DOKTORAL (PhD) THESIS

SZÉCHENYI ISTVÁN UNIVERSITY
FACULTY OF AGRICULTURAL AND FOOD SCIENCES
MOSONMAGYARÓVÁR

Leader of the Doctoral School:
Dr. Vince Ördög DSc

Supervisors:
Dr. Ágnes Bali Papp PhD **Dr. András Gáspárdy PhD**

MODERN POPULATION GENETIC EXAMINATION OF CIKTA SHEEP

Written by:
Endre Kovács DVM

MOSONMAGYARÓVÁR

2018

1. OBJECTIVES

1.1. Evaluation of population structure

The author computes currently used population genetic parameters (e.g. inbreeding coefficient, pedigree completeness, generation interval) in this processing, as well as identifies the families (*maternal lineages*) of the Cikta breed in order to discover its variability based on the whole national herd book data.

1.2. Survey of resistance to scrapie

He also has a goal to determine the frequency of prion haptotypes and -genotypes, and risk groups of the today living total population of Cikta, then to compare these results to the results of 10 years ago to control the effectiveness of the prevention program introduced in the meantime.

1.3. Determining microsatellite diversity

Here, his aim is the extension of genetic evaluations in that breed based on the microsatellite polymorphisms recommended by FAO. He is also looking for the genetic identity among its sub-populations (families), that is, whether flock-specific differences can be demonstrated.

1.4. Evaluation of mitochondrial DNA-sequence diversity

In this processing his goal is to assess the diversity of Cikta's maternal genetic background derived from past of that breed by use of sequenced cytochrome b gene region.

2. MATERIAL AND METHODS

2.1. Evaluation of population structure

The Hungarian Sheep and Goat Breeders' Association (MJKSZ) gave him the national database of Cikta's (2000-2014).

In his processing he determines the number of maternal lineages (families), evaluates the pedigree completeness, inbreeding coefficient, individuals which are mostly responsible for the total genetic variability, effective population size and the generation interval.

For the statistical processing computer software like Pedigree Viewer (Kinghorn and Kinghorn, 2010), Endog (Gutiérrez and Goyache, 2005) and Poprep (Groeneveld et al., 2009) were used.

2.2. Survey of resistance to scrapie

For the statistical analysis the MJKSZ handed over the individual scrapie genotypes which is covering the whole Cikta population (2013-2015), and determined by Agrobiogen GmbH. The first home investigation served as a control (Fésüs et al., 2004 and 2008).

He estimates the number and frequency of prion haplotypes and genotypes as well as risk categories in the whole population and sub-populations by gender. Using the Chi^2 test, the current condition (2013–2015) is compared to the former one, furthermore the male sub-population to the female one, then he controls the Hardy-Weinberg genetic equilibrium.

The required data are taken from the Microsoft Excel database and are statistically evaluated using the Dell statistics program (Dell Inc., 2015).

2.3. Determining microsatellite diversity

The seventy-two individuals selected for the present examination are representatives of the oldest families. The experimental animals with 6-5-4 ancestral generations belonged to 36 families (maternal lineages).

The blood samples are collected from three farms (Duna-Dráva National Park station Nagydorog, Pénzesgyőr and Szécsénke) in fall 2015. The DNA is extracted from the blood samples using Wizard Genomic DNA Purification Kit. The amplification of the DNA is done by means of a programmable Perkin Elmer PCR machine. Nine microsatellites are analysed in two sets of multiplex reactions. Touchdown PCR is applied for the amplification of the microsatellite loci according to the manufacturer's instructions. PCR product detection and analysis are carried out using an ABI PRISM 3130XL Genetic Analyzer controlled by ABI PRISM 310 Data Collection Software. For the identification of microsatellite alleles, the 310 GeneScan Analysis Software 3.1 is applied, that uses internal standards for length determination.

