Panel of informative SNP markers for two genetic lines of European bison: Lowland and Lowland–Caucasian

M. Wojciechowska, Z. Nowak, A. Gurgul, W. Olech, W. Drobik & T. Szmatoła

Wojciechowska, M., Nowak, Z., Gurgul, A., Olech, W., Drobik, W. & Szmatoła, T., 2017. Panel of informative SNP markers for two genetic lines of European bison: Lowland and Lowland–Caucasian. *Animal Biodiversity and Conservation*, 40.1: 17–25.

Abstract

Panel of informative SNP markers for two genetic lines of European bison: Lowland and Lowland–Caucasian.— As the result of a population bottleneck, the present population of European bison shows a high level of inbreeding, and a significant loss of genetic variability. In studies on such specific species there is a need to apply methods that obtain as much information about the genome as possible in a short time. The aim of the study was to define a panel of SNP (single nucleotide polymorphism) markers that could serve in genetic diversity analysis of European bison from two lines: Lowland (LB) and Lowland–Caucasian (LC). The study used 57 individuals from the LB line and 72 from the LC line. To identify well–performing SNPs in European bison, we used two microarrays with different markers densities: BovineSNP50 v2 BeadChip and BovineHD BeadChip. As a result of the adopted criteria, 1,421 and 22,122 markers, respectively, were selected. On the basis of statistical analysis (allele frequencies, Fisher's exact test, and the Z–test), a panel of 1,536 informative SNP markers was ultimately selected for further study; 26 of these with private alleles for the LB line and 611 with private alleles for the LC line. The data obtained in this study could further enrich and support breeding programs in the context of relatedness between particular specimens and herds from captive breeding centres.

Key words: *Bison bonasus*, European bison, Genetic marker, Microarray, Single nucleotide polymorphism

Resumen

Conjunto de marcadores de tipo PSN para dos líneas genéticas de bisonte europeo: Lowland y Lowland–Caucasiana.— Debido al cuello de botella demográfico, la población actual de bisonte europeo muestra un elevado grado de endogamia y una pérdida significativa de variabilidad genética. Es necesario que en los estudios realizados con estas especies específicas se apliquen métodos que permitan obtener tanta información genómica como sea posible en un tiempo breve. El objetivo de este estudio era definir un conjunto de marcadores de tipo PSN (polimorfismo de un solo nucleótido) que pudiera servir para analizar la diversidad genética de las dos líneas de bisontes europeos: Lowland (LB) y Lowland–Caucasiana (LC). En el estudio se analizaron 57 individuos de la línea LB y 72 de la línea LC. Para caracterizar bien el rendimiento de los PSN en el bisonte europeo, se usaron dos micromatrices multigénicas (genochip) con densidades diferentes de marcadores: BovineSNP50 v2 BeadChip y BovineHD BeadChip. Como consecuencia de los criterios adoptados, se seleccionaron 1.421 y 22.122 marcadores, respectivamente. Sobre la base del análisis estadístico (frecuencias alélicas, prueba exacta de Fisher y prueba Z), en última instancia se seleccionó un conjunto de 1.536 marcadores informativos de PSN para los estudios adicionales, 26 de los cuales tienen alelos privados para LB y 611, para la línea LC. La información obtenida en este estudio podría enriquecer aún más y apoyar a los programas de reproducción en un contexto de parentesco entre especímenes particulares y manadas que viven en cautividad en centros reproductivos.

Palabras claves: *Bison bonasus*, Bisonte europeo, Marcador genético, Micromatriz multigénica (genochip), Polimorfismo de un solo nucleótido

ISSN: 1578–665 X
eISSN: 2014–928 X

Received: 4 VI 15; Conditional acceptance: 28 X 15; Final acceptance: 18 VII 16

Marlena Wojciechowska, Zuzanna Nowak, Wanda Olech, Wioleta Drobik, Dept. of Genetics and Animal Breeding, Warsaw Univ. of Life Sciences, Ciszewskiego 8, 02–786 Warsaw, Poland.– Artur Gurgul & Tomasz Szmatoła, Lab. of Genomics, National Research Institute of Animal Production, Krakowska 1, 32–083 Balice, Poland.

Corresponding author: Marlena Wojciechowska. Email: marlena_wojciechowska@wp.pl

Introduction

By the 1920s, *Bison bonasus* were extinct in the wild. The only remaining European bison were kept in managed enclosures and amounted to just 54 individuals (29 males and 25 females). The current population of the species is derived from 12 founders: 11 individuals of the subspecies *Bison bonasus bonasus* and the last representative of the subspecies *Bison bonasus caucasicus*. After successful reintroductions, there are now two genetic lines: Lowland (LB) and Lowland–Caucasian (LC). The Lowland line, also called the Bialowieza line, is derived from seven *B. b. bonasus* individuals and is a closed line, meaning that only offspring of Lowland European bison may be classified as belonging to it. The Lowland–Caucasian line includes European bison whose pedigree includes the last and only male representative of *B. b caucasicus* (Pucek, 1991; Olech, 1999)

As the result of past bottlenecks, the present population (5,249 specimens registered in 2013 in the European Bison Pedigree Book) shows a high level of inbreeding, and a significant loss of genetic variability (Olech, 2010). The European Bison Pedigree Book was created in 1924 and is published to this day. Pedigree data now provide a basis for carrying out breeding, and make it possible, among other things, to estimate the coefficient of inbreeding and kinship of the animals in the European bison breeding centres. However, in such a specific population, the pedigree, though being extremely valuable, cannot constitute the only source of information concerning the genetic value of the animals. In studies on European bison there is a need to apply methods that ensure that as much information about the genome as possible is obtained in a short time.

