

Genetic diversity of Czech Sport Pony based of genealogical information

Lubos Vostry^{1*}, Hana Vostra-Vydrova², Barbora Hofmanova¹, Zdenka Vesela²

¹Czech University of Life Sciences Prague, Prague, Czech Republic

²Institute of Animal Science, Praha, Czech Republic

Article Details: Received: 2016-11-07 | Accepted: 2016-11-18 | Available online: 2017-12-31

<http://dx.doi.org/10.15414/afz.2017.20.04.84–89>

Population study of Czech Sport Pony breed was carried out based on pedigree information of animals registered in the Studbook. Pedigree records collected from the year 1972 to 2016 comprised information on 12,548 animals used in the analyses. The pedigree depth of the analysed individuals was relatively low (3.7 generations). The mean value of inbreeding coefficient was 0.3% (with maximum value 26%). The proportion of non-inbred animals was high (80%). The average rate of inbreeding in the reference population was lower than 1%, and the estimates of effective population sizes were relatively high (789). The presented paper is indicating that genetic diversity in the Czech Sport Pony breeds is still relatively high. However the available genetic variability in the Czech Sport Pony breed as an open population with continuous migration and gene flow was lower than was expected. Active management of the future rate of inbreeding is necessary for this breed.

Keywords: inbreeding, rate of inbreeding, effective populations, open population

1 Introduction

The Czech Sport Pony is a relatively new pony breed in the Czech Republic. The Association of Horse Breeders' Unions of the Czech Republic began breeding for the sport pony in the 2000s. Its origin is from a diverse group of breeds, because the Sport pony is not derived from specific bloodlines, but rather is a conformation type, akin to the Warmblood horse breed.

An important aim of breeding programs in small population is to maintain the highest possible genetic diversity, the level of biodiversity, and the minimum inbreeding within the considered population (Lacy and Ballou, 1998). This also applies, to some extent, to breeding programs, as the increase of inbreeding usually implies undesirable effects of inbreeding depression on fitness, productive or morphological traits and the loss of genetic diversity could compromise a future response on the trait of interest or on new breeding objectives.

Genetic diversity refers to the total number of genetic characteristics in the genetic makeup of species. Genetic diversity in animal populations can be divided into within breed and between breeds genetic diversity. This study will focus on within breed genetic diversity because it directly affects the genetic improvement of Czech Sport Pony populations. Evaluating the inbreeding and loss of genetic diversity within a breed is a requisite for sustainable improvement in the long term.

Fortunately, many breeds have pedigree information, moreover, pedigree information has already been used in monitoring and evaluating genetic diversity of breeds based on the effective population size and the probabilities of gene origin (e.g., Sørensen et al., 2005). Moreover pedigree information is considered a useful measure to identify important circumstances that affect the genetic history of a population (Valera et al., 2005).

The objective of the present study was to describe the level of genetic variability within Czech Sport Pony and to estimate the inbreeding trend in this breed based on genealogical information.

2 Material and methods

Data from pedigrees of the registered horses in the studbook of the Czech Sport Pony contained information from the year 1972 to 2016. Data were provided from the Association of Horse Breeders' Unions of the Czech Republic ($n = 12548$). The pedigree analysis was performed on one reference population. The reference population was defined as the whole active populations – individuals (stallions and mares) born in the years 1996–2013 ($n = 501$).

The pedigree completeness and the number of ancestral generations influence the estimated inbreeding coefficients and the relationship coefficients between animals. The pedigree completeness level was assessed

*Corresponding Author: Luboš Vostrý, Czech University of Life Science Prague, Kamýčká 129, 165 00 Prague6-Suchdol, Czech Republic, e-mail: vostry@af.czu.cz

as the proportion of ancestors known per generation for each offspring (MacCluer et al., 1983). The number of equivalent complete generations (t) in the pedigree was computed using the formula $\sum(1/2)^n$ (Maignel et al., 1996), where n is the number of generations between an animal and its ancestor.