The Genotyper Software is used for data interpretation and microsatellite length verification. Basic population genetics parameters – such as number of different alleles (N_a), number of effective alleles (N_e), observed, expected, and unbiased expected heterozygosity (H_o and H_e), inbreeding coefficients (F_{IS}, F_{ST}, F_{IT}) – are calculated for Cikta populations by means of the Microsatellite Analyser (Dieringer and Schrötterer, 2003) and the GenAlEx 6.503 (Peakall and Smouse, 2012) population genetic software. Hardy–Weinberg equilibrium (HWE) is tested by Chi^2 test of observed and expected microsatellite allele

frequencies. The comparison of the flocks is done by discriminant analysis. Categorization is done with Bayes classification; the comparison of the groups is based on the centroids for the groups. The deviation of the given observation from the centroid is measured by the Mahalanobis distance (Dell Inc., 2015).

2.4. Evaluation of mitochondrial DNA-sequence diversity

Blood samples are taken by jugular venepuncture using collection tubes containing EDTA as anticoagulant (from the same individuals and as in the microsatellite investigation) in fall 2015. These are stored at -20°C pending processing.

The DNA is isolated by SIGMA GenElute Blood Genomic DNA Kit according to the manufacturer's instructions. For amplification of cytochrome b mitochondrial genome a primer (1140 bp) is designed after Meadows (2005). Amplification of DNA is done by means of programmable Thermal Cycler 2720 PCR machine.

The sequencing is carried out in the Taxonomy Laboratory of Hungarian Natural History Museum.

In his analysis he is using the Tajima D-test (1989) and the Fu and Li (1993) test.

The Jukes and Cantor (1969) method is applied for correction of number of within sequence base substitutions. Based on informative mutations of haplotypes the connections among them are graphically presented on a median-joining network (Bandelt et al., 1999) with the help of software PopART (PopART) taking Genbank sequences (Hiendleder et al., 2002; Lancioni et al, 2013) as control in account.

3. RESULTS

3.1. Evaluation of population structure

The author states that the herd book of Cikta consists in total of 3648 individuals. From them 3176 individuals are with known parents, while 472 individuals with unknown parents. The number of families (maternal lineages) is 445, and from the founder generation, a continuous decline in families' number has been observed.

In terms of pedigree completeness, it can be seen that the longer the descent and more the generation number, the more critical the knowledge of the whole origin, that is, the lower degree of pedigree completeness.

The average inbreeding coefficient of the total population is 1.00%.

The genetic variability of the herd book population can be given by significantly less number of animals than the factual number of animals: only 476 individuals are responsible for the total genetic variability in the total population which consists of 3648 animals. He counted fewer numbers in the reference population which is suggesting gene losses. There are nine ancestors with the largest contribution (over 3% individual partial contribution) to the total genetic variability, however, none of them exceed 10% of the genetic variability.

The average generation interval is as long as 4 years, and the shortest one (3.80 years) is estimated for the ram-to-son pathway. Between the two merged paths (ram-progeny and ewe-progeny) there is experienced a significant difference ($P<0.05$) using a two-sample t test.

3.2. Survey of resistance to scrapie

The ARQ is the most frequent (74.93%) haplotype, which is followed by the ARR (14.19%) and AHQ (10.70%). The Chi^2 test did not show significant differences (P=0.519) for the frequencies between the current and the previous analysis.

In the present-day Cikta population the less favourable ARQ carriers are the most frequent genotypes. These are followed by the more favourable ARR- and AHQ carrier genotypes. The most sensitive homozygous genotype VRQ/VRQ does not occur. In regard to the genotypes, there is no significant deviation (P=0.083) in the present population compared to the past one, despite the fact that the least favourable genotype increased by 10%. A Chi^2 test has proven that the whole current Cikta population is in a full Hardy-Weinberg genetic equilibrium ($Chi^2 = 0.269$).

The risk group R4 is missing, and the R5 is represented by one individual only. Because of the high frequency of ARQ the R3 is appeared in almost 74%, and for the breeding most suitable risk group (R1) is only about 2.5% of the whole population. The change, regarded both the increase of R1 and R3 in the Cikta breed has been statistically proven (P=0.031) over time which is undoubtedly an alteration of the breed.

