Polymorphism analysis of prion protein gene in 11 Pakistani goat breeds

Mohammad Farooque Hassan\textsuperscript{a,\#}, Sher Hayat Khan\textsuperscript{a,\#}, Masroor Ellahi Babar\textsuperscript{b}, Lifeng Yang\textsuperscript{a}, Tariq Ali\textsuperscript{c}, Jamal Muhammad Khan\textsuperscript{d}, Syed Zahid Ali Shah\textsuperscript{a}, Xiangmei Zhou\textsuperscript{a}, Tanveer Hussain\textsuperscript{b,e}, Ting Zhu\textsuperscript{f}, Tariq Hussain\textsuperscript{a}, and Deming Zhao\textsuperscript{a}

\textsuperscript{a}Key Lab of Animal Epidemiology and Zoonosis, Ministry of Agriculture, National Animal Transmissible Spongiform Encephalopathy Laboratory, College of Veterinary Medicine, China Agricultural University, Beijing, China;  
\textsuperscript{b}University of Veterinary and Animal Sciences, Lahore, Pakistan;  
\textsuperscript{c}College of Veterinary Medicine, China Agricultural University, Beijing, China;  
\textsuperscript{d}National Animal Protozoa Laboratory & College of Veterinary Medicine, China Agricultural University, Beijing, China;  
\textsuperscript{e}Virtual University of Pakistan, Lahore, Pakistan;  
\textsuperscript{f}Fujian Agriculture and Forestry University, Fujian, China

ABSTRACT. The association between caprine PrP gene polymorphisms and its susceptibility to scrapie has been investigated in current years. As the ORF of the PrP gene is extremely erratic in different breeds of goats, we studied the PrP gene polymorphisms in 80 goats which belong to 11 Pakistani indigenous goat breeds from all provinces of Pakistan. A total of 6 distinct polymorphic sites (one novel) with amino acid substitutions were identified in the PrP gene which includes 126 (A \textsuperscript{\rightarrow} G), 304 (G \textsuperscript{\rightarrow} T), 379 (A \textsuperscript{\rightarrow} G), 414 (C \textsuperscript{\rightarrow} T), 428 (A \textsuperscript{\rightarrow} G) and 718 (C \textsuperscript{\rightarrow} T). The locus c.428 was found highly polymorphic in all breeds as compare to other loci. On the basis of these PrP variants NJ phylogenetic tree was constructed through MEGA6.1 which showed that all goat breeds along with domestic sheep and Maufflon sheep appeared as in one clade and sharing its most recent
common ancestors (MRCA) with deer species while Protein analysis has shown that these polymorphisms can lead to varied primary, secondary and tertiary structure of protein. Based on these polymorphic variants, genetic distance, multidimensional scaling plot and principal component analyses revealed the clear picture regarding greater number of substitutions in cattle PrP regions as compared to the small ruminant species. In particular these findings may pinpoint the fundamental control over the scrapie in *Capra hircus* on genetic basis.

**KEYWORDS.** Goat, neurodegenerative, PrP, PrP<sup>C</sup>, PrP<sup>Sc</sup>, Prion protein, scrapie, TSEs

**INTRODUCTION**

The accumulation of protein aggregates is an amalgamating feature of a large class protein conformational disorders that ultimately target the central nervous system and lead to progressive neurodegeneration e.g. Alzheimer disease and Prion diseases. The prion theory forecasts that certain mutations in PrP gene may result in unstable PrP molecules with transitional conformations, thus initiating the disease process and are accountable for much of the disparity in pathology and disease transmission. Transmissible spongiform encephalopathies (TSEs), or prion diseases, have cases that arise sporadically (Creutzfeldt-Jakob disease [CJD]) or are inherited (familial CJD). Transmission mechanism of prion diseases across the mammalian species is not as such effective as within species but being a zoonotic and fatal potential, scrapie prions indicated transmission efficacy comparable to BSE causing CJD.

Scrapie is a mortal, incurable and transmissible neurodegenerative disease of sheep and goats which occur due to the conversion of normal cell surface glycoprotein (PrP<sup>C</sup>) into conformationally altered isoform (PrP<sup>Sc</sup>) in the absence of nucleic acids. The key difference flanked by the 2 forms is that PrP<sup>C</sup> is entirely sensitive to proteases, while PrP<sup>Sc</sup> is only degraded in part and its amino-terminal end removed. As a natural host of scrapie both in sheep and goats showed analogous incubation period while the clinical manifestation perhaps slightly different. Moreover, the supplementary transmissible spongiform encephalopathies (TSEs) which comprise of bovine spongiform encephalopathy (BSE), Creutzfeldt–Jakob disease (CJD) and transmissible mink encephalopathy (TME), usually portrayed by accretion of an anomalous and disease specific isoform (PrP<sup>Sc</sup>) of an usual host-encoded cellular prion protein (PrP<sup>C</sup>) in the central nervous system. PrP<sup>C</sup> is also found in number of peripheral cell types along with glial cells of the CNS but in low level. Scrapie research has centered, for many years on developing more realistic live animal tests to make a diagnosis of infected animals; the identification of the genetic basis; and its probable application in breeding plans. Previously, it has been reported in Pakistan that caprine PrP polymorphisms was present in different breeds of sheep and goats such as in Beetal, Tedy and Pak Angora and it was associated with resistance to some important zoonotic and fatal group of diseases which may cross species barrier. Monitoring and investigation about the transmission of these agents via blood play vital role and has substantial concern regarding the safety of blood and its products.

