

ORIGINAL ARTICLE

A missense mutation in *AGTPBP1* was identified in sheep with a lower motor neuron disease

X Zhao¹, SK Onteru¹, KE Dittmer², K Parton², HT Blair², MF Rothschild¹ and DJ Garrick^{1,2}

A type of lower motor neuron (LMN) disease inherited as autosomal recessive in Romney sheep was characterized with normal appearance at birth, but with progressive weakness and tetraparesis after the first week of life. Here, we carried out genome-wide homozygosity mapping using Illumina Ovine SNP50 BeadChips on lambs descended from one carrier ram, including 19 sheep diagnosed as affected and 11 of their parents that were therefore known carriers. A homozygous region of 136 consecutive single-nucleotide polymorphism (SNP) loci on chromosome 2 was common to all affected sheep and it was the basis for searching for the positional candidate genes. Other homozygous regions shared by all affected sheep spanned eight or fewer SNP loci. The 136-SNP region contained the sheep ATP/GTP-binding protein 1 (*AGTPBP1*) gene. Mutations in this gene have been shown to be related to Purkinje cell degeneration (*pcd*) phenotypes including ataxia in mice. One missense mutation c.2909G>C on exon 21 of *AGTPBP1* was discovered, which induces an Arg to Pro substitution (p.Arg970Pro) at amino-acid 970, a conserved residue for the catalytic activity of AGTPBP1. Genotyping of this mutation showed 100% concordant rate with the recessive pattern of inheritance in affected, carrier, phenotypically normal and unrelated normal individuals. This is the first report showing a mutant *AGTPBP1* is associated with a LMN disease in a large mammal animal model. Our finding raises the possibility of human patients with the same etiology caused by this gene or other genes in the same pathway of neuronal development.

Heredity (2012) **109**, 156–162; doi:10.1038/hdy.2012.23; published online 16 May 2012

Keywords: *AGTPBP1*; lower motor neuron disease; missense mutation; Purkinje cell degeneration; sheep

INTRODUCTION

Motor neuron diseases (MNDs) are a group of disorders involved with selective degeneration of upper motor neurons, and/or lower motor neurons. The incidence of MND varied from 0.6 per 100 000 person-years to 2.4 per 100 000 person-years in different ethnic human populations (Lee, 2012). Upper motor neurons originate in the cerebral cortex or the brain stem, and provides indirect stimulation of the target muscles, while LMN connect the brain stem and spinal cord to muscle fibers, and provide nerve signals between the upper motor neurons and muscles. Common characteristics of MND are muscle weakness and/or spastic paralysis. Signs of LMN lesions are represented as muscle atrophy, fasciculations and weakness when associated LMN degenerate. Symptoms such as spastic tone, hyperreflexia and upgoing plantar reflex signs appear as upper motor neurons degenerate (Lee, 2012). The weakness of respiratory or bulbar muscle can even cause death of MND patients due to respiratory insufficiency (Borasio *et al.*, 1998). Based on the time of onset, types of motor neuron involvement and detailed clinical features, MND in humans can be classified into six types of diseases including amyotrophic lateral sclerosis, hereditary spastic paraplegia, primary lateral sclerosis, spinal bulbar muscular atrophy, spinal muscular atrophy and lethal congenital contracture syndrome. MNDs are considered incurable and no effective treatment is currently available. However, the gene therapy strategy of performing gene transfer to patients by utilizing adeno-associated virus vectors is a

promising therapeutic approach (Nizzardo *et al.*, 2011). Previous evidence has shown genetic variants within >30 genes might be responsible for MND (Dion *et al.*, 2009). For example, mutations in the Cu/Zn-binding superoxide dismutase (*SOD1*) gene have been identified in about 20% of patients with familial amyotrophic lateral sclerosis (Rosen *et al.*, 1993). However, genetic causes still remain unclear for more than half of human patients with MND.

In 1999, an extended family of Romney lambs was found to have a suspected hereditary MND, characterized by poor muscle tone and progressive weakness from about 1 week of age, leading to severe tetraparesis, recumbency and muscle atrophy. Lambs remained alert up to 4 weeks of age when they were euthanized. The muscles were atrophied and pale, and histologically had variation in muscle fiber diameter and small angular fibers. Large foamy macrophages were present in cerebrospinal fluid, and there were degenerate neurons in the spinal cord leading to a decrease in the number of motor neurons, consistent with LMN diseases. No histological changes were seen in the cerebellum (Anderson *et al.*, 1999). Outcross and backcross breeding trials were consistent with this LMN disease having simple autosomal recessive inheritance.

To find the causative locus responsible for this disease, we performed a genome-wide homozygosity mapping process and then fine mapped the mutation using a positional candidate gene approach within the largest homozygous genomic region found in the affected sheep. Our goals of this research were to discover the genetic basis of

¹Department of Animal Science and Center for Integrated Animal Genomics, Iowa State University, Ames, IA, USA and ²Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Palmerston North, New Zealand
Correspondence: Dr DJ Garrick, Department of Animal Science and Center for Integrated Animal Genomics, Iowa State University, 225D Kildee Hall, Ames, IA 50011, USA.
E-mail: dorian@iastate.edu

Received 17 November 2011; revised 1 March 2012; accepted 9 March 2012; published online 16 May 2012

this disease, to help sheep breeders avoid matings between carrier animals and to provide a potential animal model for gene therapy treatment of human LMN disease.

METHODS

Animals

Approvals dealing with DNA or tissue collection from animals were obtained from the Massey University Animal Ethics Committee (approval number 99/135). Iowa State University Animal Care and Use Committee did not require additional approvals.