The Chi^2 test shown that there are no differences in the frequency haplotypes, -genotypes and -risk groups according to the gender.

At the same time, the Chi^2 test informs about an existing genetic equilibrium in both male and female sub-populations ($Chi^2 = 1.5947$ and $Chi^2 = 0.1162$, respectively).

3.3. Determining microsatellite diversity

The factual and effective number of alleles is 5.63 and 3.76 on average, respectively.

The values of relative Shannon's Information Index are placed in a larger range (from cca. 40 until 60%); the largest entropy (59.2%) is typical of OarCP49 microsatellite (despite having only the third largest Shannon value).

The averages of observed and expected heterozygosity are always over 0.50 and lie in a broader range of 0.52-0.92 and 0.65-0.87, respectively.

The F_{IS} with an average value of -0.18 is reflecting an elevated level of genetic differentiation within the all individuals of the breed. The highest F_{IS} value is exactly 0.00 for OarHH41, whereas the lowest one is -0.40 for CSSM47. F_{ST}-values varied between 0.02 and 0.10. Its overall value is 0.04 which shows that approximately 4% of genetic diversity is explainable by genetic variation among flocks.

The polymorphism information content (PIC) values are always less than pair values of the expected heterozygosity.

In regards to the microsatellite length, differences are found in three cases (OarCP49, CSSM47 and OarHH41) only which mean that the flocks have a high degree of identity, and the families of Cikta breed are genetically close to each other.

The common Wilks' Lambda of the discriminant functions is 0.606 ($P < 0.001$). This higher value indicates the significant discriminating ability of the functions.

Not all of the independent variables are significant ($P > 0.05$), consequently not each microsatellite plays role in isolating the flocks and

families per flocks. There are only three microsatellites (OarCP49, CSSM47 and OarHH41) which significantly increases the difference between the flocks compared to the others.

As classification result the overall accuracy is 64.2%. The flocks' accuracies are as it follows: Nagydorog 85.2%, Pénzesgyőr 21.4%, and Szécsénke 60.0%. Low precision values indicate that the flocks (and individuals representing families) are similar, and the breed is almost equally represented by the three flocks.

The squared Mahalanobis distances from group centroids are as follow: 9.6 (P = 0.56), 9.7 (P = 0.257), and 11.1 (P = 0.184). These lower and non-significant values confirm that there is no pairwise separation of the centroids of the flocks being compared. The individuals of the smaller populations of Pénzesgyőr (n=22) and Szécsénke (n=10) are placed among the individuals of the most populous Nagydorog (n=40) flock, even if function (Root) 1 is to discriminate mostly Szécsénke from the others (P <0.001).

3.4. Evaluation of mitochondrial DNA-sequence diversity

The number of monomorphic and polymorphic (mutant) sites is 920 and 16, respectively, in the 990 bp long sequences of cytochrome b gene. For the latter, it is found 1 singleton variable site and 15 parsimony informative sites. The total number of haplotypes is 10. The overall value of nucleotide diversity (π) is $2.93*10^{-3}$. The Szécsénke shares with the Pénzesgyőr and Nagydorog populations on 11 and 12 polymorphic sites, respectively; so the resemblance among them is considered as high.

The Szécsénke population has the greatest diversity, and it is characterized by the largest number of nucleotide differences (k = 4.067) among the individuals and the largest nucleotide diversity (π = $4.34*10^{-3}$). The population in Nagydorog comes afterwards and the Pénzesgyőr population has the smallest diversity.

In comparison of Szécsénke and Pénzesgyőr the average number of nucleotide differences between populations is 2.898. Here, the average number of nucleotide substitution per site between populations with Jukes and Cantor (Dxy(JC)) is $3.27*10^{-3}$ with a standard deviation of $1.36*10^{-3}$.