In Pakistan goats are kept mainly for milk and meat purpose and play a critical role in the economy of country as well as farmers (Table 1). Presently there are about 58.3 million heads of goats in Pakistan and their population is increasing at rate of 3% per annum. As small ruminants are the only animals which are migrating in industrialized and developing countries, different zoonotic diseases spotted in small ruminants are mainly transmitted to human beings as occupational infections which badly affect the breeders and veterinarians. Therefore, the investigation of biological activity of PrP<sup>C</sup> remained crucial to understand the pathogenesis of prion associated diseases through contaminated food. Though no scrapie case has been documented yet in any of Pakistani goats, identifying the PrP gene polymorphism spectrum in these species may be useful for future breeding plans in terms of scrapie resistance.
MATERIALS AND METHODS

Study Animals and Management Practices

Pakistan is located in South Asia possessing tropical and subtropical environment that is known to be hot (35 + 13°C) and humid (45.38% ± 13.11%) in some areas.13 During summer in Punjab the temperature may go up to 48°C with humidity level (39%). The goats and other livestock animals in Pakistan are kept under similar natural climatic conditions.

TABLE 1. Phenotypic characteristics of goat breeds in Pakistan under study.

| Breed     | Type          | Body Color                        | Size  | Average body weight (kg) | General Description                                                                                                                                                                                                 |
|-----------|---------------|-----------------------------------|-------|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Barbari   | Milk          | White with small light brown patches | Large | Male(40)/female(24)      | Short size, tubular and erect ears which mainly found in Gujrat, Jhelum and Sargodha districts in Punjab Province.                                                                                                         |
| Beetal    | Milk/Meat     | Golden                            | Large | Male(55)/female(45)      | Massive head, Roman nose, long broad and pendulous ears; spiraled horns, longer in males; long stout legs; short tail; udder well-developed and long teats, milk yield 190 liters in a 150 days lactation; more than 50% twin or triplet births and Found in Punjab, Pakistan and India. |
| Bari      | Meat          | White with tan or black patches    | Medium| Male(30)/female(24)      | Long haired, all males and small number of females are bearded                                                                                                                                                        |
| Damani    | Milk          | black hair coat with tan head and legs | Medium| Male(35)/female(30)      | a long hair coat and well developed udders mainly found in the Bannu and Dera Ismail Khan districts in NWF Province, Pakistan.                                                                                         |
| Jattal    | Milk/Meat     | White and black or white with gray, red or black spots | Small | Male(23)/female(19)      | mainly found in AJK                                                                                                                                                                                                   |
| Kaghani   | Meat          | a white, gray, brown or black body coat | Large | Male(37)/female(32)      | used for the production of cashmere fiber and meat turned back horns and fine second hair coat in winter while Khurasan, Toba Kakari, Suleiman mountains region of Zhob and Sherani districts, Loralai, Ziarat, Chaghai and Pishin districts are its main niche |
| Khurasani | Milk/Meat     | black long hair coat               | Medium| Male(30)/female(25)      |                                                                                                                                                                                                                      |
| Labri     | Milk/Meat     | White long hairy coat on body      | Large | Male(45)/female(35)      | Chilas valley in Diamir district in Northern Areas                                                                                                                                                                   |
| Lehri     | Meat          | Brown and white spots on body      | Large | Male(43)/female(39)      | Found in Balochistan province and in Kachi area of Sibi                                                                                                                                                               |
| Pahari    | Milk/Meat     | White long hairy body              | Large | Male(50)/female(43)      | Found in Punjab and Balochistan provinces                                                                                                                                                                             |
| Tapri     | Milk/Meat     | White coated body                  | Medium| Male(33)/female(25)      | Found in Sindh province                                                                                                                                                                                               |
Adequate feed and water are bequeathed for their maintenance and production. To protect animals from harsh climatic conditions and predators semi closed traditional shelters system were employed where routine vaccination and de-worming were routinely carried out.

**Sample Collection**

A total of 80 blood samples (3ml from each), having specific phenotypic characteristics were cautiously collected from 11 Pakistani goat breeds as of Jattal (11), Barbari (07), Lebri (09), Tapri (03), Kaghani (05), Khurasani (08), Beetal (11), Barhi (03), Pahari (07) and Damani (06) from all 4 provinces of Pakistan along with Azad Jammu & Kashmir region (AJK) and stored at $-20^\circ$C till further processing. All the selected goat breeds were identified by an expert and their history about the age; breed and sex were also obtained from the farmers.

**Genetic Analysis**

Genomic DNA was extracted from EDTA-treated blood buffy coat using a standard DNA isolation protocol as previously described.\(^{16}\) A PCR was standardized for the specific amplification of 876 bp of exon 3 of the PrP gene and for this, a reaction mixture of a total volume of 30 $\mu$l was prepared containing nuclease free water, 1x PCR buffer (NH\(_4\)_2 SO\(_4\)), 50 mM MgCl\(_2\), 10 mM dNTPs mix, 20 pmol each of the forward (5'-1CTTTAAGTGATTTTTACGTGG21-3') and reverse primers (5'-854TGGCAAAGATTAA-GAAGATAATG876-3'), 0.2 unit of Taq DNA Polymerase (Fermentas, Cat #EP0402) and 1 $\mu$l of DNA as template. The standard amount of PCR reaction mixture, DNA samples, Primers and water was used for amplification of PrP gene. Amplification reactions were performed with a Thermal Cycler (iCycler BioRad, USA) conditions comprised of initial denaturation at 94°C for 05 minutes, followed by denaturation at 94°C for 45 sec, annealing at 60°C for 02 min and extension was at 72°C for 45 sec. Final extension was maintained at 72°C for 10 minutes and storage at 4°C. All amplicons were visualized by staining with ethidium bromide through the electrophoresis of a 05 $\mu$l reaction mixture on 2% agarose gels.

**Sequence Analysis**

Prior to Sanger sequencing, amplicons were cleaned using TIANquick Midi Purification Kit (Tiangen Biotech Beijing Co., Ltd.) while both sense and antisense strands of the amplified DNA were sequenced by using Big Dye terminator cycle sequencing chemistry v 3.1 (Applied Biosystems, USA) and electrophoresis was done by an automated DNA sequencer (ABI Prism 3130xL Genetic Analyzer, Applied Biosystems, USA) from 1st Base Laboratories Singapore. Sequences were analyzed through Codon Code aligner manually followed by the identification of novel SNPs detection.

