



## Original Research

## Runs of Homozygosity and Population History of Three Horse Breeds With Small Population Size

Gertrud Grilz-Seger<sup>a</sup>, Matjaz Mesarič<sup>b</sup>, Marko Cotman<sup>c</sup>, Markus Neuditschko<sup>d</sup>, Thomas Druml<sup>a,\*</sup>, Gottfried Brem<sup>a</sup><sup>a</sup> Institute of Animal Breeding and Genetics, Department for Biomedical Sciences, University of Veterinary Medicine Vienna, Vienna, Austria<sup>b</sup> Clinic for Reproduction and Large Animals, University of Ljubljana, Veterinary Faculty, Ljubljana, Slovenia<sup>c</sup> Institute for Preclinical Sciences, University of Ljubljana, Veterinary Faculty, Ljubljana, Slovenia<sup>d</sup> Agroscope, Swiss National Stud Farm, Avenches, Switzerland

## ARTICLE INFO

## Article history:

Received 19 August 2018

Received in revised form

12 September 2018

Accepted 13 September 2018

Available online 20 September 2018

## Keywords:

Inbreeding

Runs of homozygosity (ROH)

ROH island

Bosnian mountain horse

Posavje horse

Haflinger

SNP

## ABSTRACT

Long consecutive homozygous genotype segments, runs of homozygosity (ROH), are a result of parents transmitting identical haplotypes, which can be used to estimate autozygosity. Based on 612K single-nucleotide polymorphisms, we computed three ROH parameters (genome length covered by ROH,  $S_{ROH}$ ; number of ROH,  $N_{ROH}$ ; and autozygosity,  $F_{ROH}$ ) to investigate different scenarios in contemporary horse breeding: limited census (Bosnian mountain horse), conservation breeding (Posavje horse), and selection within closed studbook (Haflinger). The ROH parameters revealed well-defined differences between breeds.  $S_{ROH}$  was highest in the Bosnian mountain horse with 296.32 Mb, followed by the Haflinger sample ( $S_{ROH} = 270.35$  Mb) and the Posavje sample with 192.68 Mb. The highest number of ROH segments (ROHs) was observed within the Haflinger sample followed by the Posavje sample.  $F_{ROH}$  ranged at a population level from 8.59% in Posavje, over the Haflinger (mean  $F_{ROH} = 12.05\%$ ) to 13.21% in the Bosnian mountain horse breed. Bottlenecks were detected for Bosnian mountain horse and Haflinger, whereas for the Posavje, a positive effect of the conservation breeding program was documented. Investigating the distribution of ROHs across the genome, we detected four common ROH islands on equine chromosomes ECA 6, ECA 11, and ECA 17, which were present in all breeds. On breed level, the Bosnian mountain horses contained 10, the Posavje, four, and the Haflinger, 11 distinct ROH islands (containing the MC1R locus on ECA 3). With this analysis, we were able to compare genomic levels of inbreeding between breeds differing in management, pedigree completeness, and genes under selection.

© 2018 Elsevier Inc. All rights reserved.

## 1. Introduction

With the availability of genomewide single-nucleotide polymorphism (SNP) genotyping arrays, genomic inbreeding estimation based on the analysis of runs of homozygosity (ROH) has been established for several livestock species including the horse. A

range of studies have proven that the computation of ROH represents a valuable source of information to describe genomic inbreeding ( $F_{ROH}$ ) [1–6]. Furthermore, it was demonstrated that the identification of ROH segments (ROHs) can be successfully applied to investigate complex population histories and structures and to identify associations with quantitative and qualitative phenotypes [7–12]. Although methodical and technological aspects for the identification of ROH are not fully resolved [13], it is generally hypothesized that long consecutive homozygous segments are a result of identical haplotypes, which were inherited by a common ancestor [3,9,12,14,15].

Compared with pedigree-based inbreeding estimation, the analysis of ROHs provides a range of advantages including the following.  $F_{ROH}$  is the quotient of autozygous regions and total genome length; through this definition, it becomes feasible to derive reliable inbreeding coefficients from animals/populations

*Animal welfare/ethical statement:* The data used in this study (hair samples) were collected in the context of routine procedures during the studbook registration of horses by the Institute for Breeding and Health Care of Horses of the Veterinary Faculty, Ljubljana.

*Conflict of interest statement:* None of the authors has conflicts of interest.

\* Corresponding author at: Thomas Druml, Institute of Animal Breeding and Genetics, Department for Biomedical Sciences; University of Veterinary Medicine Vienna, Veterinärplatz 1, A-1210 Vienna, Austria.

E-mail address: [thomas.druml@vetmeduni.ac.at](mailto:thomas.druml@vetmeduni.ac.at) (T. Druml).

without pedigree records or incomplete pedigree information. In pedigree analyses, it is assumed that the founder population is unrelated. Runs of homozygosity segments of different length categories reflect to inbreeding events in different time frames, and thus ROH analyses enable a better comparison between fragmented populations and between breeds with different pedigree quality (generation equivalents can vary within horses from 19 in Lipizzans [16] to 3.5 and even less in local breeds, e.g., the Posavje horse [17]).

Based on the high-density SNP genotype information of the 700k Affymetrix Axiom Equine Genotyping Array, we aimed to investigate three different scenarios of contemporary horse breeding in three Slovenian horse populations (Fig.1), which are representatives for a population with extremely limited census, a conservation breeding program, and selection within closed stud-book. The Bosnian mountain horse is the oldest autochthonous horse breed of the entire Balkan Peninsula and underwent critical genetic bottlenecks within the period from 1991 to 2018. During these years, the population number decreased due to following reasons: the Yugoslav wars (from 1991 to 1999), the political and economic fragmentation of the original breeding regions, and the privatization and the shutdown of state stud farms in Bosnia and Macedonia. Although horses expressing the same type of the Bosnian mountain horse can be found in Bosnia and Herzegovina, Serbia, Macedonia, and Albania, the number of purebred and registered animals comprises 141 animals and 14 animals in the B-book. The Slovenian Bosnian mountain horse population originates from the former state stud farm Borike in Bosnia and Herzegovina. To increase the effective population size, several Bosnian mountain horses from the country breeding, which matched in type and conformation, were integrated into the Slovenian Bosnian mountain horse population. For those horses, no pedigree documentation exists [18]. The second horse population included in this study is the Slovenian Posavje horse, which belongs to the Posavina breed, originating in the northern Sava flatlands of Croatia. The breeding region also includes the southern part of Slovenia (especially the districts Krsko and Brezice). In 1993, a Slovenian breeding and conservation program was initiated. This was the commencement of pedigree documentation; whereas in the past, only breeding stallions were registered. In contrast to the Bosnian mountain horse, which is threatened by extinction, the Posavje population in Slovenia experiences a period of expansion, as the population increased from 79 breeding mares in 1993 to 627 breeding mares in 2017 [17,19]. Finally we included a sample of the Slovenian Haflinger horse in this study. The Slovenian Haflinger population was established in the 1960s based on imported horses from Austria, and the population constantly increased up to the 1990s. Within the last 2 decades, this breed was exposed to intense selection pressure according to modernized breeding objective including the introgression of foreign stallions. Furthermore, a decrease of the breeding population can be observed in the