The average number of nucleotide differences between populations Szécsénke and Nagydorog is 3.148, and the average number of nucleotide substitution per site between populations with Jukes and Cantor (Dxy(JC)) is $3.47*10^{-3}$ with a standard deviation of $1.12*10^{-3}$.

In the third comparison (Pénzesgyőr and Nagydorog) the average number of nucleotide differences between populations is 2.754. The average number of nucleotide substitution per site between populations with Jukes and Cantor (Dxy(JC)) is $2.77*10^{-3}$ with a standard deviation of $0.68*10^{-3}$.

According to the Tajima's test the average number of pairwise nucleotide differences (k) and the nucleotide diversity (π) were 2.926 and $3.13*10^{-3}$, respectively.

The Tajima's D as the difference between two measures of genetic diversity is -0.3751, and is statistically not significant (P > 0.10).

The Fu and Li's D* test statistic and Fu and Li's F* test statistic are as follow 1.1378 (P > 0.10) and 0.7179 (P > 0.10); both are not significant.

The median-joining network based on informative mutations shows that the Cikta haplotypes are located close to each other. The haplotypes of South-Italian breeds are creating a different, but in themselves related group. The haplotypes of Asiatic Arkal and Argali are located far from them.

4. NEW SCIENTIFIC RESULTS

1. In the frame of pedigree analysis of the whole Cikta herd book data the author was the first to determine the number of families, pedigree completeness, inbreeding coefficient, number of individuals which are contributing to the genetic variability and the generation interval. Based on this, he stated that the heterogeneity of the breed is high, but the variability is decreasing in time.

2. He has done the comparative study of the molecular genetic background being responsible for resistance/susceptibility to scrapie in the whole current population. He has found that the ARQ is still nowadays the most frequent haplotype. His results draw attention to the increased genetic sensitivity of the Cikta breed to scrapie.

3. Based on the oldest families, he was the first to evaluate the microsatellite polymorphisms in the Cikta breed. He has found out that the families are genetically very similar, and the breed is almost equally represented by the three flocks. At the same time, the degree of heterozygosity can be considered as high.

4. He was the first who has sequenced the cytochrome b region of mtDNA in the ancient families of the Cikta. He has proven that the genetic identity between families and flocks is remarkable. At the same time, the population is in a genetic equilibrium and is not threatened by genetic narrowing or disruption to sub-populations. This provides a good opportunity for conservation.

5. LIST OF PUBLICATIONS

Publications related to the topic of dissertation

Papers:

Posta J, **Kovács E**, Tempfli K, Sáfár L, Bali Papp Á, Gáspárdy A (2019): A kis létszámban átmentett Cikta juh származási adatainak értékelése különös tekintettel a családokra. *Magyar Állatorvosok Lapja*. *Accepted for publication*
IF=0.196 [Q4]

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. *Journal of Animal and Plant Sciences*, 29:1.
IF=0.407 [Q3]

Gáspárdy A, Holly V, Zenke P, Maróti-Agóts Á, Sáfár L, Bali Papp Á, **Kovács E** (2018): Response of prion genic variation to scrapie resistance program in Hungarian indigenous sheep breeds. *Acta Veterinaria Hungarica*, 66:4. 562-572.
IF=1,042 [Q2]

Kovács E, Mitro S, Tempfli K, Zenke P, Maróti-Agóts Á, Sáfár L, Bali Papp Á, Gáspárdy A (2017): A specific selection programme is required in the autochthonous Cikta Sheep which is endangered by own frequent ARQ prion haplotype? *Applied Agricultural and Forestry Research* (*Landbauforschung*), 67:3-4. 141-146.
IF=0,152 [Q4]

Annus K, **Kovács E**, Sáfár L, Gáspárdy A (2015): A magyar cigája behelyezése az európai juhok közé az anyai eredet szerint (mtDNA CR). *Animal Welfare Ethology and Housing Systems*, 11:2. 59-64.

