

CALCULATION OF PRP GENOTYPE AND NSP TYPE PROBABILITIES IN SLOVENIAN SHEEP

Gregor GORJANC^{a)}, Marko PLOHL^{b)} and Drago KOMPAN^{c)}

^{a)} Univ. of Ljubljana, Biotechnical Fac., Dept. of Animal Science, Groblje 3, SI-1230 Domžale, Slovenia, e-mail: gregor.gorjanc@bfro.uni-lj.si.

^{b)} Same address as ^{a)}

^{c)} Same address as ^{a)}, Ass.Prof., Ph.D.

Received August 20, 2008, accepted December 06, 2008.

Delo je prispelo 20. augusta 2008, sprejeto 06. decembra 2008.

ABSTRACT

The PrP genotype probabilities in non genotyped Slovenian sheep were calculated. Altogether 36 083 ewes and rams of various breeds were included into analysis. The PrP genotype was known for 10 504 animals. Five different PrP alleles were present in the data set. Pedigree and genotype data structure differed between breeds. Iterative allelic peeling with incomplete penetrance model was used for the calculation of genotype probabilities for each animal given the genotype data of relatives. Analyses were performed for each breed separately. Additionally, NSP (National Scrapie Plan) type probabilities and the average NSP value were calculated from the genotype probabilities. Results were presented for live animals only. There were no animals with additionally identified PrP genotype or NSP type with certainty. The PrP genotype was additionally identified with 95% probability for 0.0 to 5.7% animals of different breeds. NSP type was additionally identified with the same probability for 0.0 to 34.9% animals of different breeds. We assume that the low number of additional identifications was due to: a large number of alleles, intermediate allele frequencies, data structure, a uniform prior, and the use of incomplete penetrance model. Additional identifications provided some cost savings, but did not prove useful in the selection for scrapie resistance of the entire population. The average NSP value should be used instead, since it can be calculated for all animals and encompasses all information from genotype probabilities.

Key words: sheep / genotype probabilities / PrP genotype / NSP type / scrapie / Slovenia

IZRAČUN VERJETNOSTI GENOTIPOV PrP IN SKUPIN NSP PRI OVCAH V SLOVENIJI

IZVLEČEK

Izračunali smo verjetnosti genotipov PrP za negenotipizirane ovce v Sloveniji. V analizo smo zajeli 36 083 ovc in ovnov različnih pasem. Genotip PrP je bil znan za 10 504 živalih. V podatkih je bilo prisotnih pet različnih alelov PrP. Struktura porekla in podatkov o genotipu PrP se je med pasmami razlikovala. Z metodo alelnega lupljenja in modelom nepopolne penetrance, smo na podlagi genotipa sorodnikov iterativno izračunali verjetnosti genotipov PrP za vse živali. Analize smo opravili za vsako pasmo posebej. Iz verjetnosti genotipov smo preračunali verjetnosti za skupine NSP (National Scrapie Plan) ter iz teh povprečno vrednost NSP. Rezultate smo predstavili le na živih živalih. Prav nobeni živali nismo uspeli z gotovostjo dodatno določiti genotipa PrP ali skupine NSP. Za 0,0 do 5,7 % živali različnih pasem smo lahko dodatno določili genotip PrP s 95 % verjetnostjo, medtem ko smo lahko z enako verjetnostjo dodatno določili skupino NSP za 0,0 do 34,9 % živali različnih pasem. Menimo, da so vzroki za majhno število dodatnih določitev genotipa PrP sledeči: veliko število alelov, intermediarne frekvence alelov,

struktura podatkov, uniformna apriorna verjetnost in uporaba modela nepopolne penetrance. Dodatne določitve genotipa PrP in skupine NSP predstavljajo prihranek za rejski program. Zaradi majhnega števila dodatnih določitev metoda ni uporabna za selekcijo celotnih populacij na odpornost proti praskavcu. V ta namen lahko uporabimo povprečno vrednost NSP, saj ta parameter združuje vse verjetnosti genotipov PrP in ga lahko izračunamo za vse živali.

Ključne besede: ovce / verjetnosti genotipov / genotip PrP / skupina NSP / praskavec / Slovenija

INTRODUCTION

Scrapie is a transmissible spongiform encephalopathy (TSE) disease of sheep, like BSE in cattle. In the past, scrapie was not known in sheep in Slovenia, but in the last few years some cases have been identified. Transmissible spongiform encephalopathies are a group of diseases, where a changed form of prion protein accumulates in the central nervous system (Prusiner, 1998). The primary cause of these diseases has not been discovered yet. Research results have shown that in the same flock only some animals die due to scrapie, while others do not. This led to the detection of PrP genotype (also marked as PrnP), which plays a major role in genetic predisposition for susceptibility to scrapie (Hunter, 1997). Differences at codons 136, 154 and 171 codon of the PrP gene are strongly associated with the resistance or susceptibility to scrapie and with the age of the disease onset (Hunter, 1997). Because the mentioned codons are very close, the polymorphisms in these three codons (haplotypes) are often reported as an allele named by the initial letters of amino acids, such as A136R154R171, or just ARR. The most frequent alleles are: ARR, AHQ, ARH, ARQ and VRQ. New alleles are still being discovered (Lühken *et al.*, 2008; Ulvund, 2008). Allele ARR is associated with the highest resistance to scrapie, while allele VRQ is associated with the highest susceptibility to scrapie infection (Hunter, 1997). PrP genotypes are grouped into 5 groups based on scrapie susceptibility. These groups are usually called NSP types after National Scrapie Plan in UK (Dawson *et al.*, 2008). NSP types (Table 1) follow from 1 (the highest resistance) to 5 (the highest susceptibility). Ulvund (2008) provides a recent review about scrapie in sheep.