Slovenian Haflinger, as the new management decisions do not correspond with the attitudes of traditional Slovenian breeders. The current Slovenian Haflinger population comprises in total 680 registered animals, including 25 breeding stallions and 200 breeding mares.

The aim of this study was to characterize genomic levels of inbreeding based on ROH analysis and to investigate the ROH structure and distribution to analyze population history and to evaluate differences in breeding and selection programs. Furthermore, we identified overlapping ROH regions (ROH islands) within the respective breeds.

## 2. Materials and Methods

### 2.1. Sampling

All horses included in this study were selected to represent the genealogical population structure of the breeds based on pedigree information. From the 23 systematically selected Bosnian mountain horses (BMH), 18 were offspring from the Bosnian stud farm Borike, whereas five horses are the offspring of country-bred mares with missing pedigree information. The 28 Posavje horses (POS) and the 18 Haflinger horses (HAF) were sampled according to the genealogical structure (sire lines and mare families) of the two breeds, taken into account the relatedness of animals.

### 2.2. Single-Nucleotide Polymorphism Genotyping

The SNP genotypes for the 69 horses were determined using the Affymetrix Axiom Equine Genotyping Array [20]. The chromosomal position of the SNPs was determined based on EquCab2.0 reference genome. We did not consider the SNPs positioned on the sex chromosomes (X: 28,017 SNPs and Y: 1 SNP) and the SNPs without known chromosomal position (30,864 SNPs). The SNPs with more than 10% missing genotypes were excluded. This resulted in a total of 611,914 SNPs for each horse.

### 2.3. Genetic Diversity and Runs of Homozygosity Analysis

To illustrate the population structure, we applied principal component analysis (PCA) on the basis of the genetic relationship matrix (G) with pairwise identities by state between horses as provided by PLINK v.1.7 [21]. The PCA plot was performed using the R platform and statistical analyses, and graphical representations were performed using the software packages SAS v.9.1 [22] and R ([www.r-project.org](http://www.r-project.org)).

Runs of homozygosity segments were determined with an overlapping window approach implemented in PLINK v.1.7 [21] based on the following settings: minimum SNP density was set to one SNP per 50 kb, with a maximum gap length of 100 kb. The final



Fig. 1. Representative horses for the three breeds Bosnian mountain horse (A), Posavje horse (B), and the Slovenian Haflinger (C) (images by Matjaz Mesarič).

segments were called ROH if the minimum length of the homozygous segment was greater than 500 kb and comprised more than 80 homozygous SNPs, whereas one heterozygote and two missing genotypes were permitted within each segment.

The total number of ROH ( $N_{ROH}$ ), average length of ROH ( $L_{ROH}$ ), and the sum of all ROH segments ( $S_{ROH}$ ) of each horse were summarized according to breed and ROH length category. To analyze the ROH distribution, ROH segments were divided into the following seven length classes: 0.5–1, 1–2, >2–4, >4–6, >6–8, >8–10, and >10 Mb. The genomic inbreeding coefficients ( $F_{ROH}$ ) were calculated following the method described in [1]:

$$F_{ROH} = \sum \frac{L_{ROH}}{L_{AUTO}}$$

where the length of the autosomal genome ( $L_{AUTO}$ ) was set to 2,243 GB.

The distribution of ROHs across the genome was visualized using the R-package to detect ROHs ([www.r-project.org](http://www.r-project.org)). Putative ROH islands were determined based on overlapping homozygous regions within more than 50% of the horses. The map viewer of the equine ensemble database Equcab2 was used to identify genes located in ROH regions, available at [www.ensemble.org](http://www.ensemble.org). For the determination of Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of identified genes, we used the open source Database for Annotation, Visualization, and Integrated Discovery v.6.8 package—available at <https://david.ncifcrf.gov>. For the final analysis, we used the Equus caballus annotation file and a significance threshold of  $P > .05$ .

### 3. Results

Principal component analysis of the SNP data revealed a well-structured picture regarding the three different investigated breeds, whereas within the Bosnian mountain horse sample a higher variation was observed (Fig. 2). The first PC explaining 33% of the genetic variation separated the Bosnian mountain horses from the Posavje and Haflinger samples, PC2 (explaining 11%) differentiated the Posavje from the Haflinger samples, whereas PC3