A single Romney putative carrier ram was mated in New Zealand to 130 crossbred Finnish Landrace ewes to produce 115 daughters that were subsequently backcrossed to their sire. The backcrosses produced 5 affected sheep out of 47 lambs in the first lambing, and 11 out of 71 in the second lambing. The observed rate of affected sheep was consistent with the one in eight expectation for a homozygous recessive disorder from this mating scheme. Any affected offspring and their dams now proven to be carriers, formed the basis of the experimental population. High-density SNP genotypes were obtained from 19 affected, which included 3 affected lambs from the original Romney flock, and 11 known carrier sheep. Fine mapping using novel SNPs was conducted on 52 Romney or Romney/Finn crosses including the above mentioned 30 sheep with clear disease statuses and 22 potential carriers, which were phenotypically normal but sired by the carrier ram. Moreover, 85 control sheep consisted of 14 normal Romney sheep, 25 crossbred Texel sheep and 46 crossbred Corriedale sheep were used for validation of the causative mutation. These sheep were produced from unrelated populations without evidence of LMN disease.

SNP genotyping

DNA was extracted from blood using the QIAamp DNA Blood Mini Kit (QIAGEN, Valencia, CA, USA) and the Roche MagNA Pure automated analyzer (Roche Diagnostics N.Z., Ltd. Mt, Wellington, Auckland, New Zealand). The genotyping of 54241 evenly distributed polymorphic SNP markers across the genome was performed using the Ovine SNP50 BeadChip (Illumina, San Diego, CA, USA) at the commercial company GeneSeek Inc., Lincoln, NE, USA. Standard procedures were performed with a PCR and ligation-free protocol, which routinely achieves high-average call rates and accuracy.

Genome-wide homozygosity mapping

To define the molecular genetic basis for the LMN disease in these Romney sheep, a search for putative homozygous-by-descent (also named IBD) regions was initiated by performing a genome-wide scan using an in-house UNIX script as described in our previous publication (Zhao *et al.*, 2011). The SNP genotypes were ordered according to their genomic location as provided in the Illumina manifest, and used to identify regions of consecutive homozygous SNP loci that were common to all affected animals. The same procedure was applied to known carriers to ensure that the identified region was not simply fixed across all animals. A threshold of 10 SNPs defining a candidate consecutive homozygous region was chosen after calculating expected IBD fragment size for 19 affected lambs produced through the outcross and backcross experiments assuming exactly one crossover event occurred every meiosis on the chromosome carrying the causal mutation with uniform probability along the chromosome expressed in linkage units (cM). The homozygous fragment containing the causative mutation would have exceeded 0.5 cM or about 10 consecutive SNP with probability > 95%. The expected size of the IBD region that contains the causative mutation depends in part on the number of meioses that separate the common ancestor from the affected lambs, and the number of affected individuals. The IBD region that included the causative mutation had 23% probability of being shorter than 2 cM and a 32% probability being longer than 5 cM. The mean expected length of the IBD region was about 4 cM. Genes in the identified IBD regions were examined for potential involvement using comparative homology between ovine and bovine genomes as provided through the Ovine Genome Assembly v1.0 (<http://www.livestockgenomics.csiro.au/perl/gbrowse.cgi/oar1.0/#search>).

Positional candidate genes sequencing

Primer3 software (version 0.4.0, Whitehead Inst, Cambridge, MA, USA and Howard Hughes Medical Inst, Chevy Chase, MD, USA) and the reference

sequences from the Ovine Genome Assembly v1.0 were used to design PCR primers flanking the coding regions of ovine *AGTPBP1* (Supplementary Table S1). Among the many genes in the identified homozygous regions, *AGTPBP1* was considered to be the best candidate gene because of its role in neural functions as reported by several publications (Harris *et al.*, 2000; Fernandez-Gonzalez, *et al.*, 2002). PCR was performed using a 10- μ l cocktail mixture and a standard program as described previously (Zhao *et al.*, 2011), with an annealing temperature of 58 or 60 °C. Negative controls were performed to check for the presence of contaminants in all reactions. PCR products were verified by 1.5% agarose gel electrophoresis, and treated with ExoSAP-IT (Affymetrix, Cleveland, OH, USA), before pooling them from two affected and two carrier sheep. The pooled PCR products were sequenced commercially using a 3730xl DNA Analyzer (Applied Biosystems, Pleasanton, CA, USA). The Sequencher software (version 3.0, GeneCodes Corp., Ann Arbor, MI, USA) was used for sequence alignment and comparison.

Mutation analyses

As annotation of the ovine genome has not yet been completed, the bovine *AGTPBP1* reference mRNA sequence (GenBank: XM_001252536.3) was used as a query in cross-species BLAST searches to identify sheep expressed sequence tag. Both the bovine *AGTPBP1* mRNA and corresponding ovine expressed sequence tag served as references to search for the coding regions within ovine *AGTPBP1* gene. The *AGTPBP1* protein sequence was predicted using successfully amplified and resequenced exonic regions of the gene from the sheep population, the bovine *AGTPBP1* reference mRNA sequence and ovine expressed sequence tag. To understand the functional annotation of this protein, the conserved domains within this sheep protein sequence were searched by blasting the complete protein sequence against the Conserved Domains database (database: CDD-40526 PSSMs) in NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) while setting an option of 500 as the maximum number of hits. Multiple sequence alignments provide a basis for conserved domain models. The specific hits and super families were given for the queried protein. Moreover, the conserved protein residues involved in conserved features such as binding or catalysis were annotated when the query sequence residues were in the Entrez Protein database. The sequence alignments of protein *AGTPBP1* from multiple species, including human, mouse, chicken, frog, purpuratus, sea squirt, trichoplax, zebrafish and sheep, were displayed to show the conserved feature sites. The fold structure of the conserved domain of *AGTPBP1* in sheep (residues 850–1150) and mouse (residues 800–1110) were predicted using the 3D_PSSM program (Kelley *et al.*, 2000; <http://www.sbg.bio.ic.ac.uk/3dpssm/index2.html>). The known crystal structures of carboxypeptidases in the PDB database help to generate the predicted structure of *AGTPBP1*. Sequence alignment was performed on the ClustalW2 server (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) among sheep and mouse *AGTPBP1*, human (PDB ID: 1D1D) carboxypeptidase A2 and cattle (PDB ID: 1HDU) carboxypeptidase A. This helped us to be more confident about the conserved features of *AGTPBP1* secondary structure and the amino acids critical for carboxypeptidase catalysis. A three-dimensional macromolecular structure of the conserved domain was searched and downloaded using homologous sequences with confirmed crystal structure (<http://www.ncbi.nlm.nih.gov/Structure/MMDDB/mmdb.shtml>). Each SNP variant was analyzed to check whether it could induce missense or nonsense mutations. The identified SNPs were then genotyped using restriction enzyme digestion (PCR-restriction fragment length polymorphism) on the affected and carrier sheep as well as on other normal controls.