Table 1. Classification of PrP genotypes by NSP type (Dawson *et al.*, 2008)
Preglednica 1. Razvrstitev genotipov PrP v skupine NSP (Dawson in sod., 2008)

NSP type Skupina NSP	PrP genotype Genotip PrP
1	ARR/ARR
2	ARR/ARQ, ARR/AHQ, ARR/ARH
3	AHQ/AHQ, AHQ/ARH, AHQ/ARQ, ARH/ARH, ARH/ARQ, ARQ/ARQ
4	ARR/VRQ
5	AHQ/VRQ, ARH/VRQ, ARQ/VRQ, VRQ/VRQ

Scrapie in sheep has been known for more than 250 years, with no evidence of harmful effects on humans so far. The European Commission has accepted the regulation 2003/100/EC (European Commission, 2003), due to the incidence of BSE in cattle, possible spontaneous transmission of scrapie from sheep to cattle. At that time there was also a lack of tools to distinguish between BSE and scrapie in sheep. The regulation requires the acceptance of selection programme for increasing scrapie resistance in sheep in all member states. The aim of the programme is the elimination of the VRQ allele and the increase of the ARR allele

frequency. In Slovenia, such programme was initiated in the year 2005. The programme is based on the selection of performance recorded rams and ewes genotyped for the PrP gene.

Both, genotyping as well as tissue collection represent a considerable cost for the breeding programme. However, the genotype of an individual can be (partially) determined from the genotype of his relatives, given that the pedigree and genotype data are accurate. If the parents are homozygous, the genotype of an offspring can be determined with certainty, i.e., parents with ARQ/ARQ and ARR/ARR genotypes will definitely have offspring with ARR/ARQ genotype, given that there are no mutations or data errors. Definite genotype identification from the genotypes of relatives is possible only in special cases. However, genotype probabilities for the individual can be calculated from the genotype data collected on relatives (e.g. Jacquard, 1974). Elston and Stewart (1971) developed a general method for the calculation of genotype probabilities. They used their method in the context of human genetics. Livestock pedigrees are usually much larger and more complex due to inbreeding, multiple matings, and a larger number of offspring per parent (see Cannings *et al.*, 1978). Several methods were developed to handle these complexities (e.g. Van Arendonk *et al.*, 1989; Janss *et al.* 1995; Kerr and Kinghorn, 1996; Fernández *et al.*, 2001; Thallman *et al.*, 2001a; Henshall and Tier, 2003; Gengler *et al.*, 2007). The aim of this paper is to present the results of the calculation of PrP genotype probabilities and NSP type probabilities in Slovenian sheep using allelic peeling method (Thallman *et al.*, 2001a) as proposed by Gorjanc and Kompan (2008).

MATERIAL AND METHODS

Material

Seven breeds from the Slovenian sheep breeding programme were included in the analysis: Jezersko-Solcava (JS), Improved Jezersko-Solcava (JSR), Bela Krajina Pramenka (BP), Texel (T), Bovec (B), Improved Bovec (VFB) and Istrian Pramenka (IP). Data on PrP genotype were collected in the frame of the selection programme for TSE (scrapie) resistance in Slovenia. More than 10 000 ewes and rams (Table 2) were genotyped (at codons 136, 154 and 171 of the PrP gene) by the end of the year 2007.

Pedigree data was also included into the analysis to provide relationship information. Initially, the complete pedigree for each breed was obtained, considering all living and dead animals registered in the herdbook. Some animals did not contribute information for the calculation of genotype probabilities due to missing genotype data or non informative links within the pedigree. Non-informative animals were excluded (pruned) iteratively generation by generation in the direction from ancestors to descendants. The following criteria had to be met for the exclusion of an animal: known date of death or culling, only one offspring, and no PrP genotype data. All the data (PrP genotype data and pedigree) were obtained from the sheep and goat database of the Centre for extension service at the Animal Science Department of Biotechnical Faculty.

The largest pedigree was obtained for JSR breed (15 054; Table 2), followed by JS breed with 10 429 animals and B breed with 5 101 animals. Pedigrees of other breeds were smaller, according to the size of breed populations in Slovenia. The highest percentage of genotyped animals was in BP breed (61.2%; Table 2). The percentage of genotyped animals was also high in IP (43.9%), B (36.5%), and JS breed (35.2%). In other breeds the percentage of genotyped animals was lower than 30%. Animals with many relatives are very informative for the calculation of genotype probabilities. The highest percentage of genotyped sires was in IP breed (41.0%). Almost every third sire was also genotyped in JS (28.7%) and BP (27.0%) breed. The percentage of genotyped dams was generally much higher than for sires (Table 2). However,

dams of T breed mostly did not have PrP genotype data, because tissue collection programme has not been introduced yet in the flocks of T breed. The highest number of live non-genotyped animals was in the most numerous breed, the JSR breed (3795 animals; Table 2). JS breed followed with 2673 live non-genotyped animals. However, JS breed had much narrower (better) ratio (0.7 for JS and 1.4 for JSR breed) between the number of live non-genotyped animals and the number of all genotyped animals. In other breeds the number of non-genotyped animals was lower and the ratio did not exceed the value of 0.3, with the exception of T breed (Table 2).