Several measures of genetic variability were calculated from the pedigree data. The total number of founders (f) for each breed was determined as the total number of ancestors with unknown parents. The effective number of founders (f_e) is defined as the number of founders explaining the same level of genetic diversity as that observed in the reference population (Lacy, 1989)1989. This parameter is the reciprocal of the probability that two randomly drawn genes in the population under study originated from the same founder (James, 1972). The effective number of founders was estimated using the formula:

$$f_e = \frac{1}{\sum_{i=1}^f q_i^2}$$

where:

q_i – the genetic contribution of the i -th founder to the reference population

The founder genome equivalent (f_{ge}) is defined as the effective number of founders with a non-random loss of founder alleles resulting in identical genetic variability to that defined in the reference population (Lacy, 1989). The founder genome equivalent accounts for both unequal contributions of founders and the random loss of alleles caused by genetic drift (Lacy, 1989)1989. The founder genome equivalent was calculated using the method of Caballero and Toro (2002):

$$f_{ge} = \frac{1}{2f_g}$$

where:

f_g – the average coancestry coefficient for the reference population. According to Caballero and Toro (2000), the average coancestry coefficient was estimated using the formula:

$$f_g = \frac{\sum_{i=1}^n \sum_{j=1}^n a_{ij}}{2n^2}$$

where:

a_{ij} – represents the elements of the relatedness matrix and n is the number of individuals in the pedigree. The effective number of non-founders (N_{enf}) accounts for the effects of genetic drift in non-founder generations. This parameter

was computed using the following expression according to Caballero and Toro (2002):

$$\frac{1}{f_{ge}} = \frac{1}{f_e} + \frac{1}{N_{enf}}$$

Coefficients of inbreeding and relatedness coefficient (AR)

The inbreeding coefficient of each individual was estimated using a tabular method (Falconer and Mackay, 2009), based on VanRaden's method (1992):

$$F_i = a_{ii} - 1$$

where:

F_i – the inbreeding coefficient of individual i

a_{ii} – the additive genetic relationship between individual i and itself

The average relatedness coefficient of each individual (AR) was computed as the average coefficient integrating the row from the individual in the numerator relationship matrix **A**. This coefficient indicates the probability a randomly selected allele in the population occurs in a selected individual or amongst a group of individuals (Goyache et al., 2003).

The realized effective population size (N_e) represents the number of unrelated individuals which would cause the same increase of inbreeding as that of the reference population.

Two different approaches for effective population size calculation were compared.

1. The effective populations size calculated from individual increase in inbreeding (N_{ef}) according to Gutiérrez et al. (2008):

$$N_{ef} = \frac{1}{2\Delta F}$$

where:

$$\Delta F_i = 1 - t_i^{-1} \sqrt{1 - F_i}$$

F_i – the inbreeding coefficient of individual i

t_i – their equivalent complete generations (Maignel et al., 1996)

ΔF – increase in inbreeding

2. The effective population size calculated from individual increase in coancestry (N_{ec}) according to Cervantes et al. (2011):

$$N_{ec} = \frac{1}{2\Delta C_{ij}}$$

where:

$$\Delta C_{ij} = 1 - \frac{t_i + t_j}{2} \sqrt{1 - C_{ij}}$$

t_i and t_j – equivalent complete generations (Maignel et al., 1996) of animals i and j

ΔC_{ij} – increase in coancestry between any pair of individuals i and j

C_{ij} – the inbreeding of a descendent from both animals i and j

The loss of genetic diversity (gene identity) in the reference population due to genetic drift or unequal founder contribution was derived from f_e and f_{ge} . The total loss of genetic diversity caused by genetic drift or bottleneck was computed using the function (Lacy, 1995):

$$1 - GD$$

where:

$$GD = 1 - \frac{1}{2f_{ge}}$$

The loss of genetic diversity (gene identity) resulting from an unequal founder contribution was estimated using the function (Caballero and Toro, 2000):

