www.nature.com/hdy

ORIGINAL ARTICLE

Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: the case of European bison

M Tokarska[1], T Marshall[2], R Kowalczyk[1], JM Wójcik[1], C Pertoldi[1,3], TN Kristensen[3,4], V Loeschcke[3], VR Gregersen[4] and C Bendixen[4]

[1]Mammal Research Institute, Polish Academy of Sciences, Białowieża, Poland; [2]Field Genetics Limited, London, UK; [3]Department of Biological Sciences, Ecology and Genetics, Aarhus University, Ny Munkegade, Aarhus C, Denmark and [4]Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, University of Aarhus, Tjele, Denmark

The European bison (*Bison bonasus*) has recovered successfully after a severe bottleneck about 90 years ago but has been left with low genetic variability that may substantially hinder parentage and identity analysis. According to pedigree analysis, over 80% of the genes in the contemporary population descend from just two founder animals and inbreeding coefficients averaged almost 0.5, whereas microsatellite heterozygosity does not exceed 0.3. We present a comparison of the effectiveness of 17 microsatellite and 960 single nucleotide polymorphism (SNP) markers for paternity and identity analysis in the European bison. Microsatellite-based paternity and identity analysis was unsuccessful because of low marker heterozygosity and is not a practical approach in this species. Simulations using SNP markers suggest that 80–90 randomly selected loci, or just 50–60 of the most heterozygous loci, would be sufficient to ensure successful paternity and identity analysis in this species. For the purpose of standardizing future analysis, a panel of 50–60 bovine SNPs characterized by high heterozygosity and an even distribution in the genome could be selected. This panel of markers could be typed using VeraCode (Illumina) or similar SNP genotyping systems. The low cost of these SNP genotyping methods compared with a 16 locus microsatellite survey means that off-the-shelf SNP genotyping systems developed for domestic species represent powerful tools for genetic analysis in related species, and can be effective even in bottlenecked species in which heterozygosity of other markers such as microsatellites may be very low.

Heredity (2009) **103**, 326–332; doi:10.1038/hdy.2009.73; published online 22 July 2009

Introduction

The European bison (*Bison bonasus*) became extinct in the Białowieża Forest in 1919, when the last of the animals were probably poached. Out of the 54 remaining animals scattered in zoological gardens and private breeding centres, only 7 became founders of the contemporary Lowland line of the European bison (*B. bonasus bonasus*) population ('pure' European bison). The species was reintroduced into the Białowieża Forest in the northeast of Poland in the early 1950s. Although there were three female and four male founders of the population (Pucek *et al.*, 2004), more than 80% of the contemporary gene pool of the Lowland European bison originates from only two founders: bull Plebejer and cow Planta. Contemporary Lowland bison bulls are all descendants of Plebejer, and hence there is only one Y-chromosome variant in the whole living population (Pucek *et al.*, 2004). The dramatic decline in numbers before extinction in the wild (Krasińska and Krasiński, 2007) and the small number of founders of the captive breeding programme resulted in a high inbreeding coefficient (mean $f = 0.48$ among 1876 animals born after 1945 with full known pedigrees, Olech, 2003). Surprisingly, the high level of inbreeding in the European bison is not manifested by any visible or measurable signs of inbreeding depression (Olech, 2003). Although the number of European bison in the world today is relatively high (including over 1800 animals in the pure-bred Lowland line, and over 3800 animals when hybrid populations are included, Raczyński, 2007), it is considered to be a vulnerable species (IUCN Red List of Threatened Species, 2008) because of its low genetic variability and loss of its natural forest environment.

Parentage testing is a powerful tool in studies of life histories of animals (for example, estimation of breeding success of animals, Wilson *et al.*, 2002) and estimation of the individual inbreeding level (Ritland, 1996). It is also important in conservation management for estimation of effective population size and for reducing the level of inbreeding (Quader, 2005). In fact, in captive breeding, translocation and restocking programmes, substantial efforts are made to ensure that matings between close relatives are

minimized. In cases in which animals have an unknown pedigree, molecular markers can be used to create a relative ranking of degrees of relatedness (Nielsen et al., 2007).

Microsatellite DNA markers have become the most widely used class of genetic markers for a wide range of applications, distinguishing groups and quantifying differences between them from the species level right down to the level of individuals. Microsatellites typically have high heterozygosity and hence provide relatively high statistical power per locus. However, heterozygosity may be substantially reduced in species that have experienced a population bottleneck, and this is the case in the European bison, which has very low microsatellite heterozygosity ($H_E = 0.28$ among 12 microsatellite loci, Tokarska et al., in press). Individual-based analyses such as parentage assignment require relatively high heterozygosity and these types of analyses, in particular, may be expected to be less successful in species that have gone through a recent population bottleneck.