Over the years, a considerable number of studies have been conducted on European bison. The analyses included, among others, allozymes (Hartl & Pucek, 1994), blood groups (Sipko et al., 1995), the genes from the group of the MHC (major histocompatibility complex) (Udina & Shaikhaiev, 1998; Łopieńska et al., 2003, 2011) and microsatellites (Gralak et al., 2004; Roth et al., 2006; Nowak & Olech, 2008a). One of the more recent techniques used to estimategenetic variation involves microarrays, an approach used for several years to determine SNP (single nucleotide polymorphism) genotypes in various sites of the genome. This method allows the analysis of hundreds of thousands of markers at the same time, significantly reducing the time required to achieve a huge amount of data (Illumina®). Tokarska et al. (2009) compared effectiveness of 17 microsatellite and 960 SNP markers for paternity and identity analysis in the Lowland line of the European bison. Oleński et al. (2015) used for the first time the BovineHD microarray to find SNP markers associated with *posthitis* in the same genetic line. The first SNP analysis using the BovineSNP50 microarray which included both genetic lines (five individuals from LB and five individuals from LC) was performed by Kamiński et al. in 2012.

The aim of the present study was to identify a panel of SNP markers (among those assayed on the Illumina BovineSNP50 and BovineHD arrays) that could be used to analyse genetic structure, identify individuals and control the origin and relatedness of the European bison, as well as identify alleles specific to the two genetic lines: Lowland (LB) and Lowland–Caucasian (LC). This is the first study using BovineHD microarray to compare genetic structure of two genetic lines of European bison.

The data obtained in this study will further supplement and confirm analysis carried out on the basis of pedigree data in the context of relatedness between particular specimens and herds from captive breeding centres, making it possible, among other things, to estimate inbreeding based on multiple sources of information. Proper management of the breeding program is important for protection of the species against increasing inbreeding and its negative impact. Currently, this program is being conducted in both the European bison breeding centres and in the wild. Most of the animals from captive breeding have a known pedigree, which contributes to the control of their origins and aids population management. However, with such a low genetic variability as occurs in European bison, pedigree information may be insufficient. In addition, in the case of animals from free roaming herds, the information on their relatedness is incomplete or impossible to determine. For this reason, the assignment of SNP markers characteristic for particular genetic lines, as well as the populations within them, is clearly a great advantage, not only in future research but also to enrich and support breeding programs.

Material and methods

The animals

The study used 144 samples of European bison DNA (*Bison bonasus*), including eight samples that were analysed on two types of microarrays (BovineSNP50 v2 BeadChip–54,609 SNPs and BovineHD BeadChip–777,962 SNPs) to control the repeatability of the results and to increase the initial pool of markers. Additional control samples constituted the DNA of domestic cattle *Bos taurus*. A positive genotyping result was obtained from 129 individuals (57 individuals from LB and 72 from LC). The Lowland line included 32 males and 25 females, while the Lowland–Caucasian line included 36 animals of each sex. The biological material was collected from European bison from Polish and other breeding centres, as well as from free roaming herds. To select the most representative samples of European bison species, we were guided by inter alia, the genetic line, parental lines and the participation of ancestors. In addition, to exclude P–C (parent–child) errors and P–P–C (parent–parent–child) errors, the research included related animals —one full family: mother (sample ID K233), father (sample ID K238) and offspring (sample ID K294), as well as eight pairs of father–offspring and 14 pairs of mother–offspring. These animals were selected on the basis of the pedigree book (see table 1s in supplementary material).

DNA isolation

The test material consisted of whole blood samples and soft tissues collected by the European bison Gene Bank in the Animal Genetics and Breeding Department (Warsaw University of Life Sciences) according to decision (WPN–I.6401.90.2014.EB.1) of Regional Directorate of Environmental Protection in Warsaw. DNA from blood was isolated by the magnetic method using a MagMAX™ Express (Applied Biosystems) and a MagMAX™ Total Nucleic Acid Isolation Kit (Ambion), as well as with use of QIAamp DNA Mini Kit (QIAGEN). DNA from soft tissues was isolated using a QIAamp DNA Mini Kit (QIAGEN). The quality and concentration of the isolates was checked using NanoDrop2000 (Thermo Scientific). The DNA was then normalized to a 50ng/µl concentration. Samples with a low concentration of DNA were subjected to concentration in a Concentrator 5301 (Eppendorf AG).

Genotyping

BovineSNP50 v2 BeadChip and BovineHD BeadChip microarrays (Illumina, Inc. San Diego, CA) were used for genotyping 54K and 777K SNPs. Analyses were performed according to the Infinium Ultra and Infinium Super protocols (Illumina), and the microarrays were scanned using HiScanSQ (Illumina). The resulting intensity reading was analysed in the GenomeStudio (Illumina) software. Using the BovineSNP50 v2 BeadChip microarray we tested 91 samples. We then genotyped 46 samples using the BovineHD BeadChip microarray, including eight samples that we repeated to verify reproducibility of the results (see table 1s in supplementary material).

Criteria of markers selection

In the selection of markers for further analysis, we took into account: call rate ≥ 90%, only those markers that were genotyped in at least 90% of individuals were included; polymorphic markers whose frequency of minor allele amounted to ≥ 0.01 —adopting such a low MAF value as a criterion ensues from the specifics of the species, whose genetic variation is extremely low; no deviations from HW (Hardy–Weinberg) equilibrium at a significance level of 0.01, no P–C errors or P–P–C errors. Markers meeting the above criteria were individually tested in the GenomeStudio (Illumina) software and re–verified for proper cluster assignment, by the analysis of the GenCall Score value in SNP Graphs (figs. 1, 2). GenCall Score is a quality metric that indicates the reliability of each genotype call (GenomeStudio™, Genotyping Module v1.0 User Guide, Illumina). Automatic verification was insufficient because the measure of reliability was developed for cattle.

We then made a listing and comparison of correctly clustered markers, selected after verification, obtained from both microarrays: BovineSNP50 v2 BeadChip and BovineHD BeadChip. For the analysis of allele frequencies, we used the number of private alleles and PE (probability of exclusion) GenAlEx 6.5 (Peakall & Smouse, 2012). Using the R environment version 2.15.3, we carried out the Fisher exact test and the Z–test on two unrelated proportions for large samples to determine the statistical significance of the differences in allele frequency between the Lowland and Lowland–Caucasian lines. The final choice was made from SNP markers for which allele frequencies were significantly different, according to both tests, between the genetic lines of the European bison.

