

## Article

# From Wisent to the Lab and Back Again—A Complex SNP Set for Population Management as an Effective Tool in European Bison Conservation

Marlena Wojciechowska \*, Karol Puchała , Zuzanna Nowak-Życzyńska , Magdalena Perlińska-Teresiak , Marta Kloch , Wioleta Drobik-Czwaro and Wanda Olech 

Department of Animal Genetics and Conservation, Warsaw University of Life Sciences-SGGW, Ciszewskiego 8, 02-786 Warsaw, Poland

\* Correspondence: marlena\_wojciechowska@sggw.edu.pl

**Abstract:** Proper management and genetic monitoring of the modern European bison (*Bison bonasus*) population is one of the most important responsibilities for this species' conservation. Up-to-date, complex genetic analysis performed using a consistent molecular method is needed for population management as a tool to further validate and maintain the genetic diversity of the species. The identification of the genetic line when pedigree data are missing, as well as the identification of parentage and individuals, are crucial for this purpose. The aim of our research was to create a small but informative panel of SNP (single-nucleotide polymorphism) markers that can be used for routine genotyping of the European bison at low cost. In our study, we used a custom-designed microarray to genotype a large number of European bison, totaling 455 samples from two genetic lines. The results of this analysis allowed us to select highly informative markers. In this paper, we present an effective single nucleotide polymorphism set, divided into separate panels to perform genetic analyses of European bison, which is needed for population monitoring and management. We proposed a total of 20 SNPs to detect hybridization with *Bos taurus* and *Bison bison*, a panel of 50 SNPs for individuals and parentage identification, as well as a panel of 30 SNPs for assessing membership of the genetic line. These panels can be used together or independently depending on the research goal and can be applied using various methods.

**Keywords:** SNP—single-nucleotide polymorphism; European bison; conservation; genetic lines; individual identification; parentage control



**Citation:** Wojciechowska, M.; Puchała, K.; Nowak-Życzyńska, Z.; Perlińska-Teresiak, M.; Kloch, M.; Drobik-Czwaro, W.; Olech, W. From Wisent to the Lab and Back Again—A Complex SNP Set for Population Management as an Effective Tool in European Bison Conservation. *Diversity* **2023**, *15*, 116. <https://doi.org/10.3390/d15010116>

Academic Editor: Michael Wink

Received: 9 November 2022

Revised: 11 January 2023

Accepted: 12 January 2023

Published: 14 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The European bison (*Bison bonasus*), also known as the wisent, is Europe's largest terrestrial mammal. Population genetics and management of the modern European bison's gene pool, after its extirpation in nature in the 1920s, is considered an important task and major issue for this species' conservation [1]. As a result of past bottlenecks, the current population has experienced a significant loss of genetic variability and a high degree of inbreeding [2]. The first detailed analysis of the pedigree was conducted by Slatis [3] and included animals that were living in the 1950s. Slatis found that all European bison were derived from 17 ancestors, but some of these ancestors were represented by only a single descendant. The author decided to refer to the 12 unique genotypes as the founders: 11 individuals of the subspecies *Bison bonasus bonasus* and one of the subspecies *Bison bonasus caucasicus*.

The species is divided into two genetic lines: a Lowland line (LB) which is pure lowland subspecies (*Bison bonasus bonasus*) and a Lowland-Caucasian line (LC) which is a mixture of two subspecies. The gene pool of the present Lowland line (LB) includes only seven founders (BEGRÜNDER pedigree no. 15, PLAVIA pedigree no. 16, PLANTA pedigree no. 42, PLEBEJER pedigree no. 45, BILL pedigree no. 87, BILMA pedigree no. 89, and BISMARCK pedigree no. 147) [1,4,5]. This line is closed because both parents should belong

to the LB line. In addition, in the current population, all males in this line have a copy of the Y chromosome from the founder, named Plebejer [5]. The other 11 individuals, PLEWNA pedigree no. 35, PLACIDA pedigree no. 46, GARDE pedigree no. 95, and GATCZYNA pedigree no. 96, although also belonging to the lowland subspecies *Bison bonasus bonasus*, left only hybrid offspring with the Caucasian subspecies *B. b. caucasicus*. Their descendants established the second line (LC), containing the genes of a Caucasian founder (KAUKASUS pedigree no. 100), so all 12 founders formed the gene pool of this genetic line [4,5].

Most of the papers published thus far in the field of European bison genetics have mainly been concerned with the Lowland line, which cannot be considered a reference for the whole species. Analyses of genetic variability have shown a high degree of homogeneity in the Lowland line compared to the Lowland–Caucasian line [6]. The genetic lines are kept and bred separately in captivity. According to an International Union for Conservation of Nature (IUCN) study on European bison [1,3], the isolation of lines is important for counteracting the loss of genetic variability within the Lowland–Caucasian line. Other authors have pointed out that one of the threats to the genetically pure Lowland subspecies is possible hybridization with the Lowland–Caucasian line [7]. In both cases, there is the need for a tool that can help distinguish animals belonging to these lines if pedigree data is missing.

Another important issue is the proper assignment of parentage. The European Bison Pedigree Book (EBPB), established in 1923, contains information on the captive part of the population. Although these data are extremely valuable, they should not be used as the only source of information. Genetic monitoring of free-ranging herds cannot be based on pedigrees, so it is necessary to use molecular methods for the species. In addition, in some herds, it might be difficult to assign offspring to their mothers due to the large area of enclosure or the high relationship between females within a herd. For example, individuals in the Pszczyna breeding center are very closely related. The European Bison Pedigree Book shows that, in a given year, calves are born from one or two sires, and the mothers are often sisters or half-sisters [8]. When many calves are born within a short period, parentage is judged mainly from observations of cow and calf behavior. Another aspect is the identification of individuals to avoid duplication of data. Sometimes an animal is sampled multiple times in different periods. If this is conducted in a captive herd, it is easy to obtain the correct animal ID; but when wisents belonging to a free-ranging population are sampled, identification can pose some difficulties, especially for animals without collars or transponders.