Single nucleotide polymorphisms (SNPs) represent a more recently developed class of markers and are the most abundant class of polymorphic genetic markers in most genomes (Morin et al., 2004). Genome-wide screening based on SNP detection is a new and very promising tool in conservation genetics, particularly for those species with domestic relatives. For example, the BovineSNP50 Genotyping BeadChip (Illumina Inc., San Diego, CA, USA), developed for use in domestic cattle (Bos taurus), allows the analysis of almost 54 000 SNPs across the entire bovine genome. The European bison is closely enough related to domestic cattle such that this off-the-shelf system can be applied directly to amplify SNPs in the European bison. SNPs can be used to survey both neutral (non-coding regions) variation as well as genes under selection (coding regions) in natural populations, providing broader genome coverage compared with mitochondrial DNA or microsatellites (Vignal et al., 2002; Morin et al., 2004). An additional advantage of SNPs compared with microsatellites is that the target DNA sequence in SNP-based genotyping is appreciably shorter (for example, 50–70 bp) than that in microsatellite-based genotyping (80–300 bp), making investigations dealing with degraded DNA easier (Morin et al., 2004). However, typically SNP loci are biallelic, in which case heterozygosity cannot exceed 0.5. Such low heterozygosity is disadvantageous for parentage and identity analysis, which require high statistical power. In this paper, we describe and compare the effectiveness of microsatellite and SNP markers for paternity and identity analysis in the Lowland line of European bison.

Materials and methods

Tissue sampling and DNA extraction
Blood and soft tissues (from culled and immobilized bison), biopsy dart (Dan-Inject, Børkop, Denmark) samples (from live animals) and skull scrapings (from the mammal collection of the Mammal Research Institute in Białowieża, Poland) were collected as DNA sources from animals of the Lowland line of European bison in the Białowieża Primeval Forest, Poland. Nuclear DNA was extracted from blood using BioSprint 96 (Qiagen, Venlo, The Netherlands), or from soft tissue using the DNeasy Blood and Tissue Kit (Qiagen), or from blood

and muscle tissue by treatment with proteinase K followed by sodium chloride precipitation (Sambrook et al., 1989) or from skull scrapings using Chelex 100 (BioRad, Hercules, CA, USA, Cat. No. 142-1253) and the DNeasy Blood and Tissue Kit (Qiagen, Cat. No. 69506).

Genotyping
Microsatellites: Microsatellite genotypes were obtained for 276 bison born between 1985 and 2006. They were typed for 21 microsatellite markers: 20 isolated in domestic cattle and one isolated in reindeer (Rangifer tarandus). These 21 microsatellite markers were chosen as they were previously used for parentage analysis in the North American wood bison (B. bison athabascae) (Wilson et al., 2002). Genotyping was carried out following the procedures described in Tokarska et al. (in press). Genotypes were detected using an ABI3100 automated genetic analyser and analysed using GENEMAPPER 3.5 (Applied Biosystems, Drive Forrest City, CA, USA) and GENEMARKER 1.51 (SoftGenetics, State College, PA, USA).

SNPs: Single nucleotide polymorphism genotypes were obtained for a panel of 50 bison born between 1980 and 2006. Most of these bison were not included in the set of bison genotyped for microsatellite markers. Processing and genotyping of the SNPs were performed using BovineSNP50 BeadChip (Illumina) and BeadStudio software according to the manufacturer's protocol (Infinium II Multi-Sample), as described in Pertoldi et al. (submitted), results available on request from the authors. A total of 52 978 bovine SNP loci were successfully amplified in the European bison. Alleles from scanned bead intensities were scored using BeadStudio software. All the bison SNPs were called using bovine data for cluster separation. Unreliable samples were removed after sample call frequency analysis. All SNPs segregating in the European bison were checked manually to ensure correct calls of clusters. Only bison clusters located within the same range of intensity as for cattle were accepted.

Data analyses
Microsatellites: Basic genetic parameters including allele frequencies, expected heterozygosity (H_E), deviation from Hardy–Weinberg equilibrium (HWE) and null allele frequency were estimated using CERVUS 3.0.3 (Kalinowski et al., 2007). Of the 21 loci screened in the Lowland European bison, 4 were excluded from subsequent analyses: locus NVHRT30 (the one reindeer microsatellite included in the panel) did not follow a Mendelian inheritance pattern, whereas three bovine loci (BM3507, CSSM022 and AGLA269) were monomorphic. This left a total of 17 polymorphic microsatellite loci.

