Incorporating Mouse Genome
Mammalian Genome
Genes and Phenotypes

# Prion gene sequence variation within diverse groups of U.S. sheep, beef cattle, and deer

**Michael P. Heaton,[1] Kreg A. Leymaster,[1] Brad A. Freking,[1] Deedra A. Hawk,[2] Timothy P.L. Smith,[1] John W. Keele,[1] Warren M. Snelling,[1] James M. Fox,[1] Carol G. Chitko-McKown,[1] William W. Laegreid[1]**

[1]USDA, ARS, U.S. Meat Animal Research Center (MARC), State Spur 18D, P.O. Box 166, Clay Center, Nebraska 68933-0166, USA
[2]Wyoming Game and Fish Department, P.O. Box 3312, University Station, Laramie, Wyoming 82071, USA

*Received: 27 March 2003 / Accepted: 18 June 2003*

***Abstract***

Prions are proteins that play a central role in transmissible spongiform encephalopathies in a variety of mammals. Among the most notable prion disorders in ungulates are scrapie in sheep, bovine spongiform encephalopathy in cattle, and chronic wasting disease in deer. Single nucleotide polymorphisms in the sheep prion gene (*PRNP*) have been correlated with susceptibility to natural scrapie in some populations. Similar correlations have not been reported in cattle or deer; however, characterization of *PRNP* nucleotide diversity in those species is incomplete. This report describes nucleotide sequence variation and frequency estimates for the *PRNP* locus within diverse groups of U.S. sheep, U.S. beef cattle, and free-ranging deer (*Odocoileus virginianus* and *O. hemionus* from Wyoming). DNA segments corresponding to the complete prion coding sequence and a 596-bp portion of the *PRNP* promoter region were amplified and sequenced from DNA panels with 90 sheep, 96 cattle, and 94 deer. Each panel was designed to contain the most diverse germplasm available from their respective populations to facilitate polymorphism detection. Sequence comparisons identified a total of 86 polymorphisms. Previously unreported polymorphisms were identified in sheep (9), cattle (13), and deer (32). The number of individuals sampled within each population was sufficient to detect more than 95% of all alleles present at a frequency greater than 0.02. The estimation of *PRNP* allele and genotype frequencies within these diverse groups of sheep, cattle, and deer provides a framework for designing accurate genotype assays for use in genetic epidemiology, allele management, and disease control.

Correspondence to: M.P. Heaton; E-mail: heaton@email.marc.usda.gov

Transmissible spongiform encephalopathies (TSEs) are a heterogeneous group of fatal neurodegenerative disorders characterized by changes in trafficking and conformation of the mammalian prion protein (PrP) [for review see (Prusiner 1998)]. Natural prion diseases may occur as genetic, infectious, or sporadic disorders in a variety of mammals, most notably in humans, mink, sheep, cattle, and deer. Cattle with bovine spongiform encephalopathy (BSE) have been implicated in one human TSE, variant Creutzfeldt-Jakob disease (vCJD) (Scott et al. 1999), through the consumption of beef from affected animals. Consequently, many countries are developing policies aimed at eliminating TSE-affected animals from their food chains. Key components of TSE eradication programs include: excluding ruminant meat and bone meal from animal feed, identifying and disposing of animals with clinical disease, tracing infected animals to their population source, culling in endemic areas, and restocking with animals from TSE-free regions. Individuals with genetic resistance to TSE are desired for restocking because disease recurrence from exposure to environmental prion contamination may occur in endemic areas (Thorgeirsdottir et al. 1999). Thus, identifying genetic variation correlated with TSE resistance is an important step in plans to eliminate TSEs from the food chain.

Various PrP isoforms may influence TSEs susceptibility. In sheep, 14 single nucleotide polymorphisms (SNPs) in the prion gene (*PRNP*) coding

region have been published and/or reported in GenBank. Nucleotide variants affecting the translation of codons 136, 154, and 171 are the most often-studied polymorphisms associated with variation in susceptibility to scrapie, the TSE of sheep. *PRNP* haplotype alleles encoding alanine, arginine, and arginine (ARR) at the respective 136, 154, and 171 positions are correlated with increased scrapie resistance, whereas valine, arginine, and glutamine (VRQ) haplotypes are correlated with increased scrapie susceptibility (Belt et al. 1995; Hunter et al. 1996; Baylis et al. 2002). The most resistant individuals are those with homozygous *ARR/ARR* genotypes, and the most susceptible are the VRQ/VRQ individuals. The scrapie susceptibility for four other known *PRNP* haplotype alleles (*ARQ*, *AHQ*, *ARH*, *ARK*) is less clear because susceptibility appears to vary somewhat among breeds and populations. Nevertheless, individuals with two alleles encoding glutamine at codon 171 (i.e.,171QQ) are susceptible to natural scrapie (Goldmann et al. 1994; Westaway et al. 1994; O'Rourke et al. 1997). *PRNP* alleles encoding arginine at position 171 are dominant for increased scrapie resistance, and thus, heterozygous individuals (171QR) are considered resistant, with few exceptions (Ikeda et al. 1995; Tranulis et al. 1999; Baylis et al. 2002). The apparent dominant resistance conferred by the 171R allele implies that flocks derived from 171RR founders are expected to remain scrapie-free, even in environments where exposure to the infectious form of PrP is likely.

BSE in cattle and chronic wasting disease (CWD) in deer are similar prion diseases in species where relatively few *PRNP* polymorphisms have been characterized. A comparison of *PRNP* coding sequence (CDS) within *Bos taurus* or *Odocoileus* spp. in GenBank shows five and nine nucleotide differences, respectively, including insertion/deletion (indel) polymorphisms. The polymorphisms in cattle have not been significantly correlated with BSE (Hunter et al. 1994; Neibergs et al. 1994; Hernandez-Sanchez et al. 2002), and the correlation of deer polymorphisms with CWD is unresolved. In the human *PRNP* gene, both intronic and upstream regulatory regions appear to influence susceptibility to CJD (McCormack et al. 2002). However, polymorphisms have not been described in the *PRNP* promoter region in ungulates. Three bovine SNPs have been reported in the 5′ untranslated regions containing exons 1a and 1b (Humeny et al. 2002). Informative markers spanning both the promoter and coding regions may be useful for extended haplotype analysis of the *PRNP* gene locus.

Although TSEs occur in goats (Billinis et al. 2002), elk (Williams and Young 1992; Williams and Miller 2002), moufflon (Wood et al. 1992), kudu, and many other exotic captive ungulates (Kirkwood and Cunningham 1994), cattle and free-ranging deer populations were chosen as the focus of the present study because an extended set of polymorphic markers is needed for epidemiologic studies involving BSE and CWD. The aim of this study was to describe the nucleotide diversity in the promoter and coding regions of the *PRNP* locus in diverse populations of healthy U.S. sheep, U.S. cattle, and free-ranging deer (*O. virginianus* and *O. hemionus* from Wyoming).

## Materials and methods

***Animal groups and genomic DNA samples.*** Three different panels of ruminant DNAs were analyzed in the present study: the U.S. Meat Animal Research Center (MARC) Sheep Diversity Panel (MSDP) version 1.1 (Freking et al. 2002); the MARC Beef Cattle Diversity Panel (MBCDP) version 2.1 (Heaton et al. 2001); and a sample of white-tailed deer (*O. virginianus*) and mule deer (*O. hemionus*) populations from Wyoming (MARC *Odocoileus* panel version 1.0, this report). Each panel was designed to contain the most diverse germplasm available and, where possible, represents wide ranges of performance for a variety of economically important traits.

The sheep panel consisted of DNA from 90 individuals representing nine genetically diverse breeds of sheep. Ten rams were sampled for each breed, with no rams sharing a common sire. The breeds were divided into four classifications: 1) general purpose breeds including Dorset, Rambouillet, and Texel; 2) terminal-sire breeds including Suffolk and Composite III [1/4 Suffolk, 1/4 Hampshire, and 1/2 Columbia (Leymaster 1991)]; 3) prolific breeds including Finnsheep and Romanov; and 4) hair-shedding breeds including Dorper and Katahdin. These breeds are presently being evaluated for a wide range of performance traits at MARC and represent a diverse cross section of popular U.S. sheep germplasm. Germplasms from other popular breeds, such as Hampshire and Columbia, are significant components of the Composite III breed and thus contribute alleles to the group.

The cattle panel consisted of 92 sires from 16 popular beef breeds and four sires from the Holstein dairy breed. Sires within each breed were selected for pedigrees with minimal relationships between ancestors, to maximize the total number of unshared haploid genomes. The beef breeds in this panel comprise greater than 99% of the germplasm used in the U.S. beef cattle industry, based on the number of registered progeny for each breed.