Restriction enzyme digestion (PCR-restriction fragment length polymorphism)

The sequences of PCR primer 'LMN_new' (Supplementary Table S1) were 5' TGTTAGCACCTCTCTTTGCT 3' and 5' AGCAGACCCTTGGCATGATA 3'. DNA fragments from the PCR reactions were subjected to restriction enzyme digestion with *Acil* (New England Biolabs Inc., Ipswich, MA, USA). The missense mutation in *AGTPBP1* removes a recognition site of the *Acil* enzyme. The 10- μ l system reaction mix contained 1 μ l of NEB buffer 3, 0.5 units of *Acil* and 2 μ l of PCR product. The reaction mix was incubated at 37 °C overnight, and the products were analyzed on a 2.0% (w/v) ultra-pure agarose gel.

RESULTS

Detection of identical by descent (IBD) regions

The whole-genome scan discovered 209 homozygous fragments consisting of three or more consecutive single-nucleotide polymorphisms (SNPs) shared by all 19 affected lambs. The size distribution of common homozygous fragments in these lambs is shown (Figure 1). Only one large homozygous segment exceeded the 10-SNP threshold

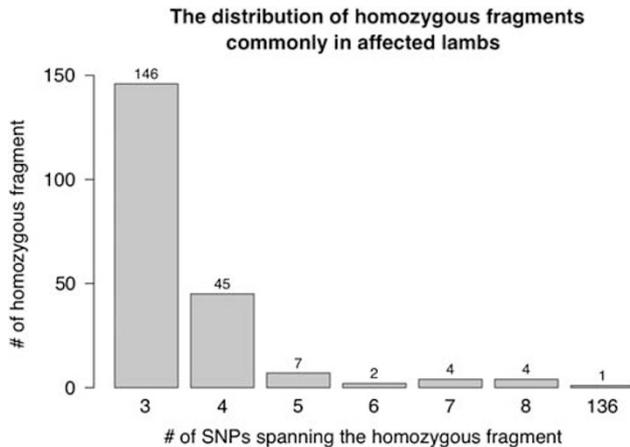


Figure 1 The distribution of identical homozygous fragments in affected lambs. Based on the different number of SNPs spanning the homozygous fragment, the counts of identical homozygous fragments in all affected lambs were shown. Only one homozygous fragment spanning 136 SNPs was identified to be common to all 19 affected lambs when the 'larger than 10-SNP' threshold was applied.

and it consisted of 136 consecutive SNP loci identified in all 19 affected sheep, whereas the 11 carriers exhibited heterozygosity in this segment. The region started from SNP s56017.1 and extended to SNP OAR2_36428027.1. This segment is physically positioned on the small arm of ovine chromosome 2 (OAR 2) from 29 284 415 to 36 428 027 bp, covering a region of 7 Mb (Figure 2a and b). We found this segment likely encompassed 36 genes (Figure 2d and Supplementary Table S2) based on the bovine reference sequences. A gene called ATP/GTP-binding protein, transcript variant 1 (*AGTPBP1*), also named *Nna1* (nervous system nuclear protein induced by axotomy protein 1 homolog) was the most plausible candidate because of its involvement in axonal regeneration and its causal relationship with mouse Purkinje cell degeneration (*pcd*) phenotype (Harris *et al.*, 2000, Fernandez-Gonzalez *et al.*, 2002). The ovine *AGTPBP1* gene spans about 0.17 Mb and encodes a 1225-amino-acid intracellular protein that contains a predicted zinc carboxypeptidase domain near its C-terminus (Harris *et al.*, 2000). This gene was initially cloned from regenerating spinal cord neurons of the mouse (supplied by OMIM 606830).