Among these 17 loci, 5 (ETH152, BM1824, BOVFSH, BMC1222 and BM1225) showed highly significant deviations from HWE and/or had an estimated null allele frequency exceeding 0.1. We therefore created two microsatellite data sets: (i) the full set of 17 polymorphic loci, and (ii) a reduced set of 12 'well-behaved' polymorphic loci with no heterozygote deficit.

Paternity tests were performed using CERVUS for 92 offspring born in the years 2004–2006. There were 35 genotyped candidate fathers for each offspring, representing ~50% of the adult bulls potentially participating in reproduction across the whole of the population during the ruts of 2003–2005 when these offspring were

conceived. Six sampled offspring had mothers known from behavioural observation. In the simulation of paternity analysis, the proportion of loci typed was 0.931 for the full set of 17 loci and 0.951 for the reduced set of 12 loci, and the simulated genotyping error rate was set at 0.01 in both cases. Critical values of Delta were determined for 80% and 95% confidence levels based on simulations of 100 000 offspring.

As even the full set of 17 loci had low power to assign parentage, we conducted simulations of paternity analysis to determine how many additional loci would be required to assign paternity reliably. First, we carried out simulations using additional identical sets of 17 loci up to a maximum of 85. Second, we selected the four most heterozygous loci among the 17 and carried out simulations using additional identical sets of these 4 loci in pairs, that is 8, 16, and so on, up to a maximum of 40.

We also used CERVUS to carry out an identity analysis on the microsatellite data set ($n = 276$ animals genotyped at up to 17 polymorphic microsatellite loci), counting the number of pairs of genotypes that were identical at all loci.

SNPs: Basic genetic parameters were computed using CERVUS as for microsatellites. Of the 52 978 cattle SNP loci on the BovineSNP50 BeadChip that amplified in the European bison, 960 loci were polymorphic.

Just 28 significant deviations from HWE at the 5% level were detected among the 623 polymorphic loci for which the HWE test could be carried out, no more than would be expected by chance (the Hardy–Weinberg tests were only carried out on loci for which the expected frequency of rare allele homozygotes exceeded five). We therefore used the whole set of 960 loci in subsequent analysis. The SNP allele frequencies were used to run simulations of paternity analysis using CERVUS. The same parameters were used as for the simulations based on microsatellite markers, except that the proportion of loci typed was 0.996

and the number of offspring in each simulation was reduced to 10 000 to ensure reasonable computation times, given the very large number of loci involved.

As 960 polymorphic SNPs turned out to provide extremely high power to assign paternity, we also evaluated how many loci would be sufficient to ensure the assignment of paternity with 95% confidence. We selected subsets of 480, 240, 120, 60, 30 and 15 loci in two ways: (i) selecting the most polymorphic loci, and (ii) taking a random subset of loci. We repeated the simulations using these subsets.

Although the set of animals in the SNP study was not suited to large-scale paternity analysis, there were three known mother–father–offspring trios in this data set. We used these known relationships to assess the frequency of errors in the SNP genotyping process.

We also used CERVUS to carry out an identity analysis on the SNP data set ($n = 50$ animals), counting the number of pairs of genotypes that matched exactly, and repeated this analysis with progressive smaller subsets of SNP loci as for the paternity simulations. Finally, for six animals, we had two independent DNA samples derived from different tissues (muscle and blood) that we used to assess reproducibility of the SNP genotypes.

Results

Genetic diversity

Microsatellites: Among 17 polymorphic microsatellite loci, we detected 52 alleles among 276 Lowland European bison, an average of 3.06 alleles per locus (range 2–5). The H_E ranged from 0.008 (AGLA232) to 0.654 (BM1225). The mean H_E for the full set of 17 loci was 0.31, and for the reduced set of 12 well-behaved loci it was 0.26 (Table 1). Allele sizes and frequencies are given in Supplementary Table S1.