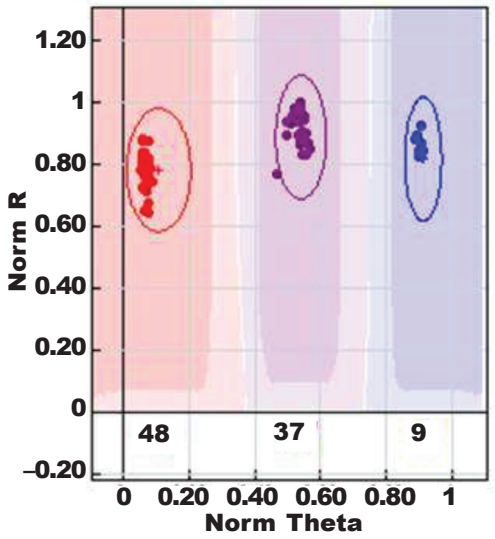

Fig. 1. SNP Graph showing the division of individuals into clusters corresponding to the genotypes at a given locus. The X–axis represents normalized theta (the angle deviation from a pure A signal, where 0 represents a pure A signal and 1 represents a pure B signal), and the Y–axis represents the distance of the point to the origin. Samples are divided according to their genotype. Samples lying within the left region are called AA; samples within the middle region are called AB and samples lying within the right region are called BB.

Fig. 1. Gráfico relativo a los PSN en el que se muestra la división de individuos en aglomerados correspondientes a los genotipos de un locus determinado. El eje de abscisas representa theta normalizada (la desviación del ángulo desde una señal A pura, donde 0 representa una señal A pura y 1 representa una señal B pura). El eje de ordenas representa la distancia del punto al origen. Las muestras se han dividido en función de su genotipo. Las muestras que quedan en la región izquierda se denominan AA; las que quedan en la región central se denominan AB; y las que quedan en la región derecha se llaman BB.

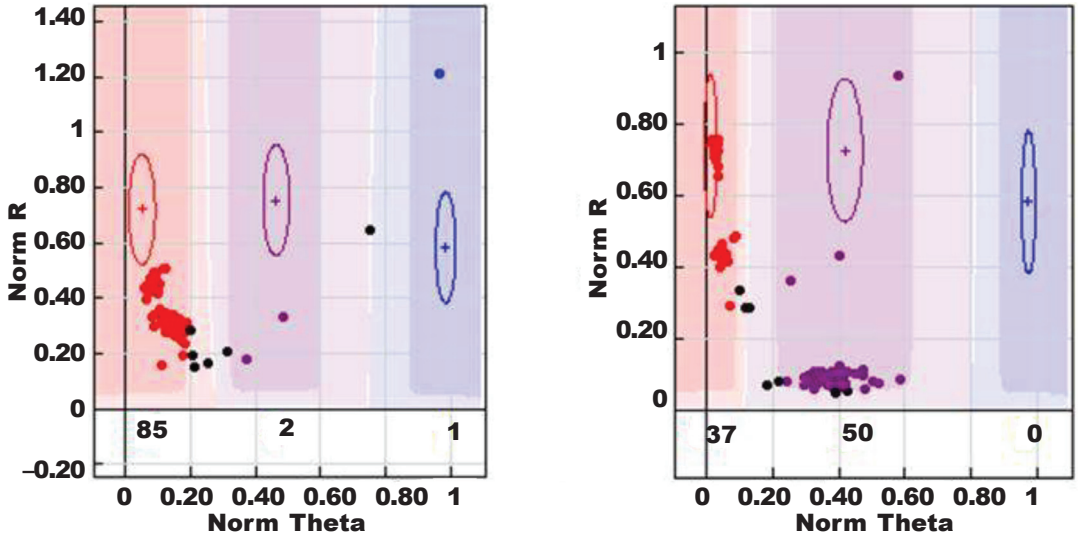

Fig. 2. SNP Graphs showing an abnormal division into clusters. The X–axis represents normalized theta (the angle deviation from pure A signal, where 0 represents pure A signal and 1 represents pure B signal), and the Y–axis represents the distance of the point to the origin. Samples are coloured according to their genotype. Samples marked in black are classified as 'no calls'. Any ambiguous division into clusters excluded a marker from further analysis.

Fig. 2. En los gráficos relativos a los PSN se muestra una división anómala en conglomerados. El eje de abscisas representa theta normalizada (la desviación del ángulo desde una señal A pura, donde 0 representa una señal A pura y 1 representa una señal B pura). El eje de ordenas representa la distancia de punto al origen. Las muestras se han coloreado en función de su genotipo. Las marcadas en color negro se clasifican como "sin determinar". Las divisiones ambiguas en conglomerados excluyeron un marcador de los análisis posteriores.

The population structure of all tested European bison was evaluated on SNPs common to both microarrays using Bayesian clustering analysis in the software STRUCTURE 2.3.4 (Pritchard et al., 2000; Falush et al., 2003). Analysis was performed under the Correlated Allele Frequencies Model and Admixture Model with 30,000 burn–in steps and 100,000 Marcov–chain Monte Carlo (MCMC) replicates for K = 1–6. Tests were conducted five times for each value of K. To determine the most likely value of K, we used the ΔK statistic (Evanno et al., 2005) Structure Harvester software (Earl & vonHoldt, 2012).

Results

In rounds of arrays preparation, the analysed *Bos taurus* samples performed well and showed call rates close to 99%. This assured us that the assay performance was essentially optimal and no flaw in the laboratory procedure would affect the results for bison.

From the 54,609 probes included on the BovineSNP50 v2 BeadChip, in correctly genotyped individuals there were 51,609 markers with a call rate equal to or greater than 90%, of which 5,997 were polymorphic in European bison (MAF ≥ 0.01). Only 1,421 SNP markers met all the aforementioned criteria.

The BovineHD BeadChip has 777,962 types of probes on its surface. A total of 735,667 SNPs showed a call rate equal to or greater than 90%, and 22,122 of these markers fulfilled all the conditions set.