Modern parentage and individual identification employing comprehensive genetic analysis performed with a reliable molecular method are needed as a population management tool to further validate and maintain the genetic diversity of the species.

Commonly used genetic markers in the study of population genetics are Single Nucleotide Polymorphisms (SNPs). SNP is defined as a difference in a single base pair at a particular location in the genome between individuals of the same species. They occur both in non-coding and coding regions of the genome. SNP detection technologies have evolved from labor-intensive and time-consuming processes to some of the most automated and relatively inexpensive methods [9]. Over the years, many methods have been used to analyze these markers, including PCR-RFLP—restriction fragment length polymorphism, or PCR-SSCP—single-strand conformation polymorphism. However, they only allow SNP analysis within a single amplified DNA fragment during a single test [10]. Another method for SNP analysis is allele-specific PCR. This technique is based on the amplification of a product that depends on an SNP in the primer binding site. The occurrence of a given allele results in the amplicon or its absence. Detection of the product depends on the technique used, e.g., thermocyclers for Real-Time PCR, whereby the product increment can be read in real-time, which shortens procedures [11].

A breakthrough in the study of single-nucleotide polymorphism turned out to be DNA microarrays. The use of microarrays has allowed hundreds of thousands of SNP markers to be studied simultaneously [12]. DNA microarrays are defined as ordered collections of molecular probes, immobilized on different types of surfaces in a well-defined arrangement.

Microarrays with short oligonucleotide probes are used in SNP analysis, but the way they are synthesized, labeled, and detected depends on the company's technology.

Our research aimed to create a small but complex and informative panel of SNP markers that can be used for routine genetic testing of the European bison, maintaining its efficiency at a low cost. For this study, we developed a custom microarray for the European bison and used it to analyze a large pool of samples, which then allowed us to propose a comprehensive set of markers that could serve as a tool for the conservation of the European bison. The panels are designed for the routine use of medium-throughput single nucleotide polymorphism (SNP) genotyping techniques. Based on the panels described in this article, it is possible to perform analyses using different techniques such as KASP, TaqMan [13], and others, depending on the capabilities of the laboratory. In this paper, we present an effective SNP set to perform the genetic analyses of European bison that are needed for population monitoring and management.

## 2. Materials and Methods

The work was divided into three phases. The first was the design of microarrays with markers selected in preliminary studies [6,14], whereas the second stage involved laboratory procedures and genotyping of European bison on the previously designed platform. The third step was to select a smaller number of markers based on the genotyping results without compromising their efficacy.

### 2.1. Microarray Design

The project was created in DesignStudio—Microarray Assay Designer 1.1.2 (Illumina, CA, USA) and the Infinium<sup>®</sup> Commercial + platform using the Infinium<sup>®</sup> BovineLD-24 + v2.0 microarray. This microarray contained a base set of 7931 SNPs, selected for different breeds of domestic cattle [15]. To this set, 2000 markers were added based on genotyping results from previous studies [6,14]. The markers were divided into various panels: Line: for assessing membership of the Lowland (LB) or Lowland–Caucasian (LC) lines; Profile: for individuals and parentage identification assessment; hybridization with cattle; and hybridization with American bison.

### 2.2. Samples and Genotyping

The biological material came from a blood and soft tissue collection conducted by the Department of Animal Genetics and Conservation and was carried out independently of these studies; the Regional Director of Environmental Protection in Warsaw gave permission for preserving the tissues of a protected species (ref. no: WPN-I.6401.90.2014.EB.1). Samples were collected from wisents that were immobilized or had died. Samples for analysis were selected according to several criteria, including genetic line (LB or LC) and parental line, place of origin, and contribution of founders, so that the pool of individuals studied was representative of the entire European bison population. The DNA was isolated using GeneMATRIX Quick Blood DNA Purification Kit (EURx, Gdańsk, Poland) and GeneMATRIX Tissue DNA Purification Kit (EURx, Gdańsk, Poland) according to the protocols provided by the manufacturer. Then, the purity and concentration of the obtained DNA were measured using a NanoDrop2000 spectrophotometer (Thermo Scientific<sup>™</sup>, Wilmington, NC, USA), and the samples with a minimum concentration of 50 ng/ $\mu$ L and ratio parameters  $A_{260}/A_{280} = 1.8\text{--}2.0$  and  $A_{260}/A_{230} = 2.0\text{--}2.2$  were classified to the next stage of analysis. Furthermore, samples whose DNA was degraded were rejected, with the degree of degradation being determined by horizontal electrophoresis. In addition to the wisent samples, six American bison (*Bison bison*) individuals and one hybrid *Bison bonasus*  $\times$  *Bos taurus* were also subjected to genotyping. The concentration of all DNA samples was normalized to 50 ng/ $\mu$ L. A total of 480 samples were genotyped on a dedicated microarray, including 10 samples that were analyzed twice to check the repeatability of the results. Genotyping was performed according to the Infinium<sup>®</sup> Super (Illumina, CA, USA) protocol. Quality control was performed using GenomeStudio 2.0 (Illumina, CA, USA) software. Individuals

whose call rate was less than 0.9 (meaning that the readings for these samples were obtained for less than 90% of SNP markers) or whose clustering results were significantly different from the rest of the samples were removed from the sample pool. Ultimately, samples from 455 wisents (Table 1), six American bison, and one *Bison bonasus* × *Bos taurus* hybrid (named Dima) remained for further analysis, with duplicate samples omitted.