The deer panel consisted of DNA from 50 white-tailed and 43 mule deer sampled from wildlife management areas across Wyoming from 1996 to 1998. White-tailed deer were sampled from five management areas in northeast Wyoming (Region A, Areas 2–5, and Region B, Area 7; htttp://gf.state.wy.us/HTML/afs/pdf/NR03deermap.pdf), and mule deer were sampled from 14 management areas across the state (Regions A, B, D, K, and T). These 14 areas included four of the same areas where white-tailed deer were sampled (Region A, Areas 2, 3, 5; Region B, Area 7). One mule deer from south-central Nebraska was also included in this panel. DNA was extracted from tissues and arrayed in 96-well plates as previously described (Heaton et al. 2001).

Estimating the minimum allele frequency required for detection in the above panels was based on the probability of observing the allele at least once in an animal group, as previously described (Heaton et al. 2001). Briefly, the probability of observing an allele at least once is $1 - (1 - p)^n$ where "p" is the frequency of the allele and "n" is the number of independent samplings or, in this case, the number of unshared haploid genomes for diploid organisms. This assumes that samplings (haploid genomes) are independent and identically distributed (e.g., the same p applies to all animal subpopulations). Setting power, or the probability of observing the allele at least once, to 0.95 results in the equation: $0.95 = 1 - (1 - p)^n$. Solving this equation for p yields $p = 1 - (0.05)^{1/n}$ for all p between 0 and 1. Based on ancestors in pedigrees with at least four to seven generations present, the cattle panel was estimated to contain 187 unshared haploid genomes and is expected to allow a 95% probability of detecting any allele with a frequency greater than 0.016 in the panel (Heaton et al. 2001). Individuals in the sheep and deer panels were also selected for minimal relationships; however, pedigree information was not available for estimating the respective number of unshared haploid genomes. If one assumes that 10% of the haploid genomes are shared among the individuals in the respective sheep and deer panels, the minimum allele frequency that would allow 95% probability of detection is less than 0.02 for each panel.

***Selection of PRNP regions for analysis, PCR amplification, and DNA sequencing.*** PCR cocktails were designed to generate *PRNP* fragments containing as much CDS as possible for cattle, sheep, and deer. Primers used for amplification of genomic DNA are listed in Table 1. Oligonucleotide primers used in PCR have the effect of "remodeling" the template in the binding site, causing polymorphisms in those regions to be unreadable. Thus, it was desirable to have minimum overlap between the primer binding sites and the CDS. The sense amplification primer binding sites overlapped the first 5 nt in cattle and sheep and 12 nt in deer on the 5′ end of the CDS. Consequently, any polymorphisms present in these first few nucleotides corresponding to the N-terminus of the PrP leader signal peptide (MV in cattle and sheep; MVKS in deer) were not measured in our experiments. A larger fragment was not designed for amplification because a number of stem-loop structures with high melting temperatures in the adjacent 5′ non-coding region were predicted to interfere with PCR primers. Beyond the stop codon on the 3′ end of the *PRNP* gene, there were 49 nt in sheep, 33 nt in cattle, and 31 nt in deer that were included in the PCR products and did not overlap the binding sites for the antisense amplification primer. Thus, nucleotide variation in the 3′ end of the *PRNP* CDS was measured in these experiments. In the *PRNP* promoter region, PCR cocktails were designed to generate an approximately 600-bp fragment centered in a 4.3-kb region of the bovine prion gene. This 4.3-kb region was previously shown to have promoter activity (Lemaire-Vieille et al. 2000). With the location of this bovine amplicon as a reference point, additional oligonucleotides were designed to generate the corresponding amplicons in sheep and deer.

A standard amplification reaction contained 50 ng of genomic DNA, 0.5 μM of each amplification primer, 200 μM of each dNTP, 1.5 mM $MgCl_2$, 1.25 U of either HotStarTaq DNA polymerase (Qiagen, Inc, Valencia, Calif.) or Thermo-Start DNA polymerase (ABgene, Epsom, UK) and 10% vol/vol reaction buffer provided by the manufacturer in a total volume of 55 μL. PCR was performed with either the PTC-200, the PTC 220 Dyad, or the PTC225 Tetrad thermal cycler chassis (MJ Research, Watertown, Mass.). Reactions were denatured at 94°C for 15 min and subjected to 45 cycles of denaturation at 94°C for 20 s, annealing at the appropriate temperature (Table 1) for 30 s, and a 72°C extension for 1 min. After cycling, an additional 3-min incubation at 72°C was included before storage at 4°C. A 5-μL portion of each amplified product was analyzed by agarose gel electrophoresis in buffer containing 90 mM Tris-borate (pH 8.0), 2 mM ethylenediamine tetraacetic acid, and 0.1 μg/mL ethidium bromide.

Both DNA strands of each amplicon were sequenced at least once with either the amplification primers or nested sequencing primers as previously described (Grosse et al. 1999). Sequencing reactions for the 280 individuals of three panels were

**Table 1. Oligonucleotides for *PRNP* amplification and DNA sequencing**

| *Organism* | *DNA segment* | *Gene region* | *Oligonucleotide* | *DNA sequence*[a] | *Orientation* | *Function* | *PCR annealing temp. (°C)* | *Amplicon length (bp)*[b] |
|---|---|---|---|---|---|---|---|---|
| Sheep | OARPRNPDS4 | Promoter | U67922:2976U21 | ctc acc att tca gaa tac ctc | Sense | Amplification/ sequencing | 58 | 593 |
| Sheep | OARPRNPDS4 | Promoter | U67922:3552L17 | gca aat ggc cct aca tc | Antisense | Amplification/ sequencing | 58 | 593 |
| Sheep | OARPRNPDS1 | Coding | U67922:22262U21 | ttg cag aga agt cat cat ggt | Sense | Amplification/ sequencing | 58 | 836 |
| Sheep | OARPRNPDS1 | Coding | U67922:23082L16 | ccc caa cct ggc aaa g | Antisense | Amplification/ sequencing | 58 | 836 |
| Sheep | OARPRNPDS1 | Coding | U67922:22572U20 | cac agt cag tgg aac aag cc | Sense | Sequencing | –[c] | – |
| Sheep | OARPRNPDS1 | Coding | U67922:22711L21 | aac ggt cct cat agt cat tgc | Antisense | Sequencing | – | – |
| Cattle | BTAPRNPDS4 | Promoter | AF163764:1653U20 | ctc acc att tcc gaa tac at | Sense | Sequencing | 58 | 596 |
| Cattle | BTAPRNPDS4 | Promoter | AF163764:2232L17 | gca aat ggc cct aca tc | Antisense | Sequencing | 58 | 596 |
| Cattle | BTAPRNPDS4 | Promoter | AF163764:1675U18 | tct acc acc aag ccc cat | Sense | Sequencing | – | – |
| Cattle | BTAPRNPDS4 | Promoter | AF163764:2205L24 | cct cat ctt agt ctc tac cat cac | Antisense | Sequencing | – | – |
| Cattle | BTAPRNPDS1 | Coding | AB001468:144U23 | cag cag ata taa gtc atc atg gt | Sense | Amplification/ sequencing | 55 | 863 |
| Cattle | BTAPRNPDS1 | Coding | AB001468:990L17 | ccc cca acc tgg taa ag | Antisense | Amplification/ sequencing | 55 | 863 |
| Cattle | BTAPRNPDS1 | Coding | AB001468:446U18 | aca tgg tgg tgg agg ctg | Sense | Sequencing | – | – |
| Cattle | BTAPRNPDS1 | Coding | AB001468:667L21 | gcc tgt agt aca ctt ggt tgg | Antisense | Sequencing | – | – |
| Deer | ODOPRNPDS4 | Promoter | U67922:2976U21 | ctc acc att tca gaa tac ctc | Sense | Amplification/ sequencing | 58 | 567 |
| Deer | ODOPRNPDS4 | Promoter | U67922:3552L17 | gca aat ggc cct aca tc | Antisense | Amplification/ sequencing | 58 | 567 |
| Deer | ODOPRNPDS5 | Promoter | ODOPRNPDS4:28U17 | cac caa gcc cca cca tt | Sense | Amplification/ sequencing | 55 | 512 |
| Deer | ODOPRNPDS5 | Promoter | ODOPRNPDS4:529L22 | att cct cat ctc agt ctc tac c | Antisense | Amplification/ sequencing | 55 | 512 |
| Deer | ODOPRNPDS5 | Promoter | ODOPRNPDS4:32U16 | aag ccc cac cat tgc c | Sense | Sequencing | – | – |
| Deer | ODOPRNPDS1 | Coding | AF156185:21U22 | ata agt cat cat ggt gaa aag c | Sense | Amplification/ sequencing | 58 | 831 |
| Deer | ODOPRNPDS1 | Coding | AB001468:988L19 | ccc cca acc tgg taa aga t | Antisense | Amplification/ sequencing | 58 | 831 |
| Deer | ODOPRNPDS1 | Coding | ODOPRNPDS1:364U19 | atg aag cat gtg gca gga g | Sense | Sequencing | – | – |
| Deer | ODOPRNPDS1 | Coding | ODOPRNPDS1:391L18 | ccc tac cac tgc tcc agc | Antisense | Sequencing | – | – |

[a] Oligonucleotide primer sequences are listed in the 5′ to 3′ direction. Primers for amplifying DNA segments OARPRNPDS4/DS1, BTAPRNPDS4, and BTAPRNPDS1 were derived from GenBank reference sequences U67922 (Lee et al. 1998), AF163764 (Lemaire-Vieillle et al. 2000), and AB001468 (Yoshimoto et al. 1992), respectively. Primers for amplifying ODOPRNPDS4 were derived from sheep and cattle and subsequently redesigned specifically for deer based on the preliminary sequence generated. Similarly, both deer sequence (AF156185; Raymond et al. 2000) and cattle sequence (Yoshimoto et al. 1992) were used to design the primers for amplifying DNA segment ODOPRNPDS1.

