

The First In-deep Pedigree Analysis of Repatriated Gyimes Racka Sheep for a Sustainable Preservation of its Genetic Resource

Edina Kárpáti^{a,*}, László Gulyás^a, János Posta^b, László Sáfár^c, András Gáspárdy^d

^aSzéchenyi István University, Albert Kázmér Faculty of Mosonmagyaróvár, Department of Animal Sciences, Vár tér 2, 9200. Mosonmagyaróvár, Hungary

^bUniversity of Debrecen, Faculty of Agricultural and Food Sciences and Environmental Management, Institute of Animal Science, Biotechnology and Nature Conservation, Böszörményi út 138, 4032 Debrecen, Hungary

^cHungarian Sheep and Goat Breeders' Association, Lőportár utca 16, 1134 Budapest, Hungary

^dUniversity of Veterinary Medicine Budapest, Institute for Animal Breeding, Nutrition and Laboratory Animal Science, István utca 2, 1078 Budapest, Hungary
edina.karpati@gmail.com

The preservation of endangered domestic animal breeds is a specific area of sustainable livestock farming. In the southern part of the Carpathians, a variant of Zaupel-type (Racka) sheep, which was small in size and resistant to the cold mountain weather, developed at the beginning of the 19th century. In order to preserve this historical variant, it was introduced in Hungary in the early 1990s. The aim of this work is to study the pedigree of the repatriated population of the Gyimes Racka to understand its population structure and to support its future breeding work. The pedigree data (2005-2020) show that the number of founding individuals (N_f) was 3,838, within which the number of maternal lineages was 2,255. The effective number of founders (f_e) and the effective number of significant ancestors (f_a) in the total pedigree population were higher than in the reference population ($n=2,591$; 67 and 56 vs. 20 and 14, respectively). The pedigree traced back to a maximum of 2.51 generations with an average of 1.34 complete generations and 1.78 equivalent complete generations on average (in the reference population: 6.05, 2.73 and 3.82, respectively). Of the 16,947 animals registered in the herd book (with an average normalised COI of 1.43 % obtained by log transformation), 3,828 were inbred (6.30 %). As regards the maternal generations, it can be observed that the inbred stage increases steadily and significantly from the beginning to the present eighth generation (from 0.00 % to 9.54 %). The average generation interval between breeding animals was 3.29 years (obtained by log transformation). The sire-progeny paths had shorter (2.92-3.02) and the dam-progeny paths had longer (3.62-3.70) intervals. The Gyimes Racka population is already showing signs of genetic narrowing during its short conservation period, which calls for attention to stop undesirable trends.

1. Introduction

The preservation of endangered domestic animal breeds is a specific area of sustainable livestock farming. In order to set and implement future plans, it is essential to know the current structure and performance of the population bred, which, of course, is based on a well-managed herd-booking (Černá et al., 2021). Modern processing of its pedigree part, even with direct genetic data (Cortellari et al., 2022), will lead to the computation of many new population genetic parameters. The pedigree analysis, the determination of the degree of relatedness between individuals, is necessary not only to increase the producing ability of intensive breeds but also to preserve the genetic diversity of rare, heritage ones (Špehar et al., 2022).

Groups of Zaupel-type (Racka) sheep include the indigenous sheep breeds of Central and South-Eastern Europe. In the southern part of the Carpathians, its variant, which was small-sized with a fleece of mixed wool and resistant to the cold mountain weather, developed at the beginning of the 19th century (Dunka, 2001). Characteristics of an endangered breed can regain a useful economic role. According to the investigation of

Barrientos Contreras et al. (2022), mixed wool, such as Gyimes Racka, can be used as a biosorbent for the removal of hydrocarbons. In many cases, the feeding of endangered ruminant breeds is based on hay. A study of Benites-Alfaro et al. (2023) showed that the methane emission of an individual kept on hay is lower than that of an individual fed with ensilage fodder. It can be deduced from this that the traditional feeding of old breeds, such as the Gyimesi Racka, does not adversely contribute to the increase of harmful greenhouse gases.