**Table 1.** Number of wisents and their origins included in the final analysis.

<b>Lowland line</b>	<b>327</b>
Belarus	76
Denmark	2
Netherlands	8
Poland	241
<b>Lowland–Caucasian line</b>	<b>128</b>
Austria	1
Belgium	1
Finland	1
France	2
Germany	60
Poland	28
Sweden	35
<b>Total</b>	<b>455</b>

### 2.3. Marker Selection and Validation

Samples of American bison and the *Bison bonasus* × *Bos taurus* hybrid were used to select markers independently of the other panels. The selection criteria differed depending on the purpose for which the markers were to be used. Since the goal was to select markers that only worked for one species, the minor allele frequency (MAF) parameter was ignored, and the initial selection was based on the call frequency parameter. Markers were verified based on the SNP graph. For the bovine hybridization panel, markers that scored for samples other than Dima were removed. Then, 10 of the markers with the best quality parameters that produced a result for a hybrid but were not present in the European bison were selected. These markers can also be used to distinguish between wisents and cattle samples. Next, for the Hybridization with American bison panel, 10 markers were selected that provided a result in all six American bison samples that were not present in the European bison. In each case, the positions of the markers on the chromosomes were compiled, and where possible, markers located on different chromosomes were selected for the panel. In other cases, only markers located more than 100 kb apart were considered.

Markers included in the Line and Profile panels had to meet the following criteria: call frequency  $\geq 0.9$  and MAF  $\geq 0.005$ . The required value for this factor (MAF), which qualified a given SNP for further analysis, was set so low because of the large number of individuals and the assumption that there may have been private alleles in the genetic lines. This criterion also allowed monomorphic markers to be removed from the analysis; in this case, non-informative markers. Other criteria included no PC or PPC (Parent–Parent–Child) errors, and correct clustering. SNP marker readings were clustered according to an algorithm developed for domestic cattle. For this reason, clustering verification was performed manually based on the SNP graph. The chromosomal position of SNPs was verified, and markers without mapped positions were removed. GenAEx 6.5 [16] utilized chi-square to test for Hardy–Weinberg equilibrium (HWE). Pairwise linkage disequilibrium analysis (LD) was performed in PLINK v1.9 [17,18], resulting in the removal of SNPs that were linked.

For the Profile panel, markers were selected based on matching the allele and genotype frequencies as closely as possible:  $f(A) = 0.5$ ,  $f(B) = 0.5$ ,  $f(AA) = 0.25$ ,  $f(AB) = 0.5$ ,  $f(BB) = 0.25$ . These could be used regardless of the genetic line. GenAEx 6.5 was used to perform the following: Multilocus Matches Analysis (MS), Probability of Identity (PI), Probability of

Identity Sibs (PIsibs), and Probability of Exclusion (PE) (PE1, when the other parent is known; PE2, when one of the parents is known but the genotype of the other is not available; and PE3, to exclude the putative parental pair). The PI parameter provided an estimate of the average probability that two unrelated individuals will have the same multilocus genotype at a certain number of markers, and PIsibs estimated the probability of identity among siblings.

Based on these results, the minimum number of markers needed for the panel was estimated. Markers were then screened with respect to the same parameters produced in preliminary experiments. Those that met the criteria in the best way were selected for a smaller SNP set, and a recheck of PI, PE, and MS was performed. Then the SNP set was validated for parentage assignment power using CERVUS 3.0.7 [19]. The software allowed for the calculation of pairwise parentage likelihood, giving an overall log order difference (LOD) score. In the parentage analysis, the LOD score was determined by the natural logarithm of the overall likelihood ratio. The assignment of candidate parents to offspring was based on the highest LOD score. The analysis was performed with three data sets: (1) a known trio of sire, dam, and offspring and an additional 10 individuals from the same breeding center (Sweden, Avesta), including the sire's full brother and the parents' half-siblings; (2) two different pairs of sires and offspring, and 11 randomly selected wisents from Białowieża (Poland); and (3) two different pairs of dams and offspring, and 13 randomly selected wisents from Pszczyna (Poland). Two approaches were chosen for the first dataset: paternity (known mother—candidate fathers) and a parent pair (sex unknown). For datasets 2 and 3, parentage analyses of paternity and maternity, respectively, were performed without a known second parent.

For the Line panel, allele frequencies and the Z-test for two unrelated proportions (for large samples) were performed in MS Excel 2019 to determine the statistical significance of the allele frequency differences between the Lowland and Lowland–Caucasian lines. Only those markers with highly significant differences ( $p$ -value < 0.001) in their allele frequency between genetic lines and the private alleles of the lines were selected. Principal Coordinates Analysis (PCoA) was performed using GenAlEx 6.5 to verify the selected markers.

### 3. Results

The SNPs selected as markers for bovine hybridization provided results for the hybrid sample only. Of 82 markers, 10 with the best quality parameters were selected, but further confirmation of their effectiveness was needed on an even larger sample of European bison. Two of these markers were located on the same chromosome (1) but were approximately 8.6 Mbp apart. The markers selected for the panel to distinguish between European bison and American bison only provided a result in the case of the American bison samples. In total, 10 of the 18 markers were selected for the panel. The two SNPs were located on the same chromosome (4), approximately 28.9 Mbp apart (Supplement Table S1).

After quality control for call frequency, MAF, PC, PPC, and correct clustering, 1836 SNP markers remained. The list of markers from the designed microarray, which was selected based on the preliminary tests [5,9], was compared to the markers that had passed the controls after genotyping a larger number of wisent samples, totaling 455. Subsequently, 713 candidate SNP markers for the Profile panel and 299 SNPs for the Line panel remained for further analysis. LD and HWE analyses allowed the number of markers to be narrowed down. The multilocus match analysis of samples, genotyped twice in independent reactions, allowed the compatibility of genotypes to be checked and confirmed the correctness of laboratory analyses.

