Origin
Generale	Gray	1787	Old Kladruber	Czech Republic
Generale - Generalissimus	Gray	1938	Old Kladruber	Slovakia
Favory	Gray	1779	Old Kladruber	Slovakia
Favory - Generalissimus	Gray	1965	Old Kladruber	Czech Republic
Sacramoso	Gray/Black	1800	Italo-Spanish	Italy
Solo	Black	1927	Old Kladruber	Czech Republic
Siglavi Pakra	Black	1946	Lipizzaner	Slovenia
Romke	Black	1966	Friesian	Netherlands
Rudolfo	Gray	1968	Lusitanian	Portugal

Table 2. Characteristics of 13 microsatellite loci analysed in nine sire lines (n=265). Number of successfully genotyped individuals (N), number of alleles (N_A), polymorphism information content (PIC), observed heterozygosity (H_O), gene diversity, Wright's F_{ST} and HWE statistics

Locus	N_A	PIC	H_O	Gene diversity	F_{ST}	F_{IS}	HWE
VHL20	9	0.784	0.8484	0.8517	0.0748	-0.0766	**
HTG4	7	0.707	0.7615	0.7732	0.0889	-0.0809	*
AHT4	8	0.654	0.7558	0.8173	0.0831	-0.0086	**
HMS7	6	0.462	0.5850	0.6720	0.0927	0.0405	***
HTG6	6	0.153	0.4787	0.4719	0.0687	-0.0892	***
AHT5	7	0.484	0.7450	0.7505	0.0849	-0.0848	***
HMS6	6	0.487	0.7127	0.6842	0.0497	-0.0962	***
ASB2	8	0.720	0.8101	0.8039	0.0386	-0.0482	
HTG10	9	0.611	0.6132	0.6611	0.0475	0.0262	*
HTG7	6	0.289	0.4676	0.5525	0.1178	0.0405	***
HMS3	8	0.580	0.6058	0.6734	0.0440	0.0590	***
HMS2	7	0.662	0.7464	0.7387	0.0406	-0.0530	
HMS1	5	0.330	0.5025	0.5443	0.0436	0.0347	

where: * P<0.05, **P<0.01,***P<0.001

(VHL20). These values indicate a relatively good polymorphism level according to Botstein et al. (1980). PIC is calculated with the total number of alleles and allele frequencies in a population. If PIC value is higher than value 0.75 the locus becomes much more informative. The observed heterozygosity between loci was from 0.47 (HTG7) to 0.85 (VHL20). The lowest values of genetic diversity were revealed in the HTG6 locus (0.47). Similarly, as for observed heterozygosity, the highest value of gene diversity was found in the VHL20 locus (0.85). The higher value of expected heterozygosity then observed heterozygosity estimated in this study can be explained by the fact that there is mixing between sire lines due to alternative mating plan, which is used for the control of inbreeding using rotating mating between the lines. Overall information about differences and total statistics are shown in Table 2. Statistically significant deviations from the Hardy-Weinberg equilibrium were found in more than two thirds of the microsatellite loci.

All studied microsatellite markers showed high variability, and only in the markers ASB2, HMS2, HMS1 were not deviations from the Hardy-Weinberg equilibrium discovered. Generally, the genetic diversity of microsatellite loci can be affected by many factors, including genetic drift, impact of selective breeding, effect of individual stallions and random effects (Petersen et al., 2013). The determined values of genetic diversity were lower than, e.g., those in Polish cold-blood horses (Iwańczyk et al., 2006) or Polish Konik (Szwaczkowski et al., 2016). The observed lower proportion of gene diversity distributed within autochthonous breeds can be explained mainly by the fact that the local breeds generally show higher degrees of diversity than breeds with more limited numbers of stallions and mares in the breeding scheme. Observed heterozygosity and the coefficient of inbreeding measured by Wright's F_{IS} index, which are used as general indicators of the genetic diversity level, indicate a sufficient level of heterozygosity in the studied population. These results are not in agreement with the values of inbreeding coefficient estimated by pedigree analysis. The average values of the inbreeding coefficient estimated by pedigree analysis were 13% (with a maximum value of 29%) for the reference population (individuals that can currently take part in reproduction) (Vostrá-Vydrová et al., 2016). The molecular inbreeding measured with a handful of molecular markers is not necessarily a good predictor of the genealogical or genomic inbreeding, that there are problems in estimation genomic heterozygosity using only a few molecular markers (Toro et al., 2009).