To preserve this historical Hungarian breed, among others, the breed was introduced in Hungary in the early 1990s. Because of the later appearance of the Gyimes Racka, the goals for its preservation were therefore not as mature as for other native Hungarian breeds (e.g., in Tsigai: Annus et al., 2015; in Cikta: Posta et al., 2019). As previously articulated by Földi et al. (2017), the period was too short to draw far-reaching conclusions for breeding from the reliable data collected so far. With this processing, the authors want to contribute to the missing breeding decisions by evaluating the breed's flock book data until 2020. The aim of this work is to analyse the pedigree of the repatriated population of the Gyimes Racka to understand its population structure and to support future effective diversity conservation.

2. Materials and methods

For the processing, the Excell database of the Hungarian Sheep and Goat Breeders Association (MJKSZ) is used, which covers the entire country's population (2005-2020, $n=16,947$). The reference population was formed from the individuals born 2017-2020 ($n_{Ref}=2,591$) because this time span corresponds to the generation interval, which is about 4 years in sheep. In this in-depth evaluation, 16 parameters describing the population structure were defined.

After data preparation (coding of maternal ancestry, creation of pedigree file) completeness of pedigree was assessed by the number of maximum generations, complete generations and complete generations equivalent, as well as the percentage of the known ancestral rows by year and by generation. To characterise the probability of gene origin, the total number of founders (N_f), the effective number of founders (f_e), the effective number of significant ancestors (f_a) and their proportion (f_a/f_e), as well as the number of ancestors explaining the genetic variability, were determined. These evaluations were carried out with the software Endog (Gutiérrez and Goyache, 2005) and Poprep (Groeneveld et al., 2009). Wright's coefficient of inbreeding (COI) was calculated using Pedigree Viewer software (Kinghorn and Kinghorn, 2010). The generation interval (GI) is the average age of the parents at the birth of the offspring (interpreted broadly, Lush, 1945). This, taking into account the gene transfer, is more precisely the average age of the parents at the birth of the offspring, which are breeding animals involved in the formation of the next generation (interpreted narrowly, McManus et al., 2019). The generation interval (broadly and narrowly interpreted) was calculated by dividing the birth dates of the registered individuals (parents and their offspring) into different paths (sire and dam to breeding and non-breeding offspring or sire-son, SS; sire-daughter, SD; dam-son, DS and dam-daughter, DD).

At the beginning of processing, using the Kolmogorov-Smirnov one-sample test and Lilliefors test, it was found that the baseline COI and GI data do not follow a normal distribution. To normalise the distribution, a natural logarithm transformation was applied. Data preparation and statistical processing of normalised data was done by use of Statistica ver. 13. (TIBCO Software Inc., 2020). Here, for both COI_{back} and GI_{back} the geometric mean and the upper and lower 95 % confidence levels (-95 % CI and +95 % CI) are presented after the back-transforming. Tukey HSD post hoc test was used to detect significant differences among classes of categories.

3. Results

The average number of generations by three variants of pedigree completeness is given in Table 1. Higher values were obtained in the reference population. The pedigree traced back to a maximum of 2.51 generations with an average of 1.34 complete generations and 1.78 complete generations equivalent (these figures were higher in the reference population: 6.05, 2.73 and 3.82, respectively).

Table 1: Average number of generations by variants of the pedigree completeness

Pedigree completeness	Total population	Reference population
Maximum generations	2.51	6.05
Complete generations	1.34	2.73
Complete generations equivalent	1.78	3.82

The fluent and tendentious increase in the proportion of ancestral knowledge as a function of degree of descent and time is clearly visible (Figure 1). As of 2010, 100 % of the parents were known, and as of 2015, all 4 grandparents were known. The youngest individuals are now in the 6th and 7th generations (and sporadically in the eighth).