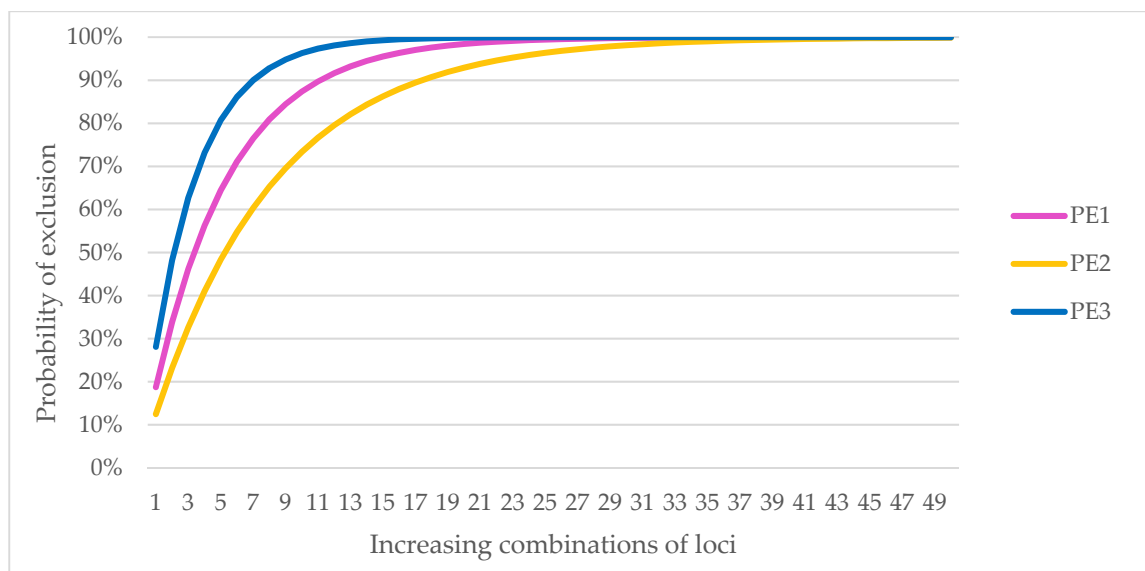
#### 3.1. "Profile"—SNP Panel for Individuals and Parentage Identification Assessment

Analysis using an increasing number of loci showed that the probability of identity for unrelated samples had already fallen below 1% at five SNP markers. For PIsibs, 10 SNPs were required to reach values below 1% for a given population size. The estimated expected number of unrelated individuals with the same multilocus genotype for increasing locus

combinations (calculated as respective probability  $\times$  population size) is less than one when combining at least seven markers, and 13 SNPs when based on PIsibs.

The results of the PE analysis for an increasing number of loci showed that to exclude parenthood at a probability greater than 99%, at least 23 SNP markers would need to be genotyped if the other parent was known (PE1), 38 SNP markers if the genotype of one of the parents was unknown (PE2), and 15 SNP markers if the test was to exclude parentage by a specific pair of individuals (PE3).

Because of the possible differences in results for individuals from different genetic lines, the PE analyses were also performed separately for the LB and LC lines. Parental exclusion at a probability above 99% was possible for LB with 24 markers for PE1, 40 for PE2, and 15 for PE3; for LC, the number of markers was 23 for PE1, 37 for PE2, and 15 for PE3 (Figure 1). The selection of markers during the preliminary studies eliminated differences arising from the genetic lines so that these markers could be used for individual identification regardless of the origin of the sample.



**Figure 1.** Probability of exclusion for an increasing number of loci. PE1, when the other parent is known; PE2, when one of the parents is known, but the genotype of the other is not available; and PE3 for the exclusion of the putative parental pair.

Based on the results of PE, we assumed that a minimum number of 40 markers was required for individual identification and parentage control. An optimally expanded panel of 50 SNPs was proposed (Supplement Table S2), selected on the markers' frequency of genotypes and alleles across all tested samples, regardless of the genetic line. For the combined 50 SNPs, the values were  $5.6 \times 10^{-22}$  for PI and  $5.2 \times 10^{-12}$  for PIsibs. Exclusion of parentage is possible in all cases with the selected markers at a probability of more than 99%. Details are shown in Figure 1. The MS analysis confirmed that there were no matches in the multilocus genotype for all 455 samples, even in closely related individuals from Pszczyna (PL subline).

The accuracy of the ancestry assignment obtained from the CERVUS software was high. At a strict 95% confidence level the assignment rate was 100% in all cases. The first dataset for a known trio (offspring, father, and mother), from Avesta (Sweden), was subjected to two different methods. In the paternity analysis with a known mother and candidate fathers, the highest pair LOD score (7.56) was obtained for the real father and offspring; also, for the given mother and real father, the Trio LOD score (10.19) was higher than for the other combinations of mother–candidate father. The values of the LOD scores for other candidates were negative, which means that these animals were less likely to be real parents than non-real parents. The second parent non-exclusion probability was  $3.21 \times 10^{-5}$ . The full brother of

the known father had a pair LOD score equal to  $-2.76$ . The second method used in the given data set was the determination of the parent pair (sex unknown). A total of 66 combinations of triplets (offspring, presumed first parent, presumed second parent) were analyzed. The obtained results unequivocally indicated the correct parent pair, even though the studied pool also contained closely related individuals. The best Trio LOD result (20.53) was found when both the true parent and the offspring were combined. The five best Trio LOD results are shown in Table 2. The parent pair non-exclusion probability was low at  $1.86 \times 10^{-8}$ .