Genetic structure and level of admixture

To infer the population genetic structure and to assess the level of admixture, discriminant analysis of principal components was applied to genotyping data. The distribution of individuals according to the BIC analysis showed that inferred clusters do not correspond to actual groups (Figure 1). For this reason, the clusters in DAPC were inferred in line with the prior assumption of population distribution (K=10). The agreement between prior and posterior assignment was 78.10%. Based on the α-score (Jombart and Collins, 2015), which indicates the number of

Figure 1.
Genetic clusters determined using discriminant analysis of principal components where 1 – Generalle, 2 - Generale-Generalissimus, 3 - Favory, 4 - Favory-Generalissimus, 5 – Sacramoso, 6 – Solo, 7 - Siglavi Pakra, 8 – Romke, 9 - Rudolfo

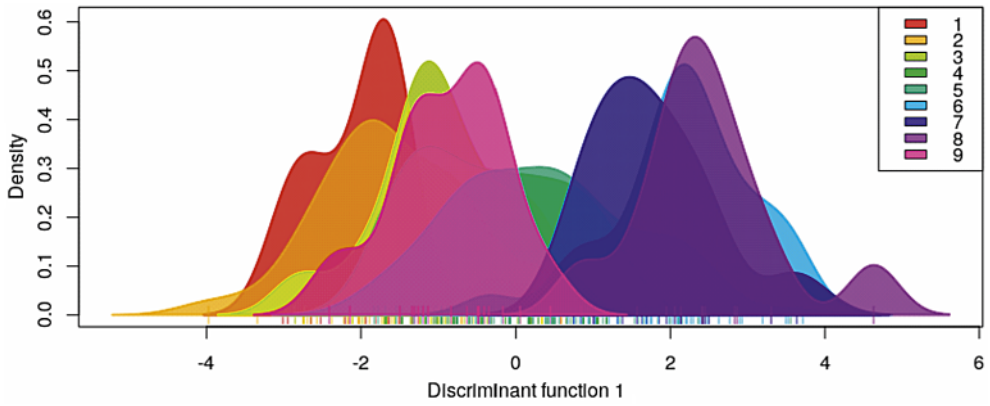

Figure 2. Genetic clusters determined using discriminant analysis based on the first discriminant function where 1 – Generalle, 2 - Generale-Generalissimus, 3 - Favory, 4 - Favory-Generalissimus, 5 – Sacramoso, 6 – Solo, 7 - Siglavi Pakra, 8 – Romke, 9 - Rudolfo

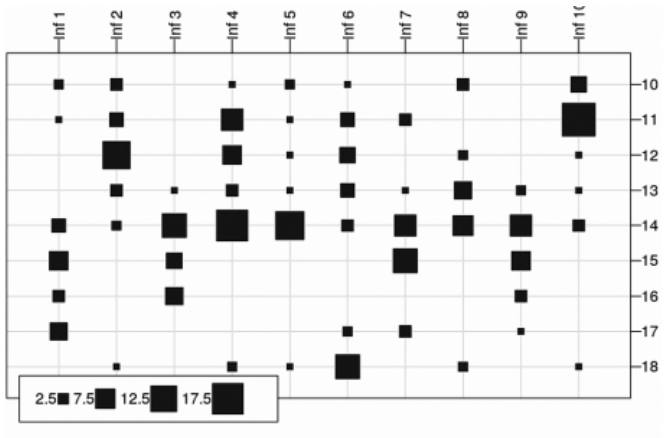

Figure 3. The BIC statistic results referring to differentiation between inferred and original clusters, where 10 – Generalle, 11 - Generale-Generalissimus, 12 - Favory, 13 - Favory-Generalissimus, 14 – Sacramoso, 15 – Solo, 16 - Siglavi Pakra, 17 – Romke, 18 - Rudolfo