$$1 - GD^*$$

where:

$$GD^* = 1 - \frac{1}{2f_{fe}}$$

in which the difference between GD and GD^* is the loss of genetic diversity caused by the genetic drift across generations of non-founders and corresponds to the inverse value of $2N_{enf}$:

$$GD - GD^* = \frac{1}{2N_{enf}}$$

3 Results and discussion

The number of animals that were registered in the Studbook between 1980 and 2016 are show in Figure 1. Between 1990 and 1998, there was increase to maximum level of registered animals (80 animals per year). From 1999 to 2003 the number of registered individuals has decreased. In last thirteen years, the number of registered individual was relatively constant between 30 and 45 individuals.

The average equivalent of the know generations in the analysed breed was 3.72 and ranged from 0 to 9.18. Pedigree completeness levels were also consistent with the preceding values. The completeness level declined to

less than 40% after 3 generations and less than 25% after 10 generations. The equivalent of completely known generations of breeds in the present study is significantly lower than in the other horse populations: Slovak Sport Pony – 4.97 (Kadlecik and Kasarda, 2014), Lipizzaner – 10.25 and 15.2 (Zechner et al., 2002; Pjontek et al., 2012), Austrian Noriker – 12.3 (Druml et al., 2009), Andalusian – 8.3 (Valera et al., 2005), Lusitano – 11.2 (Vicente et al., 2012), Spanish Arabian horse - 7.9 (Cervantes et al., 2008) and Old Kladruber – 15 (Vostrá-Vydrová et al., 2016). Similar values were determined in German Paint horse – 4.77 (Siderits et al., 2013).

Parameters describing the genetic variability of a population derived from the probability of gene origin are presented in Table 1. For the reference population the effective number of founders was only 35% of the total number of founders. The effective number of ancestors reached 18% of the effective number of founders. The ratios of the effective number of ancestors to effective number of founders (f_d/f_e) and between founders genome equivalent to the effective number of founders (f_{ge}/f_e) illustrate that studied population is affected by bottleneck and genetic drift. The value of each of these ratios was lower than 1. The f_e statistic is an important parameter used to assess if the contribution of the founders is balanced (Lacy, 1989). If all founders contributed equally, then the effective number of founders would be equal to the total number of founders. However, f_e is typically lower than f because f_e reflects unequal contributions of founders to the current population as a result of selection and variations in family size. The f_e and f values less than one indicate a loss of genetic diversity resulting from unequal contributions of the founders. The advantage of some individuals is indicated as f_d/f_e and f_{ge}/f_e ratios describing the extent of bottleneck and genetic drift, respectively (Boichard et al., 1997). The ratios calculated in the present analyses had identical or lower values compared with other reports, including the Austrian Noriker (Druml et al., 2009) – 0.25 and 0.09, Lusitano (Vicente et al., 2012) – 0.47 and 0.21, and Spanish Arabian horses (Cervantes et al., 2008) – 0.46–0.72 and 0.23–0.42. The cumulative marginal contributions of the most important ancestors are reported in Table 2. Three ancestors could explain approximately 10% of the diversity in the reference population, and teen ancestors could explain approximately 20% of the diversity in the reference population. Fifty percent of total genetic diversity could explain 128 ancestors.

The average F_i by year for the whole analysed population are illustrated in Figure 1. Value F_i close to zero were estimated in the birth years until 2002. A higher increase has been recorded since 1998. The 20% of individuals from reference population were inbred ($F_i > 0$), and 5% of