**Table 2.** The parentage assignment LOD scores for parent pairs (sex unknown).

First Candidate	Pair LOD Score	Second Candidate	Pair LOD Score	Trio LOD Score
K439	7.56	K425	10.34	20.53
K435	$-2.76$	K425	10.34	6.06
K420	$-2.69$	K425	10.34	2.03
K439	7.56	K729	$-8.74$	$-0.68$
K417	$-5.23$	K425	10.34	$-1.97$

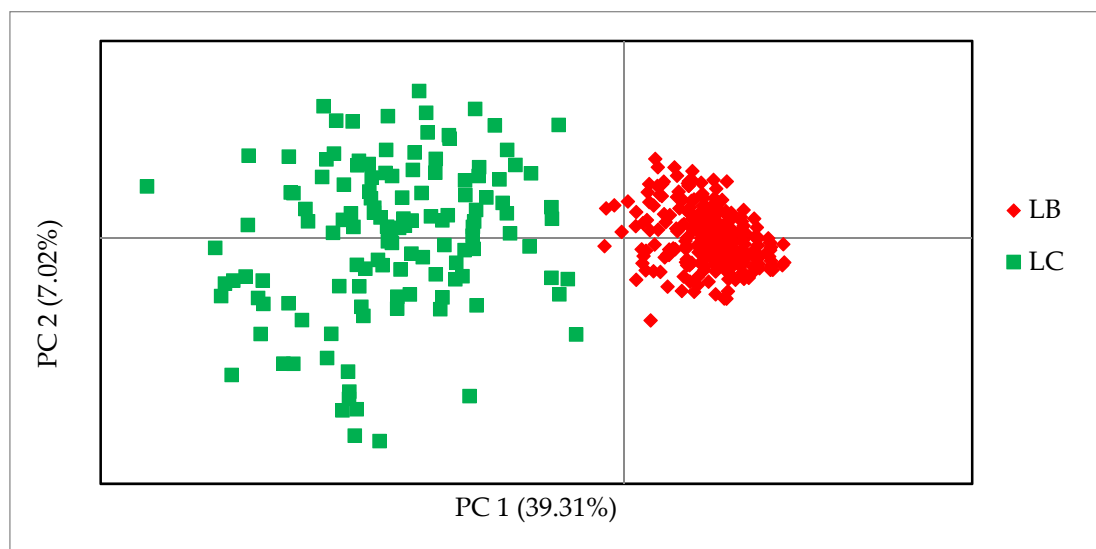
Legend: K439—true father; K425—true mother; K435—full brother of the father; K420—offspring of the father; K417—half-sister of the father; K729—half-sister of the mother.

The next dataset for offspring–candidate parents was performed on the Białowieża herd (Poland). The offspring were half-siblings (the common mother was not included in the database). The assignment of the putative second parent was made from a set of 13 wisents that included the true fathers of two of the offspring each and a grandfather of one of the offspring. The first father–son pair had a pair LOD score of 8.10 whereas the other wisents had negative scores. In the case of the second offspring, the highest pair LOD score was also obtained for the father–descendant pair (9.21) with a confidence level of 95%, whereas the grandfather–offspring scored 8.22 for their pair LOD. The probability of a correct parentage assignment increases with the value of the LOD score.

The last set of data, offspring and candidate parents, concerned animals from the breeding center in Pszczyna. The analyses were performed for two offspring and 15 candidate parents, including one true parent for each offspring. In both cases, the highest values for the pair LOD scores were observed for the true parents; this result is promising.

### 3.2. “Line”—SNP Panel for Assessing Membership of the Lowland (LB) or Lowland–Caucasian (LC) Lines

Of the markers with highly significant differences in allele frequency, a set of 30 SNPs and a set of 17 SNPs with private alleles for LC were selected. Although we found three SNPs with private alleles for the Lowland line, these should not be used to distinguish between lines, primarily because LC is an open line and is a mix of all 12 ancestors, including the seven wisents that created the LB line. The lists of the two sets of selected markers with allele frequencies are shown in Tables S3 and S4 in the Supplement. Analysis of the principal coordinates of the 30 SNPs proved the effectiveness of the markers and their usefulness in differentiating the lines. Figure 2 shows that the first coordinate (PC1) separated the samples of the two lines and was responsible for 39.31% of the genetic variability. The second coordinate (PC2) also shows a greater differentiation between the animals of the LC line. Individuals from the LB line were tightly clustered and separated from the more diverse LC individuals, regardless of their origin.



**Figure 2.** Principal Coordinates Analysis (PCoA) based on 30 SNPs. LB—Lowland line and LC—Lowland-Caucasian line.

#### 4. Discussion

Our research aimed to provide complex sets of SNPs that could be used as valuable genomic tools for monitoring genetic diversity and the management of European bison populations. An optimally reduced number of markers will allow the cost of analyzing single individuals to be reduced without having to use standard microarrays every time.

Currently, control of the origin of European bison is mainly based on pedigree data published in the European Bison Pedigree Book (EBPB) and on information provided by breeders and people who take care of the animals on a daily basis. However, this information is not always 100% certain or is not verified. Another problem is the identification of individuals in large or semi-free-ranging herds, where it is difficult to constantly observe the animals. Tokarska et al. [20] compared the effectiveness of using microsatellite and SNP markers for parentage testing. Different numbers of animals were used in simulations in the CERVUS program. In the case of SNPs, 50 animals were tested using the BovineSNP50 microarray, whereas microsatellites were used in the analysis of 276 wisents. It should be mentioned that all the examined animals belonged to the Lowland line. With 17 microsatellites, the mean expected heterozygosity was  $H_e = 0.31$ . To maintain heterozygosity at this level, 68 markers would be required to ensure a 95% probability of correct paternity assignment where the mother is unknown. On the other hand, the authors concluded that a random selection of markers would require 80–90 SNPs, and considering the highest level of heterozygosity, a set of 50–60 markers would be sufficient for selection. The markers (50 SNPs) selected for this purpose in our study had a mean expected heterozygosity of  $H_e = 0.498$ . The result of the CERVUS software gives us confidence that the selected markers can successfully be used for parentage analysis. We performed an evaluation of parentage candidates for offspring on different data sets from both genetic lines, including closely related animals. Thus, confirmation of paternity was unequivocal, even for male full siblings or for the father and grandfather of offspring considered to be potential parents in the data.