[b] The amplicon length listed corresponds to the longest alleles known for the species; i.e., it includes the insert sequences when indel polymorphisms are known.

[c] Not applicable.

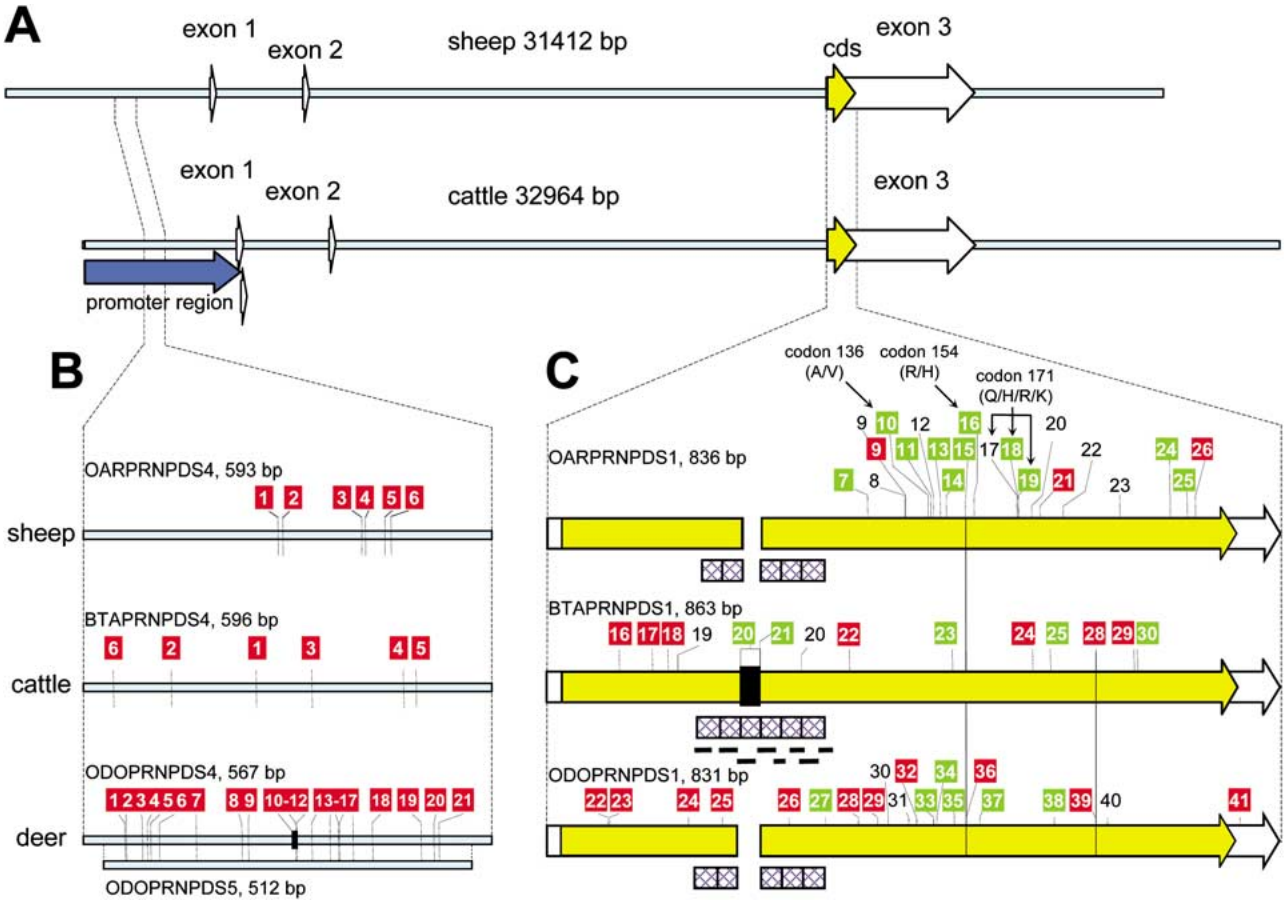


**Fig. 1.** Physical maps of *PRNP* polymorphic sites in sheep, cattle, and deer. DNA sequences were obtained and analyzed as described in the Materials and methods. Panel A: *PRNP* genomic regions derived from GenBank accessions U67922 (Lee et al. 1998), AF163764 (Lemaire-Vieille et al. 2000), D10612 (Yoshimoto et al. 1992), and AJ298878 (Hills et al. 2001). Panels B and C: DNA segments amplified from the promoter and CDS regions by PCR. The relative position of the SNPs are indicated with vertical tick marks, and the numbers correspond to the last two digits in the SNP identifier column of Tables 2 through 4. The symbol legend for the feature maps is as follows: red squares, polymorphisms identified in this report but not previously reported in GenBank or the published literature; green squares, polymorphisms observed both in this report and GenBank/published literature; black numbers, polymorphisms identified in GenBank or published literature but not observed in the present study; vertical lines connecting two numbers, SNPs present in two species; yellow arrows, CDS regions; white arrows, non-coding exon regions; blue arrow, promoter region; thin light blue rectangles, intron or intergenic regions; crosshatched squares, octapeptide repeat regions; thick black lines beneath octapeptide repeat regions; significant stem loop regions (72–99°C Tm); and black rectangles in promoter or CDS regions, indels.

performed according to the manufacturer's instructions with BigDye terminator chemistry (version 2.0) and resolved on an ABI PRISM 3700 DNA analyzer (PE Applied Biosystems, Foster City, Calif.). For amplicons spanning the *PRNP* CDS, additional pairs of outward-facing sequencing primers were designed to hybridize to conserved regions in the center of the amplicons to ensure that DNA sequences near the ends of the CDS were accurately determined. All of the resulting sequences were analyzed with the assistance of PolyPhred software (Nickerson et al. 1997) in conjunction with Phred/Phrap/Consed software (Ewing and Green 1998; Ewing et al. 1998; Gordon et al. 1998), and consensus sequences were constructed for each group of sheep, cattle, and deer as previously described (Heaton et al. 2001). Because the geographic ranges of white-tailed and mule deer overlap, and they are known to interbreed in the wild, a single consensus sequence was constructed for these *Odocoileus* spp.

The consensus sequence for each group of sheep, cattle, or deer contained monomorphic sites that were unique for the respective group and thus provided a control for inadvertent amplification of contaminating DNA from other species. Annotated consensus sequences for the amplicons have been deposited in GenBank (see Tables 2 through 4 for accession numbers).