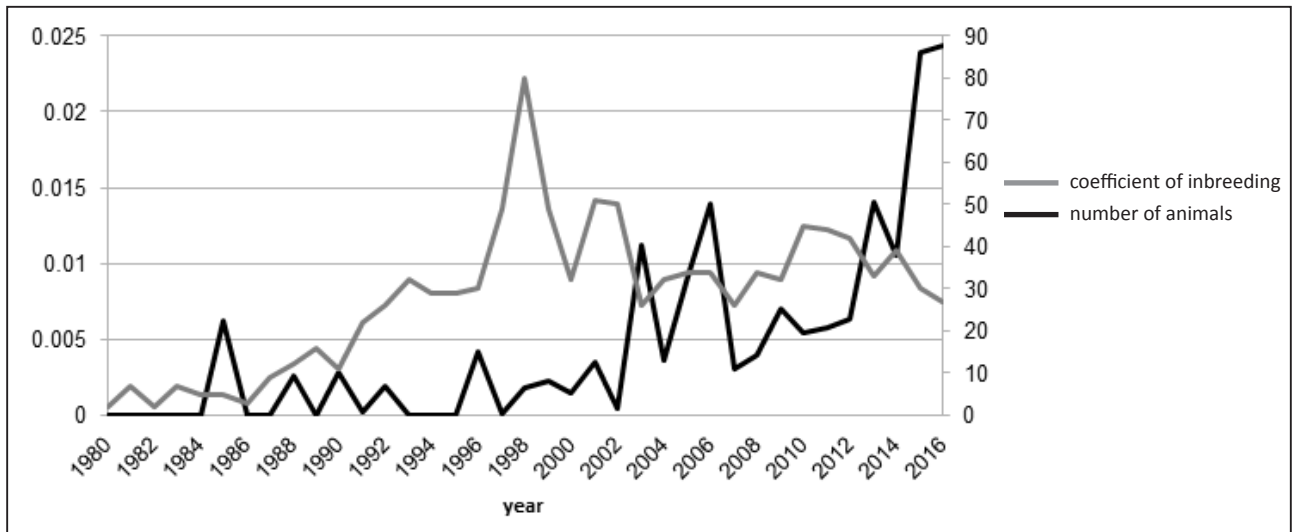


Figure 1 Mean inbreeding coefficient and number of animals registered in studbook across years of birth between 1980 and 2016 of the Czech Sport Pony

individuals from reference population had F_i higher than 1% and 2% of individual from reference population had F_i higher than 6.25%. Maximum level of F_i in the reference population was 26%. The average inbreeding coefficients in the reference population were: $F_i = 0.32\%$ (s.d. = 1.71%). The average relatedness coefficient (AR) was 1.2% in the reference population (Table 1). The AR was four times higher than the F_i , which points to preferable mating in the reference population. The average relatedness coefficient was found higher than the double of the average inbreeding coefficient, from which an increase in inbreeding coefficient in the next generations may be derived. The average inbreeding value implies a loss of genetic variability that may negatively influence fitness characteristics and increase occurrence of phenotypic defects. The influence of inbreeding depression on performance traits was neither confirmed by Curik et al. (2003) nor by Wolc and Balińska (2010). However, inbreeding depression was observed in morphological traits (Gómez et al., 2009). Rate of inbreeding (ΔF) is one of the main parameters of genetic diversity monitoring. Based on the positive values of ΔF in the reference population, an increase in F_i values can be expected in further generations of the Czech Sport Pony. The Food and Agriculture Organization of UN (FAO, 1998) stated that the average value of ΔF should not exceed 1%. The estimated average value of $\Delta F = 0.001$ is substantially lower than this recommended maximal value. The effective population sizes (Table 1) reached the value $N_{ef} = 789.30$ and $N_{ec} = 139.06$ individuals in the reference population. The values of both N_e , derived from ΔF and ΔC_{if} , were higher than the recommended minimum of N_e (50) for the conservation of genetic diversity (FAO, 1998). A difference between N_{ec} and N_{ef} provides information about non-random mating of individuals in the analysed

populations and a decrease in level of genetic diversity in future generations can be expected in relation to the parental combination. The N_{ef} and N_{ec} are asymptotically equivalent in an idealized population, the disagreement between these parameters mainly falls on their ability to assess the effect of preferential mating. In other words, the comparison between N_{ec} and N_{ef} would characterize