Table 2. Data structure
Preglednica 2. Struktura podatkov

Breed Pasma	Animals Živali	Known genotype, % Znan genotip, %			Number of live non-genotyped animals Št. živih negenotipiziranih živali	Ratio ¹ Razmerje ¹
		Animals Živali	Sires Očetje	Dams Matere		
JS	10 429	35.2	28.7	40.9	2673	0.7
JSR	15 054	18.4	11.8	20.0	3795	1.4
BP	1297	61.2	27.0	70.0	159	0.2
T	555	16.8	15.4	1.0	189	2.0
B	5101	36.5	12.8	42.4	522	0.3
VFB	1711	26.7	8.9	23.8	128	0.2
IP	1936	43.9	41.0	48.7	263	0.3

JS – Jezersko-Solcava; JSR – Improved Jezersko-Solcava; BP – Bela Krajina pramenka; T – Texel; B – Bovec; VFB – Improved Bovec; IP – Istrian pramenka; ¹ Ratio between the number of live non-genotyped animals and the number of all genotyped animals / ¹ Razmerje med številom živih negenotipiziranih živali in vseh genotipiziranih živali

Methods

The presented PrP genotype data and pedigree data were used in the calculation of PrP genotype probabilities for the non-genotyped animals. These calculations are based on probability laws of gene segregation given the known genotype data and pedigree structure. We used the allelic peeling method (Thallman *et al.*, 2001a) as implemented in the GenoProb programme (Thallman, 2002), where genotype probabilities of an individual animal are calculated as a function of genotype probabilities in ancestors and offspring. A uniform prior distribution was used for allele probabilities in founders (animals with unknown ancestors). Exact calculations were not possible, due to complex pedigrees. We used the iterative procedure as presented by Thallman *et al.* (2001b). Exploratory runs revealed that 100 iterations were enough to achieve convergence in all breeds, except in BP breed where around 250 iterations were needed.

Pedigree and genotype data might contain errors due to various reasons. We used the incomplete penetrance model to enable the use of all data (Thallman *et al.*, 2001b). With incomplete penetrance model the observed genotype data was conceptually treated as “phenotype” data (Lincoln and Lander, 1992). Inconsistencies between “phenotype” data and calculated genotype probabilities were governed by the penetrance function (Thallman *et al.*, 2001b). We used a uniform penetrance function with the error rate set to 0.1.

We obtained a vector \mathbf{g} with 15 PrP genotype probabilities [1] for each animal. These probabilities were summed into a vector \mathbf{n} of five NSP type probabilities [2] with regard to the resistance/susceptibility to scrapie infection (Table 1).

$$\mathbf{g}' = [\text{Pr}(ARR/ARR), \text{Pr}(ARR/AHQ), \dots, \text{Pr}(VRQ/VRQ)] \quad [1]$$

$$\mathbf{n}' = [\text{Pr}(NSP_1), \text{Pr}(NSP_2), \dots, \text{Pr}(NSP_5)] \quad [2]$$

Bearing in mind the resistance to scrapie it is desirable that the animals are of as low NSP type as possible. As shown by Gorjanc and Kompan (2008) the average NSP value was calculated as an average of NSP type values weighted with NSP type probabilities [3]. We also calculated the average NSP value for the entire population, where the NSP type probabilities were derived from the frequencies of PrP genotype in the entire population.

$$\overline{NSP} = \mathbf{a}' * \mathbf{n} \quad [3]$$

RESULTS AND DISCUSSION

Allele frequency

Besides the number of animals with genotype data and pedigree structure, allele frequency is also important in the calculation of genotype probabilities (Tier and Henshall, 2005). Six alleles (haplotypes) of a PrP gene have been identified so far in Slovenia: ARR, AHQ, ARH; ARQ; VRQ and VRR. The VRR allele has been identified in one case only and it has been excluded from the analysis. Allele frequencies (Table 3) were calculated as simple percentages, i.e., relationship dependencies among the collected genotypes of relatives were not taken into the consideration (e.g. Bohenke, 1991). For all breeds, except for T, a high frequency of allele ARQ was observed (from 50.0 to 69.7%). The frequency of VRQ allele (associated with the high susceptibility to scrapie) was less than 5% in all breeds. The ARR allele (associated with the high resistance to scrapie) was more frequent than the VRQ allele. The frequency of this favourable allele was 52.7% in T breed, around 35.0% in BP and IP breed, and around 20% in JS, JSR, B and VFB breed. The frequency of AHQ and ARH alleles was quite diverse and ranged between 0.1 and 25.5%. Lühken *et al.* (2008) estimated allele frequencies for 56 breeds from 15 European and Near East countries. They showed that there is a considerable variation between breeds. Allele frequencies in breeds in Slovenia are within the range of values reported by Lühken *et al.* (2008).