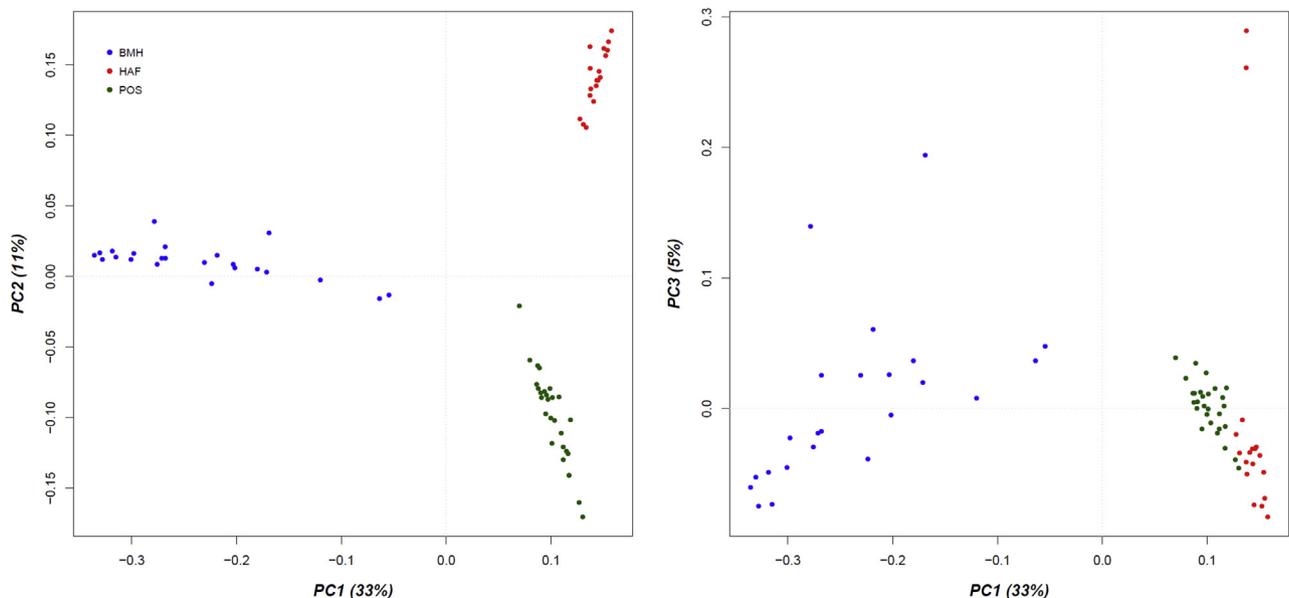
**Table 1**

Mean values, standard deviation (SE), and minimum and maximum values for the variables  $S_{ROH}$  (in Mb),  $N_{ROH}$ ,  $L_{ROH}$  (in Mb), and  $F_{ROH}$  for the samples Bosnian mountain horse, Haflinger, and Posavje.

Bosnian Mountain Horse	n	Mean	SE	Min.	Max.
$S_{ROH}$	23	296.32	201.06	42.06	913.02
$N_{ROH}$	23	137.70	36.37	61.00	213.00
$L_{ROH}$	23	1.98	0.99	0.69	4.37
$F_{ROH}$	23	0.13	0.09	0.02	0.41
Slovenian Haflinger					
$S_{ROH}$	18	270.36	113.15	1.59	447.73
$N_{ROH}$	18	155.56	57.69	3.00	213.00
$L_{ROH}$	18	1.61	0.45	0.53	2.28
$F_{ROH}$	18	0.12	0.05	0.001	0.20
Posavje Horse					
$S_{ROH}$	28	192.68	54.56	116.60	348.77
$N_{ROH}$	28	152.25	17.73	117.00	185.00
$L_{ROH}$	28	1.25	0.26	0.93	1.94
$F_{ROH}$	28	0.09	0.02	0.05	0.16

(explaining 5%) did not contribute to breed separation but highlighted outbred animals according to studbook information. Runs of homozygosity analysis of the three samples resulted in population-specific metrics. The mean genome length covered by ROH ( $S_{ROH}$ ) was highest in the Bosnian mountain horse with 296.32 Mb,  $\pm 201.06$ , followed by the Haflinger sample ( $S_{ROH} = 270.35$  Mb,  $\pm 113.15$ ) and was lowest in the Posavje sample with 192.68 Mb,  $\pm 54.56$ . The highest number of ROHs ( $N_{ROH}$ ) was observed in the Haflinger sample with  $155.56 \pm 57.69$  segments followed by the Posavje sample, where on average  $152.25 \pm 17.72$  segments were counted. The sample of Bosnian mountain horses was characterized by the lowest number of ROHs ( $N_{ROH} = 137.70 \pm 36.37$ ). Mean values, standard deviation, and maxima and minima values for the parameters  $S_{ROH}$ ,  $N_{ROH}$ , and  $L_{ROH}$  of the three samples are given in Table 1. The highest variance regarding the distribution of the parameters  $S_{ROH}$  and  $N_{ROH}$  was observed in the Bosnian mountain horse sample, whereas in the Posavje sample, less variance of  $S_{ROH}$  occurred (Fig. 3).

The plot of  $S_{ROH}$  versus  $N_{ROH}$  (Fig. 4) represents a clear structure between the breeds and describes population history. The Posavje cluster (green dots) is placed along the diagonal at a medium  $S_{ROH}$



**Fig. 2.** Visualization of the data set on the first three principal components (PCs). PC1 explains 33%, PC2, 11%, and PC3, 5% of the variation (blue dots = Bosnian mountain horses [BMH]; red dots = Haflinger [HAF]; green dots = Posavje [POS]).

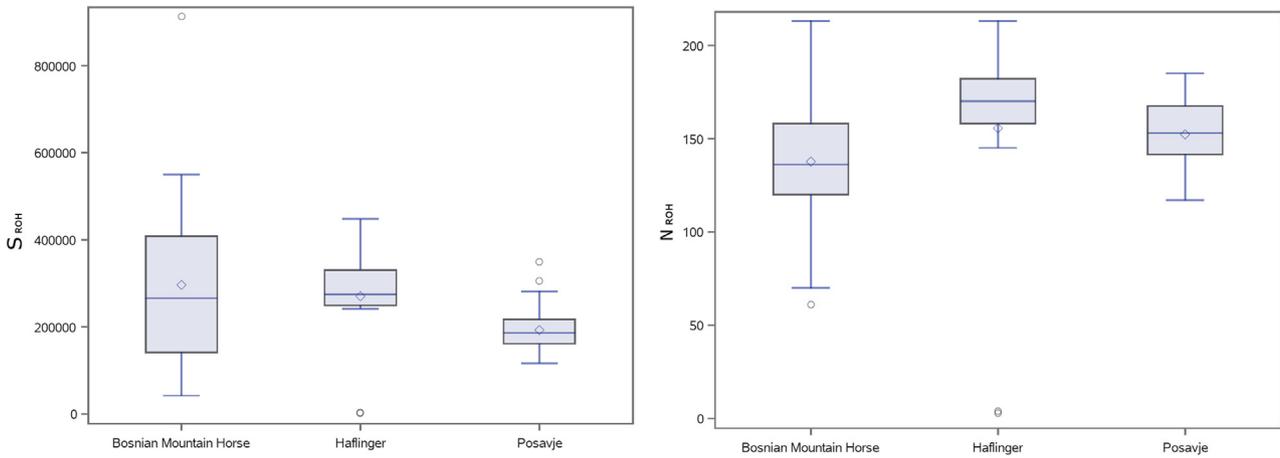


Fig. 3. Box plots to illustrate the distribution of  $S_{ROH}$  (left side) and  $N_{ROH}$  (right side) within the samples Bosnian mountain horses, Haflinger, and Posavje horses.

level, whereas the Haflinger sample (red dots) has an upward shift toward higher  $N_{ROH}$  and a right shift toward higher  $S_{ROH}$ . The Bosnian sample (blue dots) is characterized by a high variability and simultaneously by a strong right shift toward higher  $S_{ROH}$  (sub-cluster including individuals with extreme long genomewide ROH). The second Bosnian cluster includes horses with the lowest  $N_{ROH}$  and  $S_{ROH}$  values.