Gene coding region resequencing and mutation analyses

Functional mutations in candidate genes were examined by resequencing all available exons using primers based on the adjacent intronic sequences of these genes. The ovine *AGTPBP1* gene was predicted to have 25 exons based on the bovine reference gene and protein sequences. All the exons except exons 11, 12 and 17 for this gene were successfully amplified and three exonic SNPs were identified. The first exonic SNP (called c.2909G>C) was located at the sixth base pair of exon 21 of the sheep *AGTPBP1* gene. The other two were located at

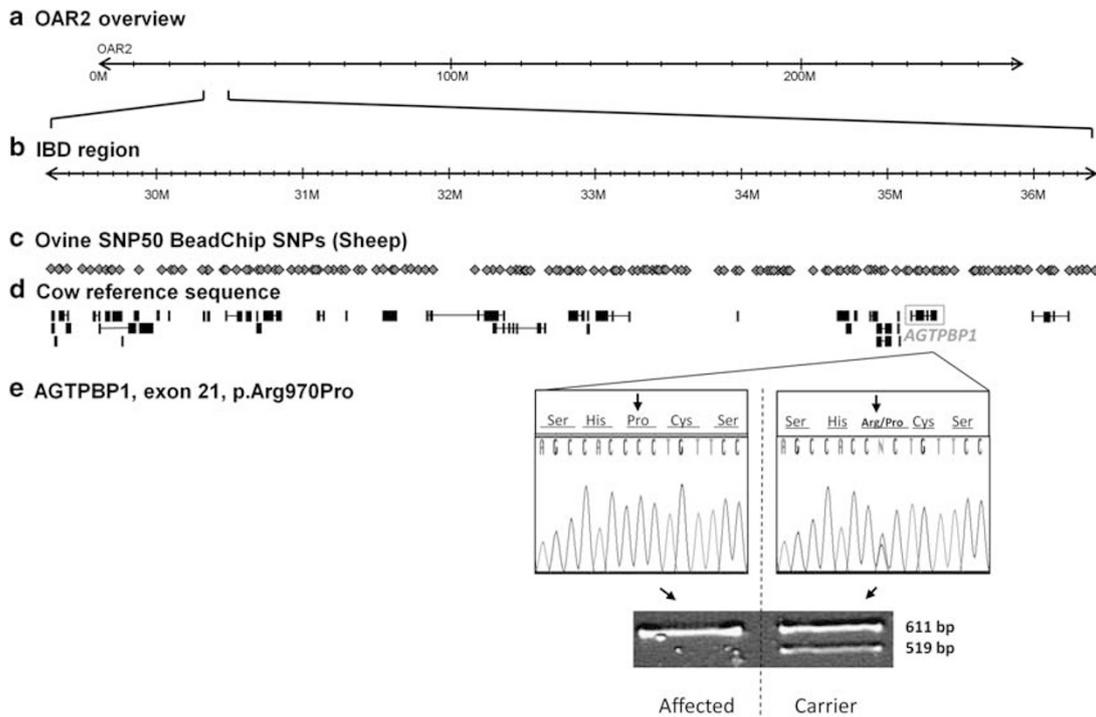


Figure 2 Work flow chart for discovering the mutation locus of low MND in Romney sheep. (a) The sheep chromosome 2 (OAR 2). (b) The IBD region containing the causative locus is located on OAR 2 between 29 and 36 Mb. (c) SNP locations on the Ovine SNP50 BeadChip in the IBD region (d) The black boxes show reference genes encompassing the sheep IBD region based on the comparative genomic maps. The gene *AGTPBP1* was depicted by a red lined box. (e) Sequence analyses of *AGTPBP1* showed a point mutation in exon 21, resulting in an Arg to Pro substitution. This missense mutation is located in a highly conserved residue for the catalytic activity of this protein. The PCR-restriction fragment length polymorphisms show different patterns for affected and carrier groups. A full color version of this figure is available at the *Heredity* journal online.

the ninety-third and ninety-ninth base pair of exon 22. An additional 17 SNPs were identified in either introns or the 3'-untranslated region on this gene. These were immediately excluded as causative loci as none were located either in the splicing donor/receptor regions or other important regulatory regions and were not in concordance with sheep phenotypes (data not shown). After comparing our findings with the bovine coding sequence of *AGTPBP1*, we observed that the first exonic SNP (c.2909G>C) introduced a missense transition (Arg>Pro) at amino-acid 970 (p.Arg970Pro; Figure 2e), and the third exonic SNP (c.3133T>C) also resulted as a missense mutation causing a Pro to Ser substitution (p.Pro1045Ser). The second exonic SNP (c.3126C>T) was a synonymous mutation and did not change any amino acid. A conserved domain named M14_Nna1 (CDD: cd06906) spanning from amino-acid residue 859–1136 on *AGTPBP1* in sheep was predicted in this study using the 'CD-search' option available in NCBI. This domain is a zinc-binding carboxypeptidase domain, which hydrolyzes the peptide bonds at the C-terminal part of certain amino acids in polypeptide chains. The sequence alignment among multiple species demonstrated two specific feature hits including metalloproteinase activity and zinc binding. Nine conserved amino-acid residues were involved in these two features hits (Supplementary Figures S1 and S2; Gomis-Ruth *et al.*, 1999; Aloy *et al.*, 2001). Three conserved residues (His920, Glu923 and His1017) expressed as a motif 'HXXE...H' are responsible for zinc binding, and

all nine conserved residues (His920, Glu923, Arg970, Asn979, Arg980, His1017, Gly1018, Met1027 and Glu1102) are important for substrate binding and catalysis of this protein (Supplementary Figure S1 and Figure 3a). In terms of the size and relative positioning of β -sheets and α -helices, the overall predicted secondary structures of *AGTPBP1* in sheep and mouse were the same as that of other metalloproteinases in human and cattle by using the 3D-PSSM program. Features and conserved amino acids and motifs are located in similar positions. For example, the well-known zinc-binding HXXE...H motif lies between the second β -sheet and the first α -helix as reported (Wang *et al.*, 2006). Six residues (His920, Glu923, Arg970, Asn979, Arg980 and Glu1102) remain conserved among these proteins (Supplementary Figure S3). The first exonic missense mutation (c.2909G>C) identified on exon 21 is the middle nucleotide within a codon 'CGC', which was translated and substituted from Arg970 (codon 'CGC'), one of the conserved residues responsible for substrate binding and catalysis, to Pro970 (codon 'CCC'; Figure 3b). Molecular modeling based on the crystal structure of the homologous protein, putative carboxypeptidase (PBDID: 3K2K), also demonstrated that Arg970 forms a part of a loop (the short yellow section of the middle loop in Figure 3c), which might be a key site for substrate binding or catalytic activity. The amino-acid residue encoded by the exonic missense mutation on exon 22 (c.3133T>C; p.Pro1045Ser) varies considerably among different species (Figure 3b).