Whole proceedings:

Kovács E, Posta J, Tempfli K, Sáfár L, Becskei Zs, Bali Papp Á, Gáspárdy A (2018): Verarbeitung der Abstammungsangaben des vollen Zuchtbuches bei gefährdeten Schafrasse Cikta. (vollständiger Beitrag). "Role of animal and plant genetic resources in ecosystems" 29th Annual Meeting of DAGENE, from 24th to 27th of June 2018, Kozárd, Hungary, *Danubian Animal Genetic Resources*, Vol. 3. 45-50.

Kovács E, Zenke P, Mitro S, Tempfli K, Sáfár L, Becskei Zs, Bali Papp Á, Gáspárdy A (2017): Die Ergebnisse der molekulargenetischen Untersuchungen der Traberkrankheit bei ungarischen Cikta Schafen (vollständiger Beitrag). "Tradition and innovation in preservation of autochthonous breeds" 28th Annual Meeting of DAGENE, from 26nd to 29th of April 2017, Pazin, Croatia, *Danubian Animal Genetic Resources*, Vol. 2. 91-97.

Kovács E, Zenke P, Sáfár L, Cenkvári É, Bersényi A, Bali-Papp Á, Gáspárdy A (2016): A surlókór molekuláris genetikai vizsgálatának eredményei a cikta juhban. Előadás. Hagyomány és innováció az agrár- és élelmiszergazdaságban, XXXVI. Óvári

Tudományos Nap, Mosonmagyaróvár, 2016. november 10. (Szerk.: Bali-Papp Ágnes és Szalka Éva) ISBN 978-615-5391-79-8:254-261.

Kovács E, Rabe A, Annus K, Sáfár L, Gáspárdy A (2016): Ursprung und Entwicklung des autochthonen Cikta Schafes bis zum Gegenwart in Ungarn. "Innovative approaches in biotechnology and genetic engineering applied in rare breed preservation" 27th Annual Meeting of DAGENE, from 22nd to 24th of April 2016, Hilgertshausen-Tandern, Germany, *Danubian Animal Genetic Resources*, Vol. 1. 23-28.

Oral presentations and posters:

Kovács E, Posta J, Gáspárdy A (2018): Pedigree analysis of a population bottlenecked with special regard to its maternal lineages. Poster. IV. Conference of Central and Eastern European Laboratory Animal (CEELA), 2018. június 2., Állatorvostudományi Egyetem, Aula, 1078 Budapest, István utca 2.

Pásztor K, Annus K, Szabó K, **Kovács E**, Gáspárdy A (2016): Egy őshonos juhfajta, a cigája, anyai vonalainak genetikai változatossága (mt-Cytb) (The genetic diversity of the Hungarian Tsigai herds (mt-Cytb). Oral presentation: A Magyar Laborállat-tudományi Egyesület (MLTE) Éves Tudományos Konferenciája, 2016. június 9., SZIE ÁOTK Sebészeti Előadó, 1078 Budapest, István utca 2.

Annus K, **Kovács E**, Sáfár L, Gáspárdy A (2015): A magyar cigája behelyezése az európai juhok közé az anyai eredet szerint (MTDNA CR). Oral presentation: V. Gödöllői Állattenyésztési Tudományos Napok Nemzetközi Konferencia, Állattenyésztés-tudomány Szekció, 2015. október 21-22.

Other periodicals:

Gáspárdy A, **Kovács E** (2019): The question of the scrapie resistency in a native sheep breed, the Cikta - Dagene research. SAVE Journal

Supervisor of thesis students:

Rabe Anna (2017: Genetic Investigations in Hungarian Zaupel Sheep. Thesis. ÁTE Budapest, VIth year student (**Dr. Kovács Endre** together with Dr. Gáspárdy András)

Shannon Adrian (in progress): Analysis of microsatellite diversity in the indigenous Cikta sheep breed. Thesis work. ÁTE Budapest, Vth year student (Dr. Kovács Endre together with Dr. Gáspárdy András)