Oleński et al. [21] created an interesting subset of 100 SNPs for parental identification in the European bison Lowland line. However, when the list of markers is compared to the data from our research, it is found that of the 1836 SNP markers that passed basic quality control, only 17 markers were identical; furthermore, our final set of 50 SNPs contained only four identical markers to their subset of SNPs. This drastic difference may be due to the fact that we included individuals from the LC line as well as animals from different LB populations in our analysis. In their study, Oleński et al. [21] focused only on the



Białowieża population, which belongs to the LB line. Nevertheless, parameters such as PI and PE were sufficient for both identifying individuals and parentage control in both cases. Hara et al. [22] developed a panel of markers for origin control in Japanese Black cattle by selecting 87 SNPs that reached a PE of 0.999998 where one parent was known. A similar power for exclusion probability was obtained in American bison [23]. Out of 191 SNPs tested, 50 markers resulted in a PE value close to 100%. According to Waits [24], a PI (Probability of Identity) value of  $10^{-3}$ – $10^{-4}$  is sufficient for wildlife analysis. Of course, if the history of the European bison species is taken into account, this value should be even lower. The combined PI for 100 SNPs in a study by Oleński et al. [21] was  $1.14 \times 10^{-40}$ , whereas for our Profile panel with 50 SNPs, we obtained PI values of  $5.6 \times 10^{-22}$  and PIsibs of  $5.2 \times 10^{-12}$ . In the studies on Japanese cattle mentioned earlier [21], a panel with 87 markers produced a PI of  $2.01 \times 10^{-34}$ . Considering the number of markers, these results are comparable.

In the Yang et al. [23] study, 17,018 markers were used for subspecies composition analysis to distinguish between the American wood bison and American plains bison. This was a large set of markers, requiring the use of microarrays or other high-throughput methods that may incur additional costs for a large number of samples. In our preliminary studies [6], the allele frequencies of the studied SNP markers for the LB and LC lines were compared. In this first phase, the marker pool was narrowed down to 1536 SNPs, of which almost 50% were monomorphic in the Lowland line; in the Lowland–Caucasian line more than 80% had a MAF greater than 0.2. These results clearly show the problem of very low genetic variability within the Lowland line. In this study, with only 30 SNPs, it was possible to distinguish the genetic lines of the European bison and to select additional markers with allele characteristics only of the Lowland–Caucasian line. The markers in the Line set separated the genetic lines, but still showed the complex structure of the LC line. The PCoA method resulted in a compact cluster of Lowland line individuals and a more dispersed and diversified group of Lowland–Caucasian individuals.

The purity of the species is as important as membership in a genetic line. In animals threatened with extinction or at risk of losing genetic variability, it is extremely important to control the origin and avoid crosses with other species. In the group of individuals studied, the markers selected for each panel met all the assumed criteria. The markers that allowed for the molecular differentiation of the European bison and its bovine hybrids provided a result only for the Dima sample, the only *Bison bonasus* × *Bos taurus* hybrid among the analyzed specimens. It is necessary to confirm these results for a larger number of hybrids. However, these markers also allow wisents to be differentiated from cattle, making them a useful tool, for example, in detecting poaching. Additionally, in the case of distinguishing wisents from American bison, it was possible to select good markers; this, however, requires further analysis. Genotyping of these markers in wisents may indicate hybridization, but it does not have to be unambiguous hybridization with the American bison. Recall that the markers were selected from SNPs that were derived from the Illumina Bovine50SNP and BovineHD microarrays, i.e., markers developed for cattle.

Transferring a known method to a new species is more efficient and less costly than searching for new markers or developing methods de novo. However, when using this type of method transfer from one species to another, one has to reckon with the need to optimize procedures and the possibility of obtaining fewer data or poorer data quality. Miller et al. [25] compared the applicability of microarrays developed for three domestic animal species, the OvineSNP50 BeadChip, the BovineSNP50 BeadChip, and the EquineSNP50 BeadChip in genome studies of other species. He determined the relationship between the effective use of microarrays and the time of phylogenetic separation of species. A linear decrease in genotyping efficiency, as measured by the call rate parameter, of about 1.5% per million years separating the compared species was observed, whereas a much steeper decrease was observed for polymorphism. In a preliminary study [6], it was shown that of the markers available on the microarrays Bovine50SNP and BovineHD, as many as 94.5% of SNPs were genotyped in European bison (with call rates equal to or greater than 90%), A

similar call rate was reported for American bison in a study by Decker et al. [26], whereas lower values were obtained by the same author for Water buffalo *Bubalus bubalis* 86.4%, or Kudu *Tragelaphus strepsiceros* 76.6%. For other species, such as deer, by using BovineSNP50 it was possible to successfully genotype a much lower number of SNPs, 38.7–53.8% of markers [27,28]. Nevertheless, microarrays have been successfully used in wildlife research. In the case of animals at threat of extinction or at risk of losing genetic variability, it is extremely important to carry out origin control and avoid interbreeding with other species. One threat can be the occurrence of individuals of two related species or subspecies in the same area. A good example is the situation of the mountain chamois (*Rupicapra rupicapra*), whose two subspecies Tatra (*R. r. tatrica*) and Alpine (*R. r. rupicapra*) occur in close proximity to each other [29]. Thus, there is a potential threat of hybridization between these subspecies, but the difficulty in obtaining biological material results in studies of a small number of individuals. Demontis et al. [29] used Bovine 50SNP to study 4 Tatra and 6 Alpine chamois. Only 151 SNPs met the criteria for analysis, and a genetic structure analysis was conducted based on them, showing a split into two groups. Analogously, in European bison, control of genetic lines is an important part of population management.