**Bioinformatics Tools**

Codon Code aligner software was used to edit and align the PrP sequences while MEGA v.6 software\(^{17}\) was used to construct the phylogenetic trees and have insight of phylogeny of Pakistani goat breeds with other mammalian species (Fig. 1). We used homologous PrP sequences ranging from primate members to jawless vertebrate sequences. Moreover, the PrP gene coding sequence of all animals was sequenced and analyzed to identify polymorphisms in selected individuals. Single nucleotide polymorphism distributions were tabulated according to every nucleotide site (Table 2). Furthermore, variation profile was summarized in Table 3 where percentages with respect to being a breed a heterozygous, homozygous, or cumulative variation were calculated. Previously at the protein level, 40 6 amino acid polymorphisms have been described (Table 4).

**Genetic Distance**

Genetic distances between populations were calculated using Barry and Hartigan evolutionary model called ‘BH87’, asynchronous distance between homologous DNA sequences based on the observed proportions of changes among the 4
bases whereas “Bioconductor” statistical package in R was used to generate genetic distance matrix for selected species.

**Multidimensional Scaling Plot**

Multidimensional scaling method convert the structure in the similarity matrix to a simple geometrical picture: the larger the dissimilarity between 2 samples (evaluated through genetic distance profiling), the further apart the points representing the experiments in the picture may be. Furthermore, genetic relationship between groups were examined using an MDS plot of the genetic distance matrix in R.

**Principal Component Analysis**

Principal component analysis (PCA), a powerful multivariate investigative tool to identify patterns in specifically many dimensions, interrelated data sets and express the data sets by pinpointing their similarities and differences which is based on Eigen analysis. Therefore, PCA used here as a means of constructing an informative graphical representation of the data set by projecting the data onto a lower dimensional space. In the study, control and training data was presented in a 2 dimensions (2D) subspace of the first 2 PCs. The PCs derived by the Eigen analysis of correlation matrix ($R_{ij}$) was a linear combination of the original $p$ variables (the species) and each PC uncorrelated with the other, meaning these were the new transformed data expressed in terms of the patterns existing in the original data set. The total PCs derived were equal to the number of original variables present in the dataset. The PCs formed were with decreasing order of magnitude of variance of the total variation in the data sets. Thus the first 2 PCs capturing most of the variation in the data set is visualized in a 2D representation. The coefficient of the variables in each of the linear combination, i.e. the PC was defined as loadings. The magnitude of these loadings represented the importance of each variable present. Thus a 2D representation of the loadings of the first 2 PCs were identified as any cluster structure present in the variables (80 PrP homologs), exhibiting the co-occurring pattern of individuals in the control and training data sets.

**RESULTS**

**Observed SNP Profiles**

Total six polymorphic sites (at positions 126, 304, 379, 414, 428 and 718) were observed in 80 individuals of 11 breeds which were most
TABLE 2. Polymorphism observed in PrP gene in Pakistani goat breeds of meat and milk potential.

| Breed    | Sample ID | c.126 | c.304 | c.379 | c.414 | c.428 | c.718 |
|----------|-----------|-------|-------|-------|-------|-------|-------|
| Barbari  | B-32      | A     | C     |       |       |       |       |
|          | B-30      |       |       |       |       |       |       |
|          | B-36      | A/G   |       |       | C/T   | A/G   |       |
|          | B-56      | A/G   |       | C/T   | A/G   |       |       |
|          | B-57      |       |       |       |       |       |       |
|          | B-58      |       | A     | C     |       |       | T     |
|          | B-61      |       |       |       | A/G   |       |       |
| Lebri    | L-08      | A/G   | C/T   | A/G   | C/T   |       |       |
|          | L-09      | A/G   | C/T   | A/G   | C/T   |       |       |
|          | L-02      | A/G   | G/T   | C/T   | A/G   | C/T   |       |
|          | L-01      |       |       |       | G     |       |       |
|          | L-06      |       |       |       | G     |       |       |
|          | L-07      |       |       |       | A/G   |       |       |
|          | L-11      |       |       |       | A/G   |       |       |
|          | L-12      |       |       |       | A/G   |       |       |
|          | L-13      |       |       |       | G     |       |       |
| Tapri    | T-53      | A/G   | C/T   |       |       |       | C/T   |
|          | T-70      |       |       |       |       |       | A/G   |
|          | T-51      |       |       |       |       |       | G     |
| Bari     | Br-07     | A/G   |       |       |       |       |       |
|          | Br-02     | A/G   |       |       |       |       |       |
|          | Br-01     |       |       |       |       |       | G     |
| Khurasani| K-42      | A/G   | C/T   |       |       |       | C/T   |
|          | K-33      |       |       |       |       |       | A/G   |
|          | K-34      |       |       |       |       |       | A/G   |
|          | K-27      |       |       |       |       |       | G     |
|          | K-48      | A/G   | C/T   | A/G   | C/T   |       |       |
|          | K-31      | A/G   | C/T   | A/G   | C/T   |       |       |
|          | K-36      |       |       |       |       |       | C/G   |
| Beetal   | B-41      | A/G   |       |       |       |       |       |
|          | B-03      |       |       |       |       |       |       |
|          | B-04      |       |       |       |       |       |       |
|          | B-04      | A/G   |       |       |       |       |       |
|          | B-14      |       |       |       |       |       |       |
|          | B-20      |       |       |       |       |       |       |
|          | B-28      |       |       |       |       |       |       |
|          | B-46      |       |       |       |       |       | G     |
|          | B-73      | A/G   |       | A/G   | C/T   | A/G   |       |
|          | B-26      | A/G   |       |       |       | A/G   | C/T   |
|          | B-49      | A/G   |       |       |       |       |       |
| Jattal   | J-38      |       |       |       |       |       |       |
|          | J-39      |       |       |       |       |       | A/G   |
|          | J-43      |       |       |       |       |       | A/G   |
|          | J-47      |       |       |       |       |       | G     |
|          | J-49      | A/G   | G/T   | C/T   | A/G   |       |       |
|          | J-48      |       |       |       |       |       | A/G   |
|          | J-46      |       |       |       |       |       | G     |
|          | J-41      |       |       |       |       |       | A/G   |
|          | J-50      |       |       |       |       |       |       |
|          | J-51      |       |       |       |       |       |       |
|          | J-11      | A/G   | G/T   | C/T   | A/G   | C/T   |       |
| Kaghani  | Kg-06     | A/G   |       |       |       |       |       |
|          | Kg-04     |       |       |       |       |       | G     |
|          | Kg-07     |       |       |       |       |       | A/G   |
|          | Kg-02     |       |       |       |       |       | A/G   |