Table 3. PrP allele frequencies, %
Preglednica 3. Frekvence alelov PrP, %

Breed Pasma	Allele / Alel				
	ARR	AHQ	ARH	ARQ	VRQ
JS	17.4	7.4	8.3	63.2	3.7
JSR	17.8	7.9	0.9	69.7	3.8
BP	35.0	4.3	0.6	58.6	1.4
T	52.7	4.8	12.4	26.3	3.8
B	16.9	17.4	7.6	57.1	1.0
VFB	22.0	24.6	2.5	50.0	0.9
IP	32.5	7.5	0.1	57.2	2.6

JS – Jezersko-Solcava; JSR – Improved Jezersko-Solcava; BP – Bela Krajina pramenka; T – Texel; B – Bovec; VFB – Improved Bovec; IP – Istrian pramenka

PrP genotype and NSP type identification

From here onwards, the results are presented for live animals only, because only the living animals are of interest for the selection process. However, all known animals (dead and alive) were used in the process of the calculation of genotype probabilities. We calculated genotype probabilities for all 15 possible PrP genotypes for each animal. The success of analysis was evaluated with the number of additionally identified or excluded PrP genotypes and NSP types. If the probability of a particular genotype was 100%, then that genotype was identified with certainty. The success was assessed also for the probability of identification of 99 and 95%. In case of five alleles, fifteen genotypes are possible. Therefore, the ratio between the number of confirmed and excluded genotypes was 1:14. This means that for a particular animal we could confirm one genotype and exclude 14 genotypes.

Table 4. Number of additional PrP genotype identifications or exclusions
Preglednica 4. Število dodatno potrjenih ali ovrženih genotipov PrP

Breed Pasma	Additional PrP genotype identifications Dodatno potrjenih genotipov PrP			Additional PrP genotype exclusions Dodatno ovrženih genotipov PrP		
	Probability Verjetnost	1.00	0.99	0.95	1.00	0.99
JS						
No. / Št.	0	66	101	0	2045	2673
%	0	2.5	3.8	0	5.1	6.7
JSR						
No. / Št.	0	105	145	0	2243	3795
%	0	2.8	3.8	0	3.9	6.7
BP						
No. / Št.	0	0	9	0	137	159
%	0	0.0	5.7	0	5.7	6.7
T						
No. / Št.	0	0	0	0	161	189
%	0	0.0	0.0	0	5.7	6.7
B						
No. / Št.	0	11	22	0	478	522
%	0	2.1	4.2	0	6.1	6.7
VFB						
No. / Št.	0	7	7	0	122	128
%	0	5.5	5.5	0	6.4	6.7
IP						
No. / Št.	0	10	14	0	219	263
%	0	3.8	5.3	0	5.6	6.7

JS–Jezersko-Solcava; JSR–Improved Jezersko-Solcava; BP–Bela Krajina pramenka; T–Texel; B–Bovec; VFB–Improved Bovec; IP–Istrian pramenka

We were not able to identify any additional PrP genotype or NSP type with certainty not even for one animal in any breed (Table 4 and 5). This is inherently due to the use of incomplete penetrance model as discussed by Gorjanc and Kompan (2008). Results were only a bit better (data not shown) when the success was evaluated on all (alive and dead) animals in the pedigree. However, those results are of no practical significance, since dead animals are not interesting for selection. The number of animals with additionally identified or excluded PrP genotype was just slightly higher at 99 or 95% probability of identification (Table 4). The percentage of animals with additionally identified PrP genotype with 95% probability ranged from 0.0 to 5.7% of animals (Table 4). For all breeds together, we additionally identified PrP genotype (with 95% probability) in 298 animals. There were about 4 to 6% of excluded PrP genotypes with 99%

probability and surprisingly 6.7% of excluded PrP genotypes with 95% probability for all breeds. Success of the analysis practically did not differ between breeds, even though the data structure was diverse (Table 2).

The success was better with the additional identification of NSP types (Table 5). The highest percentage of animals with additionally identified NSP type with 95% probability were achieved in B breed (34.9%) and VFB breed (25.8%). In other breeds the percentages were mostly lower than 10% (Table 5). Considering all breeds together, we additionally identified NSP type in 561 animals with 99% probability or 818 animals with 95% probability. The percentage of excluded NSP type with 95% probability varied between 15 and 20%, i.e., around three to four times higher in comparison to excluded PrP genotypes (Table 4 and 5).

The absolute number of additionally identified PrP genotypes and NSP types is not negligible, as there is no need to genotype these animals. Since tissue collection and genotyping are still of considerable cost, additional identifications provide economical savings for the breeding programme. Nevertheless, we can conclude that the percentage of additional genotype identifications was very low. Different data structure between breeds (Table 2) had practically no effect. In what follows, we discuss reasons for such a low number of additional identifications. We also discuss a novel parameter (average NSP type), which encompasses all information from genotype probabilities and can be used effectively in the selection for scrapie resistance.