Table 1 Parameters characterizing the probability of gene origin, effective population size and parameters of inbreeding for Czech Sport Pony

Animals total	12,548
Total number of founders (f)	2,605
Effective number of founders (f_e)	898.18
Effective number of ancestors (f_a)	161.81
Founder genome equivalent (f_{ge})	83.53
Ratio of effective number of founders to effective number of ancestors (f_e/f_a)	0.180
Ratio of founder genome equivalent to effective number of founders (f_{ge}/f_e)	0.093
Realized effective population size (N_{ef})	789.30
Realized effective population size (N_{ec})	139.06
Average inbreeding coefficient (F_i)	0.003
Average relatedness coefficient (AR)	0.012
ΔF per generation	0.001
Loss of genetic diversity	–
Unequal founder contribution and random genetic drift	0.01
Unequal founder contribution	0.00
Random genetic drift	0.01

Table 2 Marginal genetic contributions of ten most influential ancestors of the Czech Sport Pony

I.D.	Sex	Year of birth	Breed	Con. (%)
11180	M	1993	137	3.49
10820	M	2000	112	3.28
8013	M	1980	211	3.22
7226	M	1992	193	2.14
10412	M	2000	118	1.92
10185	F	1991	151	1.60
9489	M	1972	118	1.18
9119	M	1976	118	1.16
8391	F	1995	211	1.09
9169	M	1983	151	1.02

I.D. – identification number, con. (%) – contribution, M – male, F – female
 Breed: 112 – Welsh Mountain Pony, 118 – Welsh Pony, 137 – Slovak Sport Pony, 151 – welsh part-bred, 193 – New forest pony, 211 – Czech Sport Pony

the influence of preferential mating in the population (Cervantes et al., 2011). It is well known that many factors lead real populations to increases in inbreeding higher than those expected in the idealized population as a consequence of preferential mating (Falconer and Mackay, 2009). However, increase in coancestry basically reflects the drift caused by the finite size of the population (Caballero, 1994). Thus, discrepancies between increase in inbreeding and in coancestry can be interpreted as cryptic population subdivision (Cervantes et al., 2011). Minimum coancestry mating systems lead to lower increase in inbreeding than in coancestry that would be expected under random mating. The estimated values of N_{ef} were higher than N_e estimated in the other horse breeds including Lipizzaner ($N_{ef} = 102$, Zechner et al., 2002), Austrian Noriker ($N_{ef} = 157$, Druml et al., 2009) and higher than in Lusitano breed ($N_{ef} = 28$, Vicente et al., 2012).

The total amount of genetic diversity (GD) loss in the analysed breeds during the last 44 years is affected by various reasons. The average relative GD loss in the reference populations was 1%. The loss of GD due to genetic drift accumulated over non-founder generations ($GD^* - GD$) was more important than the loss resulting from the unequal contribution of founders. In general, genetic diversity was reduced in reference populations of the Czech Sport Pony breed. The genetic diversity loss results from random genetic drift. Fernandez et al. (2005) stated that the loss of genetic diversity within breeds is influenced, to a large extent, by genetic drift, resulting in increased homozygosity and allele fixation in the population. Álvarez et al. (2008) reported that losses of genetic variability occurred soon after setting up the

breeding programme, as a significant number of founders did not produce progeny in the next generation.

4 Conclusions

A complex pedigree analysis of the Czech Sport Pony breed has been performed. The results of the analysis of pedigree show a low pedigree completeness level. The low values of inbreeding coefficient were estimated. However, increased inbreeding level in the analysed time period was observed. This study indicates that genetic diversity in the Czech Sport Pony breed is still relatively high. However the available genetic variability in the Czech Sport Pony breed as an open population with continuous migration and gene flow was lower than was expected. Active management of the future rate of inbreeding is necessary for this breed.

Acknowledgments

This work was funded by a grant from the Ministry of Agriculture of the Czech Republic (Project no. QJ1510141).