(Continued on next page)
notably raised for meat and milk production in Pakistan. In general, c.428 position was found highly polymorphic in all breeds with homozygous G, A and A/G heterozygous genotypes (Table 2). In total, 81.25% of individuals were found polymorphic out of them 71.01% individuals were heterozygous for this particular site. Lowest percentage of polymorphism was achieved at c.379 as 3.75%. Similarly, significant percentage of polymorphisms were obtained at c.126, c.304, c.414 and c.718 sites as 35%, 7.5%, 36.25% and 28.75% respectively (Fig. 2). Heterozygous sites were also evaluated in proportion to their relative cumulative polymorphic variation. All individuals showed their preference for being highly heterozygous as compared to be homozygous (Table 3).

Furthermore, distance analyses (Fig. 3), MDS (Fig. 4), PCA (Fig. 5) and phylogenetic tree (Fig. 1) was re-constructed with the help of these PrP variables to have an insight of Pakistani goat breeds phylogeny to deepen our understanding in this field of great concern. The results showed that our local goat breeds were in the clade suggesting their genetic relatedness depending on PrP gene. However, deer and antelope were closer relative of *Capra hircus* on the basis of PrP gene as compared to buffalo,

### TABLE 2. Polymorphism observed in PrP gene in Pakistani goat breeds of meat and milk potential. (Continued)

| Breed   | Sample ID | c.126 | c.304 | c.379 | c.414 | c.428 | c.718 |
|---------|-----------|-------|-------|-------|-------|-------|-------|
| Lehri   | Kg-03     | A/G   | G/T   | A/G   | A/G   | C/T   |
|         | Lh-43     | A     | A/G   | C     |       |
|         | Lh-36     |       |       |       |       | G     |
|         | Lh-33     |       |       |       |       | A/G   |
|         | Lh-30     |       |       |       |       | A/G   |
|         | Lh-35     |       |       |       |       | G     |
|         | Lh-40     |       |       |       |       | A/G   |
|         | Lh-41     |       |       |       |       | A/G   |
|         | Lh-37     | A     |       |       |       | C     |
|         | Lh-38     |       |       |       |       | C     |
| Pahari  | P-01      | A/G   | C/T   |       |       | A/G   |
|         | P-08      | A/G   | G/T   | C/T   |       |       |
|         | P-03      | A     | G     | C     |       |
|         | P-12      | A/G   | G/T   | C/T   | A/G   |
|         | P-09      |       |       |       |       | G     |
|         | P-04      |       |       |       |       | A/G   |
|         | P-06      |       |       |       |       | G     |
|         | P-07      |       |       |       |       | G     |
|         | P-08      |       |       |       |       | A/G   |
|         | P-09      |       |       |       |       | G     |
|         | P-10      |       |       |       |       | A/G   |
|         | P-11      |       |       |       |       | G     |
|         | P-12      |       |       |       |       | G     |
|         | P-13      |       |       |       |       | A/G   |
|         | P-14      |       |       |       |       | A/G   |

### TABLE 3. Level of polymorphism summarized according to percentage observed at each nucleotide position.

| Sample ID       | c.126   | c.304   | c.379   | c.414   | c.428   | c.718   |
|-----------------|---------|---------|---------|---------|---------|---------|
| Homozygous (%)  | 17.85714| 16.66667| 33.33333| 20.68966| 29.23077| 13.04348|
| Heterozygous (%)| 82.14286| 83.33333| 66.66667| 79.31034| 70.76923| 86.95652|
| Polymorphic (%) | 35      | 7.5     | 3.75    | 36.25   | 81.25   | 28.75   |
while domestic yak, European bison, gazelles and cattle comes after these species. Similarly, dromedary camel, llama and Bactrian camel from UK, China, Germany, Iran and Pakistan appeared as in a separate clade (Fig. 1).

**Variations at Protein Sequence Level**

Different softwares were used to check out the effect of nucleotide change on primary and secondary structures of protein.\(^5\) The sequence having nucleotide changes at all 6 positions was selected as representative sequence and translated by Expasy Translate tool. As the individuals in this study showed close relatedness with *Capra hericus* (Fig. 1) so the predicted ORF was then aligned with reference prion protein of *Capra hircus* by using Clustal Omega which resulted in 2 conservative, 2 semi-conservatives and 1 non-conservative variation (Figure S1).

Secondary structure of reference caprine prion protein (Figure S2), as well as altered prion protein was predicted by Phyre2 software which has shown absence of alpha helix at 113-117 and break down of single alpha halix (147-159) of PrPC into 2 (150-152 and 155-159) in major prion protein (Figure S3). These variations may cause change in folding pattern of protein which may convert PrPC to PrPSC.\(^6\) This shift could be a major cause of scrapie in goats.