Table 1 Number of genotypes (N), number of alleles (A), mean expected heterozygosity (H_E), test for Hardy–Weinberg equilibrium (HWE) chi-squared value (χ^2), degrees of freedom (d.f.), P-value (P) and estimated null allele frequency (f_{null}) for 17 microsatellite loci genotyped in a total of 276 European bison from Białowieża Forest (Poland)

| Locus | N | A | H_E | HWE | | | f_{null} |
				χ^2	d.f.	P	
AGLA232	246	2	0.008	Not done			−0.0002
BBJ24	270	2	0.029	Not done			−0.0019
BM723	245	3	0.293	0.8651	1	0.3523	+0.0373
BM757	271	4	0.508	3.5551	1	0.0594	−0.0557
BM1225	266	3	0.654	7.7257	3	0.0520	+0.0527
BM1824	270	3	0.365	0.0100	1	0.9203	+0.0042
BM4440	268	2	0.157	1.3424	1	0.2466	−0.0373
BM5004	274	2	0.097	Not done			−0.0170
RT29	272	3	0.168	0.0481	1	0.8264	+0.0159
TGLA57	269	3	0.176	0.0045	1	0.9467	−0.0105
TGLA126	270	3	0.460	5.1464	3	0.1614	+0.0032
TGLA261	230	4	0.167	2.0709	1	0.1501	+0.0613
Mean (12 loci)	263	2.83	0.257				
BM2830	271	4	0.604	17.2005	3	0.0006	+0.0164
BMC1222	248	3	0.114	Not done			+0.1221
BOVFSH	198	5	0.507	11.0475	1	0.0009	+0.1308
ETH152	272	3	0.466	58.3139	3	<0.0001	+0.0730
TGLA53	247	3	0.478	8.4218	1	0.0037	+0.1010
Mean (17 loci)	258	3.05	0.309				

The first twelve loci are classed as well-behaved loci, whereas the remaining five show significant deviations from HWE and/or a high null allele frequency. Tests for HWE were not done if rare alleles could not be combined in such a way that the minimum expected genotype frequency in the χ^2-test exceeded 1.

SNPs: Among the 960 polymorphic SNP loci, all of which have just two alleles, the mean H_E was 0.31. Minor allele frequencies were almost evenly spread from 0 to 0.5 (mean 0.23) with only a slight tendency towards rarer minor alleles (Supplementary Figure S1).

Paternity analysis

Microsatellites: Paternity analyses were carried out with all 17 polymorphic microsatellite markers and with the reduced set of 12 well-behaved loci that had null allele frequencies of <10% and showed no significant deviation from HWE. With 17 loci, we were able to assign only two paternities with 80% confidence, and with the reduced set of 12 loci we were unable to assign any paternities even at this low level of confidence. The actual paternity analyses were slightly less successful than the predictions of the simulated paternity analysis, but even the success rates predicted by the simulation were very low (Table 2), a consequence of the low heterozygosity of these microsatellite loci in the European bison.

We also conducted simulations to determine how many additional loci would be required to allow paternity analysis to be reasonably successful. If additional loci had the same H_E as the existing set of 17 loci (mean $H_E = 0.31$), a total of 68 loci would be needed to ensure that paternity could always be assigned with 95% confidence in 50% of cases in which the true father was sampled (Figure 1a). On the other hand, if the panel of loci was as heterozygous as the four most heterozygous loci (mean $H_E = 0.57$), a total of 40 loci would be sufficient to give the same level of success (Figure 1b). These figures relate to paternity analysis without known mothers; fewer loci would be required in each case if the mother was known.

For the six known mother–offspring pairs, no errors were detected in the microsatellite genotypes. However, on account of the small number of known relationships and the low heterozygosity of the markers, the power of this analysis to detect errors was low.

SNPs: Simulation of paternity analysis was carried out with 960 SNP loci, and with successively smaller subsets of these loci down to a minimum of 15 loci. When subsets were created using the most heterozygous loci, the mean H_E increased from 0.31 for the full data set to 0.50 among the 15 most heterozygous loci. When subsets were created using randomly selected loci, the mean H_E remained constant at 0.31.

Selecting the most heterozygous SNPs, a total of 50–60 loci would be needed to ensure that paternity could always be assigned with 95% confidence in 50% of cases in which the true father was sampled (Figure 2a). If SNPs were selected at random instead, a total of 80–90 loci would be needed to give the same level of success (Figure 2b). These figures relate to paternity analysis without known mothers; fewer loci would be required in each case if the mother was known.

Among the three known mother–father–offspring trios, no errors were detected in the SNP genotypes. Although the power to detect errors at individual loci was low, the fact that no errors were detected among 960 loci suggests that SNPs can be typed reliably and that the true error rate may be lower than the 1% error rate assumed in the paternity simulations. If the error rate was assumed to be zero, a slightly smaller panel of SNPs would be sufficient for successful paternity analysis (data not shown).

Identity analysis

Microsatellites: Identity analysis revealed 23 pairs of genotypes among 276 animals that were identical across at least 8 loci (mean 13.3) and had no mismatching loci. These matches comprised 20 separate pair-wise combinations of genotypes and one three-way genotype match. Therefore, 43 of the 276 genotypes (16%) were not unique.