After manual verification of SNP graphs in the GenomeStudio (Illumina) software for all the markers obtained after automatic analysis from both microarrays, we selected 806 SNPs and 15,062 SNPs respectively, of which 505 markers were present on both platforms. For these 505 markers, the genotypes of all eight samples analysed on both microarrays were identical**.** We found highly significant differences in allele frequency between two European bison lines in the case of 1,904 SNPs from both arrays. Finally 1,536 markers were selected for the design of a microarray specific to bison and further analyses: 47 from BovineSNP50 v2, 1,463 from BovineHD, and 26 common to both microarrays; 1,505 selected markers were distributed on autosomes and 31 SNPs on chromosome X. The number of markers on each chromosome ranged from 8 to 136 (table 1). Assuming a similar distribution of the studied SNPs in *Bison bonasus* and *Bos taurus* genomes, based on UMD3.1 cattle genome assembly, we found that the mean genomic distance for the selected SNPs was 1,443 kbp and the median distance was estimated for 211 kbp. We found that the highest median distance

Table 1. The number of markers per chromosome, genomic distances (according to the UMD3.1 cattle genome build): Chr. Chromosome; N. Number of SNPs per chromosome; MD. Mean distance; SD. Standard deviation.

Tabla 1. Número de marcadores por cromosoma, distancias genómicas (según la versión UMD3.1 del genoma de vacuno): Chr. Cromosoma; N. Número de PSN por cromosoma.; MD. Distancia media; SD. Desviación estándar.

Chr	N	MD (kbp)	SD (kbp)	MD (kbp)	Chr	N	MD (kbp)	SD (kbp)	MD (kbp)
1	108	1,456	3,170	253	17	21	2,210	4,859	243
2	53	2,479	5,584	328	18	8	8,524	10,295	1,092
3	110	952	1,718	229	19	30	1,844	3,785	539
4	86	1,356	3,715	258	20	46	1,297	3,147	174
5	69	1,733	7,736	115	21	33	1,987	3,579	374
6	136	822	2,175	116	22	20	3,147	6,224	242
7	118	955	1,797	166	23	46	1,126	2,866	179
8	66	1,691	4,483	388	24	20	2,906	7,550	502
9	22	1,238	1,719	579	25	30	1,271	2,648	272
10	42	2,442	5,582	529	26	30	1,079	2,591	164
11	70	1,470	4,271	153	27	27	1,373	2,986	250
12	96	932	1,421	220	28	33	1,109	1,681	216
13	32	2,206	6,160	248	29	21	1,148	1,360	582
14	33	2,452	6,432	398	X	31	295	813	31
15	41	2,017	3,714	424	All	1,536	1,443	3,852	211
16	58	1,037	1,901	189					

between SNPs was for chromosome 18 (1,092 kbp) and the lowest was for chromosome X (31 kbp).

The number of private alleles in the Lowland–Caucasian line was considerably higher than in Lowland line (611 and 26 respectively). Selected 1,536 SNPs were plotted against cattle chromosomes in figure 1s in supplementary material. We calculated the probabilities of exclusion coefficients (PE, both parents known; PE1, only one parent known; PE2, exclude both parents) were calculated for combined loci from each microarray, for pooled genetic lines, and separately. For 47 SNPs from Bovine SNP50, all coefficients obtained for the LB line were significantly lower than in the LC line and in the pooled samples. The statistical difference between the value of this rate for LC and the whole population was found for PE1 only (table 2). In contrast, all analyses of PE for 1,489 markers from Bovine HD gave a result of 1,000.

The minor allele frequency (MAF) was calculated separately for both lines to pre–estimate the genetic variability among the European bison analysed. Figure 3 shows the distribution of minor allele frequency within Lowland and Lowland–Caucasian lines. In order to minimize miscalculation arising from the different

Table 2. Probability of exclusion combined for all loci from Bovine SNP50 microarray (47 SNPs): PE. Probability of exclusion (both parents known); PE1. Probability of exclusion (only one parent known); PE2. Probability of exclusion (neither parents known); * $P \leq 0.05$; ** $P \leq 0.01$

Tabla 2. Probabilidad de exclusión combinada para todos los "loci" de la micromatriz multigénica Bovine SNP50 (47 PSN): PE. Probabilidad de exclusión (ambos progenitores conocidos); PE1. Probabilidad de exclusión (solo un progenitor conocido); PE2. Probabilidad de exclusión (ningún progenitor conocido); * P ≤ *0,05;* ** P ≤ *0,01.*

	PE	PE1	PE2
LB	0.741**	0.441**	0.897**
LC	0.997**	0.875**,*	1.000**
LB + LC	0.986**	0.720*,**	0.999**

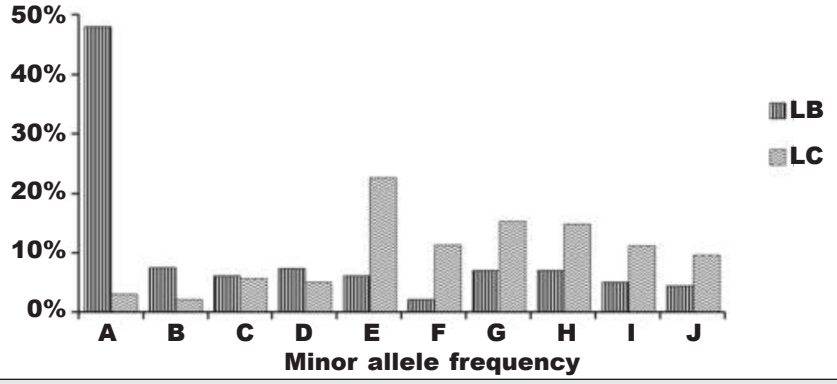

Fig. 3. Distribution of minor allele frequency (MAF) within the Lowland line (LB) and the Lowland–Caucasian line (LC). The 0 value of MAF indicates that loci were polymorphic overall, but monomorphic within one of the genetic lines: A. 0.00–0.05; B. 0.05–0.10; C. 0.10–0.15; D. 0.15–0.20; E. 0.20–0.25; F. 0.25–0.30; G. 0.30–0.35; H. 0.35–0.40; I. 0.40–0.45; J. 0.45–0.50.