High-throughput genotyping methods, such as microarrays, have a growing number of proponents because of their advantages, but the high cost of analysis and the need for good quality and high concentration DNA samples are increasingly leading researchers to use microarrays only for the screening and selection of the most valuable markers for a particular analysis, and then to base research on another, cheaper method. An example of this solution is the development of a set for the individual identification and control of the origin of Japanese Black cattle [22]. The screening was performed using BovineSNP50 microarrays for 50 animals, after which, candidate markers were selected. PCR-RFLP reactions were designed for 60 SNPs, thanks to which another 111 animals were genotyped. Similarly, in the search for markers to distinguish Japanese cattle from foreign breeds, Sasazaki et al. [30] performed preliminary analyses on 110 animals using the BovineSNP50 BeadChip. After the selection of 29 markers, further tests to confirm the correct selection were performed on 554 individuals, also using PCR-RFLP. Assuming that the number of selected markers is small, a cheaper and quite simple method, namely PCR-RFLP, is the obvious choice. However, this is a multi-step and time-consuming method, so it is not suitable for large SNP panels. In this situation, it may be helpful to use allele-specific PCR methods, which can be performed in a single reaction using standard Real-Time PCR instruments. In a preliminary study [14], to test the possibility of using this method in European bison research, reactions were designed using the KASP™ technique—KBiosciences Competitive Allele-Specific PCR (LGC Genomics). The advantage of this method is the fact that the reaction can be performed on samples with lower DNA concentrations than on microarrays, which is extremely important in the case of wisents, especially when hair roots are the only available biological material. Despite the much lower number of markers compared to the first phase of the study [6], the obtained results show that it is possible to use the medium-throughput technique as a tool for routine genotyping.

In this work, we proposed sets of SNP markers for routine genotyping in the population management of European bison. They can be used together or separately, depending on the demand. The sets include markers for detecting hybridization with domestic cattle and American bison, and two separate panels: “Profile”—a panel of 50 SNPs for individuals and parentage identification assessment, as well as “Line”—a panel of 30 SNPs for assessing membership of the Lowland (LB) or Lowland–Caucasian (LC) lines.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/d15010116/s1>. Table S1: markers selected to distinguish between European bison and American bison and to detect hybridization with cattle. Table S2: markers selected for the “Profile” panel. Table S3: markers selected for the “Line” panel. Table S4: private alleles.