**Table 2.** Allele and genotype frequencies of *PRNP* gene polymorphisms in the MARC Sheep Diversity Panel Version 1.1

| *OARPRNP DNA segment* | *SNP identifier* | *PRNP* region | *Alleles (1,2)* | *Allele frequencies*[a] | | *Genotype frequencies* | | | *Genbank accession* | *Reference* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *1* | *2* | *1,1* | *1,2* | *2,2* | | |
| 4 | MOS001-1 | Promoter | atgaa-(C,T)-gaaat | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | AY326428 | This report |
| 4 | MOS001-2 | Promoter | gaaat-(A,T)-tttct | 0.64 | 0.36 | 0.49 | 0.30 | 0.21 | AY326428 | This report |
| 4 | MOS001-3 | Promoter | tgtgg-(G,A)-agatR | 0.96 | 0.04 | 0.91 | 0.09 | 0.00 | AY326428 | This report |
| 4 | MOS001-4 | Promoter | Ragat-(G,A)-agaag | 0.88 | 0.12 | 0.79 | 0.19 | 0.02 | AY326428 | This report |
| 4 | MOS001-5 | Promoter | gtctt-(C,T)-atggt | 0.96 | 0.04 | 0.91 | 0.09 | 0.00 | AY326428 | This report |
| 4 | MOS001-6 | Promoter | tgcca-(T,C)-aacct | 0.96 | 0.04 | 0.91 | 0.09 | 0.00 | AY326428 | This report |
| 1 | MOS001-7 | CDS, M112T | caaca-(T,C)-gaagc | 0.95 | 0.05 | 0.91 | 0.09 | 0.00 | AJ000735 | Laplanche et al. 1993 |
| 1 | MOS001-8 | CDS, G127AVS[b,c] | taggg-(G,A)-Bcctt | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | n/a[d] | Gombojav et al. 2003 |
| 1 | MOS001-9 | CDS, G127AVS[c,e,f] | agggR-(G,C)-ccttg | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | AY326330 | This report |
| 1 | MOS001-9 | CDS, G127AVS[b,c,f] | agggR-(G,T)-ccttg | 0.99 | 0.00 | 0.99 | 0.00 | 0.00 | n/a | Gombojav et al. 2003 |
| 1 | MOS001-10 | CDS, A136V | aagtg-(C,T)-caYga | 0.97 | 0.03 | 0.94 | 0.06 | 0.00 | AJ000738.1 | Goldmann et al. 1991 |
| 1 | MOS001-11 | CDS, M137T | tgYca-(T,C)-gaRca | 0.97 | 0.03 | 0.95 | 0.04 | 0.01 | AJ000679 | Bossers et al. 1996 |
| 1 | MOS001-12 | CDS, S138N[b] | caYga-(G,A)-caggc | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | n/a | Thorgeirsdottir et al. 1999 |
| 1 | MOS001-13 | CDS, L141F | ggcct-(C,T)-ttata | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | AJ000680 | Bossers et al. 1996 |
| 1 | MOS001-14 | CDS, H143R | tatac-(A,G)-ttttg | 0.93 | 0.07 | 0.89 | 0.09 | 0.02 | AF180389 | O'Rourke et al. 2000 |
| 1 | MOS001-15 | CDS, R151C | aggac-(C,T)-gttac | 0.98 | 0.02 | 0.96 | 0.02 | 0.01 | n/a | Thorgeirsdottir et al. 1999 |
| 1 | MOS001-16 | CDS, R154H | ctatc-(G,A)-tgaaa | 0.97 | 0.03 | 0.94 | 0.04 | 0.01 | AJ000737 | Goldmann et al. 1991 |
| 1 | MOS001-17 | CDS, Q171RHK[b,c] | tggat-(C,A)-RKtat | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | n/a | Gombojav et al. 2003 |
| | MOS001-18 | CDS, Q171RHK[c] | ggatM-(A,G)-Ktata | 0.67 | 0.33 | 0.47 | 0.40 | 0.13 | M31313 | Goldmann et al. 1990 |
| 1 | | | | | | | | | AJ000736 | Goldmann et al. 1990 |
| 1 | MOS001-19 | CDS, Q171RHK[c] | gatMR-(G,T)-tatag | 0.97 | 0.03 | 0.94 | 0.04 | 0.01 | AJ000734 | Belt et al. 1995 |
| 1 | MOS001-20 | CDS, N176K[b] | cagaa-(C,A)-aactt | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | AF195247 | Vaccari et al. 2001 |
| 1 | MOS001-21 | CDS, H180T[e] | ttgtg-(C,T)-atgac | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | AY326330 | This report |
| 1 | MOS001-22 | CDS, G189L[b] | tcaac-(A,T)-acaca | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | n/a | Gombojav et al. 2003 |
| 1 | MOS001-23 | CDS, H211Q[b] | ggagc-(G,A)-agtgg | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | AJ000681.1 | Belt et al. 1996 |
| 1 | MOS001-24 | CDS, R231 synonymous | accaa-(A,C)-ggggg | 0.90 | 0.10 | 0.82 | 0.17 | 0.01 | AB060289.1 | Seo et al. 2001 (mouflon) |
| 1 | MOS001-25 | CDS, L237 synonymous | atcct-(C,G)-ttttc | 0.90 | 0.10 | 0.82 | 0.17 | 0.01 | AB060289.1 | Seo et al. 2001 (mouflon) |
| 1 | MOS001-26 | CDS, P241S[g] | cttcc-(C,T)-ctcct | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | AY326330 | This report |

[a] Allele or genotype frequencies that add up to 0.99 or 1.01 are the result of rounding errors.
[b] Not observed in MSDP1.1, but reported in other population samples.
[c] Codon contains multiple adjacent polymorphic sites and thus, the combined haplotype determines which amino acid is translated.
[d] Not available in GenBank.
[e] Observed in some individuals of the Dorper breed.
[f] Three different nucleotide alleles observed at this site (c, g, and t).
[g] Observed in some individuals of the Dorset breed.

**Table 3. Allele and genotype frequencies of *PRNP* gene polymorphisms in the MARC Beef Cattle Diversity Panel version 2.1**

| *BTAPRNP* | | | | *Allele frequencies*[b] | | *Genotype frequencies* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *DNA segment* | *SNP identifier* | *PRNP region*[a] | *Alleles (1,2)* | *1* | *2* | *1,1* | *1,2* | *2,2* | *Genbank accession* | *Reference* |
| 4 | AH25-1 | Promoter | atcta-(G,A)-ttcac | 0.51 | 0.49 | 0.31 | 0.41 | 0.28 | AF465161 | This report |
| 4 | AH25-2 | Promoter | gcttc-(C,T)-tatca | 0.96 | 0.04 | 0.93 | 0.07 | 0.00 | AF465161 | This report |
| 4 | AH25-3 | Promoter[c] | acttc-(G,A)-ttagc | 0.98 | 0.02 | 0.97 | 0.03 | 0.00 | AF465161 | This report |
| 4 | AH25-4 | Promoter | ccagg-(T,C)-tccag | 0.65 | 0.35 | 0.47 | 0.36 | 0.17 | AF465161 | This report |
| 4 | AH25-5 | Promoter | aaccc-(C,T)-caaac | 0.76 | 0.24 | 0.62 | 0.28 | 0.09 | AF465161 | This report |
| 4 | AH25-6 | Promoter | catca-(G,T)-tgccc | 0.51 | 0.49 | 0.31 | 0.41 | 0.28 | AF465161 | This report |
| 1 | AH25-16 | CDS, L23 synonymous | ggcct-(C,T)-tgcaa | 0.98 | 0.02 | 0.96 | 0.04 | 0.00 | AY335912 | This report |
| 1 | AH25-17 | CDS, T36 synonymous[d] | aacac-(T,A)-ggggg | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | AY335912 | This report |
| 1 | AH25-18 | CDS, P42 synonymous[d] | taccc-(A,G)-ggaca | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | AY335912 | This report |
| 1 | AH25-19 | CDS, S46I[e] | gggca-(G,T)-tcctg | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | AF117327 | Wopfner et al. 1999 |
| 1 | AH25-20 | CDS, octapeptide indel (71-78)[f] | ttatcca-(6R,5R)-ggtggta | 0.95 | 0.05 | 0.90 | 0.10 | 0.00 | AJ132392 | Schlapfer et al. 1999 |
| 1 | AH25-20 | CDS, octapeptide indel (78-86)[f] | ttatcca-(6R,7R)-ggtggta | 0.95 | 0.00 | 0.90 | 0.00 | 0.00 | AF455119 | Naharro et al. 2003 |
| 1 | AH25-21 | CDS, Q78 synonymous (6R)[g] | ttatcca-[74 nt]-(G,A)-[75 nt]-ggtggta | 0.74 | 0.26 | 0.59 | 0.31 | 0.11 | AJ132392 | Schlapfer et al. 1999 |
| 1 | AH25-22 | CDS, P105 synonymous | aaacc-(C,T)-agtaa | 0.86 | 0.14 | 0.76 | 0.20 | 0.04 | AY335912 | This report |
| 1 | AH25-23 | CDS, S146N[d] | tggca-(G,A)-tgact | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | AF117327 | Wopfner et al. 1999 |
| 1 | AH25-24 | CDS, N177 synonymous | aacaa-(C,T)-tttgt | 0.97 | 0.03 | 0.94 | 0.06 | 0.00 | AY335912 | This report |
| 1 | AH25-25 | CDS, N184 synonymous | gtcaa-(C,T)-atcac | 0.90 | 0.10 | 0.81 | 0.18 | 0.01 | AB001468 | Yoshimoto et al. 1992 |
| | | | | | | | | | D10612 | Yoshimoto et al. 1992 |
| 1 | AH25-28 | CDS, T202 synonymous | ttcac-(C,T)-gaaac | 0.98 | 0.02 | 0.97 | 0.03 | 0.00 | AY335912 | This report |
| 1 | AH25-29 | CDS, C217 synonymous[d] | atgtg-(C,T)-atYac | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | AY335912 | This report |
| 1 | AH25-30 | CDS, 1218 synonymous[d] | tgYat-(T,C)-accca | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | AF117327 | Wopfner et al. 1999 |

[a] The bovine PrP isofom with five octapeptide repeats was used as a reference for the codon position number.
[b] Allele or genotype frequencies that add up to 0.99 or 1.01 are the result of rounding errors.
[c] Observed in some individuals of the Brahman and Beefmaster breeds.
[d] Observed in some individuals of the Santa Gertrudis breed.
[e] Not observed in MBCDP2.1, but reported in other population samples.
[f] This is an ''octapeptide repeat'' region that consists of either five, six. or seven in-frame 24- or 27-bp repeats. The nulceotide sequences flanking the octapeptide region are unique (e.g., ttatcca-[repeat region]-ggtggta) and the polymorphism involves the presence or absence of one complete octapeptide repeat. The ''6R'' allele is the most common in this panel and corresponds to the presence of all six octapeptide elements. The ''5R'' allele corresponds to the absence of the third octapeptide repeat. The ''7R'' allele corresponds to an insertion of a 24-bp repeat unit between the 5th and 6th repeat units.
[g] Polymorphic in the haplotypes with six octapeptide repeats. The SNP is in the 24th nt of the third repeat (i.e., 3R-1).