From the distribution of the single ROH length classes within the three samples, distinct differences could be detected within the medium and high categories (Fig. 5). Although within the category of small ROHs (0.5–1 Mb), all breeds did not show essential deviations, the Posavje sample had the highest proportions of ROHs shorter than 4 Mb (80.6%), followed by the Haflinger (65.8%) and the Bosnian mountain horse (57.2%). In the category of ROHs >10 Mb Bosnian horses had the highest proportion (21.2%), followed by the Haflinger (10.3%) and the Posavje (4.6%) samples. Three Bosnian mountain horses showed a maximum coverage of ROHs longer than 10 Mb, which ranged from 45.4% to 48.4%.

Genomic inbreeding expressed by  $F_{ROH}$  ranged at population level from 8.59% ( $\pm 2.43\%$ ) in the Posavje horse, over the Haflinger (mean  $F_{ROH} = 12.05\%$ ;  $\pm 5.04\%$ ), to 13.21% ( $\pm 8.93\%$ ) within the Bosnian mountain horse breed. At individual level, the highest

inbreeding coefficients were identified for four Bosnian mountain mares ( $F_{ROH}$ : 40.71%; 24.5%; 22.5%; 22.1%). More recent inbreeding ( $F_{ROH}$  based on segments >10 Mb) in those four animals amounted to 27.9%, 14.4%, 13.6%, and 13.7%, respectively, indicating closer consanguinity within the last 4–5 generations. The lowest  $F_{ROH}$  values were documented for two Haflingers ( $F_{ROH}$  of 0.07% and 0.11%) and five outbred Bosnian mountain horses ( $F_{ROH}$  ranging from 1.9% to 4.0%). All these animals lacked ROHs longer than 6 Mb. The distribution and the breed overlap of  $F_{ROH}$  values are illustrated in Fig. 6.

We identified three equine chromosomes (ECA 6, ECA 11, ECA 17), which contained four ROH islands that were shared by more than 50% of individuals in the total sample (Fig. 7). An overview of the known genes in these regions is given in Table 2, for example, *HLF* (ECA 11) or *ERC1* (ECA 6) both playing a role as transcription factors; *RAD52* (ECA 6) a DNA repair protein; *B4GALNT3* (ECA 6) a catalyst of galactose transfer; *COX11* (ECA 11) involved in oxidative phosphorylation process [23].

Applying the aforementioned threshold  $\geq 50\%$  on a breed level, we detected differences in the number and localization of ROH islands. Within the Posavje sample, we identified the lowest number of ROH islands ( $n = 4$ ) located on ECA 6, ECA 8, ECA 9, and ECA 28. On average, these ROH islands were shared by 50% up to 57% of horses (Supplementary Table 1). Close to the defined threshold (ROH frequency = 46, 43%), an island on ECA 11 contained several genes of the homeobox B cluster (*HOXB*

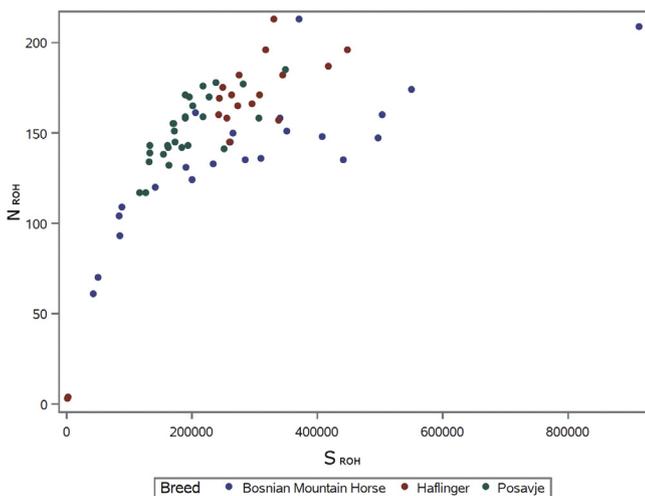


Fig. 4. Scatter plot of the number of ROH segments ( $N_{ROH}$ ) and genomewide sum of ROH in kb ( $S_{ROH}$ ). Each individual is marked as dot (blue = Bosnian mountain horse, green = Posavje; red = Haflinger).

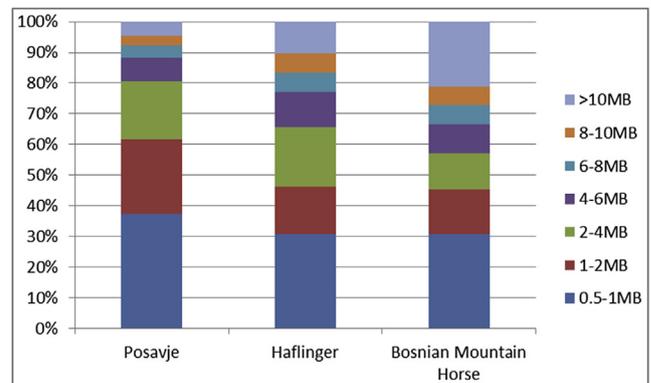
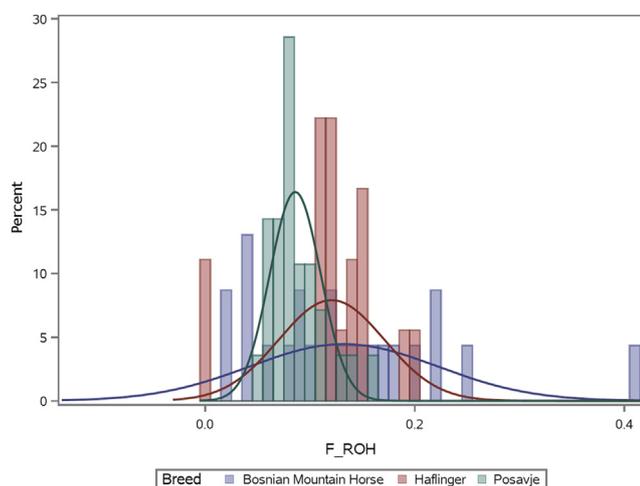


Fig. 5. Distribution of ROH segments in different length classes (0.5–1, >1–2, >2–4, >4–6, >6–8, >8–10, and >10 Mb) within the three samples Bosnian mountain, Haflinger, and Posavje horse.