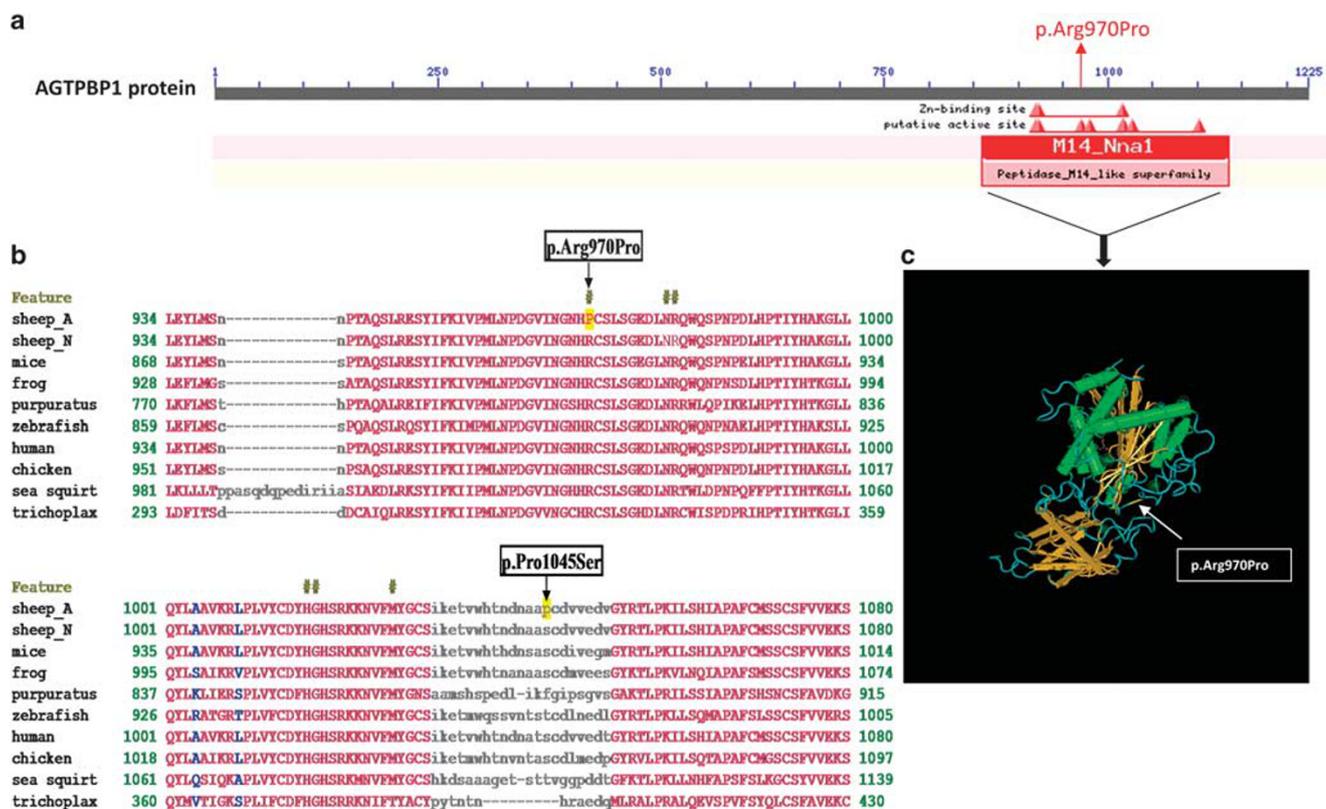


Figure 3 The conserved domain structures of partial protein sequence of *AGTPBP1*. (a) The conserved domains includes zinc-binding sites and putative active sites shared by M14_Nna1-like superfamily genes. (b) Alignment of partial sequence of sheep protein *AGTPBP1* with eight most dissimilar protein sequences from species of human, mouse, chicken, frog, purpuratus, sea squirt, trichoplax and zebrafish. Alignment of the full sequence is shown in the Supplementary information. The residues labeled by hash marks on the top are known to be involved in substrate binding or catalytic activity. The two missense mutations are labeled on the protein sequence. Upper case amino acids are aligned and a read to blue color scale represents the degree of conservation, with red showing highly conserved residues. Lower case, gray amino acids are unaligned. Dashes indicate variations in sequence length among multiple sequence alignment. (c) Three-dimensional molecular modeling of the conserved domain based on the crystal structure of putative carboxypeptidase (PBDID: 3K2K). The short yellow section in the middle loop represents the position of the mutated amino-acid residue.

It indicates that this missense mutation may not affect the function of the protein AGTPBP1 as it is neither involved in zinc binding nor in substrate binding and catalysis.