References

- ÁLVAREZ, I. et al. (2008) Relationship between genealogical and microsatellite information characterizing losses of genetic variability: Empirical evidence from the rare Xalda sheep breed. In *Livest. Sci.*, vol. 115, pp. 80–88. doi:<http://dx.doi.org/10.1016/j.livsci.2007.06.009>
- BOICHARD, D. et al., (1997) The value of using probabilities of gene origin to measure genetic variability in a population. In *Genet. Sel. Evol.*, vol. 29, pp. 5–23. doi:<http://dx.doi.org/10.1051/gse:19970101>
- CABALLERO, A. (1994) Developments in the prediction of effective population size. In *Heredity*, vol. 73, pp. 657–679. doi:<http://dx.doi.org/10.1038/hdy.1994.174>
- CABALLERO, A. and TORO, M. A. (2000) Interrelations between effective population size and other pedigree tools for the management of conserved populations. In *Genet. Res.*, vol. 75, pp. 331–343.
- CABALLERO, A. and TORO, M. A. (2002) Analysis of genetic diversity for the management of conserved subdivided populations. In *Conserv. Genet.*, vol. 3, pp. 289–299.
- CERVANTES, I. et al. (2011) Estimation of effective population size from the rate of coancestry in pedigreed populations: Effective population size from rate of coancestry. In *J. Anim. Breed. Genet.*, vol. 128, pp. 56–63. doi:<http://dx.doi.org/10.1111/j.1439-0388.2010.00881.x>
- CERVANTES, I. et al. (2008) Population history and genetic variability in the Spanish Arab Horse assessed via pedigree analysis. In *Livest. Sci.*, vol. 113, pp. 24–33. doi:<http://dx.doi.org/10.1016/j.livsci.2007.02.011>
- CURIK, I. (2003) Inbreeding, Microsatellite Heterozygosity, and Morphological Traits in Lipizzan Horses. In *J. Hered.*, vol. 94, pp. 125–132. doi:<http://dx.doi.org/10.1093/jhered/esg029>
- DRUML, T., BAUMUNG, R. and SÖLKNER J. (2009) Pedigree analysis in the Austrian Noriker draught horse: genetic diversity and the impact of breeding for coat colour on population structure. In *J. Anim. Breed. Genet.*, no. 126, pp. 348–356. doi:<http://dx.doi.org/10.1111/j.1439-0388.2008.00790.x>