**Table 4. Alelle and genotype frequencies of *PRNP* gene polymorphisms in the MARC *Odocoileus* Panel version 1.0**

| | | | | *Allele frequencies*[a] | | | | | | *Genotype frequencies* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *Combined* | | *O. virginianus* | | *O. hemionus* | | *Combined* | | | *O. virginianus* | | | *O. hemionus* | | | | |
| *ODOPRNP DNA segment* | *SNP identifier* | *PRNP* region | *Alleles (1,2)* | *1* | *2* | *1* | *2* | *1* | *2* | *1,1* | *1,2* | *2,2* | *1,1* | *1,2* | *2,2* | *1,1* | *1,2* | *2,2* | *Genbank accession* | *References* |
| 4 | MDS001-1 | Promoter | ggccc-(C,T)-aYaca | 0.98 | 0.02 | 0.97 | 0.03 | 1.00 | 0.00 | 0.96 | 0.04 | 0.00 | 0.93 | 0.07 | 0.00 | 1.00 | 0.00 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-2 | Promoter | cccYa-(C,T)-acaac | 0.99 | 0.01 | 1.00 | 0.00 | 0.99 | 0.01 | 0.99 | 0.01 | 0.00 | 1.00 | 0.00 | 0.00 | 0.97 | 0.03 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-3 | Promoter[b] | tcttt-(T,G)-aaata | 0.91 | 0.09 | 0.67 | 0.33 | 0.99 | 0.01 | 0.82 | 0.18 | 0.00 | 0.33 | 0.67 | 0.00 | 0.97 | 0.03 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-4 | Promoter | aataa-(G,A)-gtcaW | 0.93 | 0.07 | 0.88 | 0.12 | 0.99 | 0.01 | 0.86 | 0.14 | 0.00 | 0.77 | 0.23 | 0.00 | 0.97 | 0.03 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-5 | Promoter | Rgtca-(T,A)-gtcac | 0.97 | 0.03 | 0.95 | 0.05 | 1.00 | 0.00 | 0.95 | 0.04 | 0.01 | 0.92 | 0.06 | 0.02 | 1.00 | 0.00 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-6 | Promoter | accct-(G,A)-cttaa | 0.93 | 0.07 | 0.97 | 0.03 | 0.88 | 0.12 | 0.90 | 0.07 | 0.03 | 0.96 | 0.02 | 0.02 | 0.82 | 0.13 | 0.05 | AY360090, AY360092 | This report |
| 4 | MDS001-7 | Promoter | gacct-(G,A)-gtcat | 0.99 | 0.01 | 0.99 | 0.01 | 1.00 | 0.00 | 0.99 | 0.01 | 0.00 | 0.98 | 0.02 | 0.00 | 1.00 | 0.00 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-8 | Promoter[b] | cttta-(C,T)-tgtct | 0.56 | 0.44 | 0.92 | 0.08 | 0.16 | 0.84 | 0.52 | 0.07 | 0.41 | 0.91 | 0.02 | 0.07 | 0.10 | 0.12 | 0.78 | AY360090, AY360092 | This report |
| 4 | MDS001-9 | Promoter[b] | tctaa-(G,C)-gtcag | 0.92 | 0.08 | 0.85 | 0.15 | 0.99 | 0.01 | 0.84 | 0.15 | 0.01 | 0.72 | 0.26 | 0.02 | 0.98 | 0.02 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-10 | Promoter[c] | taaga-(I,Z)-gagtt | 0.96 | 0.04 | 0.94 | 0.06 | 0.98 | 0.02 | 0.91 | 0.09 | 0.00 | 0.88 | 0.12 | 0.00 | 0.95 | 0.05 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-11 | Promoter[b,c] | aaaca-(A,G)-aSagt | 0.50 | 0.50 | 0.13 | 0.87 | 0.91 | 0.09 | 0.45 | 0.11 | 0.45 | 0.04 | 0.18 | 0.78 | 0.90 | 0.03 | 0.08 | AY360090, AY360092 | This report |
| 4 | MDS001-12 | Promoter | acaRa-(G,C)-agttt | 0.97 | 0.03 | 0.98 | 0.02 | 0.96 | 0.04 | 0.96 | 0.03 | 0.01 | 0.98 | 0.00 | 0.02 | 0.93 | 0.07 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-13 | Promoter | agctt-(C,T)-tccat | 0.99 | 0.01 | 0.98 | 0.02 | 1.00 | 0.00 | 0.99 | 0.00 | 0.01 | 0.98 | 0.00 | 0.02 | 1.00 | 0.00 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-14 | Promoter | tctat-(G,A)-tgatg | 0.99 | 0.01 | 0.99 | 0.01 | 1.00 | 0.00 | 0.99 | 0.01 | 0.00 | 0.98 | 0.02 | 0.00 | 1.00 | 0.00 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-15 | Promoter | ttgac-(A,C)-Rgtaa | 0.64 | 0.36 | 0.41 | 0.59 | 0.93 | 0.08 | 0.52 | 0.24 | 0.24 | 0.21 | 0.40 | 0.40 | 0.90 | 0.05 | 0.05 | AY360090, AY360092 | This report |
| 4 | MDS001-16 | Promoter | tgacM-(G,A)-gtaac | 0.97 | 0.03 | 0.96 | 0.04 | 0.99 | 0.01 | 0.95 | 0.03 | 0.01 | 0.94 | 0.04 | 0.02 | 0.98 | 0.03 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-17 | Promoter | ggaca-(C,T)-gggag | 0.96 | 0.04 | 0.94 | 0.06 | 0.98 | 0.02 | 0.92 | 0.07 | 0.01 | 0.90 | 0.08 | 0.02 | 0.95 | 0.05 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-18 | Promoter | tcaaa-(A,G)-aatgt | 0.52 | 0.48 | 0.18 | 0.82 | 0.88 | 0.13 | 0.46 | 0.11 | 0.43 | 0.09 | 0.19 | 0.72 | 0.86 | 0.02 | 0.11 | AY360090, AY360092 | This report |
| 4 | MDS001-19 | Promoter | gcttg-(T,C)-cattc | 0.99 | 0.01 | 0.99 | 0.01 | 0.99 | 0.01 | 0.98 | 0.02 | 0.00 | 0.98 | 0.02 | 0.00 | 0.98 | 0.02 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-20 | Promoter | aaatg-(G,C)-ccttc | 0.96 | 0.04 | 0.93 | 0.07 | 0.99 | 0.01 | 0.91 | 0.09 | 0.00 | 0.85 | 0.15 | 0.00 | 0.98 | 0.02 | 0.00 | AY360090, AY360092 | This report |
| 4 | MDS001-21 | Promoter | ccttc-(C,T)-ttata | 0.99 | 0.01 | 1.00 | 0.00 | 0.98 | 0.02 | 0.98 | 0.02 | 0.00 | 1.00 | 0.00 | 0.00 | 0.95 | 0.05 | 0.00 | AY360090, AY360092 | This report |
| 1 | MDS001-22 | Signal peptide D20G | gagtg-(A,G)-Ygtgg | 0.96 | 0.04 | 0.99 | 0.01 | 0.93 | 0.07 | 0.92 | 0.08 | 0.00 | 0.98 | 0.02 | 0.00 | 0.86 | 0.14 | 0.00 | AY360089, AY360091 | This report |
| 1 | MDS001-23 | Signal peptide D20G | agtgR-(C,T)-gtggg | 0.98 | 0.02 | 0.97 | 0.03 | 0.99 | 0.01 | 0.97 | 0.02 | 0.01 | 0.95 | 0.02 | 0.02 | 0.98 | 0.02 | 0.00 | AY360089, AY360091 | This report |
| 1 | MDS001-24 | CDS R51 synonymous | aaccg-(C,T)-tatcc | 0.96 | 0.04 | 0.94 | 0.06 | 0.99 | 0.01 | 0.93 | 0.07 | 0.00 | 0.89 | 0.11 | 0.00 | 0.98 | 0.02 | 0.00 | AY360089, AY360091 | This report |

| | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | MDS001-25 | CDS, G65E[d] | atcca-[34 nt]-(G,A)-[91 nt]-Rgtggtac | 0.83 | 0.17 | 0.95 | 0.05 | 0.69 | 0.31 | 0.70 | 0.25 | 0.05 | 0.93 | 0.05 | 0.02 | 0.45 | 0.48 | 0.08 | AY360089, AY360091 | This report |
| 1 | MDS001-26 | CDS, G81 synonymous[e] | atcca-[83 nt]-(T,A)-[42 nt]-Rgtggtac | 0.99 | 0.01 | 0.99 | 0.01 | 1.00 | 0.00 | 0.99 | 0.01 | 0.00 | 0.98 | 0.02 | 0.00 | 1.00 | 0.00 | 0.00 | AY360089, AY360091 | This report |
| 1 | MDS001-27 | CDS, G96S | gtcaa-(G,A)-gtggt | 0.88 | 0.12 | 0.80 | 0.20 | 0.96 | 0.03 | 0.79 | 0.19 | 0.02 | 0.64 | 0.32 | 0.05 | 0.95 | 0.05 | 0.00 | AF091560[f] | O'Rourke et al. 1998, unpublished |
| | | | | | | | | | | | | | | | | | | | AF156184 | Raymond et al. 2000 |
| 1 | MDS001-28 | CDS, P108 synonymous | aaacc-(A,G)-aaaac | 0.96 | 0.04 | 0.93 | 0.07 | 0.99 | 0.01 | 0.91 | 0.09 | 0.00 | 0.86 | 0.14 | 0.00 | 0.97 | 0.03 | 0.00 | AY360089, AY360091 | This report |
| 1 | MDS001-29 | CDS. A116G | tgtgg-(C,G)-aggag | 0.94 | 0.06 | 0.90 | 0.10 | 0.97 | 0.03 | 0.90 | 0.08 | 0.03 | 0.83 | 0.14 | 0.02 | 0.97 | 0.00 | 0.03 | AY360089, AY360091 | This report |
| 1 | MDS001-30 | CDS, A120 synonymous[g] | gctgc-(C,A)-gctgg | 1.00 | 0.00 | 1.00 | 0.00 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | AF009181[f] | O'Rourke et al. 1998 |
| 1 | MDS001-31 | CDS, L128 synonymous[g] | ggcct-(T,C)-ggtgg | 1.00 | 0.00 | 1.00 | 0.00 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | AF009181[f] | O'Rourke et al. 1998 |
| 1 | MDS001-32 | CDS, Y131 synonymous | ggcta-(C,T)-atgct | 0.96 | 0.04 | 1.00 | 0.00 | 0.91 | 0.09 | 0.91 | 0.09 | 0.00 | 1.00 | 0.00 | 0.00 | 0.82 | 0.18 | 0.00 | AY360089, AY360091 | This report |
| 1 | MDS001-33 | CDS, N138S | catga-(G,A)-cagRc | 0.67 | 0.33 | 0.85 | 0.15 | 0.49 | 0.51 | 0.40 | 0.53 | 0.06 | 0.69 | 0.31 | 0.00 | 0.11 | 0.76 | 0.13 | U25965[f] | Cervenakova et al. 1995, unpublished |
| | | | | | | | | | | | | | | | | | | | U97331[f] | O'Rourke et al. 1997, unpublished |
| | | | | | | | | | | | | | | | | | | | AF091559[f] | O'Rourke et al. 1998, unpublished |
| | | | | | | | | | | | | | | | | | | | AF156186 | Raymond et al. 2000 |
| 1 | MDS001-34 | CDS, R139 synonymous[b] | aRcag-(A,G)-cctct | 0.79 | 0.21 | 0.96 | 0.04 | 0.62 | 0.38 | 0.60 | 0.39 | 0.01 | 0.92 | 0.08 | 0.00 | 0.26 | 0.71 | 0.03 | AF009180[f] | O'Rourke et al. 1998 |
| | | | | | | | | | | | | | | | | | | | AF009181[f] | O'Rourke et al. 1998 |
| 1 | MDS001-35 | CDS, N146 synonymous | ggcaa-(C,T)-gacta | 0.94 | 0.06 | 0.93 | 0.07 | 0.96 | 0.04 | 0.89 | 0.11 | 0.00 | 0.85 | 0.15 | 0.00 | 0.92 | 0.08 | 0.00 | AF091560[f] | O'Rourke et al. 1998, unpublished |
| | | | | | | | | | | | | | | | | | | | AF009181[f] | O'Rourke et al. 1998 |
| | | | | | | | | | | | | | | | | | | | AF156184[f] | Raymond et al. 2000 |
| | | | | | | | | | | | | | | | | | | | AF156185[f] | Raymond et al. 2000 |
| 1 | MDS001-36 | CDS, R151C | aggac-(C,T)-gttac | 0.96 | 0.04 | 1.00 | 0.00 | 0.91 | 0.09 | 0.91 | 0.09 | 0.00 | 1.00 | 0.00 | 0.00 | 0.82 | 0.18 | 0.00 | AY360089, AY360091 | This report |
| 1 | MDS001-37 | CDS, N156 synonymous | gaaaa-(C,T)-atgta | 0.78 | 0.22 | 0.94 | 0.06 | 0.61 | 0.39 | 0.57 | 0.42 | 0.01 | 0.88 | 0.12 | 0.00 | 0.24 | 0.73 | 0.02 | AF009180[f] | O'Rourke et al. 1998 |
| 1 | MDS001-38 | CDS, 1185 synonymous[b,h] | aacat-(C,T)-acagt | 0.89 | 0.11 | 0.80 | 0.20 | 0.99 | 0.01 | 0.78 | 0.22 | 0.00 | 0.61 | 0.39 | 0.00 | 0.97 | 0.03 | 0.00 | AF156184 | Raymond et al. 2000 |
| 1 | MDS001-39 | CDS, T202 synonymous[b] | ttcac-(C,T)-gaaac | 0.79 | 0.21 | 0.96 | 0.04 | 0.59 | 0.41 | 0.58 | 0.40 | 0.01 | 0.93 | 0.08 | 1.00 | 0.22 | 0.76 | 0.03 | AY360089, AY360091 | This report |
| 1 | MDS001-40[a] | CDS, 1206 synonymous[g] | gacat-(T,C)-aagat | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | AF009181 | O'Rourke et al. 1998 |
| 1 | MDS001-41 | 3′ UTR[i] | gacat-(T,C)-aagat | 0.51 | 0.49 | 0.44 | 0.56 | 0.59 | 0.41 | 0.28 | 0.45 | 0.27 | 0.31 | 0.26 | 0.44 | 0.26 | 0.66 | 0.09 | AY360089, AY360091 | This report |

[a] Allele or genotype frequencies that add up to 0.99 or 1.01 are the result of rounding errors.
[b] Differences in minor allele frequencies, suggesting that these alleles may be used to differentiate *Odocoileus* spp. (i.e., ratio of the minor allele frequencies greater than 9:1 before rounding).
[c] Marker MDS001-10 is an 8-bp indel. The I allele indicates the presence of the 8 nt (taaacaRa), whereas the "Z" allele indicates its absence. Marker MDS001-11 is an SNP that was only observed in the insertion (I) allele.
[d] i.e., present in the second octapeptide repeat: ccatg-[R]-aggtg.
[e] i.e., present in the fifth octapeptide repeat: catgg-[W]-ggtgg.
[f] The Genbank sequence was derived from animals diagnosed with CWD and contained allele 2.
[g] Not observed in the MARC *Odocoileus* Panel 1.0, but reported in other population samples.
[h] GenBank accession AF091560 (O'Rourke et al. 1998) contains a third allele at this position ("A" nucleotide) and is synonymous with I185. The "A" allele was not observed in members of our sample group.
[i] Meets criteria for highly informative animal identity and parentage marker in both *Odocoileus* spp. (Heaton et al. 2002).

## Results

***Sequence comparisons in the PRNP coding region.*** Comparing DNA sequences from the present study and those previously reported revealed a total of 53 polymorphic sites in the *PRNP* coding region: 20 in sheep, 13 in cattle, and 20 in deer (Fig. 1C). These include three, seven, and 11 previously unreported polymorphisms, respectively. Some sites were monomorphic in our panels but polymorphic in previous reports (six in sheep, one in cattle, and three in deer). There were two sites in the *PRNP* coding region that were polymorphic among more than one species. For example, the first position of codon 151 contained a C/T polymorphism (R151C) in both sheep and deer, and the third position of codon 202 contained a C/T polymorphism (T202, synonymous) in cattle and deer. In both the 151 and the 202 codons, the C nucleotide was the common allele in all three species and appears to represent the allelic state of the most recent common ancestor.