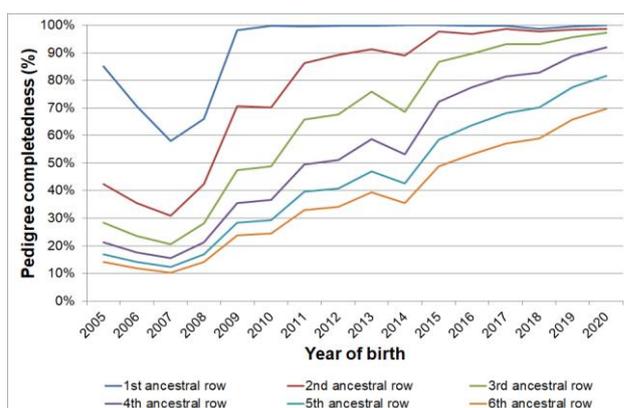


Figure 1: Percentage of the known ancestral rows by year and by generation

The pedigree data (2005-2020) show that the number of founding individuals (N_f) was 3,838, within which the number of females (maternal lineages) was 2,255 (Table 2). The effective number of founders (f_e) and the effective number of significant ancestors (f_a) in the total pedigree population were higher than in the reference population ($n=2,591$; 67 and 56 vs. 20 and 14, respectively). The f_a to f_e ratio was close to 1 from below in both the total (0.836) and the reference population (0.700). The average relatedness coefficient (AR) is 1.393, while the value of the reference population is much twice as high (3.297).

Table 2: Population parameters of the Gyimes Racka sheep breed

Parameters	Total population	Reference population
Population size	16,947	2,591
Total number of founders (N_f)	3,838	-
Number of female founders	2,255	-
Number of male founders	1,583	-
Effective number of founders (f_e)	67	20
Total number of significant ancestors (N_a)	2,208	396
Effective number of significant ancestors (f_a)	56	14
f_a/f_e ratio	0.836	0.700
Average relatedness coefficient (AR)	1.393	3.297

The number of ancestors accounting for the genetic variability to different degrees is presented in Table 3. There are 30 individuals who are responsible for half of the breed's diversity. In the reference population, this value is only 6. Furthermore, the total genetic diversity of the Gyimes Racka population comes from 2,208 individuals, compared to the 16,947 individuals recorded in the herd book.

Table 3: Number of ancestors accounting for the genetic variability

Degree of genetic variability	Total population	Reference population
50 %	30	6
60 %	57	9
70 %	113	15
80 %	318	27
90 %	775	64
100 %	2,208	396

There were 3,828 animals inbred. The geometric mean of the back-transformed COI_{back} in the total population is 1.43 % (Table 4). As regards the maternal generations, it can be observed that the inbred stage of the total population increases steadily and significantly from the beginning to the present eighth generation (from 0.00 % to 9.54 %). At the same time, within the inbred sub-population, the COI_{back} is relatively stagnant after an initial decrease, starting to increase from the fifth generation (from 5.69 % to 9.54 %, $P<0.05$). Here, the COI_{back} has risen to 6.30 %. There is no significant difference in COI_{back} between the female and male sub-populations ($P=0.301$).

Table 4: Results of back-transformed inbreeding coefficient (COI_{back} , %) according to the maternal generations

Maternal generations	Total population		Inbred sub-population	
	COI_{back}	weighted mean (n)	COI_{back} geometric mean,	95 % CI ¹ (n)
Founder	0.00	(2,255)	-	-
Second	0.05	(6,712)	21.24 ^d	17.38 – 25.95 (17)
Third	1.57	(3,934)	7.06 ^c	6.65 – 7.50 (874)
Fourth	3.48	(2,054)	5.69 ^a	5.42 – 5.99 (1,255)
Fifth	4.40	(1,188)	5.70 ^{a,b}	5.36 – 6.05 (917)
Sixth	6.71	(621)	7.11 ^c	6.68 – 7.57 (586)
Seventh	7.19	(172)	7.36 ^{a,b,c}	6.54 – 8.29 (168)
Eighth	9.54	(11)	9.54 ^{a,b,c,d}	6.93 – 13.13 (11)
Overall	1.43	(16,947)	6.30	6.13 – 6.49 (3,828)

¹ lower and upper 95 % bounds of confidence interval, n – number of individuals,

a, b, c, d - $P < 0.001$, Tukey HSD (honest significant difference) for unequal number of individuals

Table 5 shows that the average generation interval (GI_{back}) in the total population is 3.61 y, which is less than 4 years. At the same time, the generation interval for breeding animals is even lower, at 3.29 y. The sire–breeding progeny paths have significantly ($P < 0.001$) shorter (2.92-3.02 y), and the dam-progeny paths have longer (3.62-3.70 y) intervals. Here, there is no difference between male and female offspring by parentage. However, according to the parent and the utilisation of the offspring, the generation interval showed a significant difference ($P < 0.001$) for each pathway. In general, it is observed that both parent-offspring relationships are statistically proven ($P < 0.001$) shorter than parent-non-offspring relationships. This is shortest for sire–breeding offspring (3.01 y) and longest for dam – non-breeding offspring (4.13 y)