SNPs: Identity analysis with 960 SNP loci revealed no genotypes among 50 animals that matched exactly. When the analysis was repeated with either the most heterozygous subset of loci or with randomly selected loci, no matching genotypes were found even when as few as 15 loci were used. Note that as many fewer animals were genotyped for SNPs than for microsatellites, the probability of encountering matching pairs was considerably less in the SNP set independent of any difference in the statistical power of the marker panel.

Across the six animals for which we had two independent DNA samples derived from different tissues, we found no differences in genotypes between samples across the 960 SNP loci, confirming the

Table 2 Number of paternity assignments among 92 European bison offspring, 86 without known mothers and 6 with known mothers

Number of loci	Confidence level of assignment	Father alone			Father given known mother		
		Critical Delta	Observed assignments	Expected assignments	Critical Delta	Observed assignments	Expected assignments
17	95%	5.05	0 (0%)	1 (1%)	5.24	0 (0%)	0.4 (6%)
	80%	3.13	2 (2%)	4 (5%)	3.20	0 (0%)	1.0 (16%)
	Unassigned		84 (98%)	82 (95%)		6 (100%)	5.0 (84%)
	Total		86 (100%)	86 (100%)		6 (100%)	6 (100%)
12	95%	5.30	0 (0%)	0 (0%)	5.58	0 (0%)	0.1 (2%)
	80%	3.51	0 (0%)	1 (1%)	3.62	0 (0%)	0.4 (6%)
	Unassigned		86 (100%)	85 (99%)		6 (100%)	5.6 (94%)
	Total		86 (100%)	86 (100%)		6 (100%)	6 (100%)

Results are shown for the full set of 17 microsatellite loci and the reduced set of 12 loci, respectively. The critical Delta scores and expected number of paternity assignments were determined by simulation of paternity analysis using Cervus 3.0.3 (see text for details).
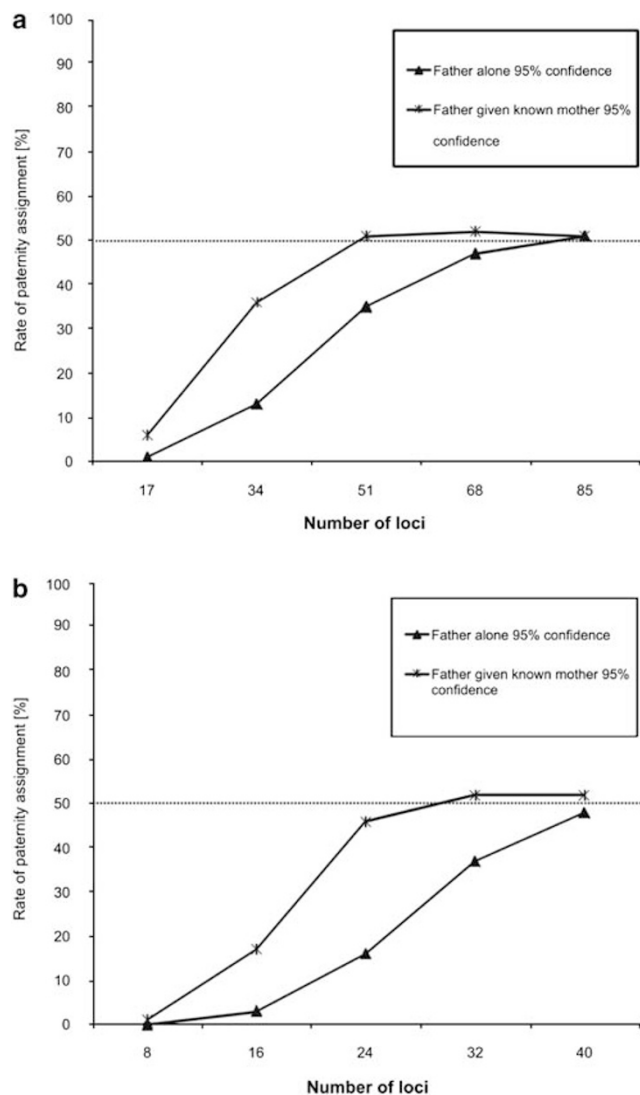
Figure 1 Simulated rates of paternity assignment in the European bison for varying numbers of microsatellite loci. (**a**) 1–5 sets of 17 loci with mean expected heterozygosity, $H_E = 0.31$. (**b**) 2–10 sets of the four most polymorphic loci with mean $H_E = 0.57$. In each panel, the assignment rate is shown separately for assignment of fathers alone and fathers given known mothers at a confidence level of 95%. As only 50% of candidate fathers were sampled in this population, the theoretical maximum assignment rate is 50%. This is shown as a dashed line. In practice, false-positive paternities in cases in which the true father is not sampled allow the assignment rate to go above 50%. As yet, more loci are added until the assignment rate settles at 50%.