With regard to sheep, 18 of the 20 SNPs in the coding region are predicted to encode amino acid differences in the translated prion protein. A summary of the known *PRNP* SNPs in sheep is presented in Table 2. Three of these predicted amino acid differences have not been previously reported (G127A, H180T, P241S), although their frequency in the sheep diversity panel was low (0.01). Five haplotypes affecting the translation of codons 136, 154, and 171 were observed (ARQ, ARR, AHQ, VRQ, ARH) with frequencies of 0.57, 0.33, 0.03, 0.03, and 0.03, respectively, in the MARC sheep diversity panel (data not shown). With regards to key genotypes in scrapie eradication programs, the frequencies of the most resistant genotype (*ARR*/*ARR*) was 0.13 ($CI_{95\%}$ = 0.09 to 0.19), and the susceptible genotypes (*ARQ*/*ARQ*, *ARQ*/*VRQ*, and *VRQ*/*VRQ*) was 0.43 ($CI_{95\%}$ = 0.36 to 0.51).

In cattle, only three of the 13 polymorphisms in the coding region are predicted to encode amino acid differences in the translated prion protein. None of the seven newly recognized SNPs is predicted to affect the amino acid sequence of PrP. The frequencies of the minor alleles for these SNPs ranged from 0.01 to 0.14 (Table 3), with some SNPs (e.g., AH25-22) having both alleles present in more than half of the breeds tested (data not shown).

In deer, seven of the 20 SNPs in the coding region are predicted to encode amino acid differences in the translated prion protein. Of the 11 newly recognized SNPs, five were predicted to encode amino acid differences, five were synonymous substitutions, and one was in the 3′ untranslated region (UTR). Two of the non-synonymous substitutions are predicted to affect the translation of codon 20 in the signal peptide region of PrP (Table 4, D20G, MDS001-22 and MDS001-23). Most of the newly recognized SNPs were observed in both *O. virginianus* and *O. hemionus* spp., although allele frequencies of some SNPs appeared to differ between species (Table 4, footnote b). These latter SNPs may be useful in estimating the extent of interbreeding in populations where the ranges of *O. virginianus* and *O. hemionus* overlap. One SNP was highly informative in both species (Table 4, MDS001-41). The frequencies of the minor alleles and minor homozygous genotypes were greater than 0.41 and 0.09, respectively, and thus may be useful for animal identity and parentage testing in both *Odocoileus* spp.

***Sequence comparisons in the PRNP promoter region.*** Sequence comparison of the DNA segments amplified from the *PRNP* promoter region revealed 33 polymorphic sites: six each in sheep and cattle and 21 in deer (Fig. 1B). No polymorphic sites were in common among the three groups. All sites were SNPs except in deer, where one complex site (MDS001-10) contained an eight-base indel, with the insertion allele having an additional SNP (MDS001-11). None of the 33 sites have been previously reported and thus represent new markers for evaluating the *PRNP* promoter regions of sheep, cattle, and deer for association with TSE susceptibility by linkage disequilibrium.

## Discussion

These results describe previously unrecognized nucleotide diversity in the *PRNP* promoter and coding regions and provide allele frequency estimates in healthy individuals from U.S. sheep, cattle, and free-ranging Wyoming deer populations. Knowledge of the minor allele frequencies is critical for 1) estimating genotyping error rates caused by SNPs that lie within a primer binding site, 2) assessing whether markers may be useful for animal identification or parentage testing, and 3) evaluating their association with TSE susceptibility. The number of individuals sampled within each animal group was sufficient to detect more than 95% of all alleles present at a frequency greater than 0.02. The total number of polymorphisms observed in approximately 1400 bp of sequence was similar between species when deer were counted as two species, i.e., 26 in sheep, 19 in cattle, and 41 in white-tailed and mule deer combined. In cattle, this is slightly higher than the 15 SNPs expected on the basis of the average from previous analyses of SNP density in the same panel (Heaton et al. 2002). Similar estimates have not yet

been made for the sheep and deer panels. A striking species difference was observed, however, in the number of polymorphisms predicted to affect the amino acid sequence. The ratio of non-synonymous to synonymous amino acid substitutions was 9:1 in sheep, 0.3:1 in cattle, and 0.6:1 in deer. The reasons for this difference are unknown.

The present report provides aggregate frequency estimates for the most resistant and susceptible genotypes in sheep breeds that contribute significantly to U.S. production. In this group of sheep, 13% ($CI_{95\%}$ = 9–19%) have the most resistant genotype (ARR/ARR), whereas 43% ($CI_{95\%}$ = 36–51%) have the most susceptible genotypes (*ARQ/ARQ*, *ARQ/VRQ*, and *VRQ/VRQ* combined). Because the overall ARR allele frequency is relatively low (33%, $CI_{95\%}$ = 27–41%), selection for this allele would substantially increase the number of animals with resistant *ARR/ARR* genotypes and significantly reduce the overall genetic risk of developing scrapie. Furthermore, this selection strategy is compliant with scrapie eradication programs.

It is important to note that in spite of significant sampling, any population, breed, or lineage that has not been analyzed may contain additional polymorphisms. The minor allele of unrecognized polymorphisms may be rare in the overall national population, yet be quite common in isolated or inbred populations. This phenomenon is a common source of error when genotype tests are developed from information in one population and applied towards other untested populations. Any previously unrecognized SNP in the binding site of any primer used in genotype assays may prevent the detection of a disease-predisposing allele. This phenomenon may significantly affect livestock disease control and eradication programs when the sires are the focus of selection and the ratio of sires to dams in breeding is 10:1 or 25:1. Both of these conditions are typical in many sheep and cattle production systems. Thus, the success of disease control and eradication programs may be significantly diminished if even a few disease carriers are incorrectly classified.

In summary, the goal of this project was to obtain informative markers spanning both the promoter and coding regions to facilitate extended haplotype analysis of the *PRNP* gene locus. Such analysis may lend power to future genetic epidemiologic studies involving BSE and CWD because disease-associated SNPs may occur independently, yet share a common haplotype. By identifying many SNPs across 20 kb of the *PRNP* locus, haplotypes and the mutational steps in the haplotype network may be identified and tested for correlation with genetic predisposition to TSE diseases. The ability to identify individuals with genetic resistance to TSE is predicted to significantly enhance the efficiency with which TSEs are eliminated from the food chain.

**Note added in proof:** While this manuscript was in press, Hills et al. reported PRNP sequence variation in European populations of cattle and sheep (Anim Genet. 2003 Jun;34(3):183–90).

## Acknowledgements

We thank Drs. J. L. Bono, B. Caughey, M. L. Clawson, B. T. Green, J. E. Keen, and A. Oommen for critical comments regarding this manuscript; J. K. Carnahan for outstanding technical assistance; R. A. Godtel, R. T. Lee, G. Schuller-Chavez, and S. J. Simcox for expert technical support; P. C. Anderson, R. R. Bradley, B. Herman, and J. E. Wray for database and network support; S. L. Kluver and J. K. Rosch for secretarial assistance. Product and company names are necessary to accurately report the methods and results; however, the USDA neither guarantees nor warrants the standard of the product, and the use of names by USDA implies no approval of the product to the exclusion of others that may also be suitable.