Fig. 3. Distribución de la frecuencia alélica mínima dentro de la línea Lowland (LB) y la línea Lowland–Caucasiana (LC). El valor 0 de la frecuencia alélica mínima indica que los "loci" eran polimorfos en general, pero monomorfos dentro de una de las líneas genéticas. (Para las abreviaturas, véase arriba.)

number of polymorphic loci in both genetic lines and to firmly demonstrate the difference between them, we also included loci polymorphic in one line but monomorphic in the other.

In the Lowland line, we found almost 50% of monomorphic SNPs, indicating a high level of inbreeding, which is unavoidable in a closed group. In contrast to the Lowland population, in the Lowland–Caucasian line, more than 80% of the markers had an MAF > 0.2; of these, about 50% were characterized by MAF as greater than 0.3, indicating a far greater variation between individuals in the Lowland–Caucasian line than in the Lowland line. Analysis of the genetic structure of the population carried out in STRUCTURE 2.3.4 on all tested samples also showed clear differences between genetic lines. The highest value ΔK pointed to the division of the population into two clusters (K = 2), dividing individuals from both lines to clearly separate groups. The results of this analysis are given in figure 4.

Discussion

The BovineSNP50BeadChip, which was designed for domestic cattle, has been successfully used to analyse the genetic structure of several wild species. In the case of goats, two subspecies: the Tatra chamois (*Rupicapra rupicapra tatrica*) and the Alpine chamois (*Rupicapra rupicapra rupicapra*), were genotyped using the above mentioned microarray (Demontis et al., 2011). In this study, 505 of among 54,000 markers were found to be polymorphic, although only 151 were correctly clustered after manual verification. Such a low number of correctly clustered markers could indicate low genetic variability, but could also be the result of species differences. In turn, a study by Haynes & Latch (2012) for the deer species *Odocoileus hemionus columbianus, O. h. hemionus* and *O. Virginianus,* obtained 21,131 genotyped markers in at least 90% of the animals tested, of which 1,068 were polymorphic. Although *Odocoileus* spp. are genetically more distant from domestic cattle than bison, use of the same microarray allowed to obtain a relatively high number of polymorphic markers. The evolution of the genus *Bison* shows that the European bison as a wild species is genetically more similar to the domestic cattle than the American bison. The reason for this is the hybridization of the aurochs, which —like the introgression of yak in *Bison bison*— influenced the distance of these two subspecies of *Bison* (Nowak & Olech, 2008b). For comparison, in studies by Tokarska et al. (2009) the 50 Lowland European bison (LB) tested gave a reading of 52,978 SNPs, of which 960 markers were polymorphic. In contrast, Kamiński et al. (2012), despite testing only 10 European bison (five LB and five LC specimens), obtained 1,337 polymorphic SNPs. The number of markers was higher due to inclusion of both genetic lines in the studies and LC is intrinsically more diverse than LB, which is also noticeable in the present study. The participation of ancestors is different in each of the genetic lines, therefore the testing of only one of them is insufficient and cannot be used to estimate the genetic structure of the entire species. Tokarska et al. (2009) presented results of paternity analysis carried out on two marker sets: the most heterozygous SNPs, and a randomly selected set of markers. They concluded that in the case of the first set, 50–60 SNPs would be needed to assign

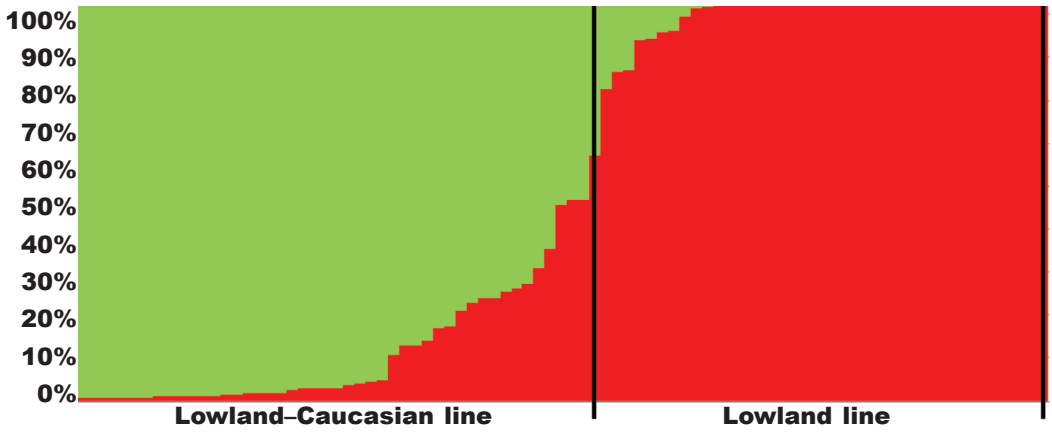

Fig. 4. Bayesian clustering analysis performed by STRUCTURE. Each individual is represented by one bar divided into segments, illustrating the proportion of estimated membership in each cluster. The vertical black lines separate a group of European bison from the Lowland line.

Fig. 4. Análisis bayesiano de conglomerados llevado a cabo por STRUCTURE. Cada individuo está representado por una barra dividida en segmentos que ilustra la proporción estimada de individuos en cada conglomerado. Las líneas negras verticales separan un grupo de bisonte europeo de la línea Lowland.

paternity with 95% confidence and 80–90 loci for the random set. Our study for the probability of exclusion (PE) partly confirms these results. Analyses were carried out for pooled genetic lines (LB + LC), and each line separately. We checked how many loci would be needed for a 99% of confidence. For the PE (both parents known) in the Lowland line —the same genetic line as in Tokarska et al. (2009)— 57 markers would be sufficient, but in the case of PE1 (only one parent known) 160 would be necessary. Other results were obtained for the Lowland–Caucasian line; only 27 SNPs in the case of PE, and 53 markers for PE1. For combined lines we estimated that 50 markers would be needed for PE, and 59 SNPs in the case of PE1. In 2015, Oleński et al. used the BovineHD BeadChip for an association study in the Lowland line. Besides reporting SNP markers significantly associated with *postitis* disease, the authors concluded that information from the subsets of SNPs could be a useful tool for the European bison breeding program, from a conservation and epizootic point of view.