Figure 2 Simulated rates of paternity assignment in the European bison for varying number of single nucleotide polymorphism loci: (**a**) most polymorphic subsets of 960 loci, and (**b**) randomly selected subsets of 960 loci. In each panel, the assignment rate is shown separately for assignment of fathers alone and fathers given known mothers. The dashed line represents the theoretical maximum assignment rate of 50%, the proportion of candidate parents sampled (see Figure 1 for further details). Note that unlike Figure 1, the *x* axis is on a log scale.

reproducibility of the data between runs and independent of source tissue type.

Discussion

A problem often associated with the use of SNPs in population studies is ascertainment bias (Nielsen and Signorovitch, 2003; Clark *et al.*, 2005). Bias may be generated by heterogeneity in the SNP discovery process, varying sample sizes or differences in sample composition, and may cause underestimation or over-estimation of the frequency of SNPs (Clark *et al.*, 2005). As the SNP markers used in our survey were selected for
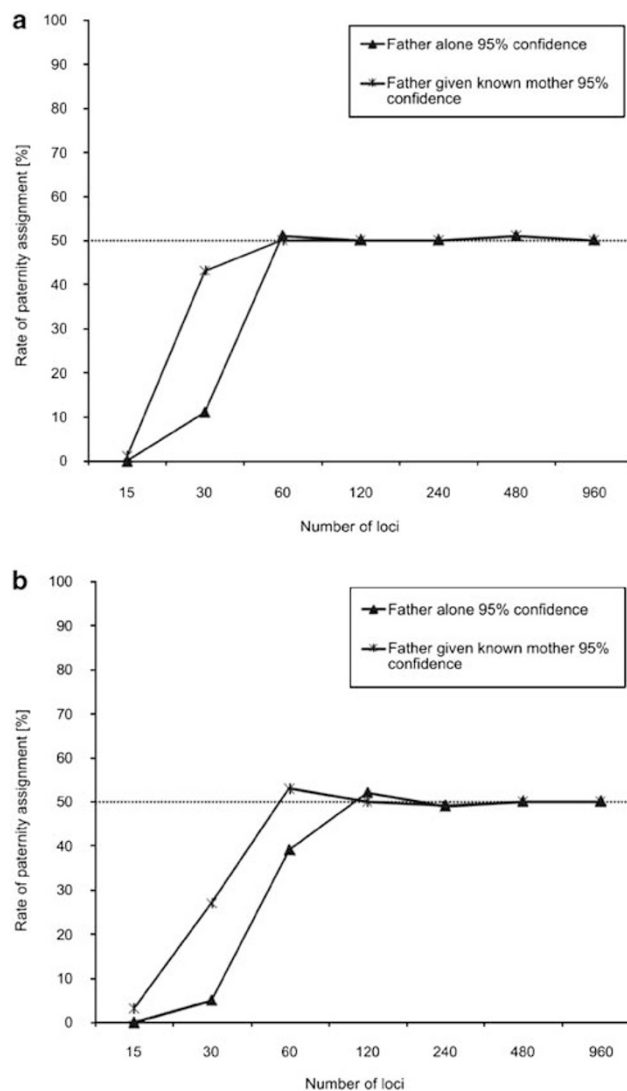
cattle, a possible ascertainment bias could appear when comparing the SNP-based genetic variability of cattle and bison. However, ascertainment bias should not constitute a problem for single-species analyses such as this study.

The aim of this study was to assess the usefulness and limitations of microsatellite markers and SNPs for paternity and identity analysis in a species with extremely low genetic variability. Our study shows that microsatellite heterozygosity is so low in the European bison that neither paternity nor identity analysis can be achieved using a panel of 17 polymorphic microsatellite loci, and that 2–4 times the number of loci would be required for successful paternity analysis. In contrast, a panel of 960 polymorphic SNP loci provided extremely high power to determine the paternity and identity in the

European bison, and just 5–10% of these SNP loci would be sufficient for successful paternity analysis (in general, a set of loci that has sufficient power for parentage analysis also has sufficient power for identity analysis).

