The aim of this study was to assess the genetic variations and relationships across evolutionary related cervid species in order to estimate the genetic diversity of the Red deer population that inhabits the forest area in the south-western part of Slovakia. The study was based on the application of cross-species SNP genotyping. The genomic data were obtained from a total of 86 individuals representing six genera (Axis, Dama, Cervus, Alces, Rangifer, and Odocoileus) using Illumina BovineSNP50 BeadChip. From 38.85% of successfully genotyped loci up to 1,532 SNPs showed polymorphism and were informative for subsequent analyses of the diversity and interspecific genetic relationships. Generally, a good level of observed heterozygosity was found across all species. The value of $F_{IS}$ (0.23±0.13) signalised the increase of a homozygous proportion within them. The application of molecular variance analysis to the hierarchical population structure showed that most of the variation was conserved within separate species (96%). The performed diversity analysis of Slovak Red deer population and comparative analysis of their phylogenic relationships among subspecies from genus Cervus did not identify a remarkable loss of genetic variability. Also, were not identified any degree of admixture that could be due to the historical background of deer farming in Slovakia or reintroduction and hybridisation by other species from genus Cervus (C.canadensis, and C. nippon) which are the major risk of loss of autochthonous Red deer populations in many areas of Central Europe. The analysis of individual’s ancestry showed consistent results with patterns of evaluated group differentiations which means low migration rates among all species.

**Key words**: deer, cross-species genotyping, genetic diversity, SNP

### INTRODUCTION

The deer family (Cervidae) is distributed throughout the northern hemisphere, as well as in South America and Southeast Asia. With 55 species recognized by the International
Union for the Conservation of Nature (IUCN) the family *Cervidae* can be considered as the second richest family of artiodactyls after *Bovidae* [1]. The Red deer (*Cervus Elaphus*) belonging to the free living hunting ruminant species is an important member of the native fauna in Slovakia. Due to the development and wider use of firearms the distribution and numbers of Red deer in the forest area of Slovakia decreased continuously from the 10th up to the 19th century [2]. By the first half of 19th century Red deer were on the verge of extinction with the remains of a native Carpathian population thriving only in remote mountain areas. This dramatic decline led to a large-scale reintroduction of other deer populations to strengthen the remains of the native population. Red deer of various origin were introduced (*Cervus Elaphus Hippelaphus*, *C. Elaphus Sibiricus*, and *C. Elaphus Canadensis*), bred in enclosures and acclimation pens and then released into the wild. Despite that, nowadays the Red deer in Slovakia are often described as belonging to the subspecies *Cervus Elaphus Hippelaphus*. The population is in practice rather a mongrel mixture of genotypes [2,3]. At the present, the size of Red deer population according to statistical data is estimated to be up to 62,000 individuals (62,781 in 2014). Moreover, due to the cattle numbers decline during the last ten years the value of Red deer as new farming species in Slovakia has been increased.

In the recent past, human activity has led to severe effects in many wild species populations that usually resulted in an increase of the genetic drift strength leading to the loss of genetic variation [4]. Through hunting, habitat degradation, domestication and translocation, humans have brought many non-native species into contact with native heterospecifics, leading to anthropogenically induced hybridisation that can cause significant losses in the economic value of native genetic resources [5,6]. Despite the world distribution of *Cervidae* family half of the species are currently considered as endangered or extinct in the world [7,8]. The genetic structure and conservation status of native European Red deer populations are threatened with extinction in some parts of Central Europe. Due to the hybridisation with the Japanese Sika deer and North American Wapiti [9-11]. For long-term conservation and development purpose, it therefore appears compulsory to manage their wildlife to maintain both species survival and within species genetic diversity [12,13].

The phylogenetic relationships and genetic diversity, as well as the level of admixture across native populations of the *Cervidae* family have been investigated primarily based on mitochondrial DNA [14,15] and microsatellite markers [11-13]. Profiling a large number of single nucleotide polymorphisms (SNPs) was only realistically possible for organisms with well-developed genomic resources, and the high cost of developing these resources has been a major impediment to studies for non-model organisms, e.g. deer [16]. However, most of the current studies assessing phylogeny or population genetics of wild animals require only a reduced set of SNPs that are spread across the genome [7,17,18]. The commercially developed genotyping array for cattle were successfully applied in relation to the novel SNPs detection and validation on two genera from the families *Cervidae*, *Odocoileus* [19] and *Rangifer* [20]. The bovine SNP
panel was used to assess the phylogenetic relationship between the species from both \textit{Bovidae} and \textit{Cervidae} families [21]. The New Zealand deer industry currently uses genomic technology to produce SNPs panels for parentage and breed assignment [22].