**Fig. 6.** Distribution of  $F_{ROH}$  values for the three breeds Bosnian mountain horse (blue), Haflinger (red), and Posavje (green).

1,2,3,5,6,7,8,13), a highly conserved family of transcription factors that are involved in morphogenesis. Gene Ontology analysis highlighted the homeobox B cluster associated with biological processes and molecular functions. The highest significance for enrichment was found for *HOXB3*, *HOXB1*, *HOXB2*, *HOXB7*, *HOXB8*, *HOXB5*, and *HOXB6* genes, which are involved in embryonic skeletal system morphogenesis (GO:0048704) (Supplementary Table 2).

In the Bosnian mountain horse sample, 10 ROH islands on ECA 1, ECA 3, ECA 6, ECA 11, ECA 18, and ECA 23 were present, which were shared by 52%–70% of individuals (Supplementary Table 1). On ECA 11, four islands were found. Of 36 identified genes, GO analysis underlined two genes (*MMD*, *VLDLR*) related to biological processes. Slightly below the level of significance ( $P < .066$ ), additional two genes (*INHBB*, *RALB*) that play a role in cellular response to starvation (GO:0009267) could be outpointed (Supplementary Table 2).

The Haflinger sample had the highest number of ROH islands ( $n = 11$ ) located on ECA 3, ECA 6, ECA 8, ECA 9, ECA 11, ECA 15, ECA 17, and ECA 26. Most islands were located on chromosome 3 (4 islands), and three of them were characterized by long stretches (2179 kb, 1123 kb, and 1365 kb). The ROH island on ECA 3 position 35,305,705–36,428,488 was shared by 85% of individuals and contained the *MC1R* locus (Fig. 8). This island also harbored a ROH region of 688.5 kb, which was present in 89% of individuals (excluding the two outbred Haflingers). The following known genes are located in this region: *CBFA2T3*, *ACSF3*, *CDH15*, *SPG7*, *RPL13*, *CPNE7*, and *DPEP1*. Overall, the proportion of Haflinger horses sharing this island was high, as in five cases, the percentage was higher than 66%. Of the 87 identified genes, GO analysis underlined two enriched terms of molecular function. Two genes (*SLC9B1*, *SLC9B2*) were involved in solute:proton antiporter activity (GO:0015299) and the two genes *CNDP1*, *DPEP1* were related to metalloproteinase activity (GO:0070573). Two KEGG pathways were significantly enriched ( $P < .05$ ) containing the genes *PABPN1L*, *NFKB1*, *IFNGR2*, and *IFNAR1* (ecb05164/Influenza A); and *LAPTM4B*, *GALNS*, *MANBA*, and *IDUA* (ecb04142/lysosome) (Supplementary Table 2).

#### 4. Discussion

From the three investigated horse populations, two breeds (Bosnian mountain horse and Posavje) are considered endangered, whereas the Bosnian mountain horse exhibits an extreme limited

consensus with 141 purebred and registered horses worldwide. In the present study, the parameters  $S_{ROH}$ ,  $N_{ROH}$ , and  $L_{ROH}$  revealed well-defined and gradual differences between the breeds. In literature,  $S_{ROH}$  is described as the best descriptor for discrimination [3,9,12,24]. In our case, all the three parameters ( $S_{ROH}$ ,  $L_{ROH}$ , and  $N_{ROH}$ ) contained useful information for the differentiation purpose, able to describe distinct breeding management scenarios: (1) population undergoing a bottleneck, (2) conservation breeding, and (3) selection within closed studbook.

According to our expectations, the Bosnian mountain horse sample showed the highest  $S_{ROH}/F_{ROH}$  values and simultaneously the longest ranging ROHs (>10 Mb). All those parameters thus indicate bottleneck effects that are due to the Bosnian war in the 1990s connected with ongoing consanguineous mating in a small population. The relation of  $S_{ROH}$  and  $N_{ROH}$  (right shift, higher variation, comp. Fig.4) correspond with the interpretation given by Ceballos et al. [14] for bottlenecked and consanguineous populations. A subcluster in the Bosnian mountain horse characterized by low  $S_{ROH}$  and low  $N_{ROH}$  can be explained by outbreeding of five purebred Bosnian mountain horses, which are country-bred mares with unknown pedigree.

A completely different scenario was found in the Posavje sample. Although the founder population of the Slovenian Posavje with 79 mares and six stallions registered in 1993 can be considered small, the horses analyzed in the present study were characterized by the lowest mean genome length covered by ROH ( $S_{ROH} = 191$  Mb), the lowest  $F_{ROH}$  of 8.6% and the lowest mean  $L_{ROH}$  of 1246 kb, and by a general smaller variation of all ROH parameters. These values are similar to those reported for the Noriker horse ( $S_{ROH} = 228$  Mb,  $L_{ROH} = 1378$  kb), an Austrian draft horse, which is bred on a broad genetic base in terms of gene pool diversity and effective population size [5,25]. Considering the different length classes of ROH, the Posavje horses showed the lowest proportion of ROHs greater than 6 Mb, indicating a relative lack of recent inbreeding. The ROH profile (comp. Fig.4) of the Posavje is an example for a small population with few consanguinity and admixture [14]. The Slovenian Posavje breed has been managed since 1993 following the principles of conservation genetics: narrow relation of stallions to mares (1 sire to 7.8 breeding mares), introgression of Croatian Posavina stallions (16 imported sires from 1993 to 2014), balanced breeding by the use of different stallions (mean number of offspring per stallion is 8) and moderate selection. Our results of the ROH analysis confirm the possibility to maintain genetic diversity in a small breeding population, if optimized conservation strategies are applied and population expansion occurs.