**Genetic Distance Evaluation**

Distance profiling in biological sciences is an important technique to reveal the level of diversity among individuals of a population/

**TABLE 4. Single nucleotide polymorphisms identified in the PrP region of animals all over the world.**

| SNPs Reported in literature | SNPs Reported in literature |
|----------------------------|----------------------------|
| W18R\(^{32}\) | R151H\(^{35}\) |
| V21A\(^{33}\) | R154H\(^{33}\) |
| L23P\(^{33}\) | P168Q\(^{33}\) |
| G37V\(^{34}\) | V179V\(^{35}\) |
| S39R\(^{a}\) | D181O\(^{35}\) |
| P42P\(^{36}\) | I185F\(^{37}\) |
| G49S\(^{33}\) | T194P\(^{37}\) |
| Q101R\(^{32}\) | F201T\(^{32}\) |
| Q101O\(^{32}\) | T202T\(^{38}\) |
| G102W\(^{35}\) | K207K\(^{35}\) |
| K107K\(^{33}\) | R211Q\(^{39}\) |
| T110T\(^{34}\) | R211G\(^{30}\) |
| V125V\(^{40}\) | I218L\(^{26}\) |
| S127Q\(^{23,26}\) | T219R\(^{40}\) |
| L133Q\(^{38}\) | T219T\(^{41}\) |
| M137I\(^{38}\) | Q220H\(^{33}\) |
| S138S\(^{36}\) | Q222K\(^{34}\) |
| I142M\(^{36}\) | Q222Q\(^{32}\) |
| I142T\(^{37}\) | K231R\(^{36}\) |
| I421I\(^{32}\) | G232W\(^{32}\) |
| H143R\(^{36}\) | G232G\(^{31}\) |
| N146D\(^{30}\) | L237L\(^{3}\) |
| N146S\(^{35}\) | S240R\(^{30,42}\) |

**FIGURE 2.** NJ Phylogenetic tree constructed using maximum likelihood method between 11 Pakistani goat breeds and 40 other mammalian species. Bootstrap value was kept 1000.
Several algorithms were available in form of software packages as well as others have been implemented in Bioconductor. In this study, we used BH87 computational model to generate distance profile against the selected individuals of different breeds/species (Fig. 3). The outcome of this model was a table containing pair wise distance values in it. Diversity pattern, constructed using these computed distance values, is shown in Figure 3 (Distance Plot).18

**MDS Plot**

Multidimensional scaling (MDS) is a tool by which researchers can obtain quantitative estimates of similarity/dissimilarity among groups of items. We provided input in the form of pair wise distance matrix generated for all species mentioned above using BH87 statistical model (Fig. 3). The outcome of MDS, conveyed us the relationships among member sequences. Similar items wherein were seemed to be located proximal to one another, and more diverse items were found proportionately further apart. For example, 3 clusters were uniquely located at variable spots in the MDS diagram (Fig. 4). Member species of each cluster were more related to each other as compared to the individuals of a different cluster as shown in MDS plot. N-1 dimensions (N for number of sequences) were implicitly computed for our total sequence data set. We screened out 1st 2 dimensions only to observe variation pattern in the current data set.19

In this analysis, slightly scattered pattern of large ruminant cluster suggests greater substitution rates in their PrP region as compared to goat and canine species.

**2D Loading PCA**

Principal components analysis can be used to explore a large data matrix for the direction of largest variation. The whole idea of principal components analysis was to find new directions in the data along with maximal variation. A direction was defined as a linear combination \( Z_k \) of the data \( Z \) by a vector \( k \) with weights. So, we employed bioconductor implementation to achieve linear combinations with the help of 2 implicitly calculated Eigen values. These Eigen values provided transformations of our genetic dataset in order to give direction of maximum source variation in the distance profile.

PrP gene along with its homologs was the targeted genomic regions of the selected species. For this purpose, previously calculated genetic distance profile using BH87 computational model (Fig. 3), was used to generate initially 2 linear combinations such as PC1 and PC2 residing in the Figure 5 as vertical and horizontal axis. In total, 70 homologous genes from different organisms were included in our data set. According to PCA, it revealed that
FIGURE 4. Multidimensional Scaling plot of genetic distance between populations (Basques – closed circles, Iberia – open circles, Europe – closed squares, Middle East – cross, Caucasus – closed triangles, North Africa – open triangles). The first 2 axes account for highest genetic variation present in the sample and can be demonstrated that the plot is an accurate representation of the genetic distance matrix. The small ruminant group cluster together on the bottom-left side of the plot, near their bigger ruminant neighbors represented on the upper left place. The canine populations are found near the bottom-right of plot, differentiated from the other 2 ruminant groups.
highest variation may be due to the presence of top most cluster which were resident very far from 2 other clusters. In a group, the members were less variable with respect to each other.  

Final identification of the cluster structures in the binding site preferences of the training and control data set was achieved with the application of Principal Component Analysis (PCA). The first 2 PCs explained maximum of the total variation present in the patterned correlated data set. The 2D loading plot of the first 2 PCs derived from PCA on control data exposed a scenario of interactive pattern of prions across the species in the training data set. The 2D loading plot exhibits clarity of 3 distinct sets of closely packed prions, i.e., 3 cluster configurations of species. The cluster configurations presented coincide with the cluster configuration perceived in $R_X = [r_{ij}^2]$. The species which co-occurring with the PrPs packed in cluster-3 were also merged in it.

**Phylogenetic Tree Reconstruction**

A NJ phylogenetic tree was reconstructed using PrP sequences of 50 individuals including 10 sample sequences. Rest of other homologs was retrieved from GenBank. The rectangular view showed that all our local goat breeds placed themselves in one clade indicating their genetic relatedness based on PrP gene. In the tree diagram (Fig. 1), our breeds can be observed as the closest relative of *Capra hircus*, then domestic yak, European bison, gazelles and cattle comes after these species. Killer whale, dolphin appeared in a separate clade. Wild ass, horse and zebra came as close relative in another separate clade. Similarly, alpaca, llama and Bactrian camel from Pakistan appeared as in a separate clade. In the last clade fox, wolf, canada lynx, bobcat, cheetah, puma, mongoose, walrus, sea lions, seal, skunk bear and lesser panda appeared.
As far as the first clade was concerned, regular trend of all goats were observed in which all breeds showed close relatedness with each other. Another important thing was also observed that domestic sheep placed itself into goat clade. Barbari-58, Khurasani-36, Lehri-43, Kaghani-03, Beetal-26 and Mouflon sheep place themselves closer to other mammals.