The Arg970Pro mutation was genotyped by PCR-restriction fragment length polymorphism (Figure 2e). The 611 bp length of the 'LMN_new' PCR fragment amplified from normal sheep had one restriction site, which could be recognized by the restriction enzyme *AcI* and cut into two fragments with lengths of 519 and 92 bp. However, the restriction site was removed due to the Arg970Pro mutation, and the 611 bp fragment amplified from the affected lambs was intact after digestion. Therefore, the PCR fragment amplified from sheep with a genotype of 'GG' was cut into two fragments with lengths of 519 and 92 bp; the one with a genotype of 'GC' had three fragments with lengths of 611, 519 and 92 bp; the one with a genotype of 'CC' had a single fragment with length 611 bp. Our genotyping results showed all 19 affected sheep were 'CC', the 11 known carriers were 'GC' and the 22 phenotypically normal but possible carrier sheep were either 'GC' or 'GG' genotypes. In contrast, 85 control (unrelated) sheep were only 'GG' genotypes (Table 1). This mutation was 100% concordant with the recessive pattern of inheritance in affected, carrier, phenotypically normal and unrelated normal individuals. The two exonic SNP (c.3126C>T and c.3133T>C) on exon 22 were genotyped by direct sequencing. However, there is no perfect concordant rate for these two SNPs based on the counts for each of the genotypes (Table 1). At this stage, we excluded these two SNP as causative mutations for this defect due to their genotyping results and mutation analyses. Taken together, our results strongly suggest that the Arg970Pro substitution is the causative locus for LMN disease in Romney sheep and it could act by decreasing the function of protein AGTPBP1 to generate LMN disease in these lambs.

DISCUSSION

The *AGTPBP1* gene encodes a protein that belongs to a member of the cytosolic carboxypeptidase subfamily (M14D subfamily; Kalinina *et al.*, 2007, Rodriguez de la Vega *et al.*, 2007). This protein has a role in protein turnover by cleaving proteasome-generated peptides into amino acids (Berezniuk *et al.*, 2010), enzymatically removing gene-coded glutamic acids or tyrosine from the C-terminus of proteins (Rogowski *et al.*, 2010, Kalinina *et al.*, 2007) and also has a role in neuronal bioenergetics at least in mouse brains (Chakrabarti *et al.*, 2010). Similar to other cytosolic carboxypeptidases, the mouse *Nna1* protein is broadly expressed in the brain, pituitary, eye, testis and

other tissues (Kalinina *et al.*, 2007). Mutations or abnormal expression of *Nna1* have been found to cause a *pcd* phenotype in mice characterized by degeneration of Purkinje cells, mitral cells of the olfactory bulb, thalamic neurons and retinal photoreceptor cells as well as degeneration of sperm (Fernandez-Gonzalez *et al.*, 2002, Mullen *et al.*, 1976). A total of 12 alleles in mouse mutants related to *Agtbbp1* have been summarized by the Jackson Laboratory (http://www.informatics.jax.org/searches/allele_report.cgi?_Marker_key=79447). Among which, loss-of-function of *Agtbbp1* was indicated as being responsible for *pcd*^{3j} and *pcd*^{5j} mouse mutants (Fernandez-Gonzalez, *et al.*, 2002; Wang and Morgan, 2007; Chakrabarti *et al.*, 2008). Expression of full-length *Nna1* successfully rescued Purkinje cell loss and ataxic behavior in *pcd*^{3j}, and retinal photoreceptor loss and *pcd* in *pcd*^{5j} mice (Wang *et al.* 2006, Chakrabarti *et al.*, 2008).

In this study, the missense mutation c.2909G>C on exon 21 of *AGTPBP1* induces an Arg to Pro substitution (p.Arg970Pro) at amino-acid 970 (R970). Based on the crystal structure of human carboxypeptidase A2, it has been reported that arginine 127 (R127) in carboxypeptidase A2 is an electrophile site crucially responsible for the catalytic activity along with glutamate 270 (Mangani *et al.*, 1994; Garcia-Saez *et al.*, 1997; Christianson *et al.*, 1989). Similarly, the homologous arginine residue R127 (numbered according to mature bovine carboxypeptidase A1) in *Caenorhabditis elegans* AGBL4 (AGP/GTP-binding-like protein 4) has been predicted to stabilize the oxyanion hole in the S1' site of the carboxypeptidase domain (Rodriguez de la Vega *et al.*, 2007). Sequence comparison and structural modeling of sheep AGTPBP1 with other carboxypeptidases predicted R970 was located in a region known as the substrate-binding pocket and catalytic center of metallo-carboxypeptidase as mouse *Nna1* (Supplementary Figure S3; Wang *et al.*, 2006). These results enforce the view that R970 in sheep is a very conserved site being critical for its function of catalytic activity. Therefore, we concluded the Arg970Pro substitution could be a candidate causative mutation possibly altering the characteristic of the AGTPBP1 protein and very likely impairing its function of catalytic activity.

Comparing to the phenotype of the original *pcd* mouse model, the affected sheep do not have apparent degeneration of Purkinje cells but show their specific characteristics including degeneration of neurons in ventral horns of the spinal cord and brain stem, Wallerian degeneration of motor nerve fibers and atrophy of skeletal muscles. The phenotype of these sheep is more similar to the symptoms in the *Sod1* transgenic mice (mSOD1), which were constructed as animal

Table 1 The genotyping results of three exonic SNPs

Sheep breed	LMND status	Exon 21				Exon 22							
		c.2909G>C (p.Arg970Pro)				c.3126C>T		c.3133T>C (p.Pro1045Ser)					
		CC	CG	GG	Rate of concordance (%)	TT	CT	CC	Rate of concordance (%)	CC	CT	TT	Rate of concordance (%)
Romney and Romney-sired crossbreeds	Affected	19	0	0	100	19	0	0	100	19	0	0	100
	Known carrier	0	13	0	100	3	10	0	76.9	10	3	0	23.1
	Phenotype normal ^a	0	13	9	100	1	13	6	95.0	18	2	0	10.0
	Normal	0	0	14	100								
Texel/Corriedale	Normal	0	0	71	100								

Abbreviations: LMND, lower motor neuron disease; SNPs, single-nucleotide polymorphisms.