When anticipate that our design of a specific SNP panel for European bison with characteristic markers for particular genetic lines (Lowland and Lowland–Caucasian) and parental lines will provide a key tool for the future analyses of the genetic structure of the species, specimen identification, and control of the origin and relatedness of European bison, both in captive breeding centres and in free roaming populations. Such knowledge is crucial for optimal management of breeding programs for these highly valued animals, and will contribute to their direct protection in the future.

Acknowledgements

We thank the veterinarians and breeders at breeding centers in Poland and other regions for providing biological material from European bison. The study was funded by the Project '*Ex situ* conservation of European bison *Bison bonasus* in Poland' under a measure 5.1 priority V of Operational Programme Infrastructure and Environment 2007–2013, in accordance with the grant agreement no. POIS.05.01.00–00–155/09–02 and the agreement on co–financing of eligible expenditures of the Project implemented under priority V of Operational Programme Infrastructure and Environment no. 615/2010/Wn–50/OP–WK–PS/D and from a grant from Warsaw University of Life Sciences no. 505–10–070200–K00253–99. The authors declare that they have no conflict of interest.

References

Demontis, D., Czarnomska, S. D., Hajkova, P., Zemanova, B., Bryja, J., Loeschcke, V. & Pertoldi, C., 2011. Characterization of 151SNPs for population structure analysis of the endangered Tatra chamois (*Rupicapra rupicapra tatrica*) and its relative, the Alpine chamois (*R. r. rupicapra*). *Mammalian Biology,* 76: 644–645.

Earl, D. A. & vonHold, B. M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2): 359–361.

Evanno, G., Regnaut, S. & Goudet, J., 2005. Detect-

ing the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14: 2611–2620.

Falush, D., Stephens, M. & Pritchard, J. K., 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, 164: 1567–1587.

Gralak, B., Krasińska, M., Niemczewski, C., Krasiński, Z. A. & Żurkowski, M., 2004. Polymorphism of bovine microsatellite DNA sequences in the lowland European bison. *Acta Theriologica*, 49(4): 449–456.

Hartl, G. B. & Pucek, Z., 1994. Genetic depletion in the European bison (*Bison bonasus*) and the significance of electrophoretic heterozygosity for conservation. *Conservation Biology*, 8: 167–174.

Haynes, G. D. & Latch, E. K., 2012. Identification of novel single nucleotide polymorphisms (SNPs) in Deer (Odocoileus spp.) using the BovineSNP50 BeadChip. *PLoS ONE,* 7(5): e36536.

Kamiński, S., Olech, W., Oleński, K., Nowak, Z. & Ruść, A., 2012. Single nucleotide polymorphisms between two lines of European bison (*Bison bonasus*) detected by the use of Illumina Bovine 50 K BeadChip. *Conservation Genet Resour*, 4: 311–314.

Łopieńska, M., Nowak, Z., Charon, K. M. & Olech, W., 2003. Estimation of polymorphism in chosen region of MHC in two genetic lines of European bison (*Bison bonasus* L.). *Applied Science Reports*: *Cattle production and breeding*, 68(1): 17–24.

– 2011. A comparison of polymorphism of DQA genes in European bison belonging to two genetic lines: Lowland and Lowland–Caucasian. *Ann. Warsaw Univ. of Life Sci.–SGGW, Anim. Sci.*, 49: 93–102.

Nowak, Z. & Olech, W., 2008a. Microsatellite variability within the sex chromosomes in European bison. *European Bison Conservation Newsletter*, 1: 72–78. [In Polish]

– 2008b. Verification of phylogenetic hypothesis concerning the evolution of genus *Bison. Ann. Warsaw Univ. of Life Sci.–SGGW, Anim. Sci.,* 45: 65–72.

Olech, W., 1999. The number of ancestors and their contribution to European bison (*Bison bonasus* L.) population. *Ann. Warsaw Univ. of Life Sci.–SGGW, Anim. Sci*, 35: 111–117.

– 2010. The Genetic Variability within *Bison bonasus* Species 90 Years after Bottleneck. In: *Restoration of endangered and extinct animals*: 48–57 (R. Słomski, Ed.). Poznań University of Life Sciences Publisher, Poznań, Poland.

Oleński, K., Tokarska, M., Hering, D. M., Puckowska, P., Ruść, A., Pertoldi, C. & Kamiński, S., 2015. Genome–wide association study for posthitis in the free–living population of European bison (*Bison bonasus*). *Biology Direct*, 10: 2.

Peakall, R. & Smouse, P., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics,* 28(19): 2537–2539.

Pritchard, J. K., Stephens, M. & Donnelly, P., 2000. Inference of population structure Rusing multilocus genotype data. *Genetics*, 155: 945–959.

Pucek, Z., 1991. History of the European bison and problems of its protection and management. In: *Global trends in wildlife management*: 19–39 (B. Bobek, K. Perzanowski & W. Regelin, Eds.). Trans. 18th IUGB Congress, Krakow 1987, Świat Press, Krakow–Warsaw, Poland.

Roth, T., Pfeiffer, I., Weising, K. & Brenig, B., 2006. Application of bovine microsatellite markers for genetic diversity analysis of European bison (*Bison bonasus*). *J. Anim. Breed. Genet*., 123(6): 406–409.

Sipko, T. P., Rautian, G. S., Udina, I. G., Ukhanov, S. V. & Berendyaeva, Z. I., 1995. Investigation of blood group polymorphism in European bison (*Bison bonasus*). *Russian Journal of Genetics*, 31: 93–100.