**DISCUSSION**

Ovine and caprine management has endured huge economic losses due to scrapie epidemics in Europe. In the majority of countries, the prevalence of natural scrapie in goats remained underrated as compared to sheep. Studies to produce information on PrP gene polymorphisms and their distribution in goat breeds domesticated in diverse countries have become progressively more important, especially after the discovery of BSE or prions in goats in the 1990s.21 Surely, identification of candidate PrP gene polymorphic alleles that may confer susceptibility or resistance to goat scrapie, was a beginning step in understanding the likelihood of any selection program. Modern gene sequence technologies have now paved the path for breeders to make sheep and goats resistant against distressing disorders including scrapie by helping them in determining the allele frequencies, which is a prerequisite to devise an appropriate breeding plan for reasons of animal welfare as well as human food safety and food security.

This study provides new evidence in Pakistani goats for an association of PrP gene polymorphisms with resistance to classical scrapie. In contrast there is no indication for gene variants studied in an association with increased susceptibility to classical scrapie, such as the one found for codon 136 valine polymorphism of European ovine PrP.22 Of the 6 PrP coding region polymorphisms that were observed in investigated herds (G102W, S127G, N135S, H143R, K231R and S240P), one N135S was described for the first time in the Pakistani goat population. In conformity with previous association studies, we found that the goat PrP gene was highly variable in all the 11 breeds studied with majority of animals showing H143R polymorphism. Among caprine PrP polymorphisms I142M, H143R, N146S/D, R154H, R211Q and Q222K which has been reported both in European23 and Asian sheep breeds24 in association to scrapie resistance, only H143R was detected for the studied goats.13 This polymorphism has also been identified in other Pakistani goat breeds like Pak Angora, Teddy and Beetal8 along with G102W25 and K231R26 in European caprine (Table 5). Moreover, F141L and R154H that were risk factors for typical scrapie and has been reported in some European27,28,29,30 and Asiatic31 breeds, were not detected in the populations of Pakistani studied goats. Our genetic analysis also revealed an association of S127G with a decreased

**TABLE 5. Comparison between studied PrP sequences of animals all over the world with indiginous genotypes of goat breeds of Pakistan.** Entries in bold and underlined format in the reported column were also determined as amino acid polymorphisms in studied breeds. Similarly, one entry placed the 3rd spot in the right most column was the novel polymorphism detected in indigenous goat population.

| Residual Polymorphism | Reported | Reported in Pakistan | In our Breeds |
|-----------------------|----------|-----------------------|---------------|
| W18R                  | R151H    | S39R<sup>a</sup>       | G102W         |
| V21A                  | R154H    |                       |               |
| L23P                  | P168Q    |                       |               |
| G37V                  | V179V    | I142M<sup>13</sup>    |               |
| S39R                  | D181O    |                       |               |
| P42P                  | I185F    |                       |               |
| G49S                  | T194P    | H143R<sup>13</sup>    |               |
| Q101R                 | F201F    |                       |               |
| Q101Q                 | T202T    |                       |               |
| **G102W**             | K207K    | N146S/D<sup>13</sup>  | N135S         |
| K107K                 | R211Q    |                       |               |
| T110P                 | R211G    |                       |               |
| V125V                 | I218L    | R154H<sup>13</sup>    | H143R         |
| S127G                 | T219I    |                       |               |
| L133Q                 | T219T    |                       |               |
| M137I                 | Q220H    | I185F<sup>a</sup>     | K231R         |
| S138S                 | Q222K    |                       |               |
| I142M                 | Q222Q    |                       |               |
| I142T                 | K231R    | R211Q<sup>13</sup>    |               |
| I421I                 | G232W    |                       |               |
| **H143R**             | G232G    |                       | S240P         |
| N146D                 | L237L    | Q222K<sup>13</sup>    |               |
| N146S                 | S240P    |                       |               |
probability to develop clinical scrapie in agreement to European breeds. It appears that domestic goat populations maintain a number of PrP polymorphisms as might be expected from balancing selection and the majority of these polymorphisms were associated with a degree of protection against prions compared to the wild type. We were interested to see if a goat population from a region which had never any confirmed scrapie outbreaks would show equally diverse PrP genetics as the UK or other scrapie affected countries (Table 5). A comprehensive list of already reported polymorphisms as well as published from Pakistani biologists was generated. Out of 6 variants, 5 named G102W, S127G, H143R, K231R and S240P were in correspondence with the variant list from all over the world. Here we reported a novel genotype N135S identified in Pakistani goats would of great interest for future correspondence. In general, c.428 position was observed as highly polymorphic in all breeds with homozygous G, A and A/G heterozygous genotypes (Table 2). With respect to that site, approximately 81.25% of individuals were found polymorphic (Fig. 2). Out of them, 71.01% individuals were heterozygous for this particular site. Lowest percentage of polymorphism was achieved at c.379 as 3.75%. Similarly, significant percentage of polymorphisms were obtained at c.126, c.304, c.414 and c.718 sites as 35%, 7.5%, 36.25% and 28.75% respectively (Fig. 2). Whereas, heterozygous sites were evaluated as proportion to the cumulative polymorphic variation. Surprisingly, all species had adopted the heterozygous trend as compared to be homozygous (Table 3) which might had made them less susceptible to scrapie.

This important PrP sequences with important variation profiles from scrapie resistant individuals were further exploited to discover its evolutionary mechanism. In distance plot, MDS, PCA and phylogenetic analysis, these species/breeds had shown their origin were of great interest. Analyzing all sequences with respect to multiple angles generated their origin association with other species/breeds and on the basis of this gene deer and antelope species were clustered with *Capra hircus*. Through sequence analyses we came up with an important finding that killer whale has the lowest rate of variation while mongoose seemed to possess the largest rate of substitutions (Fig. 1). Similarly, MDS and PCA (Figs. 4 and 5) diagrams elucidated that all canine sequences had observed the highest level of variations described by PC1 (representing 75% of variation alone).