## References

1. Baylis M, Goldmann W, Houston F, Cairns D, Chong A et al. (2002) Scrapie epidemic in a fully PrP-genotyped sheep flock. J Gen Virol 83, 2907–2914
2. Belt PB, Muileman IH, Schreuder BE, Bos-de Ruijter J, Gielkens AL et al. (1995) Identification of five allelic variants of the sheep PrP gene and their association with natural scrapie. J Gen Virol 76, 509–517
3. Belt PBGM, Bossers A, Schreuder BEC, Smits MA (1996) PrP allelic variants associated with natural scrapie. In: *Bovine Spongiform Encephalopathy. The BSE Dilemma*, Gibbs C.J., ed. (New York: Springer), pp 294–305
4. Billinis C, Panagiotidis CH, Psychas V, Argyroudis S, Nicolaou A et al. (2002) Prion protein gene polymorphisms in natural goat scrapie. J Gen Virol 83, 713–721
5. Bossers A, Schreuder BE, Muileman IH, Belt PB, Smits MA (1996) PrP genotype contributes to determining survival times of sheep with natural scrapie. J Gen Virol 77, 2669–2673
6. Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8, 186–194
7. Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8, 175–185

8. Freking BA, Murphy SK, Wylie AA, Rhodes SJ, Keele JW et al. (2002) Identification of the single base change causing the callipyge muscle hypertrophy phenotype, the only known example of polar overdominance in mammals. Genome Res 12, 1496–1506
9. Goldmann W, Hunter N, Foster JD, Salbaum JM, Beyreuther K et al. (1990) Two alleles of a neural protein gene linked to scrapie in sheep. Proc Natl Acad Sci USA 87, 2476–2480
10. Goldmann W, Hunter N, Benson G, Foster JD, Hope J (1991) Different scrapie-associated fibril proteins (PrP) are encoded by lines of sheep selected for different alleles of the Sip gene. J Gen Virol 72, 2411–2417
11. Goldmann W, Hunter N, Smith G, Foster J, Hope J (1994) PrP genotype and agent effects in scrapie: change in allelic interaction with different isolates of agent in sheep, a natural host of scrapie. J Gen Virol 75, 989–995
12. Gombojav A, Ishiguro N, Horiuchi M, Serjmyadag D, Byambaa B et al. (2003) Amino acid polymorphisms of PrP gene in Mongolian sheep. J Vet Med Sci 65, 75–81
13. Gordon D, Abajian C, Green P (1998) Consed: a graphical tool for sequence finishing. Genome Res 8, 195–202
14. Grosse WM, Kappes SM, Laegreid WW, Keele JW, Chitko-McKown CG et al. (1999) Single nucleotide polymorphism (SNP) discovery and linkage mapping of bovine cytokine genes. Mamm Genome 10, 1062–1069
15. Heaton MP, Chitko-McKown CG, Grosse WM, Keele JW, Keen JE et al. (2001) Interleukin-8 haplotype structure from nucleotide sequence variation in commercial populations of U.S. beef cattle. Mamm Genome 12, 219–226
16. Heaton MP, Harhay GP, Bennett GL, Stone RT, Grosse WM et al. (2002) Selection and use of SNP markers for animal identification and paternity analysis in U.S. beef cattle. Mamm Genome 13, 272–281
17. Hernandez-Sanchez J, Waddington D, Wiener P, Haley CS, Williams JL (2002) Genome-wide search for markers associated with bovine spongiform encephalopathy. Mamm Genome 13, 164–168
18. Hills D, Comincini S, Schlaepfer J, Dolf G, Ferretti L et al. (2001) Complete genomic sequence of the bovine prion gene (PRNP) and polymorphism in its promoter region. Anim Genet 32, 231–232
19. Humeny A, Schiebel K, Seeber S, Becker CM (2002) Identification of polymorphisms within the bovine prion protein gene (Prnp) by DNA sequencing and genotyping by MALDI-TOF-MS. Neurogenetics 4, 59–60
20. Hunter N, Goldmann W, Smith G, Hope J (1994) Frequencies of PrP gene variants in healthy cattle and cattle with BSE in Scotland. Vet Rec 135, 400–403
21. Hunter N, Foster JD, Goldmann W, Stear MJ, Hope J et al. (1996) Natural scrapie in a closed flock of Cheviot sheep occurs only in specific PrP genotypes. Arch Virol 141, 809–824
22. Ikeda T, Horiuchi M, Ishiguro N, Muramatsu Y, Kai-Uwe GD et al. (1995) Amino acid polymorphisms of PrP with reference to onset of scrapie in Suffolk and Corriedale sheep in Japan. J Gen Virol 76, 2577–2581
23. Kirkwood JK, Cunningham AA (1994) Epidemiological observations on spongiform encephalopathies in captive wild animals in the British Isles. Vet Rec 135, 296–303
24. Laplanche JL, Chatelain J, Westaway D, Thomas S, Dussaucy M et al. (1993) PrP polymorphisms associated with natural scrapie discovered by denaturing gradient gel electrophoresis. Genomics 15, 30–37
25. Lee IY, Westaway D, Smit AF, Wang K, Seto J et al. (1998) Complete genomic sequence and analysis of the prion protein gene region from three mammalian species. Genome Res 8, 1022–1037
26. Lemaire-Vieille C, Schulze T, Podevin-Dimster V, Follet J, Bailly Y et al. (2000) Epithelial and endothelial expression of the green fluorescent protein reporter gene under the control of bovine prion protein (PrP) gene regulatory sequences in transgenic mice. Proc Natl Acad Sci USA 97, 5422–5427
27. Leymaster KA (1991) Straightbred comparison of a composite population and the Suffolk breed for performance traits of sheep. J Anim Sci 69, 993–999
28. McCormack JE, Baybutt HN, Everington D, Will RG, Ironside JW et al. (2002) PRNP contains both intronic and upstream regulatory regions that may influence susceptibility to Creutzfeldt-Jakob disease. Gene 288, 139–146
29. Naharro G, Yugueros J, Temprano A, del Rio ML, Rodriguez-Ferri EF et al. (2003) Prion protein gene polymorphisms in a population of Spanish cows. Vet Rec 152, 212–213
30. Neibergs HL, Ryan AM, Womack JE, Spooner RL, Williams JL (1994) Polymorphism analysis of the prion gene in BSE-affected and unaffected cattle. Anim Genet 25, 313–317
31. Nickerson DA, Tobe VO, Taylor SL (1997) PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Res 25, 2745–2751
32. O'Rourke KI, Holyoak GR, Clark WW, Mickelson JR, Wang S et al. (1997) PrP genotypes and experimental scrapie in orally inoculated Suffolk sheep in the United States. J Gen Virol 78, 975–978
33. O'Rourke KI, Baszler TV, Miller JM, Spraker TR, Sadler-Riggleman I et al. (1998) Monoclonal antibody F89/160.1.5 defines a conserved epitope on the ruminant prion protein. J Clin Microbiol 36, 1750–1755
34. O'Rourke KI, Baszler TV, Besser TE, Miller JM, Cutlip RC et al. (2000) Preclinical diagnosis of scrapie by immunohistochemistry of third eyelid lymphoid tissue. J Clin Microbiol 38, 3254–3259
35. Prusiner SB (1998) Prions. Proc Natl Acad Sci USA 95, 13363–13383
36. Raymond GJ, Bossers A, Raymond LD, O'Rourke KI, McHolland LE et al. (2000) Evidence of a molecular barrier limiting susceptibility of humans, cattle and sheep to chronic wasting disease. EMBO J 19, 4425–4430

37. Schlapfer I, Saitbekova N, Gaillard C, Dolf G (1999) A new allelic variant in the bovine prion protein gene (PRNP) coding region. Anim Genet 30, 386–387
38. Scott MR, Will R, Ironside J, Nguyen HO, Tremblay P et al. (1999) Compelling transgenetic evidence for transmission of bovine spongiform encephalopathy prions to humans. Proc Natl Acad Sci USA 96, 15137–15142
39. Seo SW, Hara K, Kubosaki A, Nasu Y, Nishimura T et al. (2001) Comparative analysis of the prion protein open reading frame nucleotide sequences of two wild ruminants, the moufflon and golden takin. Intervirology 44, 359–363
40. Thorgeirsdottir S, Sigurdarson S, Thorisson HM, Georgsson G, Palsdottir A (1999) PrP gene polymorphism and natural scrapie in Icelandic sheep. J Gen Virol 80, 2527–2534
41. Tranulis MA, Osland A, Bratberg B, Ulvund MJ (1999) Prion protein gene polymorphisms in sheep with natural scrapie and healthy controls in Norway. J Gen Virol 80, 1073–1077
42. Vaccari G, Petraroli R, Agrimi U, Eleni C, Perfetti MG et al. (2001) PrP genotype in Sarda breed sheep and its relevance to scrapie. Brief report. Arch Virol 146, 2029–2037
43. Westaway D, Zuliani V, Cooper CM, Da Costa M, Neuman S et al. (1994) Homozygosity for prion protein alleles encoding glutamine-171 renders sheep susceptible to natural scrapie. Genes Dev 8, 959–969
44. Williams ES, Miller MW (2002) Chronic wasting disease in deer and elk in North America. Rev Sci Tech 21, 305–316
45. Williams ES, Young S (1992) Spongiform encephalopathies in Cervidae. Rev Sci Tech 11, 551–567
46. Wood JL, Lund LJ, Done SH (1992) The natural occurrence of scrapie in moufflon. Vet Rec 130, 25–27
47. Wopfner F, Weidenhofer G, Schneider R, von Brunn A, Gilch S et al. (1999) Analysis of 27 mammalian and 9 avian PrPs reveals high conservation of flexible regions of the prion protein. J Mol Biol 289, 1163–1178
48. Yoshimoto J, Iinuma T, Ishiguro N, Horiuchi M, Imamura M et al. (1992) Comparative sequence analysis and expression of bovine PrP gene in mouse L-929 cells. Virus Genes 6, 343–356