DOCTORAL THESIS ABSTRACT

Doctoral thesis title: „Phylogenetic relationships and genetic diversity of endangered romanian cattle breeds, Grey Steppe and Pinzgau”

Key words: *genetic diversity; Grey Steppe cattle breed, Pinzgau cattle breed; phylogeny; genetic variability; cytochrome b; d-loop; mitochondrial markers; gene sequencing*

In the last 20 years, about 300 breeds of approx. 6,000 belonging to different species of animals have been identified by the Food and Agriculture Organization of the United Nations (F.A.O.) as endangered.

The main risk factors are genetic erosion due to artificial insemination, economic pressure on farmers with a direct target on large milk or meat production or crossbreeding between individuals with different genetic structure (Pariset et al., 2010).

The category of endangered cattle breeds includes two local breeds, Grey Steppe and Pinzgau, which are the subject of this study.

The indigenous cattle breeds from Romania are classified by numerous researchers in primitive breeds (Grey Steppe and Mocanita) and improved breeds (Romanian Spotted, Romanian Black Spotted, Brown, Pinzgau) (Creangă, Maciuc, 2010; Maciuc, 2006).

Romanian Grey Steppe cattle breed belongs to the group of podolian breeds, still found in different areas of Europe. Its common origin, with other Podolian breeds (Iskar Gray, Bulgarian Gray, Istrian, Slavonian Podolian, Katerini, Hungarian Gray, Maremma, Podolica, Turkish Gray, etc.) (Teneva et al., 2005; Pariset, 2010) is the wild ancestor *Bos taurus primigenius*, which was become extinct around the 16th century, the last herd recorded being in Poland, in 1627 AD. (Götherström et al., 2005; Pitt et al., 2018).

It was widespread in Eurasia and North Africa during two major epochs: the *Late Pleistocene* and *Early Holocene* (Meadow, 1993; Clutton-Brock, 1990), giving birth to the two ancestors who played a crucial role in the domestication of cattle: *Bos taurus* and *Bos indicus* (Park et al., 2015) and which differs by the presence of a hump (specific for *Bos indicus*).

Bos taurus includes typical cattle from Europe, Northeast Asia and North Africa, adapted mainly to the cold climate (Upadhyay et al., 2017). Numerous studies have shown that modern cattle breeds from Europe, are characterized by the mitochondrial haplogroup T (Edwards et al., 2007; Scheu et al., 2008).

Pinzgau cattle breed is adapted for breeding and exploitation in areas with altitudes of 400-1600 m, rich in rainfall and fertile natural meadows, with high resistance and capitalizing very well on feed with a high cellulose content (Cotos, 2005; Fisteag, 1956; Fisteag, 1958). The romanian Pinzgau breed was formed by crosses between cows from the Grey Steppe breed and Pinzgauer bulls from Austria (Angelescu, 1974; Mang, 2011).

Revised bibliographic references (Bradley, Magee, 2006; Georgescu et al., 2008; Creangă et al., 2009; Bittante, 2011; Maretto et al., 2012), regarding the endangered indigenous cattle breeds in Romania, confirm that they are carriers of a reservoir of valuable genes that must be preserved by applying different reproductive biotechnologies (Han and Bobiș, 2018).

To complete this doctoral thesis, a series of researches were carried out in order to evaluate the genetic diversity of the two cattle breeds, Grey Steppe and Pinzgau, which are endangered in Romania, by analyzing two mitochondrial markers (cytochrome b and d-loop), relevant for studies of genetic diversity, phylogeny, molecular phylogeography and identification of taxonomic relationships between individuals of different species.

The first part of the thesis is made after consulting the bibliographic references in the literature, summing up a number of 3 chapters that present aspects on the genetic diversity of cattle, starting from archaeological evidence attesting their domestication and evolution, origin and phylogeny of European breeds and analysis the risk status of romanian Grey Steppe and Pinzgau.

The second part of the thesis brings a series of additions to the existing information in the literature on genetic diversity and clarification of taxonomic uncertainties and highlighting new considerations related to the evolutionary history of these breeds within the subfamily *Bovidae*, totaling a number of 4 chapters.

The aim of this research was to quantify the genetic variability of Grey Steppe and Pinzgau cattle breeds and to establish phylogenetic relationships within the *Bovinae* subfamily.

The objective plan includes several stages of research:

1. Quantification of the genetic variability of Grey Steppe and Pinzgau cattle breeds, by performing the following steps:

- identification of haplotypes present in the Grey Steppe population (S.C.D.C.B., Dancu-Iași) and in the Pinzgau population (S.C.D.C.B., Târgu-Mureș), by analyzing the nucleotide sequences of cytochrome b and d-loop mitochondrial control region;
- estimation of intra and inter-population genetic diversity based on the results obtained from the analysis of the two mitochondrial markers.