Table 5. Number of additional NSP type identifications or exclusions
Preglednica 5. Število dodatno potrjenih ali ovrženih skupin NSP

Breed Pasma	Additional NSP type identifications Dodatno potrjenih skupin NSP			Additional NSP type exclusions Dodatno ovrženih skupin NSP		
	1.00	0.99	0.95	1.00	0.99	0.95
Probability Verjetnost						
JS						
No. / Št.	0	209	283	0	1583	2488
%	0	7.8	10.6	0	11.8	18.6
JSR						
No. / Št.	0	213	293	0	1469	3402
%	0	5.6	7.7	0	7.7	17.9
BP						
No. / Št.	0	2	11	0	123	148
%	0	1.2	6.9	0	15.5	18.6
T						
No. / Št.	0	0	1	0	75	158
%	0	0.0	0.5	0	7.9	16.7
B						
No. / Št.	0	101	182	0	456	512
%	0	19.3	34.9	0	17.5	19.6
VFB						
No. / Št.	0	26	33	0	111	123
%	0	20.3	25.8	0	17.3	19.2
IP						
No. / Št.	0	10	15	0	164	223
%	0	3.8	5.7	0	12.5	17.0

JS – Jezersko-Solcava; JSR – Improved Jezersko-Solcava; BP – Bela Krajina pramenka; T – Texel; B – Bovec; VFB – Improved Bovec; IP – Istrian pramenka

First of all, five alleles of PrP gene were included into the analysis: ARR, AHQ, ARH, ARQ and VRQ. These alleles are the most frequent in our as well as in other populations of sheep (Lühken *et al.*, 2008; Ulvund, 2008). Five alleles give rise to 15 different genotypes, which is not

negligible. According to Tier and Henshall (2005), Elston and Stewart (1971) developed a method for the calculation of genotype probabilities in human genetics, where pedigrees under inspection were limited to a smaller number of families and loci with rare lethal alleles. In such cases, the aim was to determine the carriers of lethal alleles with the use of genotype probabilities. Tier and Henshall (2005) assessed the limits of additional genotype identifications for beef cattle, sheep and pigs. They concluded that better results were obtained in cases with a small number of alleles having diverse frequencies. This is in accordance with the data in human genetics that inspired Elson and Stewart (1971). In our study, the situation was opposite. We had higher number of alleles (five) with intermediate frequencies (Table 3). Kinghorn (1999) assessed the percentage of additional genotype identifications. He reported that in the case of two alleles, complete penetrance model, and 10, 20, and 80% of genotyped animals, 50, 60, and even 100% of additional identifications could be achieved, respectively. We also could merge several alleles into two groups, i.e., ARR and other alleles. However, this is not optimal as some flocks did not have any animal carrying ARR allele, but they had a substantial variation for other alleles.

The next limitation in the view of larger number of identifications was used prior for allele probabilities in founders. Prior probabilities are the initial allele probabilities for the animals at the top of the pedigree. Thallman (2002) proposed a uniform distribution for allele probabilities in founders, resulting in our data. This proposal is the handiest, as it is the least informative. Prior probabilities in founders are used in the calculation of genotype probabilities via the Bayes theorem of conditional probability and are as such subsequently used in the calculation of genotype probabilities for all relatives in the pedigree. Until the collected genotype data (the likelihood) do not dominate the genotype probabilities (the posterior) the effect of a prior does not vanish (e.g. Gelman *et al.*, 2004, Sorensen and Gianola, 2007). The uniform prior therefore adds to the uncertainty in genotype probabilities and lowers the number of additional PrP genotype identifications with certainty or high probability. Results by Lühken *et al.* (2008) and many others show that some PrP alleles are more frequent than others. For example ARQ allele is the most frequent in the majority of breeds and it seems logical to assign higher prior probability to ARQ allele than say VRQ allele. Instead of uniform allele probabilities, the sample allele frequencies could be used as a prior information. However, Tier and Henshall (2002) warned against such use, because the allele frequencies in founders can be entirely different to the ones in the collected data. Furthermore, the allele frequencies can also be markedly different between families (Tier and Henshall, 2002). The estimated allele frequencies in founders (e.g. Bohenke, 1991) could be used instead. Kerr and Kinghorn (1996) have established, using simulation, that the uniform prior probability decreases the number of erroneous genotype exclusions. Gorjanc and Kompan (2008) argued that it is of paramount importance that inference about VRQ/* genotypes is correct and recommended the uniform prior as a conservative choice.

In Slovenia, the selection programme for scrapie has started in the year 2005. Regarding this fact, we can divide animals in our study into three groups. The first group represents genotyped animals. The majority of these animals are still alive. The second group represents non-genotyped ancestors that are already dead. Finally, the third group represents live non-genotyped animals. Our aim was to identify the PrP genotype for animals in the third group. Animals of the first group were used as a primary source of genotype information, while animals of the second group provided additional link between the first and the third group through the pedigree. Kerr and Kinghorn (1996) have shown that the success is much lower in such scenario compared to the one, where genotype data spans over generations.

The use of incomplete penetrance model (Thallman *et al.*, 2001b) also added to the uncertainty and therefore lowered the number of additional PrP genotype identifications. Errors in genotype and/or pedigree data can cause deviations between laboratory reported genotype and genotype inferred from genotype data of relatives. Erroneous data can be corrected, but it is often

hard to determine which record is erroneous in the case of large and complex pedigrees with incomplete genotype data at one locus with multiple alleles. Thallman *et al.* (2001b) used a model of incomplete penetrance in order to conceptually treat genotype data as “phenotype” data via penetrance function (Lincoln and Lander, 1992). This enables the use of data with potential errors. Unfortunately, the incomplete penetrance model adds to the uncertainty of the results as manifested in small probabilities for all genotypes that are not consistent with the “phenotype”. These small probabilities are also transferred throughout the pedigree as it is the case with prior allele probabilities in the founders. Consequently, the number of additional genotype identifications is lowered (Table 4 and 5). Gorjanc and Kompan (2008) showed the effect of use of incomplete penetrance model on calculated PrP genotype probabilities and the number of additional PrP genotype identifications. Nevertheless, Thallman *et al.* (2001b), Gengler *et al.* (2007), and Gorjanc and Kompan (2008) argued that incomplete penetrance is very useful or even essential for the applied work on field data in animal breeding.