The aim of the study was to infer the genetic variations and relationships across eight evolutionary related cervid species with respect to the estimation of genetic diversity within Red deer population from Slovakia based on use of a commercially developed genotyping array for cattle.

\section*{MATERIALS AND METHODS}

Overall 29 biological samples of free living and farmed deer were collected from male progeny of sires from New Zealand and dams from Hungary born on farms and free living trophy animals from legal game hunting in Slovakia (Table 1, population 1, 2). The sample of analysed animals represents the gene pool of the Red deer population that inhabits the forest area in south-western part of Slovakia. In respect to the description of Red deer diversity and potential level of admixture due to the reintroduction of other populations in history the test dataset of 7 species from the family \textit{Cervidae} has been used. The test dataset was compiled with genotyping data from the annex of Decker et al. [21]. The detailed description of sample size and data source can be found in Table 1 (population 3-9). All of them were genotyped using Illumina BovineSNP50 BeadChip V1 or V2 (Table 1) (Illumina, Inc. San Diego, USA). The genome-wide data from in total 86 individuals of 8 cervid species were included in the final dataset. After joining of both the original and test datasets the 36,451 loci common to the applied Illumina genotyping arrays were retained in the reduced panel of SNPs.

\begin{table}[h]
\centering
\caption{Dataset summary}
\begin{tabular}{|l|l|l|l|}
\hline
\textbf{Species} & \textbf{Total number} & \textbf{BovineSNP50 BeadChip} & \textbf{Source} \\
\hline
1 & \textit{Cervus elaphus} (Red deer) & 27 & V2 & Slovakia \\
2 & \textit{Dama dama} (Fallow deer) & 2 & & \\
3 & \textit{Cervus canadensis} (Wapiti) & 8 & & \\
4 & \textit{Cervus nippon} (Sika deer) & 8 & & \\
5 & \textit{Dama dama} (Fallow deer) & 8 & & \\
6 & \textit{Axis axis} (Axis deer) & 8 & V1 & Decker et al (2009) \\
7 & \textit{Alces alces} (Moose) & 10 & & \\
8 & \textit{Rangifer tarandus} (Reindeer) & 7 & & \\
9 & \textit{Odocoileus virginianus} (White-tailed deer) & 8 & & \\
\hline
Total & & 86 & & \\
\hline
\end{tabular}
\end{table}
The quality control of genotyping data was performed according to Purcell et al. [23]. The SNPs were eliminated from further analysis if they were missing in at least 90% of analysed individuals and Hardy-Weinberg equilibrium test with limit 5x10^{-5}. Only SNPs with minor allele frequency higher than 0.01 were retained in our analysis. The physical map locations of all loci were ignored mainly due to the fact that the bovine genome is not representative to the genetic changes that have taken place within the evolution of the cervid genome, and the genetic map of the genus *Cervus* itself is not determined yet.

The basic diversity indices including measures derived from allelic frequency such as heterozygosity (observed and expected) were estimated separately for each species according to Excoffier et al. [24]. Because individuals evaluated within separate groups are not really a population, but members of one species, the Hardy–Weinberg expectation may not have been necessarily met and is therefore ignored in further analysis. The genetic differentiation level within and across species was evaluated based on pairwise $F_{ST}$ values and Nei's genetic distances using package StAMPP implemented in R software [25]. The analysis of molecular variance (AMOVA) was performed based on 10,000 permutations according to Excoffier et al. [24].

The Bayesian clustering approach adopted in Structure 2.3.4 [26] has been used to infer the population structure and estimate the number of clusters represented by all analysed individuals in relation to the individual admixture proportion. The analysis was performed using the default parameters of an admixture model and correlated allele frequencies based on burn-in period of 100,000 followed by 1,000,000 MCMC (Markov chain Monte Carlo) replications. Twenty runs were carried out from K=1 to K=20. The optimal value of K was selected based on the evaluation of the log probability of delta K ($\Delta K$) according to Evanno et al. [27].

In order to obtain additional information about the genetic structure in the evaluated group of cervids the discriminant analysis of principal components (DAPC) implemented in R package Adegenet [28] was used. Based on DAPC the optimal number of clusters by running a K-means algorithm that uses BIC (Bayesian Information Criterion) to assess the best supported model and the number and nature of clusters was recognized. Moreover, using the DAPC analysis the membership probability was obtained to present the overall genetic background of an individual.