The Lowland European bison in this study are descended from just seven founders, some of whom may themselves have been related. Worse, founder contributions are very unequal, with two founders estimated to have contributed >80% of the entire gene pool (Pucek *et al*., 2004). At any given locus, there is an absolute maximum of four distinct alleles in these two founders, and at many loci these two founders are likely to have had one or more alleles in common. It is therefore not surprising that most polymorphic microsatellite loci in the Lowland European bison have just two or three alleles, with three loci having four alleles and just one locus having five alleles. Once the number of alleles falls to two, the H_E is limited to a maximum of 0.5. Reflecting this history, the heterozygosity for Lowland European bison microsatellites (mean $H_E = 0.31$) is less than half of that found in the North American bison or domestic cattle (mean $H_E = 0.67$–0.70) estimated with the same microsatellite panel used in this study (results available on request from the first author).

In the Arabian oryx (*Oryx leucoryx*), a species with a similar population history (extinction in the wild, captive breeding from a small number of founders and subsequent reintroduction to the original habitat), paternity analysis using only six microsatellite loci (mean $H_E = 0.55$) is problematic. However, it is still possible with just six loci to uniquely identify 343 individual Arabian oryx (Marshall *et al*., 1999). In contrast with the European bison, heterozygosity at 17 microsatellite loci was so low that 16% of the 276 genotypes were not unique.

Many factors prevented successful paternity analysis in the European bison: in addition to the extremely low mean H_E, very few offspring had known mothers and only half of the candidate fathers were sampled; 7% of the genotypes were missing and several loci seemed to have null alleles at relatively high frequency (although the heterozygote deficits on which these null allele frequency estimates are based could also be caused by population substructure and/or inbreeding).

A possible solution for successful paternity analysis is simply to genotype more microsatellite loci. There is no need to clone new microsatellites as there are thousands of microsatellite sequences available as a result of the bovine genome mapping/sequencing projects (Barendse *et al*., 1994; Bishop *et al*., 1994; Bovmap Database 2.0 http://locus.jouy.inra.fr/), many of which are likely to amplify polymorphic microsatellites in the European bison. However, genotyping around 70 microsatellite loci across several hundred animals is expensive and time consuming. A refinement of this approach is to run a large number of candidate microsatellite loci in a test panel of 10–20 animals and select only those loci showing the highest heterozygosity for genotyping across the whole population. Simulations suggest that this could almost halve the number of loci that need to be genotyped across the whole population. Nevertheless, the number of microsatellite loci required would far exceed the budget for genotyping in most studies.

Field observations of maternity would also greatly increase the efficiency of paternity analysis. Simulations suggest that the number of loci required would be reduced by 35–40%. Furthermore, increasing the sampling of candidate fathers would allow paternity to be assigned to some of the 50% of offspring whose true fathers are unsampled. If necessary, DNA could be obtained from hair, faeces or post mortem. On the other hand, filling in all gaps in the genetic data only marginally increases the success of paternity analysis (data not shown), and eliminating the loci that deviate from HWE or have a high frequency of null alleles reduces the success of paternity analysis.

The reality is that microsatellite-based parentage analysis may prove almost impossible in species with such low heterozygosity unless a large number of carefully selected microsatellite loci can be combined with near-complete sampling of the parental generation.

The European bison has the great advantage of being closely related to domestic cattle, for which powerful SNP analysis systems have already been developed. Similar systems are available for a range of domestic animal and plant species, and SNP genotyping may well become a method of choice for genetic studies in taxa closely related to domestic species. In species for which off-the-shelf genotyping systems are not suitable, identifying a suitable panel of SNPs takes more effort. Initial studies of SNPs in birds (Primmer *et al*., 2002) and mammals (Aitken *et al*., 2004) suggest that relatively straightforward and inexpensive techniques could generate useful panels of around 50 SNP markers, though there is considerable variation in the success rate between taxa.

On the basis of this study, a panel of 50–60 SNP markers that are highly informative in the European bison could be designed. This SNP panel could be conveniently genotyped using VeraCode (Illumina) or similar technologies at a much reduced cost relative to typical microsatellite genotyping: a panel of 50 bovine SNPs would cost around 20 euros to genotype using the VeraCode system, 50% of the cost of genotyping 16 microsatellite loci in four four-locus multiplexes (around 40 euros: both values were estimated excluding set-up costs and salary costs, whereas microsatellite cost estimation did not also include the costs of optimizing marker panels and replicating genotyping). The usage of an SNP panel as proposed here would allow not only reliable parentage and identity analysis but would also provide a standardized panel for the exchange of genotype data between laboratories.