Average NSP value

The selection for scrapie resistance in sheep is based on the elimination of the VRQ allele and on the increase of the ARR allele frequency. In order to apply selection to the whole population of performance recorded sheep, we tried to identify the PrP genotype and NSP type for non-genotyped animals. Unfortunately, we were not able to identify (with certainty or high probability) PrP genotype and NSP type for the majority of non-genotyped animals. Instead of singly identified PrP genotype and NSP type, all PrP genotype probabilities can be used for the selection process. Van Arendonk *et al.* (1989) have given similar suggestion, but for a gene with only two alleles. Work with probabilities for all fifteen PrP genotypes for each animal is, of course, cumbersome. Therefore, we used the average NSP value as suggested by Gorjanc and Kompan (2008). The average NSP value is a weighted average of NSP type values weighted with NSP type probabilities derived from PrP genotype probabilities. Therefore, this parameter combines all information in PrP genotype probabilities into a single value.

Distribution of the average NSP value for the non-genotyped animals agreed to a great extent with the distribution of NSP type for the genotyped animals in all sheep breeds (Fig. 1). Similar distribution between non-genotyped and genotyped animals is expected because the PrP genotype probabilities and subsequently NSP types were derived from the collected genotype data. In some breeds peaks were observed at values around 1.5, 2, 2.5, 3, 3.5, 4, and 5. This can be attributed to the small number of NSP types (five) and discrete NSP type values.

To increase resistance to scrapie, breeders should select animals having low NSP type value. Apart from selecting only among genotyped animals, non-genotyped animals could also be included in the selection process with the means of average NSP value. Breeders should select animals with average NSP value below the population average NSP value (dashed line in Fig. 1). As seen in Table 6, the use of average NSP value increased the pool of animals for selection by 2034 animals in all breeds together. The increase varied between breeds; from 19.0% in T breed to 45.9% in BP breed (Table 6). The increased pool of animals inherently provides more space for selection for other economically important traits in sheep and lowers the rate of inbreeding. The later is very important for the autochthonous rare breeds: Bela Krajina pramenka, Istrian pramenka, and Bovec breed. Jezersko-Solcava breed is also autochthonous but the size of the population (performance recorded flocks and others flocks) is sufficiently large to avoid close inbreeding.