Table 5: Results of back-transformed generation interval (GI_{back} , y) according to the paths of broad and narrow interpretation

Category (paths)	Number of connections	GI_{back} geometric mean	95 % CI ¹
Parent – all offspring	15,308	3.61	3.58 – 3.64
Sire – breeding offspring	1,168	3.01 ^a	2.94 – 3.08
Sire – non-breeding offspring	7,410	3.33 ^b	3.30 – 3.37
Dam – breeding offspring	921	3.69 ^c	3.58 – 3.81
Dam – non-breeding offspring	5,809	4.13 ^d	4.08 – 4.19
Parent – breeding offspring	2,089	3.29	3.23 – 3.36
Sire – sire	99	2.92 ^a	2.72 – 3.15
Sire – dam	1,069	3.02 ^a	2.95 – 3.09
Dam – sire	80	3.62 ^b	3.28 – 4.00
Dam – dam	841	3.70 ^b	3.58 – 3.82

¹ lower and upper 95 % bounds of the confidence interval

a, b, c, d - $P < 0.001$, Tukey HSD (honest significant difference) for unequal number of individuals per categories of GI

4. Discussion

In our study, f_a/f_e of 0.836 suggests that the majority of founders played a role as significant ancestors, indicating a very mild genetic narrowing. Although, the reference population value was already 0.700. However, these values were higher and more favourable in another Hungarian heritage breed, Cikta (0.955 and 0.923, respectively) (Posta et al., 2019).

The ancestor number accounting for 50 % of the variability was 30 (i.e., 1.4 % of significant ancestors), indicating that certain animals were used more frequently, which represents a genetic bottleneck, which contributes to a reduction in genetic variability.

Comparing the untransformed and transformed data, Kárpáti et al. (2022) found that the average untransformed inbreeding coefficient in the same Gyimes Racka population was 1.99 %, while it was 8.81 % in the inbred sub-population. The geometric means obtained by log transformation in the current study have lower values (1.43 % and 6.30 %, respectively). The average inbreeding rate of the total pedigree population of the Cikta breed was only 1.00 %, while the average inbreeding rate of the reference population was 1.16 % (Posta et al., 2019). In the Brazilian breed, Santa Inês ($n=17,097$) the average inbreeding coefficient of the whole population was

2.33 %, and higher levels of inbreeding already had a detrimental effect on lamb weight gain (Pedrosa et al., 2010).

Direct genomic information based on several microsatellites (Kovács et al., 2019) and thousands of SNPs (Ferenčaković et al., 2013) can also be used today to estimate the so-called molecular homozygosity (autozygosity). Mitochondrial DNA (mtDNA) samples have also been compared to investigate the genetic distance between the Gyimes Racka and Turcana (Kusza et al., 2015). Using Geneseek Ovine SNP50 BeadChip, it was found that white and black Hortobágy Racka are well separated from each other based on ROH (Runs of Homozygosity) (Zsolnai et al., 2021). Mészárosová et al. (2022) also used ROH and quantitative trait loci (QTL) to conduct genetic studies in crossbred Valaska (genomic diversity and environmental adaptation), while Machová et al. (2023) compared the Valaska sheep with the Sumava sheep using similar parameters.

In the current processing, the narrowly defined average generation interval was 3.29 years. This is shorter for sire paths and almost 1 year longer for dam paths. In the non-normalized data of the Gyimes Racka, Schütz et al. (2023) obtained a longer average GI (3.66 years) with the same tendencies (SS, SD, DS and DD 3.14, 3.29, 4.00 and 4.14, respectively). Oravcová and Krupa (2011), in their study of the Slovak Valaska breed (strain dates 1986-2010), obtained values of 3.52 (sire-son) and 2.64 y (sire-daughter) on the paternal side. On the maternal side, the length of the generation interval was 4.70 (dam-son) and 5.36 y (dam-daughter). Our values are similar on the paternal side but higher on the maternal side. The literature provides different data on the GI of sheep, from 2.55 to 4.94 y (Goyache et al., 2003; Venkataramanan et al., 2013; Mokhtari et al., 2015; Rafter et al., 2022). If the aim of breeding is to preserve a population, a long generation interval is considered reasonable (Groeneveld, 2009). In addition, the similarity of maternal and paternal paths is of great advantage in maintaining a rare or endangered breed.