### ABBREVIATIONS

- MRCA: Most Recent Common Ancestors
- TSEs: Transmissible Spongiform Encephalopathies
- CJD: Creutzfeldt-Jakob disease
- TME: Transmissible mink encephalopathy
- GPI: Glycosyl-phosphatidyl inositol
- CDS: Coding DNA sequence
- EDTA: Ethylenediamine tetra-acetic acid
- SNPs: Single Nucleotide Polymorphism
- PrPc: Cellular prion protein
- PrPsc: Scrapie prion protein

### DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

We have no conflicts of interest to disclose; all authors approved the manuscript and this submission.

### FUNDING

This work was supported by the Natural Science Foundation of China (Project No. 31472166).

### REFERENCES

1. Aguzzi A, Lakkaraju AK. Cell Biology of Prions and Prionoids: A Status Report. Trends in cell biology. (2016); 26(1):40–51; http://dx.doi.org/10.1016/j.tcb.2015.08.007
2. Huang Z, Gabriel JM, Baldwin MA, Fletterick RJ, Prusiner SB, Cohen FE. Proposed three-dimensional structure for the cellular prion protein. Proc Natl Acad Sci 1994; 91(15):7139–43.
3. Colby DW, Prusiner SB. Prions. Cold Spring Harb. Perspect Biol 2011; 3:a006833; PMID:21421910
POLYMORPHISM ANALYSIS OF PRION PROTEIN GENE IN GOAT BREEDS

[4] Beringue V, Vilotte JL, Laude H. Prion agent diversity and species barrier. Vet Res 2008; 39(4):1–30; PMID:18519020

[5] Cassard H, Torres JM, Lacroux C, Douet JY, Bene- estad SL, Lantier F, Reine F. Evidence for zoonotic potential of ovine scrapie prions. Nat Commun 2014; 5:5821; PMID:25510416

[6] Wang J, Wang X, Gao X, Vortmeyer AO. Prion diseases and their prpsc-based molecular diag- nostics. J Neurol Neurosci 2002; 4(6):56.

[7] Westergard L, Heather M, Christensen David A, Harris . The cellular prion protein (PrPC): Its physiological function and role in disease. Biochimicaet-Biophysica Acta 2007; 1772:629–644

[8] Babar ME, Farid A, Benkel BF, Ahmad J, Sajid IA, Imran A, Hussain T. Genetic variability at seven codons of the prion protein gene in nine Pakistani sheep breeds. J Genet 2008; 87(2):187–90; PMID:18776650; http://dx.doi.org/10.1007/s12041-008-0029-z

[9] Ford MJ, Burton LJ, Morris RJ, Hall SM. Selective expression of prion protein in peripheral tissues of the adult mouse. Neuroscience 2002; 113:177–92; PMID:12123696

[10] Moser M, Colello RJ, Pott U, Oesch B. Developmental expression of the prion protein gene in glial cells. Neuron 1995; 14:509–17; PMID:7695897; http://dx.doi.org/10.1016/0896-6273(95)90307-0

[11] Babar ME, Farid A, Benkel BF, Ahmad J, Nadeem A, Imran M. Frequencies of PrP genotypes and their implication for breeding against scrapie susceptibility in nine Pakistani sheep breeds. Mol Biol Rep 2009; 36(3):561–5; PMID:18297414; http://dx.doi.org/10.1007/s11033-008-9214-7

[12] Mabbott, Neil Andrew. “Prospects for safe and effective vaccines against prion diseases.” Expert review of vaccines 14.1 (2015): 1–4. DOI: 10.1586/14706584.15.965691

[13] Hussain A, Babar ME, Imran M, Haq IU, Javed MM. Detection of four novel polymorphisms in PrP gene of Pakistani sheep (Damani and Hashnagri) and goats (Kamori and Local Hairy) breeds. Virol J 2011; 8(1):1.

[14] Harris DA, Peters PJ, Taraboulos A, Lingappa VR, DeArmond SJ, Prusiner SB. Cell biology of prions, in: SB Prusiner (Ed.), Prion Biology and Diseases. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY 2004; 483–544

[15] Hussein, Mansour F., and Saud I. Al-Mufarrej. “Prion Diseases: A Review.” Scientific Journal of King Faisal University (Basic and Applied Sciences) 5.2 (2004): 1425

[16] Maryam, J. Babar ME, Nadeem A, Hussain T. “Genetic variants in interferon gamma (IFN-γ) gene are associated with resistance against ticks in Bos taurus and Bos indicus.” Molecular biology reports 39.4 (2012): 4565–4570

[17] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA6: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011; 28:2731–9; PMID:21546353; http://dx.doi.org/10.1093/molbev/ms121

[18] Barry, Daniel, and J. A. Hartigan. “Statistical analysis of hominoid molecular evolution.” Statistical Science (1987): 191–207. http://dx.doi.org/10.1214/ss/1177013353

[19] Cox TF, Cox MA. Multidimensional scaling. Boca Raton, FL: CRC Press; 2000.

[20] Venables, WN., “Ripley,” BD. (1996). Modern Applied Statistics with S—Plus. 2nd edition. New York: Springer-Verlag; 2002

[21] Foster JD, Parnham DW, Hunter N, Bruce M. Distribution of the prion protein in sheep terminally affected with BSE following experimental oral transmission. J Gen Virol 2001; 82(10):2319–26; PMID:11562525

[22] Goldmann W, Ryan K, Stewart P, Parnham D, Xicohtencatl R, Fernandez N, Foster J. Caprine prion gene polymorphisms are associated with decreased incidence of classical scrapie in goat herds in the United Kingdom. Vet Res 2011; 42(1):110; PMID:22040234

[23] Holko I, Novackova A, Holkova T, Kmet V. PrP genotyping of sheep breeds in Slovakia. Vet Rec 2005; 157(20):628; PMID:16284332

[24] Barillet F, Mariat D, Amigues Y, Faugeras R, Caillat H, Moazami-Goudarzi K, Perrin-Chauvineau C. Identification of seven haplotypes of the caprine PrP gene at codons 127, 142, 154, 211, 222, and 240 in French Alpine and Saanen breeds and their association with classical scrapie. J Gen Virol 2009; 90(3):769–776; PMID:19218225