The Slovenian Haflinger, belonging to the internationally recognized Haflinger breed, showed a similar high mean  $S_{ROH}$  compared with the Bosnian mountain horse. In addition, the distribution of ROHs in different length classes matched that of the Bosnian mountain horse; the only difference was found in the frequency of ROHs longer than 10 Mb, where the Bosnian mountain horse reached outstanding values. According to the ROH profile plot (Fig.4), Slovenian Haflinger horses were characterized by a right shift and a tendency toward higher number of ROHs, which can be interpreted as the effect of bottleneck situations [14]. The values presented for this sample ( $S_{ROH} = 270.4$ ;  $N_{ROH} = 155.6$ ) were slightly different to the Austrian ( $S_{ROH} = 282.1$  Mb;  $N_{ROH} = 208.5$ ), and the Italian Haflinger ( $S_{ROH} = 316.7$  Mb;  $N_{ROH} = 188.3$ ) populations, from which the Italian Haflinger exhibited strong recent consanguinity [5]. Two Slovenian Haflinger horses (outbred animals from Austrian-bred sires) were characterized by extremely low  $S_{ROH}$  (1.6 Mb; 2.6 Mb)/ $F_{ROH}$  (0.07%; 0.11%) and a complete lack of ROHs longer than 1 Mb. This admixture effect has previously been reported from pedigree analysis in the Austrian Haflinger population [26]. In this study, it could be demonstrated that in

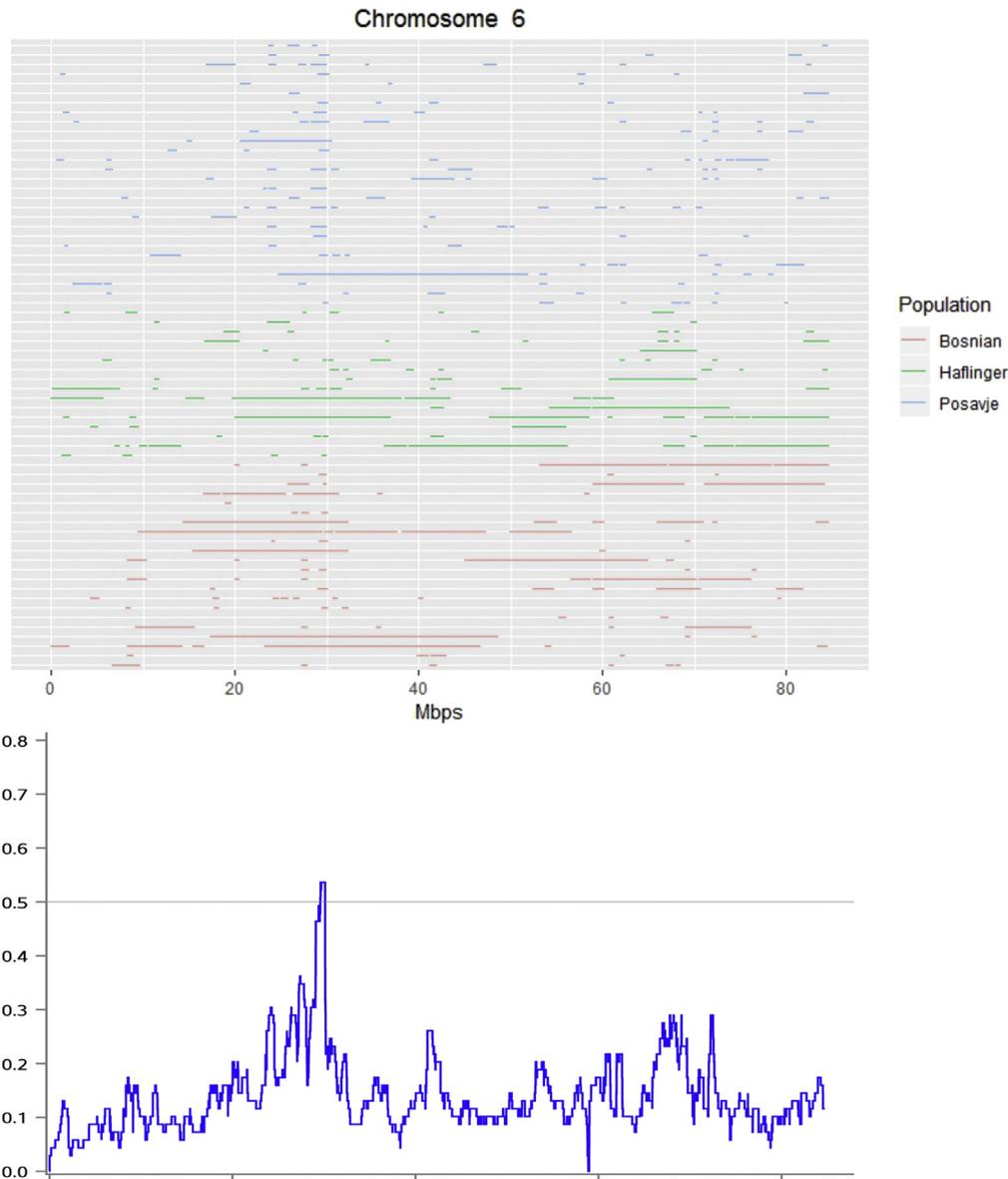


Fig. 7. Visualization of ROHs per individual on chromosome 6 (above), below the percentage of individuals sharing ROH per SNP position.

certain cases, the common ancestors were found more than 5 generations ago, resulting in zero recent inbreeding (5 generation pedigree) and moderate inbreeding on a long-term scale (pedigrees comprising 10 to 24 generations). This effect presumes a genetic drift within the national subpopulations (e.g., Italy, Austria, Slovenia). The result of genetic drift—a separation between the Austrian and Italian Haflinger—could also be illustrated by a combined network and admixture analysis in the study of Druml et al. [5]. The Haflinger, a transnational population comprising more than 20,000 breeding animals worldwide, illustrates in