Migration and genetic drift recognized as the two major factors affecting population diversity of free range animals were analysed. The estimation of recent migration rates were tested by the method implemented in Bayesass v1.3 [29] which uses individual multilocus genotypes to evaluate rates of recent immigration in the Bayesian statistical framework. The TreeMix software [30] has been then used to recognize the genetic drift and migration events at genome-wide polymorphisms level to infer the genetic relationship among analysed cervid species. The consistency of migration events were evaluated using ten times separate runs with -m set to maximum 8 migration events. The ancestry graph and relative migration level to explore the patterns of gene flow...
that have influenced a group of population samples were visualised using R software [31].

RESULTS

The quality control of genotyping data resulted in a loss of 96.55% of SNPs that retained in reduced loci panel after the joining original and test datasets. Observed call rate at level 61.23% indicated the decrease in genotyping quality. Only 38.85% from the total number of loci in the reduced panel has been successfully genotyped in at least 90% of individuals. From them up to 1,532 SNPs could be regarded as polymorphic, based on generally accepted criteria characterizing any nucleotide variation at a single position in a DNA sequence among individuals occurring in a population with a frequency higher than 1% as a single nucleotide polymorphism. If the distribution of polymorphic SNPs according to the map location on the bovine reference genome assembly (Btau 4.0) should be taken into account it could be said that the proportion of polymorphic loci varied across each autosome from 24 to 95. The highest total number of polymorphic loci was found on autosome 1 and the lowest on autosome 23 (Figure 1A). However, the predicted position of each locus may not be consistent with the real localization in the genome of cervids, mainly due to the changes that could happen during the phylogeny of families Bovidae and Cervidae. Across all polymorphic loci the average value of MAF was found at level of 0.15 with a confidence interval of (95%) 0.15 – 0.16. The distribution of the MAFs estimated within five classes is given in Figure 1B.

Due to the small sample size a high standard deviation for all genetic diversity indices was found. Among all species generally good levels of expected and observed heterozygosity were identified based on the selected informative SNPs panel. The level of observed heterozygosity ranged from 0.13±0.07 to 0.36±0.28. The species from subfamily Odocoilenae (A. alces, R. tarandus, O. virginianus) showed a higher level of expected heterozygosity (0.42±0.12) compared to other. Across species from subfamily Cervinae (C. elaphus, C. Canadensis, C. nippon, A. axis, D. dama) the expected heterozygosity at level 0.33±0.18 was found. However higher average value of

![Figure 1. Distributions of informative polymorphic SNPs across bovine autosomes (A) and minor allele frequency within MAF classes (B).](image-url)
inbreeding (0.23±0.13), measured by $F_{IS}$ index, characterizing genetic variability within each population indicated the prevalence of homozygous genotypes in particular subpopulations. Similarly, the observed $F_{IT}$ value at level 0.76 denoted the decrease of heterozygous genotype proportion across all evaluated animals.

The analysis of molecular variance (AMOVA) showed that the genetic variation within species accounted for 95.74%. Across individuals was distributed 3.29% of variations and the remaining 0.97% was explained by differences among individuals within each species. The genetic differentiation estimated based on $F_{ST}$ varied from 0.90 to 0.99 among species in the subfamily *Odocoileinae* and from 0.46 to 0.99 across all species from subfamily *Cervinae*. Among subfamily *Odocoileinae* the lowest $F_{ST}$ value was found between species *R. Tarandus* and *O. Virginianus* that are originating from the same tribe *Odocoileini*. Within the subfamily *Cervinae* the genetic differences based on $F_{ST}$ were observed at level 0.86 among separate genera and the lowest $F_{ST}$ values (in average 0.67) were found between species *C. Nippon*, *C. Canadensis* and *C. Elaphus* which are originating from the same genus *Cervus*.

The observed genetic distances averaged across analysed species indicate a division of individuals into three main separate groups according to their origin; the group formed by genera *Rangifer* and *Odocoileus*, the group formed by genus *Alces*, and the group formed from genera *Axis*, *Dama*, and *Cervus* (Figure 2A). The average genetic distance across all individuals ranged from 0.02 (between species *C. Elaphus* and *C. Canadensis*) to 0.55 (between species *A. Axis* and *A. Alcei*). Similar to the $F_{ST}$ values the highest level of genetic similarity was found between individuals from genus *Cervus* (in average 0.13). The obtained results clearly signaled the significant genetic differentiation among separate cervid species referred to their phylogeny and historical origin. As expected only the subspecies from genus *Cervus* showed a higher degree of genetic similarity mainly in respect to their evolution and history of their breeding in Slovakia.