2. Phylogenetic analysis of breeds:

- comparing the evolution rate of mitochondrial markers, with the evolution rate of the complete mitochondrial genome, by analyzing the gene sequences resulting from sequencing;
- the impact of mitochondrial markers in estimating the time of divergence.

The achievement of the objectives proposed in this research was possible by performing analysis in specialized laboratories both in the country and abroad.

To quantify the genetic diversity, a series of analyzes were performed within the *Molecular Genetics Laboratory of the Faculty of Biology*, “Alexandru Ioan Cuza”, University of Iasi.

For the sequencing of the cytochrome b and d-loop mitochondrial markers, DNA samples were purified and prepared to be sent to a specialized laboratory in Amsterdam-Netherlands, "*Macrogen-Europe*" and the DNA sequences were sent in electronic format, in the form of FASTA files and chromatograms specific to each analyzed sequence.

A series of genetic analysis programs, listed and presented below, was used to analyze the gene sequences of cytochrome b and d-loop mitochondrial control region.

- Sequencing of PCR amplicons: SANGER type sequencing within the Molecular Genetics Laboratory of the Faculty of Biology, “Alexandru Ioan Cuza”, University of Iași and within the Macrogen-Europe Laboratory in the Netherlands;
- Chromatogram alignment and sequence correction using the DNA Baser program;
- Sequence analysis:

- Alignment - Mega X program;
- Substitution model - jModelTest program;
- Haplotypes Network - PopArt program;
- Construction of phylogenetic trees (NJ-Neighbor Joing and ML - maximum likelihood) - SeaView program;
- Population genetics analysis - DnaSP 5.10 and R programs.

Pedigree analysis

The evaluation of genetic diversity of Grey Steppe cattle breed based on the pedigree analysis was performed on a number of 171 individuals**.**

The data were first evaluated for each individual and then used to calculate: generation interval (L), kinship (fIBD), native kinship (fIBD | N), native genome equivalent (NGE), inbreeding coefficient (F), effective population size (Ne), genetic diversity, pedigree-based native contribution (NPED), genetic contribution of other breeds and optimal contribution selection, using the R optiSel package (Wellmann, 2019).

The Grey Steppe population from S.C.D.C.B., Dancu was increased in 2002, by in vitro insemination of a number of 19 females, using semen from 11 bulls. The male ancestors used to build all 9 genetic lines were pure bred individuals with a pedigree and phenotype specific to Grey Steppe. Female specimens presented varying degrees of uncertainty as to their purebred status, some of them having non-breed-specific characteristics.

Initially, the population size registered a phase of accentuated growth, with a maximum of 88 individuals identified in 2007, followed by a short phase of the decrease in the number of specimens (period 2007-2009) and from 2010, it is stable at approximately 58 individuals.

Overall, the birth-death ratio was positive, with an average birth rate of 8.3 individuals/year and an average mortality rate of 6 individuals/year. The average lifespan was only 6.4 years but the average interval between generations was quite high (12.75).

The high value of the interval between generations is influenced by the age of the females, which varies between 0 - 24 years, while the males have a lifespan of no more than 5 years. Based on pedigree data, 6 generations for females and 5 for males was completely followed. The integrity of the pedigree information showed relatively high values (< 0.85) until the second generation, both for males and females.

The inbreeding coefficient showed values in the range 0 - 0.39, while the average was 3.82% for living individuals and 2.3% for total individuals analyzed.

The current population has a level of genetic diversity of ~ 94%. The genetic contribution of individuals with unclear or nonspecific origins has gradually decreased over time from ~ 50% in 2002 to less than 3% in 2019.

Gene sequence analysis

Following the sequencing of the cytochrome b gene and d-loop mitochondrial control region, it was found that these markers are of particular relevance in highlighting genetic differences and establishing phylogeny between individuals.

By analyzing and interpreting the gene sequences specific to Grey Steppe cattle breed, a number of 4 haplogroups were identified, with different frequencies (T1-3%, T2-19%, T3/T4-62% and P'QT-16 %), in which the 32 individuals studied were included. The identification of the P'QT haplogroup following the analysis of nucleotide sequences, indicates that this haplogroup is of ancestral type, being specific to the ox (*Bos taurus primigenius*), from which this breed of cattle evolved.

Within this haplogroup, a number of 5 individuals were identified, representing a percentage of 16% of the total analyzed specimens, which could be used in the selection, for crosses, in order to preserve the purebred breed.