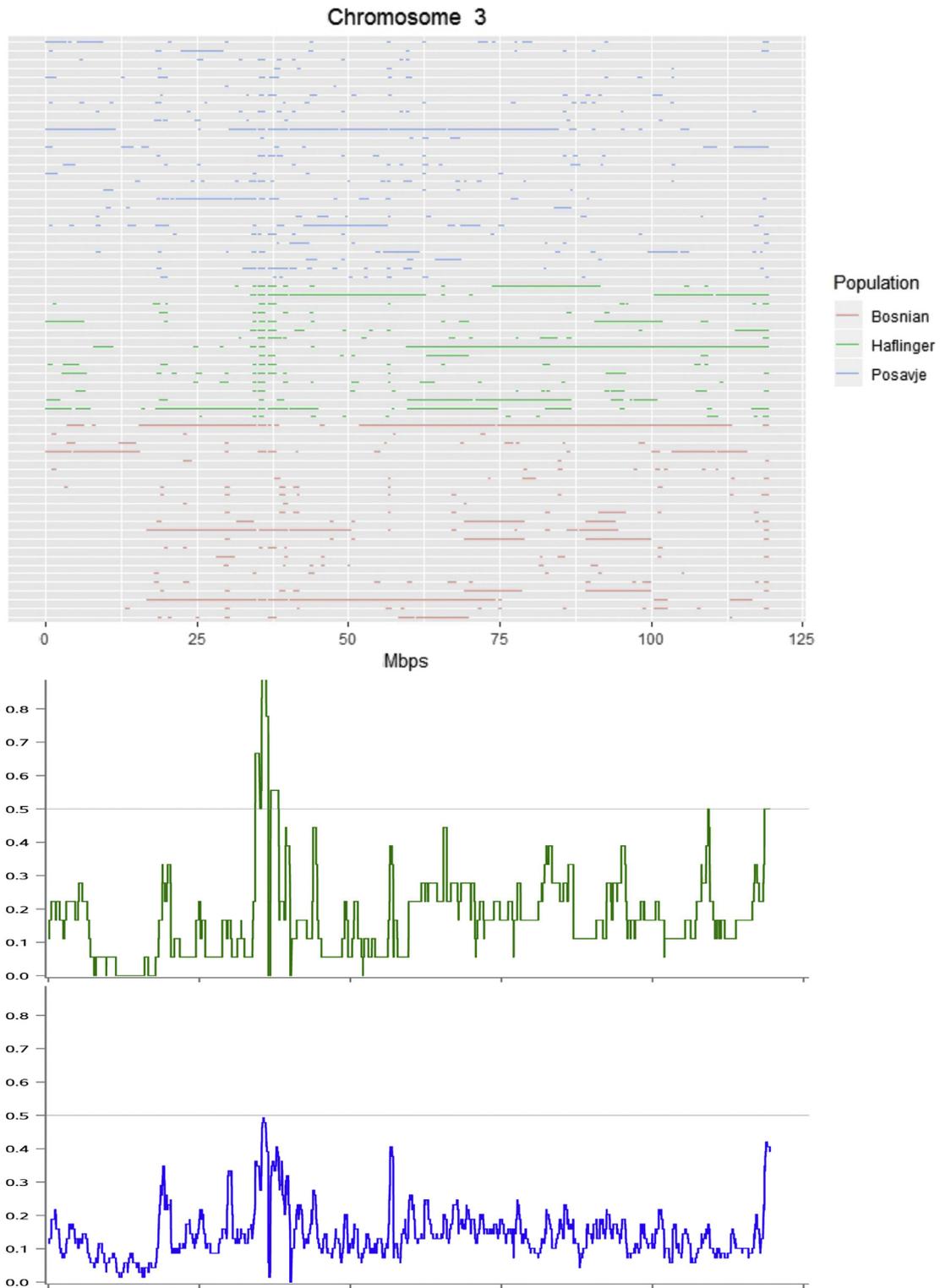
matters of ROH parameters a reduced genetic diversity, which is also the case for the Slovenian population.

Although ROHs commonly are distributed unequally over the genome, ROH islands are defined as relative short homozygous runs in a specific region shared among many individuals in one or more populations [15]. There are two hypotheses for the formation of ROH islands: one assumption is that they arise in regions of low recombination rate with a higher probability of accumulation of similar haplotypes; the second explanation is given by the driving forces of positive selection resulting in increased homozygosity

Table 2

Runs of homozygosity islands that were shared by more than 50% of individuals (ROH freq) in the total sample.

Chr.	Begin	End	Length (kb)	ROH freq.	Island Shared by n Animals/Breed	Known Genes
6	29,599.757	30,119.306	519.609	0,534	16 POS, 13 BMH, 6 HAF	<i>ERC1, RAD52, WNK1, NINJ2, B4GALNT3</i>
11	29,528.525	29,733.365	204.840	0,507	6 POS, 16 BMH, 11 HAF	<i>COX11, STXBP4</i>
11	29,831.944	30,078.455	246.511	0,510	7 POS; 16 BMH, 12 HAF	<i>HLF, MMD</i>
17	18,245.585	18,548.489	302.904	0,507	10 POS; 11 BMH, 12 HAF	—



**Fig. 8.** Visualization of ROHs per individual on chromosome 3 (above), below the percentage of individuals sharing ROH per SNP position. The highly conserved area around position 34–36 Mb in the Haflinger breed (green) includes 67–85% of animals and also the *MC1R* locus at ECA 3, position 36,259,276–36,260,354. The percentage of individuals sharing ROH per SNP position in the total sample is marked as blue line.

around the target locus (selective sweep) [15]. From our three studied breeds, selection pressure in Posavje and Bosnian mountain horse was moderate, whereas the Haflinger underwent a strong selection concerning height, caliber, weight, conformation, coat color, and muscle mass in the last 50 years. In our data, the

Haflinger exhibited not only the highest number of ROH islands, but also the longest islands exceeding the size of 2 Mb. An accumulation of islands on ECA 3, which among others contained the *MC1R* locus responsible for the chestnut coat color, was present in 55–85% of the individuals. In terms of ROH island distribution, we

were able to show that within Bosnian mountain horse and Haflinger, both affected by bottlenecks and higher autozygosity level, an increased number of islands were present that were shared by a higher percentage of animals compared with the Posavje sample. Gene Ontology analysis highlighted the homeobox B cluster in the Posavje horse. One group of these genes is involved into embryonic skeletal system morphogenesis. The Posavje horse is phenotypically characterized by an enormous muscle mass in relation to body size and by a pronounced sexual dimorphism. In the case of the Bosnian mountain horse, which traditionally has been kept semi feral under harsh environmental conditions, the GO analysis indicated an enrichment of genes involved into cellular response on deprivation of nourishment, indicating an environmental adaptation. These results provide novel insights into genomic areas that might be of relevance for further genomic studies and for conservation breeding programs.

## 5. Conclusions

In this study, the analysis of ROHs enabled us to compare genomic levels of inbreeding between breeds differing in management and pedigree completeness. Especially in fragmented populations, ROH analysis is capable to produce valuable information for designing optimal breeding and management decisions as ROH parameters facilitate a fine calibrated resolution on individual level. Furthermore, the ROH evaluation of the conservation breeding program of the Posavje horse revealed that the current applied conservation strategy successfully avoided inbreeding. The distinct differences between the number and distribution of the ROH island in the three breeds suggest that ROH can be applied to identify signatures of selection.