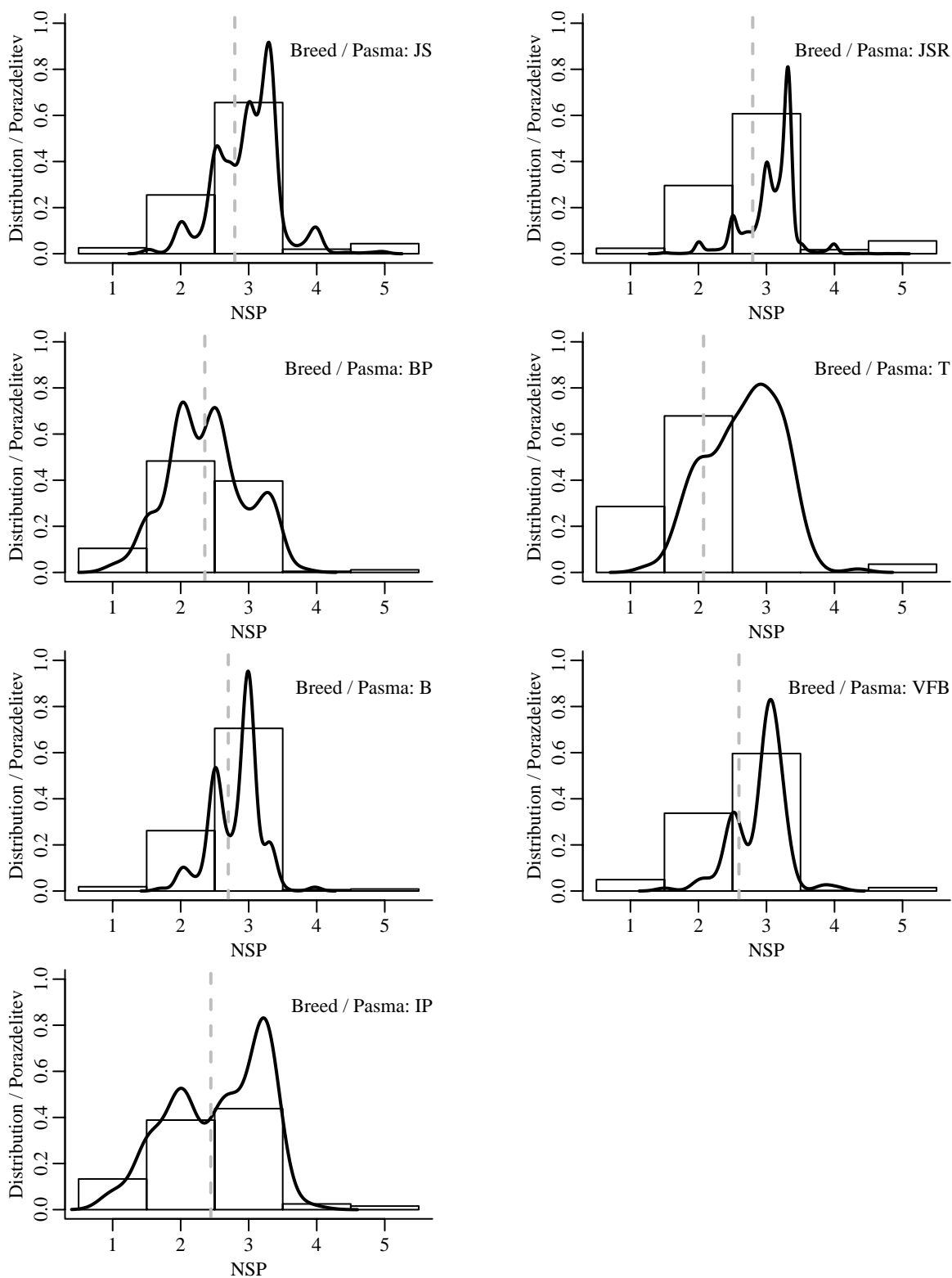


Figure 1. Distribution of animals by average NSP value bars – genotyped animals, superimposed curve – non-genotyped animals, vertical dashed line –average NSP value in population

Slika 1. Porazdelitev živali glede na povprečno vrednost NSP: stolpci – genotipizirane živali, naložena krivulja – negenotipizirane živali, navpična črtkana črta – populacijska povprečna vrednost NSP

Table 6. Live non-genotyped animals bellow population average NSP value
 Preglednica 6. Žive negenotipizirane živali pod populacijskim povprečjem vrednosti NSP

Breed Pasma	Number of live non-genotyped animals Število živih negenotipiziranih živali	$\overline{NSP}_i < \overline{NSP}_p$	
		No. / Št.	%
JS	2673	879	32.9
JSR	3795	726	19.1
BP	159	73	45.9
T	189	36	19.0
B	522	185	35.4
VFB	128	32	25.0
IP	263	103	39.2

JS – Jezersko-Solcava; JSR – Improved Jezersko-Solcava; BP – Bela Krajina pramenka; T – Texel; B – Bovec; VFB – Improved Bovec; IP – Istrian pramenka

Besides the additional genotype identifications and selection, genotype probabilities can be used for the calculation of genotype probability index (Kinghorn, 1997). This index is used to create a list of animals for the next round of genotyping in such a way that the newly identified genotypes will provide maximal gain of information in the pedigree. Use of this index can therefore help breeders to use resources in a more efficient way. Additionally, genotype probabilities can also be used in association analyses to increase the statistical power as has been, for example, shown by Vitezica *et al.* (2005).

CONCLUSIONS

PrP genotype probabilities and NSP type probabilities were calculated for non-genotyped ewes and rams. The success of PrP genotype or NSP type identification was low. We maintain that the main reasons for these results are: a high number of alleles, intermediate allele frequencies, data structure (genotype data known for animals of recent generations only), a uniform prior, and the use of incomplete penetrance model. Nevertheless, PrP genotype probabilities can be used for the calculation of the average NSP value, which is a useful and practical parameter in the selection for scrapie resistance of the entire population.

ACKNOWLEDGEMENTS

We want to express our gratitude to Jurij Pohar, Ph. D. for valuable comments on an earlier draft of this paper.

REFERENCES

- Bohenke, M. Allele frequency estimation from data on relatives. *Am. J. Hum. Gen.*, 48(1991), 22–25.
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1682747> (31. mar. 2008)
- Cannings, C./ Thompson, E.A./ Skolnick, M.H. Probability functions on complex pedigrees. *Adv. Appl. Prob.*, 10(1978), 26–61.
- Dawson, M./ Moore, R.C./ Bishop, S.C. Progress and limits of PrP gene selection policy. *Vet. Res.*, 39(2008), 25–37. <http://dx.doi.org/10.1051/vetres:2007064> (10.mar. 2008)
- European Commission, 2003/100/EC: Commission Decision of 13 February 2003 laying down minimum requirements for the establishment of breeding programmes for resistance to transmissible spongiform encephalopathies in sheep. *Official Journal*, L04(2003), 0041–0045.
<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32003D0100:EN:HTML> (31. mar. 2008)