[25] Goldmann W, Chong A, Foster J, Hope J, Hunter N. The shortest known prion protein gene allele occurs in goats, has only three octapeptide repeats and is non-pathogenic. J Gen Virol 1998; 79:3173–6; PMID:9880037; http://dx.doi.org/10.1093/molbev/msr121

[26] Zang L, Li N, Fan B, Fang M, Xu W. PRNP polymor- phisms in Chinese ovine, caprine and bovine breeds. Anim Genet 2004; 35:457–61; PMID:15566469; http://dx.doi.org/10.1111/j.1365-2052.2004.01204.x

[27] Dawson, M. Hoinville LJ, Hosie BD, Hunter N, Venables, WN. “Ripley,” BD. (1996). Modern Applied Statistics with S—Plus. 2nd edition. New York: Springer-Verlag; 2002

[28] Buschmann, Anne, Lühken G, Schultz J, Erhardt G, Groschup MH. “Neuronal accumulation of...
abnormal prion protein in sheep carrying a scrapie-resistant genotype (PrPARR/ARR).” Journal of General Virology 85.9 (2004): 2727–2733.

[29] Goldmann W, Baylis M, Chihota C, Stevenson E, Hunter N. Frequencies of PrP gene haplotypes in British sheep flocks and the implications for breeding programmes. J Appl Microbiol. 2005;98 (6):1294–302. PMID: 15916643.

[30] Saunders GC, Cawthraw S, Mountjoy SJ, Hope J, Windl O. PrP genotypes of atypical scrapie cases in Great Britain. J Gen Virol. 2006 Nov;87(Pt 11):3141–9. PMID:17030846.

[31] Lan, Z. Wang ZL, Liu Y, Zhang X. “Prion protein gene (PRNP) polymorphisms in Xinjiang local sheep breeds in China.” Archives of virology 151.10 (2006): 2095–2101

[32] Vaccari G, Panayiotidis CH, Acin C, Peletto S, Barillet F, Acutis PL, Bossers A, Langeveld J, van Keulen L, Sclaviasi T, et al. State-of-the-art review of goat TSE in the European Union, with special emphasis on PRNP genetics and epidemiology. Vet Res 2009; 40:48; PMID:19505422; http://dx.doi.org/10.1051/vetres/2009031

[33] Billinis C, Panagiotidis HC, Psychas V, Argyroudis S, Nicolau A, Leontides S, Papadopoulos O, Sclaviasi T. Prion protein gene polymorphisms in natural goat scrapie. J Gen Virol 2002; 83:713–21; PMID:11842266; http://dx.doi.org/10.1099/0022-1317-83-3-713

[34] Agrimi U, Conte M, Morelli L, Di Bari MA, Di Guardo G, Ligios C, Antonucci G, Aufero GM, Pozzato N, Mutinelli F, et al. Animal transmissible spongiform encephalopathies and genetics. Vet Res Commun 2003; 27:31–8; PMID:14535366; http://dx.doi.org/10.1023/B:VERC.0000014115.18327.26

[35] Papasavva-Stylianou P, Kleanthous M, Toumazos P, Mavrikou P, Loucaides P. Novel polymorphisms at codons 146 and 151 in the prion protein gene of Cyprus goats, and their association with natural scrapie. Vet J 2007; 173:459–62; PMID:16314132; http://dx.doi.org/10.1016/j.tvjl.2005.09.013

[36] Goldmann W, Martin T, Foster J, Hughes S, Smith G, Hughes K, Dawson M, Hunter N. Novel polymorphisms in the caprine PrP gene: a codon 142 mutation associated with scrapie incubation period. J Gen Virol 1996; 77:2885–91; PMID:8922485; http://dx.doi.org/10.1099/0022-1317-77-11-2885

[37] Acutis PL, Colussi S, Santagada G, Laurella C, Maniaci MG, Riina MV, Peletto S, Goldmann W, Bossers A, Caramelli M, et al. Genetic variability of the PRNP gene in goat breeds from Northern and Southern Italy. J Appl Microbiol 2008; 104:1782–9; PMID:18217941; http://dx.doi.org/10.1111/j.1365-2672.2007.03703.x

[38] Acutis PL, Bossers A, Priem J, Riina MV, Peletto S, Mazza M, Casalone C, Forloni G, Ru G, Caramelli M. Identification of prion protein gene polymorphisms in goats from Italian scrapie outbreaks. J Gen Virol 2006; 87, 1029–33; PMID:16528054; http://dx.doi.org/10.1099/vir.0.81440-0

[39] Wopfner F, Weidenhofer G, Schneider R, Von Brunn A, Gilch S, Schwarz TF, Werner T, Schatzl M. Analysis of 27 mammalian and 9 avian PrPs reveals high conservation of flexible regions of the prion protein. J Mol Biol 1999; 289:1163–78; PMID:10373359; http://dx.doi.org/10.1006/jmbi.1999.2831

[40] Zhou RY, Li XL, Li LH, Wang HY, Lu JG. Polymorphism of the PRNP gene in the main breeds of indigenous Chinese goats. Arch Virol 2008; 153:979–82; PMID:18369524; http://dx.doi.org/10.1007/s00705-008-0074-1

[41] Vaccari G, Di Bari MA, Morelli L, Nonno R, Chiappini B, Antonucci G, Marcon S, Esposito E, Fazzi P, Palazzi N, et al. Identification of an allelic variant of the goat PrP gene associated with resistance to scrapie. J Gen Virol 2006; 87:1395–402; PMID:16605343; http://dx.doi.org/10.1099/vir.0.81848-0

[42] Kurosaki Y, Ishiguro N, Horiuchi M, Shinagawa M. Polymorphisms of caprine PrP gene detected in Japan. J Vet Med Sci 2004; 67:321–3; PMID:15805738; http://dx.doi.org/10.1292/jvms.67.321