## Acknowledgment

Authors' contributions: G.S.G. contributed to conceptualization and writing original draft; D.T. contributed to data curation, formal analysis, and visualization; N.M. contributed to data curation, formal analysis, and visualization; M.M. and C.M. contributed to conceptualization and sampling; B.G. contributed to funding acquisition and project administration.

Financial disclosure: This work was supported by the Austrian Research Promotion Agency (FFG) [Contract number 843464] and Slovenian Research Agency program P4-0053 to M. Cotman.

## Supplementary Data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.jevs.2018.09.004>.

## References

- [1] McQuillan R, Leutenegger A, Abdel-Rahman R, et al. Runs of homozygosity in European populations. *Am J Hum Genet* 2008;83:359–72.
- [2] Ferencakovic M, Hamzic E, Gredler B, Solberg TR, Klemetsdal G, Curik I, Sölkner J. Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *J Anim Breed Genet* 2013;130:286–93.
- [3] Marras G, Gaspa G, Sorbolini S, Dimauro C, Ajmone-Marsan P, Valentini A, Williams JL, Macciotta NP. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. *Anim Genet* 2014;46:110–21.
- [4] Kamiński S, Hering DM, Jaworski Z, Zabolowicz T, Ruś A. Assessment of genomic inbreeding in Polish Konik horses. *Pol J Vet Sci* 2007;20:603–5.
- [5] Druml T, Neuditschko M, Grilz-Seger G, Horna M, Ricard A, Mesarić M, Cotman M, Pausch H, Brem G. Population networks associated with runs of homozygosity reveal new insights into the breeding history of the haflinger horse. *J Hered* 2018;109:384–92.
- [6] Lemes RB, Nunes K, Carnavall JEP, Kimura L, Mingroni-Netto RC, Meyer D, et al. Inbreeding estimates in human populations: applying new approaches to an admixed Brazilian isolate. *PLoS One* 2018;13:e0196360.
- [7] Metzger J, Karwath M, Tonda R, Beltran S, Águeda L, Gut M, Gut LG, Distl O. Runs of homozygosity reveal signatures of positive selection for reproduction traits in breed and non-breed horses. *BMC Genomics* 2015;16:764.
- [8] Kim E-S, Cole JB, Huson H, Wiggins GR, VanTassell CP, Crooker BA, Liu G, Da Y, Sonstegard TS. Effect of artificial selection on runs of homozygosity in US Holstein cattle. *PLoS One* 2013;8:e80813.
- [9] Kirin M, McQuillan R, Franklin CS, Campbell H, McKeigue PM, et al. Genomic Runs of Homozygosity Record Population History and Consanguinity. *PLoS One* 2010;5:e13996.
- [10] Lencz T, Lambert C, DeRosse P, Burdick KE, Morgan TV, Kane JM, Kucherlapati R, Malhotra AK. Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. *Proc Natl Acad Sci U S A* 2007;104:19942–7.
- [11] Manunza A, Noce A, Serradilla JM, Goyache F, Martínez A. A genome-wide perspective about the diversity and demographic history of seven Spanish goat breeds. *Genet Sel Evol* 2016;48:52–61.
- [12] Purfield DC, Berry DP, McParland S, Bradley DG. Runs of homozygosity and population history in cattle. *BMC Genet* 2012;70/13:1471–2156.
- [13] Nandolo W, Utsunomiya YT, Meszaros G, Wurzinger M, Khayadzadeh N, et al. Misidentification of runs of homozygosity islands in cattle caused by interference with copy number variation or large inter-marker distances. *Genet Sel Evol* 2018;50:43–57.
- [14] Ceballos FC, Joshi PK, Clark DW, Ramsay M, Wilson J. Runs of homozygosity: window into population history and trait architecture. *Nat Rev Genet* 2018;19:220–34.
- [15] Pemperton TJ, Absher D, Feldman MW, Myers RM, Rosenberg NA, Li JZ. Genomic patterns of homozygosity in worldwide human populations. *Am J Hum Genet* 2012;91:275–92.
- [16] Zechner P, Sölkner J, Bodo I, Druml T, Baumung R, Achmann R, Marti E, Habe F, Brem G. Analysis of diversity of diversity and population structure in the Lipizzan horse breed based on pedigree information. *Livst Prod Sci* 2002;77:137–46.
- [17] Malovrh S, Flisar T, Slavič J, Potočnik K, Mesarič M, Rus R, Kovač M. Tri slovenske lokalne pasme konj. In: Domžale Malovrh S, editor. GENETSKA pestrost na osnovi porekla pri slovenskih lokalnih pasmah do vključno leta 2016. Domžale, Slovenia: Biotehniška fakulteta, Oddelek za zootehniko; 2017. p. 87–104.
- [18] Mesarič M, Dolinšek A, Žiga E. The stud book of Bosnian Mountain Horse. Trbovlje, Slovenia: International Association of Bosnian Mountain Horse breeders; 2017.
- [19] Mesarič M, Rus J. Rodoviska knjiga posavskega konja. Ljubljana, Slovenia: Univerza v Ljubljani, Veterinarska fakulteta; 2014.
- [20] Schaefer JR, Schubert M, Bailey E, Bannasch DL, Barrey E, Kahila Bar-Gal G, Brem G, Brooks SA, Distel O, Fries R, et al. Developing a 670k genotyping array to tag ~2M SNPs across 24 horse breeds. *BMC Genet* 2017;18:565.
- [21] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR. PLINK: a toolset for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–75.
- [22] SAS Institute. SAS version 9.1. Cary, NC: SAS Institute Inc; 2009.
- [23] Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, et al. Ensembl 2018. *Nucleic Acids Res* 2018;46:D754–61.
- [24] Sillio L, Rodrigucz M, Fernandez A, Barragan C, Benitez R, Ovilo C, Fernandez A. Measuring inbreeding and inbreeding depression on pig growth from pedigree or SNP-derived metrics. *J Anim Breed Genet* 2013;140:349–60.
- [25] Druml T, Baumung R, Sölkner J. Pedigree analysis in the Austrian Noriker draught horse: genetic diversity and the impact of breeding for coat colour on population structure. *J Anim Breed Genet* 2009;126:348–56.
- [26] Druml T, Sauer K, Elsbacher J, Grilz-Seger G, Brem G. Analyse des Genpools, der genetischen Diversität und der Inzuchtverhältnisse der österreichischen Haflingerpopulation. *Züchtungskunde* 2016;88:379–94.