- Elston, R.C./ Stewart, J. A general model for the genetic analysis of pedigree data. *Hum. Hered.*, 21(1971), 523–542.
- Fernández, S.A./ Fernando, R.L./ Guldbrandtsen, B./ Totir, L.R./ Carriquiry, A.L. Sampling genotypes in large pedigrees with loops. *Genet. Sel. Evol.*, 33(2001), 337–367.
<http://dx.doi.org/10.1051/gse:2001122> (31. mar. 2008)
- Gelman, A./ Carlin, J.B./ Stern, H.S./ Rubin, D.B. Bayesian data analysis. Boca Raton, Chapman & Hall/CRC, 2004, 668 p.
- Gengler, N./ Mayeres, P./ Szydlowski, P. A simple method to approximate gene content in large pedigree populations: application to the myostatin gene in dual-purpose Belgian Blue cattle. *Animal*, 1(2007), 21–28.
<http://dx.doi.org/10.1017/S1751731107392628> (31. mar. 2008)
- Gorjanc, G./ Kompan, D. Use of genotype probabilities in selection on PrP genotype in sheep. *Liv. Sci.*, 2008, submitted.
- Henshall, J.M./ Tier, B. An algorithm for sampling descent graphs in large complex pedigrees efficiently. *Genet. Res.*, 81(2003), 205–212. <http://dx.doi.org/10.1017/S0016672303006232> (31. mar. 2008)
- Hunter, N. PrP genetics in sheep and the implications for scrapie and BSE. *Trends in Microbiol.*, 5(1997), 331–334.
[http://dx.doi.org/10.1016/S0966-842X\(97\)01081-0](http://dx.doi.org/10.1016/S0966-842X(97)01081-0) (31. mar. 2008)
- Jacquard, A. The genetic structure of populations. Berlin, Springer-Verlag, 1974, 569 p.
- Janss, L.L.G./ Van Arendonk, J.A.M./ Van der Werf, J.H.J. Computing approximate monogenic model likelihoods in large pedigrees with loops. *Genet. Sel. Evol.*, 27(1995), 567–579.
http://www.gse-journal.org/articles/gse/pdf/1995/06/GSE_0999-193X_1995_27_6_ART0007.pdf
- Kerr, R.J./ Kinghorn, B.P. An efficient algorithm for segregation analysis in large populations. *J. Anim. Breed. Genet.*, 113(1996), 457–469.
- Kinghorn, B.P. An Index of Information Content for Genotype Probabilities Derived from Segregation Analysis. *Genetics*, 145(1997), 479–483. <http://www.genetics.org/cgi/content/abstract/145/2/479> (31. mar. 2008)
- Kinghorn, B.P. Use of segregation analysis to reduce genotyping costs. *J. Anim. Breed. Genet.*, 116(1999), 175–180. <http://dx.doi.org/10.1046/j.1439-0388.1999.00192.x> (31. mar. 2008)
- Lincoln, S.E./ Lander, E.S. Systematic detection of errors in genetic linkage data. *Genomics*, 14(1992), 604–610.
[http://dx.doi.org/10.1016/So888-7543\(05\)80158-2](http://dx.doi.org/10.1016/So888-7543(05)80158-2) (31. mar. 2008)
- Lühken, G./ Lipsky, S./ Peter, C./ Erhardt, G. Prion protein polymorphisms in autochthonous European sheep breeds in respect to scrapie eradication in affected flocks. *Small. Rum. Res.*, 75(2008), 43–47.
<http://dx.doi.org/10.1016/j.smallrumres.2007.07.010> (31. mar. 2008)
- Prusiner, S.B. Prions. *Proc. Natl. Acad. USA*, 95(1998), 13363–13383.
<http://www.pnas.org/cgi/content/abstract/95/23/13363> (31. mar. 2008)
- Sorensen, D./ Gianola, D. Likelihood, Bayesian and MCMC Methods in Quantitative Genetics. New York, Springer, 2007, 760 p.
- Thallman, R.M./ Bennet, G.L./ Keele, J.W./ Kappes, S.M. Efficient computation of genotype probabilities for loci with many alleles: I. Allelic peeling. *J. Anim. Sci.*, 79(2001a), 26–33.
<http://jas.fass.org/cgi/content/abstract/79/1/26> (31. mar. 2008)
- Thallman, R.M./ Bennet, G.L./ Keele, J.W./ Kappes, S.M. Efficient computation of genotype probabilities for loci with many alleles: II. Iterative method for large complex pedigrees. *J. Anim. Sci.*, 79(2001b), 34–44.
<http://jas.fass.org/cgi/content/abstract/79/1/34> (31. mar. 2008)
- Thallman, R.M. User's Manual for GenoProb Version 2.0. Computation of genotype and phase probabilities in complex pedigrees by iterative allelic peeling. 2002, 1–18.
- Tier, B./ Henshall, J.M. Founder allele frequencies and segregation analysis. In: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, 2002-8-19/23. Castanet-Tolosan, INRA, cop, 2002, Communication, 16–17.
- Tier, B./ Henshall, J.M. Limits to genotypic probabilities for single loci. In: Application of new genetic technologies to animal breeding. Proceedings of the 16th Conference of the Association for the Advancement of Animal Breeding and Genetics, Noosa, 2005-09-25/28, Australia, 2005, 366–369.
- Ulvund, M.J. Ovine scrapie disease: Do we have to live with it? *Small Rum. Res.*, 76(2008), 131–140.
<http://dx.doi.org/10.1016/j.smallrumres.2007.12.004> (31. mar. 2008)
- Van Arendonk, J.A.M./ Smith, C./ Kennedy, B.W. Method to estimate genotype probabilities at individual loci in farm livestock. *Theor. Appl. Genet.*, 78(1989), 735–740. <http://dx.doi.org/10.1007/BF00262571> (31. mar. 2008)
- Vitezica, Z.G./ Elsen, J.M./ Rupp, R./ Diaz, C. Using genotype probabilities in survival analysis: a scrapie case. *Genet. Sel. Evol.*, 37(2005), 403–415. <http://dx.doi.org/10.1051/gse:2005008> (31. mar. 2008)