^aThese sheep are including phenotypically normal and suspected sheep, which died without clear reasons.

models for human amyotrophic lateral sclerosis patients (Gurney *et al.*, 1994). It is interesting to note that the lysyl oxidase (*Lox*) gene, encoding a protein having an important role in the formation and repairing of extracellular matrix, was found to be upregulated in both mSOD1 mice and *Agtbbp*^{pcd-sid} mice (Li *et al.*, 2004; Li *et al.*, 2010; Lucero and Kagan, 2006). The *Agtbbp*^{pcd-sid} mutant was characterized as a deletion of *Agtbbp1* exon 7 that produced a stop codon in exon 8. The absence of protein *Agtbbp1* will decrease dendrite development of cerebellar Purkinje cells by changing the expression of downstream components, including *Lox* propeptide and the NF- κ B signaling pathway in mice (Li *et al.*, 2010). Therefore, we assume the p.Arg970Pro substitution might impair the catalytic activity of AGTPBP1 and decrease its ability of protein turnover of lysyl oxidase. The overabundant lysyl oxidase would cause abnormal development of dendrite in LMN and then develop phenotypes shown in these sheep.

To ascertain that the Arg970Pro substitution was really a causative mutation, measuring mRNA expression of the genes *AGTPBP1* and *LOX* and their enzymatic activities in the spinal cord, brain stem and cerebellar tissues would be desired. It could exclude the possibility that the Arg970Pro mutation in protein AGTPBP1 is only a coincidental finding from the genetic mapping study or this mutation disrupts expression of another gene in the same region of the genome. Nevertheless, a point mutagenesis of mouse *Nna1* gene at arginine 962 (R962), the homologous residue of R970 in sheep, to generate transgenic mice and then using the intact full-length *Nna1* gene to rescue phenotypes in mice will provide further evidence for the Arg970Pro substitution being a causative mutation in sheep.

The missense mutation c.2909G>C (p.Arg970Pro) in *AGTPBP1* appears to be a spontaneous one arising from a single sheep family. To avoid spreading out this disease throughout the sheep population, a simple genetic test can be easily designed based on our results to identify and remove carriers with the defective c.(2902C) allele, and to decrease the risk of economic losses in the future.

Regenerating motor neurons expressing protein NNA1 in human nervous tissues indicated that NNA1 was a key factor for motor neuronal degeneration (Harris *et al.*, 2000). To date, no human cases have been reported with spontaneous mutations in *AGTPBP1* that would implicate the gene as a cause for a LMN disease. Neither have they been reported to occur in other large mammalian species. Mice and hamsters are convenient animal models for studies of human neural diseases (Mashimo *et al.*, 2009; Akita *et al.*, 2007). However, there are still significant structural and functional differences between rodent and large mammalian species. Even among rodents themselves, the gene distributions could be different. Hamsters with undetectable *Nna1* mRNA in the brain did not exhibit distinctive neurodegenerative features as serious as mice mutants, which could possibly be explained due to the different distributions of *Nna1* and other *Nna1*-like genes in the brains of mice and hamsters (Kalinina *et al.*, 2007; Akita *et al.*, 2007; Akita and Arai, 2009). Sheep have large brain size, long life span and long gestation period relative to rodents. Our finding in sheep, a large mammalian species, raises the possibility that human patients with similar etiology may also result from this gene or other genes in the same pathway of neuronal development. As there is no effective treatment available for MNDs, the identification of molecular pathogenetic targets will shed light on the development of gene therapy for this type of diseases (Nizzardo *et al.*, 2011). Here, we reported that LMN disease originating in a Romney ram was likely caused by a mutation in a gene locus *AGTPBP1* (*NNA1*). The affected sheep could be used as a valuable animal model for gene therapy if human patients are successfully diagnosed in the future.

DATA ARCHIVING

Data deposited in the Dryad repository: doi:10.5061/dryad.8fp2pc88.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We appreciate the financial support provided by the State of Iowa and Hatch Funds and the support from the College of Agriculture and Life Sciences at the Iowa State University. Animal studies and diagnoses were funded by the Massey University. HTB was partially funded by the National Research Center for Growth and Development.