Demographic and spatial expansion is based on the *Mismatch distribution model* and shows the distribution of the number of differences, identified between pairs of haplotypes. As a rule, this distribution can be unimodal (specific to populations that have undergone either a recent demographic expansion or a spatial expansion, characterized by a high degree of migration) or multimodal (in the case of populations in demographic equilibrium). For the Grey Steppe cattle population, a multimodal distribution resulted both from the analysis of the nucleotide sequences of cytochrome b and d-loop mitochondrial control region.

In the case of Pinzgau breed, as in the case of the Grey Steppe breed, the haplogroup frequency for the 24 individuals was identified, based on the analysis of the cytochrome b gene sequences and d-loop mitochondrial control region. The proportion of the four nitrogenous bases for both sequences and the specificity coefficient was also calculated based on the A+T/G+C ratio. The dynamics of the evolution rate of the two mitochondrial markers and the demographic expansion of the group of analyzed individuals were followed.

As a result of the analysis of gene sequences, unlike the Grey Steppe breed, in the case of Pinzgau breed, 3 haplogroups were identified, having different frequencies (T1, T2 and T3). The 24 individuals were distributed in the corresponding haplogroups, as follows: 18 individuals belonging to the T3 haplogroup (75%), 5 individuals identified within the T2 haplogroup (21%) and one individual representative for the T1 haplogroup (4%).

The first chapters of the thesis focused on data on archaeological evidence that cattle went through two major domestication events, namely the first event in southwest Asia, giving birth to cattle with offspring in *Bos primigenius* and the second of which evolved zebu or *Bos indicus* (Loftus et. al., 1994; Troy et. al., 2001).

Mitochondrial DNA analyzes in european cattle breeds showed that most of them belonged to the T haplogroup, with an estimated the time of divergence of ~ 16 thousand years. This T haplogroup is divided into two sister subclades, namely T1-T2-T3 (in which a large part of the individuals belonging to the two breeds were identified) and the T4-T5 subclades (Mannen et. Al., 2004; Bradley et. al., 1996).

Haplogroups T and P'QT, identified in the Grey Steppe breed, were involved in the same domestication process in the Fertile Crescent (Achilli et. Al., 2008; Achilli et. Al., 2009).

Pinzgau individuals presented a multimodal distribution, resulting from both the analysis of the nucleotide sequences of the cytochrome b and d-loop mitochondrial control region, which shows a demographic balance.

The results of investigations regarding the quantification of the genetic variability of the two romanian cattle breeds, Grey Steppe and Pinzgau and identification of phylogenetic relations within the subfamily *Bovinae* led to the formulation of the following conclusions:

- the spectrophotometric quantification of total DNA, validated quantitatively and qualitatively the isolation of deoxyribonucleic acid, making it possible to move on to the next stages of analysis;
- by migrating the agarose gel PCR products, it was found that there were no foreign amplifications and no contamination and the primers used for amplification showed a high degree of specificity for the two mitochondrial markers were analyzed;
- following the sequencing of the cytochrome b and d-loop mitochondrial markers, it was found that they are of particular relevance in highlighting genetic differences and establishing phylogeny between individuals;
- the molecular analyzes performed regarding the genetic diversity, the demographic and spatial expansion of the two romanian cattle breeds, demonstrate the spatio-temporal sustainability of the individuals, a favorable detail for the process of their reintroduction in freedom;
- phylogenetic research has shown that in order to correctly quantify phylogenetic relationships within the *Bovinae* subfamily and to estimate time of divergence, it is insufficient to use a tree with high topological support, which can not correctly estimate clade age but to analyze concatenated mitochondrial markers and analyzed individually for the breeds included in the study;
- through the molecular investigations performed, it was found that the cytochrome b gene represents the mitochondrial marker most able to correctly evaluate the phylogenetic relationships within the *Bovinae* subfamily.

The results of this research can contribute to the improvement of the conservation program of Grey Steppe and Pinzgau breeds and to the completion of current information on their genetic diversity, representing a valuable tool for efforts to conserve these endangered gene reserves.

Maintaining the genetic background of these breeds requires the application of additional measures and the allocation of considerable financial resources. To design appropriate and effective conservation strategies, a set of accurate knowledge about all existing individuals in a population must be taken into account. Therefore, further research is recommended to consider the analysis of the complete mitochondrial genome, the Y chromosome (paternal), microsatellites, the analysis of DNA extracted from fossils to *Bos taurus* bulls, thus ensuring a better understanding of the real genetic structure and the previous history of the studied population.

An important aspect to ensure an efficient conservation program of the two breeds is the conservation of genetic material by cryogenizing the samples (tissue and genomic DNA) and creating a database with archived samples, phenotypic parameters and genotypic profile of each individual, so to ensure the spread of the breed.

Genomic analysis, the application of reproductive biotechnology and the creation of a bank of genetic material will contribute to an active conservation and elimination of breeds from the list of endangered cattle and the development of specific guidelines, in economic, genetic and cultural terms.