![Figure 2](image_url)

Figure 2. Hierarchical plot based on Nei’s genetic distance (A), maximum likelihood tree showing inferred relationship between analysed species (B) and the representative results from STRUCTURE at K=3 (C).
The Bayesian clustering and the method according to Evanno et al. [27], evaluating the second order of change of the likelihood with respect to $\Delta K$, confirmed the local maxima of $K=3$ consistent with tribes (Figure 2C). Subsequent discriminant analysis of principal components (DAPC) validated a strong distinction among cervid species (Figure 3) at the tribe levels. The division of individuals in to the clusters produced by BIC analysis showed that the inferred clusters corresponded to the actual groups (Figure 3A) and therefore the clusters in DAPC were defined by priori assumptions of population membership. The retained ten PCA axes corresponded to more than 95% of conserved variance and obtained five discriminant functions to more than 50% of variance. The first discriminant function clearly showed division of individuals into the clusters in relation to three main group; the species from tribe Alceini (A. Alces), Odocoileini (R. Tarandus, O. Virginianus) and Cervini (A. Axis, D. Dama, C. Elaphus, C. Canadensis, C. Nippon) (Figure 3B). The evidence of admixture among evaluated individuals has been found only within species from tribe Cervini. As expected the presence of admixed individuals with membership probability higher than 90% was identified in group of C. Elaphus and C. Canadensis. But the highest frequency of admixture was found between the species A. Axis and C. Nippon with individual membership lower than 90%. Between the species from subfamily Odocoilenae no evidence of admixture between analysed individuals was detected (Figure 3C).

**Figure 3.** The BIC results referred to the differentiation between inferred and original clusters (A), genetic clusters determined based on first discriminant function (B), and membership probability of each individual (C).

For analysis of ancestry we used the approach adopted in Treemix in order to verify that the selected panel of SNPs can reflect the history of evaluated cervid species. We run the analysis with B. Taurus as outgroup species to detect the patterns of divergence mainly based on the fact that the genotyping array was prepared in respect to the nucleotide changes in his genome and the selected polymorphic SNPs are only orthologous loci to the divergence between Cervidae and Bovidae families. The
analysis showed only a low level of migration events that also confirmed the analysis of relative migration rates. The observed results allowed us to estimate the proportion of variance in relation to species that indicated the division of individuals into the branches according to previously reported results (Figure 2B). Following the analysis of relative migration rate, e.g., number of individuals that can exchange among species and could differentiate the effect of genetic drift, showed mostly consistent results with patterns of analysed group differentiation which means low migration rates in all species. The highest migration rates have been detected between the species of *C. Elaphus* and *C. Canadensis*, and the lowest proportion of migrants relative to the other groups was detected among species from subfamily *Odocoilenae* (Figure 4).

**DISCUSSION**

As most biologists are probably aware, technological advances in molecular biology during the last few years have opened up the possibility to generate genome-wide data from non-model organisms. The application of genotyping array produced for livestock to wild phylogeny related species can provide the source for novel marker discovery without the need of the whole genome sequencing [18,32]. We adopted this approach across the non-model species from the family *Cervidae* to demonstrate the suitability of bovine genotyping array for cross-species genotyping of phylogenetically related species and for the evaluation of genetic diversity within the deer population which provides a representative sample of current genetic resources of Red deer in Slovakia. The common phylogenetic origin between *Bovidae* and *Cervidae* families could result in the occurrence of orthologous mutations that can be considered as valuable source referring to the genetic variability across cervids. In most of the wild species the absence of a whole genome scan and consequently unknown physical localization of loci allows only the use of population parameters derived from allelic frequency differences. However, the discovery and validation of novel SNPs based on available genotyping arrays that are currently routinely employed in livestock seem simple and
cost effective ways for analysis of genetic diversity across various species including cervids.

The obtained call rate indicated a rapidly decrease of genotyping quality compared to the species from the family Bovidae. However it has been expected, as previous results by Miller et al. [33] demonstrated that the decrease is on the border line of 1.5% per each million year a divergence between model (target species of array) and non-model species. Based on the fact that the Bovidae and Cervidae separated from each other about 20 – 30 million years ago [34] the observed call rate can be regarded as surprisingly high. In this case the results signalised that the decrease is not so rapid and can reach a level of around 1% per million year divergence. The total number of identified informative polymorphic markers across 50K bovine loci panel at the level of 4.2% was comparable to results presented by Haynes and Latch [19] and Kharzinova et al. [20]. Both studies detected using BovineSNP50 BeadChip on approximately 5% of polymorphic loci within the genus Odocoileus and Rangifer. The detected major loss of SNPs (in order to obtain only informative markers for diversity analyses) can be attributed mainly to the fact that the cross-species genotyping is likely biased in favour of conserved genomic regions that still retain polymorphisms ancestral to the divergence between Cervidae and Bovidae families. Despite the decrease of identified informative loci number compared to bovids if only 1% of markers on a 50k SNP genotyping array are polymorphic then those 500 loci would represent a substantial increase in genetic resources available for wild species [19,33].