Tokarska, M., Marshall, T., Kowalczyk, R., Wójcik, J. M., Pertoldi, C., Kristensen, T. N., Loeschcke, V., Gregersen, V. R. & Bendixen, C., 2009. Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: the case of European bison. *Heredity*, 103: 326–332.

Udina, I. G. & Shaikhaev, O., 1998. Restriction fragment length polymorphism (RFLP) of exon 2 of Mhc–DRB3 gene in European bison *Bison bonasus*. *Acta Theriologica*, Suppl. 5: 75–82.

Supplementary material

Table 1s. Detailed information about individuals used in the study: SID. Sample ID; GL. Genetic line (LB. Lowland line, LC. Lowland–Caucasian line); PL. Paternal line (15. Begründer, 45. Plebejer, 100. Kaukasus); ML. Maternal line (16. Plavia, 35. Plewna, 42. Planta, 89. Bilma); FID. Father ID; MID. other ID; UM. Used microarray (50. BovineSNP50 v2 BeadChip, HD. BovineHD BeadChip).

Tabla 1s. Información detallada sobre los individuos empleados en el estudio: SID. Identificador de cada muestra; GL. Línea genética (LB. Línea Lowland, LC. Línea Lowland–Caucasiana); PL. Línea paterna (15. Begründer, 45. Plebejer, 100. Kaukasus); ML. Línea materna (16. Plavia, 35. Plewna, 42. Planta, 89. Bilma); FID. Identificador paaterno; MID. Otro identificador; UM. Micromatriz multigénica utilizada (50. BovineSNP50 v2 BeadChip, HD. BovineHD BeadChip).

SID	GL	PL	ML	Sex	Breeding centre	FID	MID	UM
623	LB	45	16	F	Białowieża			HD
624	LB	45	89	M	Białowieża			HD
625	LB	45	89	M	Białowieża			HD
626	LB	45	89	F	Białowieża			HD
631	LB	45	16	F	Białowieża			50
632	LB	45	16	F	Białowieża			HD
637	LB	45	89	F	Białowieża			HD
672	LB	45	16	F	Białowieża			HD
701	LB	45	89	F	Białowieża		637	50
712	LB	45	89	F	Białowieża		903	50
716	LB	45	89	M	Białowieża			50
742	LB	45		M	Białowieża (free–living herd)			50
745	LB	45	16	M	Białowieża			50
758	LB	45		F	Białowieża (free–living herd)			50
767	LB	45		F	Białowieża (free–living herd)			50
773	LB	45	42	M	Białowieża	L110		50
781	LB	45	16	F	Białowieża		632	HD
782	LB	45	42	M	Białowieża			50
785	LB	45		M	Białowieża (free–living herd)			50
806	LB	45	89	F	Białowieża	L110		HD
834	LB	45		M	Białowieża (free–living herd)			50
838	LB	45		M	Białowieża (free–living herd)			50
841	LB	45	89	M	Białowieża			50
868	LB	45		M	Białowieża (free living herd)			50
871	LB	45		M	Białowieża (free living herd)			50
877	LB	45	42	M	Białowieża			HD
878	LB	45	16	F	Białowieża		631	50
888	LB	45		M	Białowieża (free–living herd)			50
900	LB	45	16	M	Białowieża			HD
901	LB	45	42	M	Białowieża			HD
903	LB	45	89	F	Białowieża			50
915	LB	45	16	M	Białowieża		878	50
L002	LB	45	42	F	Międzyzdroje			HD

Table 1s. (Cont.)

SID	GL	PL	ML	Sex	Breeding centre	FID	MID	UM
L024	LB	45	42	F	Borki			HD
L034	LB	45	42	F	Ebeltoft			50
L110	LB	45	89	M	Białowieża			50
L111	LB	45	42	M	Borki			HD
L143	LB	45	16	M	Białowieża		781	50
L147	LB	45	16	F	Borki			50
L149	LB	45	42	F	Borki			50
L201	LB	45	42	M	Pszczyna			50
L209	LB	45	89	F	Białowieża		903	50
L227	LB	45	16	M	Białowieża		781	50
L304	LB	45	42	M	Wrocław			50
L331	LB	45		M	Borki			50
L342	LB	45	42	F	Niepołomice		L149	50
L410	LB	45	16	F	Bydgoszcz			50
L457	LB	45		M	Białowieża (free–living herd)			50
L460	LB	45		M	Białowieża (free–living herd)			50
L498	LB	45	89	F	Panevėžys			50&HD
L540	LB	45		F	Białowieża			50
L541	LB	45		F	Białowieża			50
L570	LB	45		M	Białowieża (free–living herd)			50
L584	LB	45	42	M	Pszczyna			HD
L585	LB	45	42	M	Pszczyna			HD
L619	LB	45	16	M	Gołuchów		L147	50
L640	LB	45	16	M	Smardzewice			50
K015	LC	45	89	M	Amsterdam			50
K024	LC	100	89	M	Vanatori Neamt		K026	50&HD
K026	LC	100	89	F	Vanatori Neamt			50&HD
K027	LC	45	89	F	Vanatori Neamt			50
K032	LC	100		F	Bussolengo			50
K033	LC	100		F	Bussolengo			50
K052	LC	45	89	F	Avesta			50
K106	LC	45	89	M	Bayerischer Wald			50
K107	LC	45	89	F	Praga			50
K109	LC	100	89	F	Karlsruhe			50
K110	LC	100	89	F	Karlsruhe			50
K111	LC	100	89	M	Karlsruhe			50
K153	LC	45	89	M	Damerower Werder			50
K173	LC	45		F	Gera			50
K174	LC	100	89	F	Karlsruhe			50

Table 1s. (Cont.)