- FALCONER, D. S. and MACKAY, T. F. C. (2009) *Introduction to quantitative genetics*. 4. ed., Pearson, Prentice Hall, Harlow.
- FERNANDEZ, J. et al. (2005) Efficiency of the Use of Pedigree and Molecular Marker Information in Conservation Programs. In *Genetics*, vol. 170, pp.1313–1321. doi:<http://dx.doi.org/10.1534/genetics.104.037325>
- GÓMEZ, M. D. et al. (2009) Assessment of inbreeding depression for body measurements in Spanish Purebred (Andalusian) horses. In *Livest. Sci.*, vol. 122, pp. 149–155. doi:<http://dx.doi.org/10.1016/j.livsci.2008.08.007>
- GOYACHE, F. et al. (2003) Using pedigree information to monitor genetic variability of endangered populations: the Xalda sheep breed of Asturias as an example. In *J. Anim. Breed. Genet.*, vol. 120, pp. 95–105. doi:<http://dx.doi.org/10.1046/j.1439-0388.2003.00378.x>
- GUTIÉRREZ, J. P. (2008) Individual increase in inbreeding allows estimating effective sizes from pedigrees. In *Genet. Sel. Evol.*, vol. 40, pp. 359–378. doi:<http://dx.doi.org/10.1051/gse:2008008>
- JAMES, J. W. (1972) Computation of genetic contributions from pedigrees. In *Theor. Appl. Genet.*, vol. 42, pp. 272–273. doi:<http://dx.doi.org/10.1007/BF00277555>
- KADLEČÍK, O. and KASARDA, R. (2014) Genetic diversity of Slovak Sport Pony based on genealogical information. In Book of abstracts of the 65th annual meeting of the European federation of animal science. Wageningen: Wageningen Academic Publishers, 2014, p. 378.
- LACY, R. C. (1989) Analysis of founder representation in pedigrees: Founder equivalents and founder genome equivalents. In *Zoo Biol.*, vol. 8, pp. 111–123. doi:<http://dx.doi.org/10.1002/zoo.1430080203>
- LACY, R. C. (1995) Clarification of genetic terms and their use in the management of captive populations. In *Zoo Biol.*, vol. 14, pp. 565–577. doi:<http://dx.doi.org/10.1002/zoo.1430140609>
- LACY, R. C. and BALLOU J. D. (1998) Effectiveness of Selection in Reducing the Genetic Load in Populations of *Peromyscus polionotus* During Generations of Inbreeding. In *Evolution*, vol. 52, pp. 900. doi:<http://dx.doi.org/10.2307/2411285>
- MacCLUER, J. et al. (1983) Inbreeding and Pedigree Structure in Standardbred horses. In *Heredity*, vol. 74, pp. 394–399.
- MAIGNEL, L., BOICHARD D. and VERRIER E. (1996) Genetic variability of French dairy breeds estimated from pedigree information. In *Interbull Bull.*, vol. 14, pp. 49–54.
- PJONTEK, J. et al. (2012) Pedigree analysis in four Slovak endangered horse breeds. In *Czech J Anim Sci.*, vol. 57, pp. 54–64.
- SIDERITS, M., BAUMUNG, R. and FUERST-WALTTL, B. (2013) Pedigree analysis in the German Paint Horse: Genetic variability and the influence of pedigree quality. In *Livest. Sci.*, no. 151, pp. 152–157. doi:<http://dx.doi.org/10.1016/j.livsci.2012.10.018>
- SØRENSEN, A. C., SØRENSEN, M. K. and BERG, P. (2005) Inbreeding in Danish dairy cattle breeds. In *J. Dairy Sci.*, vol. 88, pp.1865–1872.
- VALERA, M. et al. (2005) Pedigree analysis in the Andalusian horse: population structure, genetic variability and influence of the Carthusian strain. In *Livest. Prod. Sci.*, vol. 95, pp. 57–66. doi:<http://dx.doi.org/10.1016/j.livprodsci.2004.12.004>
- VANRADEN, P. M. (1992) Accounting for Inbreeding and Crossbreeding in Genetic Evaluation of Large Populations. In *J. Dairy Sci.*, vol. 75, pp. 3136–3144. doi:[http://dx.doi.org/10.3168/jds.S0022-0302\(92\)78077-1](http://dx.doi.org/10.3168/jds.S0022-0302(92)78077-1)
- VICENTE, A. A., CAROLINO, N. and GAMA, L.T. (2012) Genetic diversity in the Lusitano horse breed assessed by pedigree analysis. In *Livest. Sci.*, vol. 148, pp. 16–25. doi:<http://dx.doi.org/10.1016/j.livsci.2012.05.002>
- VOSTRÁ-VYDROVÁ, H. et al. (2016) Pedigree analysis of the endangered Old Kladruber horse population. In *Livest. Sci.*, vol. 185, pp. 17–23. doi:<http://dx.doi.org/10.1016/j.livsci.2016.01.001>
- WOLC, A. and BALIŃSKA, K. (2010) Inbreeding effects on exterior traits in Polish konik horses. In *Arch Tierz.*, vol. 53, pp. 1–8.
- ZECHNER, P. et al. (2002) Analysis of diversity and population structure in the Lipizzan horse breed based on pedigree information. In *Livest. Prod. Sci.*, vol. 77, pp.137–146. doi:[http://dx.doi.org/10.1016/S0301-6226\(02\)00079-9](http://dx.doi.org/10.1016/S0301-6226(02)00079-9)