Although the results are in accordance to previously published studies, they should be interpreted with caution because any cross-species SNP genotyping may be affected by ascertainment biases, systematic deviations of population genetic statistics from theoretical expectations that can be caused by sampling a non-random set of individuals or by biased SNP discovery protocols [35]. One of the major impediment that limited the application of genotyping array on the phylogenetically related species is the divergence of genomes that resulted in the occurrence of non-representative nucleotide changes within them. The karyotype differences caused an incorrect determination of loci physical positions within the genome which lead to the strongly biased analyses based on e.g. linkage disequilibrium or identifying runs of homozygosity. On the other hand the number of loci detecting by cross-species genotyping is overall too small for reliable analysis of homozygous segments describing the variability across whole genome. In contrast to the study of Kharzinova et al. [20] present analyses of genetic diversity and relationship within cervids were based only on parameters that reflect the allele differences observed within each individual. The power to detect the genetic effects of polymorphic loci depends also on minor allele frequency. Previous studies showed that the occurrence of rare genotypes can cause contentious results therefore generally only SNPs with MAF higher than 0.01 are included [36]. The identified level of MAF was comparable with previous studies that similarly utilized the bovine genotyping array for genotyping of non-model but phylogenetically related species [19,37,38].
The analysis of genetic variability showed that most of the variations were conserved within each species and reflected their phylogenetic history. Generally, good level of observed heterozygosity was found across all species. In the Slovak Red deer population the rate of heterozygous was ascertained at 41% and it showed lower level of heterozygosity that observed in free living Red deer Ernst et al. (2011) [12] and Maršálková et al. (2014) [13] using microsatellite markers. Even if the trend of heterozygosity had a declining tendency, the obtained value of $F_{is}$ (0.04) can be still regarded as sufficient for the maintaining of genetic diversity within them.

All of the applied methods used for the estimation of the population structure were capable of identifying the genetic differentiation among individuals mainly at level of tribes. The main tree clusters reflected genetic background of species originating from tribe Cervini, Alceini, and Odocoileini. As expected the highest degree of intraspecific genetic similarity was found among species C. nippon, C. canadensis and C. elaphus that are originating from the genus Cervus. In addition the methods did not indicate the presence of any admixture level within the genus Cervus that can result from the history of Red deer in Slovakia. The methods identified clear separation of individuals into the group reflecting their phylogeny origin in accordance to previous published studies [1,39,40].

In the situation where many areas of Central Europe are threatened by the risk of loss of the autochthonous Red deer population due to the hybridisation with Sika deer the Slovak Red deer can be a source of a valuable gene pool within them. Across the analysed population of Red deer any remarkable loss of genetic diversity or degree of admixture signifying the potential risk of hybridisation with other species from the genus Cervus were found. In practical use these results can provide a genetic basis to the development of a deer farming system in Slovakia in respect to the conservation of genetic resources, sustainable management and prevention of genetic diversity loss.

Acknowledgement

This study was supported by the Slovak Research and Development Agency (Contract No. APVV-14-0054 and APVV-0636-11). We would like to thank for cooperation the Xcell Slovakia Breeding Services s.r.o. and Ing. J. Pokorádi, Phd. for providing of the biological samples of Red deer.

Authors’ contributions

RK carried out the molecular genetic studies, participated in the sequence alignment and drafted the manuscript and conceived of the study, and participated in its design and coordination and helped to draft the manuscript. NM conceived of the study, and participated in its design and coordination and helped to draft the manuscript. AT, ZK and OK conceived of the study, and participated in
its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES


23. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Skalp P, De Bakker PL, Daly MJ, Sham PC: PLINK: a tool set for whole genome association and population-based linkage analysis. Am J Hum Genet 2007, 81:559–575.


etskih odnosa između vrsta koje spadaju u rod *Cervus*, nije identificovala značajan gubitak genetičke varijabilnosti. Isto tako, nisu uočeni poremećaji koji bi bili posledica uzgoja jelenja u farmskim uslovima. Takođe, nisu uočene ni posledice unosa ili hibridizacije sa drugim vrstama roda *Cervus* (*C. canadensis* i *C. nippon*) što bi predstavljalo veliki rizik od gubitka autohtone populacije crvenog jelenja u mnogim regionima Evrope. Analiza individualnog nasleda je pokazala konzistentnost rezultata razlika po grupama koje su ispitivane što ukazuje na nizak stepen migracije između svih ispitivanih vrsta.