**Author Contributions:** Conceptualization, M.W., Z.N.-Ż. and W.O.; methodology, M.W. and Z.N.-Ż.; validation, M.W.; formal analysis, M.W., K.P. and W.D.-C.; investigation, K.P., M.W. and M.K.; resources, M.W. and W.O.; data curation, M.W.; writing—original draft preparation, M.W. and M.P.-T.; writing—review and editing, M.W., M.P.-T., M.K. and W.O.; visualization, M.W.; supervision, W.O.; project administration, M.W.; funding acquisition, W.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Complex Project of European Bison Conservation by State Forests, which is financed by the Forest Fund (Poland), grant number OR.271.3.10.2017.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Olech, W.; Perzanowski, K. European Bison Strategic Species Status *Bison bonasus* Review 2020. IUCN SSC Bison Specialist Group and European Bison Conservation Center: Warsaw, Poland, 2022; p. 1.
- Olech, W. The Genetic Variability within *Bison bonasus* species 90 years after bottleneck. In *Restoration of Endangered and Extinct Animals*; Słomski, R., Ed.; Poznań University of Life Sciences Press: Poznań, Poland, 2010; pp. 48–57.
- Slatis, H.M. An analysis of inbreeding in the European bison. *Genetics* **1960**, *45*, 275. [[CrossRef](#)]
- Pucek, Z.; Belousova, I.P.; Krasińska, M.; Krasiński, Z.A.; Olech, W. European bison. In *Status Survey and Conservation Action Plan*; Pucek, Z., Ed.; IUCN Publication Services Unit: Gland, Switzerland, 2004; p. 54.
- Olech, W. The number of ancestors and their contribution to European bison (*Bison bonasus* L.) population. *Ann. Wars. Agric. Univ. Anim. Sci.* **2000**, *35*, 111–117.
- Wojciechowska, M.; Nowak, Z.; Gurgul, A.; Olech, W.; Drobik, W.; Szmatoła, T. Panel of informative SNP markers for two genetic lines of European bison: Lowland and Lowland–Caucasian. *Anim. Biodivers. Conserv.* **2017**, *40*, 17–25. [[CrossRef](#)]
- Kowalczyk, R.; Ławreszuk, D.; Wójcik, J.M. *Ochrona Żubra w Puszczy Białowieskiej. Zagrożenia i Perspektywy Rozwoju Populacji = European Bison Conservation in the Białowieża Forest: Threats and Prospects of the Population Development*; Zakład Badania Ssaków Polskiej Akademii Nauk: Białowieża, Poland, 2010.
- Raczyński, J.; Bołbot, M. *European Bison Pedigree Studbook 2018*; Białowiecki Park Narodowy: Białowieża, Poland, 2019.
- Marwal, A.; Gaur, R.K. Molecular markers: Tool for genetic analysis. In *Animal Biotechnology*; Academic Press: Cambridge, MA, USA, 2020; pp. 353–372.
- Hayashi, K. PCR-SSCP: A simple and sensitive method for detection of mutations in the genomic DNA. *PCR Methods Appl.* **1991**, *1*, 34–38. [[CrossRef](#)] [[PubMed](#)]
- Kusza, S.; Czyszter, L.T.; Ilie, D.E.; Sauer, M.; Padeanu, I.; Gavojdian, D. Kompetitive Allele Specific PCR (KASP™) genotyping of 48 polymorphisms at different caprine loci in French Alpine and Saanen goat breeds and their association with milk composition. *PeerJ* **2018**, *6*, e4416. [[CrossRef](#)] [[PubMed](#)]
- Kamiński, S.; Olech, W.; Oleński, K.; Nowak, Z.; Ruś, A. Single nucleotide polymorphisms between two lines of European bison (*Bison bonasus*) detected by the use of Illumina Bovine 50 K BeadChip. *Conserv. Genet. Resour.* **2012**, *4*, 311–314. [[CrossRef](#)]
- Ayalew, H.; Tsang, P.W.; Chu, C.; Wang, J.; Liu, S.; Chen, C. Comparison of TaqMan, KASP and rhAmp SNP genotyping platforms in hexaploid wheat. *PLoS ONE* **2019**, *14*, e0217222. [[CrossRef](#)]
- Wojciechowska, M. Genetic Variability Analysis of European Bison *Bison Bonasus* Based on SNP (Single Nucleotide Polymorphism) Markers. Ph.D. Thesis, Warsaw University of Life Sciences, Warsaw, Poland, 2019.
- Boichard, D.; Chung, H.; Dasonneville, R.; David, X.; Eggen, A.; Fritz, S.; Bovine LD Consortium. Design of a bovine low-density SNP array optimized for imputation. *PLoS ONE* **2012**, *7*, e34130. [[CrossRef](#)]
- Peakall, R.; Smouse, P.E. GenAEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—An update. *Bioinformatics* **2012**, *28*, 2537–2539. [[CrossRef](#)]
- Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Sham, P.C. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
- Chang, C.C.; Chow, C.C.; Tellier, L.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets. *GigaScience* **2015**, *4*, s13742-015. [[CrossRef](#)]
- Kalinowski, S.T.; Taper, M.L.; Marshall, T.C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **2007**, *16*, 1099–1106. [[CrossRef](#)] [[PubMed](#)]
- Tokarska, M.; Marshall, T.; Kowalczyk, R.; Wójcik, J.M.; Pertoldi, C.; Kristensen, T.N. Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: The case of European bison. *Heredity* **2009**, *103*, 326–332. [[CrossRef](#)] [[PubMed](#)]
- Oleński, K.; Kamiński, S.; Tokarska, M.; Hering, D.M. Subset of SNPs for parental identification in European bison Lowland-Białowieża line (*Bison bonasus bonasus*). *Conserv. Genet. Resour.* **2018**, *10*, 73–78. [[CrossRef](#)]

22. Hara, K.; Kon, Y.; Sasazaki, S.; Mukai, F.; Mannen, H. Development of novel SNP system for individual and pedigree control in a Japanese Black cattle population using whole-genome genotyping assay. *Anim. Sci. J.* **2010**, *81*, 506–512. [[CrossRef](#)] [[PubMed](#)]
23. Yang, T.; Miller, M.; Forgacs, D.; Derr, J.; Stothard, P. Development of SNP-Based Genomic Tools for the Canadian Bison Industry: Parentage Verification and Subspecies Composition. *Front. Genet.* **2020**, *11*, 585999. [[CrossRef](#)]
24. Waits, L.P.; Luikart, G.; Taberlet, P. Estimating the probability of identity among genotypes in natural populations: Cautions and guidelines. *Mol. Ecol.* **2001**, *10*, 249–256. [[CrossRef](#)]
25. Miller, J.M.; Kijas, J.W.; Heaton, M.P.; McEwan, J.C.; Coltman, D.W. Consistent divergence times and allele sharing measured from cross-species application of SNP chips developed for three domestic species. *Mol. Ecol. Resour.* **2012**, *12*, 1145–1150. [[CrossRef](#)]
26. Decker, J.E.; Pires, J.C.; Conant, G.C.; McKay, S.D.; Heaton, M.P.; Chen, K. Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18644–18649. [[CrossRef](#)]
27. Haynes, G.D.; Latch, E.K. Identification of novel single nucleotide polymorphisms (SNPs) in deer (*Odocoileus* spp.) using the BovineSNP50 BeadChip. *PLoS ONE* **2012**, *7*, e36536. [[CrossRef](#)]
28. Kasarda, R.; Moravčiková, N.; Šidlová, V.; Trakovická, A.; Kadlečík, O.; Pokorádi, J.; Židek, R. Suitability of BovineSNP50 BeadChip for the evaluation of the Cervidae family diversity. *Czech J. Anim. Sci.* **2015**, *60*, 391–399. [[CrossRef](#)]
29. Demontis, D.; Czarnomska, S.D.; Hájková, P.; Zemanová, B.; Bryja, J.; Loeschcke, V.; Pertoldi, C. Characterization of 151 SNPs for population structure analysis of the endangered Tatra chamois (*Rupicapra rupicapra tatraica*) and its relative, the Alpine chamois (*R. r. rupicapra*). *Mamm. Biol.-Z. Für Säugetierkunde* **2011**, *76*, 644–645. [[CrossRef](#)]
30. Sasazaki, S.; Hosokawa, D.; Ishihara, R.; Aihara, H.; Oyama, K.; Mannen, H. Development of discrimination markers between Japanese domestic and imported beef. *Anim. Sci. J.* **2011**, *82*, 67–72. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.