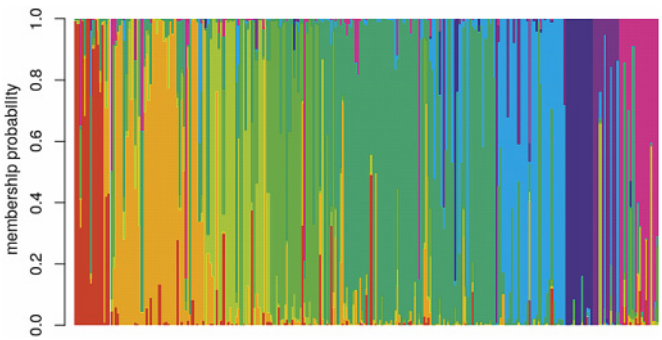

Figure 4. Membership probability resulting from DAC analysis (K=10)

principal components (PCs) adjusted for the successful repeated assignment of individuals, 26 PCA axes were left in DAPC. These axes correspond to more than 93.62% of variability. The two discriminant functions obtained correspond to 47.92% of variance. The first discriminant function separates the individuals of black and grey variants (Figure 2). For the discriminant function, however, the differentiation was not significant (Figure 3). As expected, a certain level of admixture was revealed in all studied sire lines (Figure 4). The occurrence of individuals with high level of probability of admixture of other sire lines was determined in all sire lines. The results indicate that the DAPC is not sensitive enough to determine genetic differences between sire lines of the Old Kladruber horse.

The results of DACP show that the structure of the studied sire lines of Old Kladruber horse was not sufficiently differentiated and that there was a high level of admixture between the sire lines. An admixture between the grey and black variants was proven, but to a limited extent. This admixture can be caused by the fact that individuals of the Sacramoso sire line are present in both colour varieties. The chosen methods were not able to correctly separate sire lines. These results again reflect the mating strategy in analysed breed. In such situation, this study might provide the first insight into further conservation strategies that should be considered.

Conclusion

In conclusion, this study presents a valuable insight into the genetic structure and diversity of sire lines of Old Kladruber horse. Our data suggest a low level of differentiation as well as a high gene flow between them. Hence, a revision of the breeding strategy is suggested. The modified genetic conservation program should include both classical breeding methods based on pedigree information and molecular data.

References

Bílek F. (1957): The Old Kladruber Horse (CzechLanguage). In:Speciálnízootechnika II. Chovkoní, 74–104. Státnízemědělské Nakl, Praha, p.1031.

Iwańczyk E., Juras R., Cholewiński G., Cothran E. G. (2006): Genetic structure and phylogenetic relationships of the Polish Heavy Horse. Journal of Applied Genetics, 47: 353–359.

Jombart T., Ahmed I. (2011): adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics, 27, 3070–3071.

Jombart T., Collins C., 2015. A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0. London: MRC Centre for Outbreak Analysis and Modelling, Imperial College London. http://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf (accessed 20.11.16)

Nagy S, Poczai P, Cernák I, Gorji AM, Hegedűs G and Taller J (2012): PICcalc: an online program to calculate polymorphic information content for molecular genetic studies. Biochemical Genetics Volume 50, Issues 9-10, pp 670-672. Paradis, E., 2010. pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics 26, 419–420.

Petersen J. L., Mickelson J. R., Cothran E. G., Andersson L. S., Axelsson J., Bailey E., Bannasch D., Binns M. M., Borges A. S., Brama P., da Câmara Machado A., Distl O., Felicetti M., Fox-Clipsham L., Graves K. T., Guérin G., Haase B., Hasegawa T., Hemmann K., Hill E.W., Leeb T., Lindgren G., Lohi H., Lopes M.S., McGivney B. A., Mikko S., Orr N., Penedo M.C.T., Piercy R.J., Raekallio M., Rieder S., Røed K.H., Silvestrelli M., Swinburne J., Tozaki T., Vaudin M., Wade C., McCue M.E. (2013): Genetic Diversity in the Modern Horse Illustrated from Genome-Wide SNP Data. PLoS ONE 8, e54997.

Szwaczkowski T., Greguła-Kania M., Stachurska A., Borowska A., Jaworski Z., Gruszecki T. M. (2016): Inter- and intra-genetic diversity in the Polish Konik horse: implications for the conservation program. Canadian Journal of Animal Science, 96, 570–580.

Toro, M. A., Fernandez, J., Caballero, A. (2009): Molecular characterization of breeds and its use inconservation. Livestock Science. 120. 174-195.

Vostrá-Vydrová H., Vostrý L., Hofmanová B., Krupa E., Zavadilová L. (2016): Pedigree analysis of the endangered Old Kladruber horse population. Livestock Science, 185, 17-23.

Weir B.S., Cockerham C. C. (1984): Estimating F-Statistics for the Analysis of Population Structure. Evolution 38, 1358.

acs82_57