Joakimsen (1969) had previously given a graphical representation of the age of parents at the time of birth of their offspring and called attention to its skewed distribution. However, he has not normalised the data in his processing; to our knowledge, no other author has done so. For the description of the characteristics related to the length of the generation intervals, according to Arkenberg (2014), the median is more reliable than the mean. In the future, we recommend attempting to normalise the base data of the coefficient of inbreeding and generation interval prior to data processing. The geometric mean obtained by log transformation has lower values than the arithmetic mean. It is believed that the former values are more representative of the group mean since they were determined on a data basis that is closer to the normal distribution. Lower values for the inbreeding coefficient can also be seen as more favourable for maintaining genetic diversity. However, for the generation interval, lower values tend to be more unfavourable, as they can increase the amount of genetic change per unit of time (per year) that may occur.

5. Conclusions

The pedigree of the Gyimes Racka breed was first and efficiently processed. The evaluation of several generations has already established the genetic structure of the population and some trend changes in population parameters. The Gyimes Racka population is already showing signs of genetic narrowing and an increasing homozygosity status during its short conservation period, which calls for attention to stop undesirable trends. In view of this, it is necessary to draw the attention of breeders to the importance of using appropriate mating plans in the future to preserve the initial genetic diversity along with maternal lineages of the breed. Statistical processing of non-normally distributed population genetics parameters is recommended to be performed on their normalised data.

References

- Annus K., Arkenberg H., Prikoszovich M., Oláh J., Maróti-Agóts Á., Gáspárdy A., 2015, Characterisation of the Female Tsigai Population by Use of Hungarian Herd-Book Data. In: Hajas P., Gáspárdy A. (Eds.), 25 Years with DAGENE. 1st ed. Palatia Printing and Publishing Ltd.; Győr, Hungary, 108–113.
- Arkenberg H., 2014, Calculation and evaluation of breeding parameters of the sheep breed Tsigai and the role of the females for breeding (in German). Thesis. University of Veterinary Medicine Budapest, Hungary.
- Barrientos Contreras N.V., Gonzales Tineo R.X., Castaneda-Olivera C.A., Benites Alfaro E., 2022, Use of Assaf Sheep Wool for Bioretention of Hydrocarbons (diesel) in Water Bodies. *Chemical Engineering Transactions*, 92, 217–222.
- Benites-Alfaro E., Cajaleon M.Y., Tovar S.S., Tello Mendivil V., Lizarzaburu Aguinaga D.A., 2023, Control of Food Diet in the Minimization of Greenhouse Gases Generated by Cattle Bovine. *Chemical Engineering Transactions*, 100, 67–72.
- Černá M., Milerski M., Mušková M., 2021, The effect of inbreeding on the growth ability of meat sheep breeds in the Czech Republic. *Czech Journal of Animal Science*, 66, 122–128.