- Akita K, Arai S (2009). The ataxic Syrian hamster: an animal model homologous to the pcd mutant mouse? *Cerebellum* **8**: 202–210.
- Akita K, Arai S, Ohta T, Hanaya T, Fukuda S (2007). Suppressed *Nna1* gene expression in the brain of ataxic Syrian hamsters. *J Neurogenet* **21**: 19–29.
- Aloy P, Companys V, Vendrell J, Aviles FX, Fricker LD, Coll M *et al.* (2001). The crystal structure of the inhibitor-complexed carboxypeptidase D domain II and the modeling of regulator carboxypeptidases. *J Biol Chem* **276**: 16177–16184.
- Anderson PD, Parton KH, Collett MG, Sargison ND, Jolly RD (1999). A lower motor neuron disease in newborn Romney lambs. *NZ Vet J* **47**: 112–114.
- Berezniuk I, Sironi J, Callaway MB, Castro LM, Hirata IY, Ferro ES *et al.* (2010). CCP1/ *Nna1* functions in protein turnover in mouse brain: implications for cell death in Purkinje cell degeneration mice. *FASEB J* **24**: 1813–1823.
- Borasio GD, Gelinas DF, Mechanical YN (1998). Ventilation in amyotrophic lateral sclerosis: a cross-cultural perspective. *J Neurol* **245**(Suppl 2): S7–S12.
- Chakrabarti L, Eng J, Martinez RA, Jackson S, Huang J, Possin DE *et al.* (2008). The zinc-binding domain of *Nna1* is required to prevent retinal photoreceptor loss and cerebellar ataxia in Purkinje cell degeneration (*pcd*) mice. *Vision Res* **48**: 1999–2005.
- Chakrabarti L, Zahra R, Jackson SM, Kazemi-Esfarjani P, Sopher BL, Mason AG *et al.* (2010). Mitochondrial dysfunction in *Nna1* mutant flies and Purkinje cell degeneration mice reveals a role for *Nna* proteins in neuronal bioenergetics. *Neuron* **66**: 835–847.
- Christianson DW, Mangani S, Shoham G, Lipscomb WN (1989). Binding of D-phenylalanine and D-tyrosine to carboxypeptidase A. *J Biol Chem* **264**: 12849–12853.
- Dion PA, Daoud H, Rouleau GA (2009). Genetics of motor neuron disorders: new insights into pathogenic mechanisms. *Nat Rev Genet* **10**: 769–782.
- Fernandez-Gonzalez A, La Spada AR, Treadaway J, Higdon JC, Harris BS, Sidman RL *et al.* (2002). Purkinje cell degeneration (*pcd*) phenotypes caused by mutations in the axotomy-induced gene, *Nna1*. *Science* **295**: 1904–1906.
- Garcia-Saez J, Reverter D, Vendrell J, Aviles FX, Coll M (1997). The three-dimensional structure of human procarboxypeptidase A2. Deciphering the basis of the inhibition, activation and intrinsic activity of the zymogen. *EMBO J* **16**: 6906–6913.
- Gomis-Ruth FX, Companys V, Qian Y, Fricker LD, Vendrell J, Aviles FX *et al.* (1999). Crystal structure of avian carboxypeptidase D domain II: a prototype for the regulatory metallo-carboxypeptidase subfamily. *EMBO J* **18**: 5817–5826.
- Gurney ME, Pu H, Chiu AY, Dal Canto MC, Polchow CY, Alexander DD *et al.* (1994). Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation. *Science* **264**: 1772–1775.
- Harris A, Morgan JI, Pecot M, Soumare A, Osborne A, Soares HD (2000). Regenerating motor neurons express *Nna1*, a novel ATP/GTP-binding protein related to zinc carboxypeptidases. *Mol Cell Neurosci* **16**: 578–596.
- Kalinina E, Biswas R, Berezniuk I, Hermoso A, Aviles FX, Fricker LD (2007). A novel subfamily of mouse cytosolic carboxypeptidases. *FASEB J* **21**: 836–850.
- Kelley LA, MacCallum RM, Sternberg MJ (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J Mol Biol* **299**: 499–520.
- Lee CN (2012). Reviewing evidences on the management of patients with motor neuron disease. *Hong Kong Med J* **18**: 48–55.
- Li J, Gu X, Ma Y, Calicchio ML, Kong D, Teng YD *et al.* (2010). *Nna1* mediates Purkinje cell dendritic development via lysyl oxidase propeptide and NF-kappaB signaling. *Neuron* **68**: 45–60.
- Li PA, He Q, Cao T, Yong G, Szauder KM, Fong KS *et al.* (2004). Up-regulation and altered distribution of lysyl oxidase in the central nervous system of mutant SOD1 transgenic mouse model of amyotrophic lateral sclerosis. *Brain Res Mol Brain Res* **120**: 115–122.
- Lucero HA, Kagan HM (2006). Lysyl oxidase: an oxidative enzyme and effector of cell function. *Cell Mol Life Sci* **63**: 2304–2316.
- Mangani S, Ferraroni M, Orioli P (1994). Interaction of carboxypeptidase A with anions: crystal structure of the complex with the HPO₄²⁻ anion. *Inorg Chem* **33**: 3421–3423.
- Mashimo T, Hadjebi O, Amair-Pinedo F, Tsurumi T, Langa F, Serikawa T *et al.* (2009). Progressive Purkinje cell degeneration in tambaleante mutant mice is a consequence of a missense mutation in HERC1E3 ubiquitin ligase. *PLoS Genet* **5**: e1000784.
- Mullen RJ, Eicher EM, Sidman RL (1976). Purkinje cell degeneration, a new neurological mutation in the mouse. *Proc Natl Acad Sci USA* **73**: 208–212.
- Nizzardo M, Simone C, Falcone M, Riboldi G, Rizzo F, Magri F *et al.* (2011). Research advances in gene therapy approaches for the treatment of amyotrophic lateral sclerosis. *Cell Mol Life Sci*. (doi:10.1007/s00018-011-0881-5).

- Rodriguez de la Vega M, Sevilla RG, Hermoso A, Lorenzo J, Tanco S, Diez A *et al.* (2007). Nna1-like proteins are active metalloproteinases of a new and diverse M14 subfamily. *FASEB J* **21**: 851–865.
- Rogowski K, van Dijk J, Magiera MM, Bosc C, Deloulme JC, Bosson A *et al.* (2010). A family of protein-deglutamylating enzymes associated with neurodegeneration. *Cell* **143**: 564–578.
- Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A *et al.* (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**: 59–62.
- Wang T, Morgan JI (2007). The Purkinje cell degeneration (*pcd*) mouse: an unexpected molecular link between neuronal degeneration and regeneration. *Brain Res* **1140**: 26–40.
- Wang T, Parris J, Li L, Morgan JI (2006). The carboxypeptidase-like substrate-binding site in Nna1 is essential for the rescue of the Purkinje cell degeneration (*pcd*) phenotype. *Mol Cell Neurosci* **33**: 200–213.
- Zhao X, Dittmer KE, Blair HT, Thompson KG, Rothschild MF, Garrick DJ (2011). A novel nonsense mutation in the DMP1 gene identified by a genome-wide association study is responsible for inherited rickets in Corriedale sheep. *PLoS One* **6**: e21739.

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