SID	GL	PL	ML	Sex	Breeding centre	FID	MID	UM
K175	LC	100	89	F	Karlsruhe			50
K176	LC	100	89	F	Karlsruhe			50
K189	LC	15	35	F	Damerower Werder			50
K194	LC	45	89	M	Goldau			50
K198	LC	45	89	M	Borås			50
K213	LC	45	89	M	Damerower Werder			50
K214	LC	45	89	M	Eriksberg			50
K219	LC	45		M	Gera			50
K220	LC	45		M	Gera		K173	50
K225	LC	45	89	M	Damerower Werder			50
K233	LC	45	35	F	Hardehausen			50
K234	LC	15	35	F	Hardehausen		K233	50&HD
K235	LC	15	35	F	Hardehausen			50
K236	LC	15	35	F	Hardehausen			50
K237	LC	15	35	M	Hardehausen		K244	50
K238	LC	100	35	M	Hardehausen			50&HD
K239	LC	100	35	F	Hardehausen		K244	50
K240	LC	15	35	M	Hardehausen		K242	50&HD
K241	LC	100	35	F	Hardehausen		K235	HD
K242	LC	45	35	F	Hardehausen			HD
K244	LC	15	35	F	Hardehausen			50
K245	LC	15	35	M	Hardehausen		K248	50&HD
K247	LC	15	35	M	Hardehausen		K244	50&HD
K248	LC	45	35	F	Hardehausen			50
K250	LC	45	89	M	Damerower Werder			50
K274	LC	45	35	F	Damerower Werder			50
K275	LC	45		F	Damerower Werder			50
K282	LC	45	89	M	Bern			50
K284	LC	45	89	F	Goldau			50
K286	LC	45	16	M	Thoiry			50
K289	LC	45	89	F	Damerower Werder			50
K290	LC	45		M	Damerower Werder			50
K291	LC	45	89	F	Damerower Werder			50
K292	LC	45	89	F	Damerower Werder			50
K294	LC	100	35	M	Hardehausen	K238	K233	HD
K302	LC	45		F	Damerower Werder			50
K358	LC	15	35	M	Neumünster			HD
K373	LC	100	89	M	Karlsruhe			HD
K377	LC	100	35	F	München, Hellabrunn			HD

Table 1s. (Cont.)

SID	GL	PL	ML	Sex	Breeding centre	FID	MID	UM
K378	LC	45	35	F	München, Hellabrunn			HD
K380	LC	45	35	M	München, Hellabrunn			HD
K461	LC	15	35	M	Neumünster	K358		HD
K502	LC	100	35	F	Weilburg			HD
K533	LC	100	89	F	Bad Orb			HD
K534	LC	45		M	Eulbach			HD
K545	LC	100	35	M	Hanau			HD
K559	LC	100	89	M	Sababurg	K545		HD
K560	LC	100	89	M	Sababurg	K545		HD
K563	LC	100	16	F	Edertal–Hemfurth			HD
K586	LC	100	35	M	Hardehausen	K238		HD
K591	LC	100	35	M	Hardehausen	K238		HD
KB317	LC			M	Bieszczady (free–living herd)			50
KB342	LC			F	Bieszczady (free–living herd)			50
KB344	LC			M	Bieszczady (free–living herd)			50
KB394	LC			M	Bieszczady (free–living herd)			HD
KB396	LC			F	Bieszczady (free–living herd)			HD
KB397	LC			M	Bieszczady (free–living herd)			HD

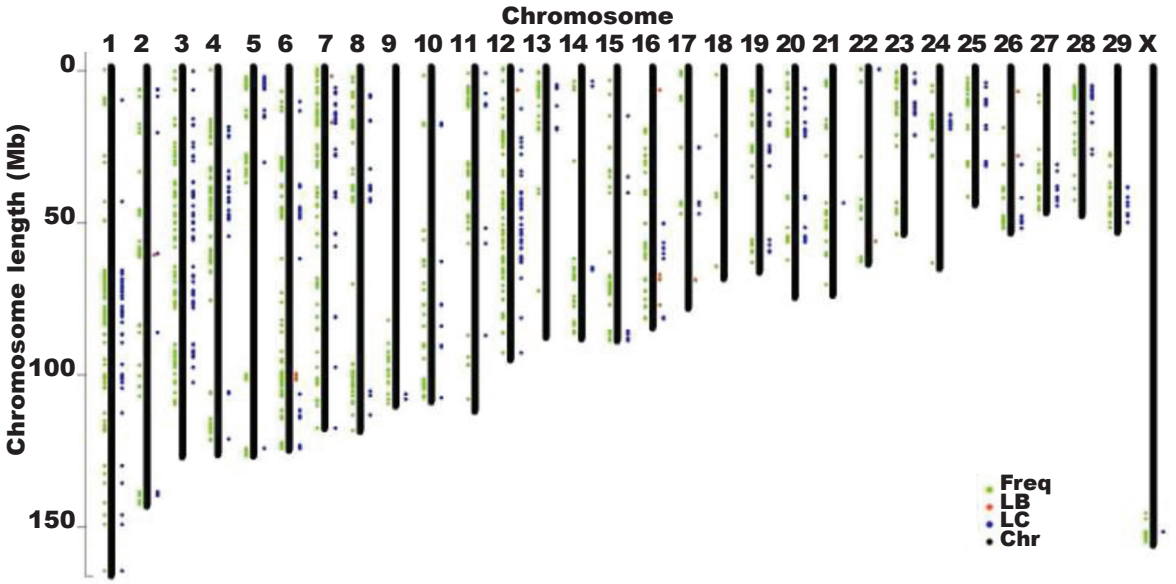

Fig. 1s. Localization of SNPs with significant differences in allele frequency between LB and LC populations and SNPs with private alleles in each population. Black line indicates chromosome length: Freq. SNPs differing in frequency between LB and LC population; LB. SNPs with private allele in LB population; LC. SNPs with private allele in LC population; Chr. Chromosome.

Fig. 1s. Localización de los PSN que presentan diferencias significativas en las frecuencias alélicas entre las poblaciones LB y LC y de los PSN con alelos privados en cada una de las poblaciones. La línea negra indica la longitud del cromosoma: Freq. PSN cuya frecuencia difiere entre las poblaciones LB y LC; LB. PSN con alelos privados en la población LB; LC. PSN con alelos privados en la población LC; Chr. Cromosoma.