As the range of off-the-shelf SNP genotyping systems grows, increasing numbers of endangered species (and rare domestic breeds) may fall within their scope. Further research is needed to determine if existing SNP genotyping systems can be applied across whole taxonomic orders or families or are only transferable at the tribe or genus level, and to what extent success depends on taxonomic distance from the source taxon or population history of the study taxon.

Acknowledgements

332

References

Aitken N, Smith S, Schwarz C, Morin PA (2004). Single nucleotide polymorphism (SNP) discovery in mammals: a targeted-gene approach. *Mol Ecol* **13**: 1423–1431.

Barendse W, Armitage SM, Kossarek LM, Shalom A, Kirkpatrick BW, Ryan AM *et al.* (1994). A genetic linkage map of the bovine genome. *Nat Genet* **6**: 227–235.

Bishop MD, Kappes SM, Keele JW (1994). A genetic linkage map for cattle. *Genetics* **136**: 619–639.

Bovmap Database (INRA, France) http://locus.jouy.inra.fr/.

Clark AG, Hubisz MJ, Bustamante CD, Williamson SH, Nielsen R (2005). Ascertainment bias in studies of human genome-wide polymorphism. *Genome Res* **15**: 1496–1502.

IUCN Red List of Threatened Species (2008): http://www.iucnredlist.org/.

Kalinowski ST, Taper ML, Marshall TC (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* **16**: 1099–1106.

Krasińska M, Krasiński Z (2007). *European Bison. The Nature Monograph*. MRI PAS: Białowieża. pp 65–90.

Marshall TC, Sunnucks P, Spalton JA, Greth A, Pemberton JM (1999). Use of genetic data for conservation management: the case of the Arabian oryx. *Anim Conserv* **2**: 269–278.

Morin PA, Luikart G, Wayne RK (2004). SNPs in ecology, evolution and conservation. *Trends Ecol Evol* **19**: 208–216.

Nielsen RK, Pertoldi C, Loeschcke V (2007). Genetic evaluation of the captive breeding program of the Asiatic wild ass, onager *Equus hemionus onager*. *J Zool* **272**: 1–9.

Nielsen R, Signorovitch J (2003). Correcting for ascertainment biases when analyzing SNP data: applications to the estimation of linkage disequilibrium. *Theor Popul Biol* **63**: 245–255.

Olech W (2003). *The influence of individual and maternal inbreeding on calf survival in the European bison (Wpływ inbredu osobniczego i inbredu matki na przeżywalność cieląt żubra, Polish text with English abstract)*. Treatises and monographs, Warsaw University of Life Sciences: Warsaw. pp 32–36.

Pertoldi C, Wójcik JM, Tokarska M, Kawałko A, Kristensen TN, Loeschcke V *et al.* (submitted). Depauperate genome variability in the European and American bison detected using the BovineSNP50 Illumina Genotyping BeadChip. *Biol Direct*.

Primmer CR, Borge T, Lindell J, Saetre GP (2002). Single nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Mol Ecol* **11**: 603–612.

Pucek Z, Belousova IP, Krasińska M, Krasiński ZA, Olech W (2004). European bison. Status Survey and Conservation Action Plan. In: Pucek Z (ed). *IUCN/SSC Bison Specialist Group*. IUCN, Gland: Switzerland and Cambridge, UK. p 8.

Raczyński J (ed) (2007). *European Bison Pedigree Book*. Białowieski Park Narodowy: Białowieża. pp 1–6.

Ritland K (1996). Estimators for pairwise relatedness and individual inbreeding coefficients. *Genet Res* **67**: 175–185.

Quader S (2005). Mate choice and its implications for conservation and management. *Curr Sci* **89**: 1220–1229.

Russello MA, Amato G (2004). Ex situ population management in the absence of pedigree information. *Mol Ecol* **13**: 2829–2840.

Sambrook J, Fritsch EF, Maniatis T (1989.) In: Ford N, Nolan, C, Ferguson, M (eds). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press: New York, USA. pp 9.16–9.19.

Tokarska M, Kawałko A, Wójcik JM, Pertoldi C (in press). Genetic variability in the European bison (*Bison bonasus*) population from Białowieża forest over fifty years. *Biol J Linn Soc*.

Vignal A, Milan D, SanCristobal M, Eggen A (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genet Select Evol* **34**: 275–305.

Wilson GA, Olson W, Strobeck C (2002). Reproductive success in wood bison (*Bison bison athabascae*) established using molecular techniques. *Can J Zool* **80**: 1537–1548.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)