- Cortellari M., Negro A., Bionda A., Grande S., Cesarani A., Carta A., Macciotta N., Biffani S., Crepaldi P., 2022, Using Pedigree and Genomic Data toward Better Management of Inbreeding in Italian Dairy Sheep and Goat Breeds. *Animals*, 12, 2828.
- Dunka B., 2001, The Gyimesi Racka. In I Bodó (ed.) *Living Heritage – Old Historical Hungarian Livestock*. Agroinform Kiadó, Budapest, Hungary, 56–57.
- 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*, 130, 286–293.
- Földi D., Földi Gy., Sáfár L., 2017, Gyimesi Racka. In L Sáfár (ed.) *Our native sheep and goat breeds*. (in Hungarian), HVG Press, Budapest, Hungary, 139–179.
- Groeneveld E., Westhuizen B.V.D., Maiwashe A., Voordewind F., Ferraz J.B.S., 2009, POPREP: a generic report for population management. *Genetics and Molecular Research*, 8, 1158–1178.
- Goyache F., Gutiérrez J.P., Gómez E., Álvarez I., Díez I., Toyo L.J., 2003, Using pedigree information to monitor genetic variability of endangered populations: the Xalda sheep breed of Asturias as an example. *Journal of Animal Breeding and Genetics*, 120, 95–105.
- Gutiérrez J.P., Goyache F., 2005, A note on ENDOG: a computer program for analysing pedigree information. *Journal of Animal Breeding and Genetics*, 122, 172–176.
- Joakimsen, Ø., 1969, Generation Interval in Norwegian Sheep, *Acta Agriculturae Scandinavica*, 19, 175–177.
- Kárpáti E., Kovács E., Gáspárdy A., Sáfár L., Gulyás L., 2022, Homozygosity status of Gyimes Racka in Hungary based on preliminary analysis of pedigree data. (in German), *Danubian Animal Genetic Resources*, 7, 57–63.
- Kinghorn B.P., Kinghorn A.J., 2010, *Pedigree Viewer 6.5*. University of New England: Armidale, Australia.
- Kovács E., Tempfli K., Shannon A., Zenke P., Maróti-Agóts Á., Sáfár L., Bali Papp Á., Gáspárdy A., 2019, STR diversity of a historical sheep breed bottlenecked, the Cikta. *The Journal of Animal and Plant Sciences*, 29, 41–47.
- Kusza Sz. Zakar E., Budai Cs., Csiszter L.T., Padeanu I., Gavojdian D., 2015, Mitochondrial DNA variability in Gyimesi Racka and Turcana sheep breeds. *Acta Biochimica Polonica*, 62, 273–280.
- Lush J.L., 1945, *Animal Breeding Plans*. Iowa State College Press, Ames, IA, USA, 443.
- Machová K., Marina H., Arranz J.J., Pelayo R., Rychtářová J., Milerski M., Vostrý L., Suárez-Vega A., 2023, Genetic diversity of two native sheep breeds by genome-wide analysis of single nucleotide polymorphisms. *Animal*, 17(1), 100690.
- McManus C., Facó O., Shiotsuki L., De Paula Rolo J.L.J., Peripolli V., 2019, Pedigree analysis of Brazilian Morada Nova hair sheep. *Small Ruminant Research*, 170, 37–42.
- Mészárosová M., Mészáros G., Moravčíková N., Pavlík I., Margetín M., Kasarda R., 2022, Within- and between-Breed Selection Signatures in the Original and Improved Valachian Sheep. *Animals*, 12(11), 1346.
- Mokhtari M.S., Miraei-Ashtiani S.R., Jafaroghli M., Gutiérrez J.P., 2015: Studying Genetic Diversity in Moghani Sheep Using Pedigree Analysis. *Journal of Agricultural Science and Technology*, 17, 1151–1160.
- Oravcová M., Krupa E., 2011, Pedigree Analysis of The Former Valachian Sheep. *Slovak Journal of Animal Science*, 44, 6–12.
- Pedrosa V.B., Santana Jr. M.L., Oliveira P.S., Eler J.P., Ferraz J.B.S., 2010, Population structure and inbreeding effects on growth traits of Santa Inês sheep in Brazil. *Small Ruminant Research*, 93, 135–139.
- Posta J., Kovács E., Tempfli K., Sáfár L., Gáspárdy A., 2019, Pedigree analysis of a population bottlenecked, the Cikta with special regard to its maternal lineages (in Hungarian). *Magyar Állatorvosok Lapja*. 141, 171–180.
- Rafter P., McHugh N., Pabou T., Berry D.P., 2022, Inbreeding trends and genetic diversity in purebred sheep populations. *Animal*, 16(8), 100604.
- Schütz L., Gulyás L., Wagenhoffer Zs., Sáfár L., Becskei Zs., Gáspárdy A., Kárpáti E., 2023, Determination of the generation interval in the Gyimes Racka and its importance in endangered sheep breeds. (in German), *Danubian Animal Genetic Resources*, 8(1), 21–28.
- Špehar M., Ramljak J., Kasap A., 2022, Estimation of genetic parameters and the effect of inbreeding on dairy traits in Istrian sheep. *Italian Journal of Animal Science*, 21, 331–342.
- TIBCO Software Inc., 2020, *Data Science Workbench*, version 14, <tibco.com>, accessed 21.11.2023.
- Venkataramanan R., Subramanian A., Sivaselvam S.N., Sivakumar T., Sreekumar C., Anilkumar R., Iyue M., 2013, Pedigree analysis of the Nilagiri sheep of South India. *Animal Genetic Resources*, 53, 11–18.
- Zsolnai A., Egerszegi I., Rózsa L., Anton I., 2021, Genetic status of lowland-type Racka sheep colour variants. *Animal*, 15(2):100080.