Institute of Animal Breeding and Genetics

Justus Liebig University Giessen, Germany

Host-pathogen genetics and the epidemiology of lentivirus infection in German and Iranian sheep

Dissertation to obtain the doctoral degree of Agriculture (Dr. agr.)

Faculty of Agricultural Science, Nutritional Science and Environmental Management,

Justus Liebig University Giessen, Germany

presented by

Vahid Molaee

From Iran

Gießen

January 2021

1st Referee: Prof. Dr. Gesine Lühken

Institute of Animal Breeding and Genetics

Justus Liebig University Giessen, Germany

2nd Referee: Prof. Dr. Gerald Reiner

Department of Veterinary Clinical Science

Justus Liebig University Giessen, Germany

List of tables……..... III

List of figures…….….. III

List of abbreviations…….….IV

1 SUMMARY……..……..1

2 INTRODUCTIONS…….……...3

2.1 Small ruminant lentiviruses (SRLVs)…….……....3

2.2 The pathologic characteristics of SRLV infection in sheep (maedi-visna)……....3

2.3 History of MV disease and two hypotheses regarding its emergence……..…..4

2.4 Organisation and classification of SRLVs…….……6

2.5 The history of sheep domestication…….…...8

2.6 Host-pathogen co-evolution in lentiviruses……..…..9

2.7 The consequences of SRLV infection……...…….…..10

2.8 Virus transmission, diagnosis, and different strategies against SRLVs……...11

2.9 Host genetic factors involved in SRLV-associated disease…….….12

2.10 Aims and hypotheses……...…….…..14

3 ORIGINAL WORKS……...……...…….……15

3.1 Study 1: First survey on association of *TMEM154* and *CCR5* variants with serological maedi-visna status of sheep in German flocks……15

3.2 Study 2: Lentivirus susceptibility in Iranian and German sheep assessed by determination of *TMEM154* E35K……...……..31

3.3 Study 3: Phylogenetic analysis of small ruminant lentiviruses in Germany and Iran suggests their expansion with domestic sheep…….46

4 DISCUSSION…….65

4.1 Research topics……...……..65

4.2 Issue i……...…….65

4.2.1 Organisation of sampling……..……...….65

4.2.2 Evaluating the association of *TMEM154* and *CCR5* variants with SRLV infection status…….……..66

4.2.3 SRLV susceptibility in German and Iranian sheep based on the frequency of allele E……..…….…….....68

4.2.4 The epidemiology of SRLV infection in Germany and Iran……..…….. 70

4.3 Issue ii…….....71

4.4 Limitations……...73

4.5 Suggestions for future research…….……..…75

5 REFERENCES……..…….76

Declaration……...83

List of tables

Table Nr.	**Title**
1	SRLV subtypes A1–20 and their geographical distribution……..…….8
2	Overview of sampling in studies 1 and 2……..66
3	Association analyses of *TMEM154* with SRLV infection status in German and Iranian sheep breeds/mixed…….68

List of figure/s

Figure Nr.	**Title**
1	The proviral genome of MVV…….…….7

APOBEC3	apolipoprotein B mRNA-editing enzyme type 3
BIV	bovine immunodeficiency virus
CA	capsid protein
CAEV	caprine arthritis encephalitis virus
CCR5	chemokine (C-C motif) receptor type 5 gene
dUTPase	deoxyuridine triphosphatase
EIAV	equine infectious anaemia virus
ELISA	enzyme-linked immunosorbent assay
ENV	envelope
FIV	feline immunodeficiency virus
GAG	group-specific antigen
GWAS	genome-wide association study
HIV	human immunodeficiency virus
IN	integrase
LTR	long terminal repeat
MA	matrix
MHC	major histocompatibility complex
MV	maedi-visna disease
NC	nucleocapsid
OPP	ovine progressive pneumonia
OvLV	ovine lentivirus
PAMPs	pathogen-associated molecular patterns
PCR	polymerase chain reaction

POL	polymerase
PR	protease
REV	regulator of virion protein expression
RIPA	radioimmunoprecipitation assay
RT	reverse transcriptase
SIV	simian immunodeficiency virus
SRLVs	small ruminant lentiviruses
SU	surface
TLRs	toll-like receptors
TM	transmembrane
TMEM154	transmembrane protein 154
VIF	viral infectivity factor

Small ruminant lentiviruses (SRLVs) cause maedi-visna disease (MV) in sheep and are widespread throughout the world, including Germany and Iran. There is no vaccine or cure available for SRLVs. Generally, there are two strategies for controlling the spread of SRLVs: (1) Diagnosis of an SRLV infection followed by culling, segregation of infected animals and feeding their offspring with virus-free colostrum/milk. SRLVs have been classified into four genotypes: A–C and E. Genotype A is found mostly in sheep and has been subdivided into 20 subtypes. Because SRLVs have high variability, to design suitable diagnostic tests, it is critical to know the genetic characteristics of the circulating virus. There is no information on the genetic composition of German and Iranian SRLVs, making SRLV diagnosis in those countries difficult. (2) Breeding sheep to lower their susceptibility to SRLV infection. Recently, an amino acid substitution (E/K) at position 35 of the transmembrane protein 154 (*TMEM154*) was found in sheep in the United States, along with a deletion in the chemokine (C-C motif) receptor type 5 gene (*CCR5*), which, reportedly, affected the serological MV status or SRLV provirus concentration. Before present studies, no efforts have been made to evaluate the association of *TMEM154* or *CCR5* variants with SRLV infections in German or Iranian sheep.

The overall objective of the present studies was to extend our knowledge of host-pathogen genetics and the epidemiology of lentivirus infection in German and Iranian sheep. The research project was divided into two main parts for simplicity: the host and the pathogen genetic. The host genetic, together with the epidemiology of SRLV infection, was examined in studies 1 and 2. The pathogen genetic was explored in study 3.

In study 1, the association of the two candidate genes *TMEM154* and *CCR5* with SRLV infection was evaluated over 500 samples in four sample sets. The association of *TMEM154* with SRLV infection status was significant in Texel and Milk sheep. The deletion of the

CCR5 promoter did not show a consistent association with SRLV infection status. In study 2, we aimed at evaluating the association of *TMEM154* with SRLV infection in Iranian and some more German sheep flocks/breeds. The following results were obtained in study 2: (1) The association of *TMEM154* E35K with SRLV infection status was tested in four sheep breeds and found to be significant in Kermani, Merinoland, and Brown Hair; (2) regression analysis showed that SRLV susceptibility in a flock/breed could be predicted based on the frequency of allele E with a variation of approximately +/ – 20%; and (3) for the first time, SRLV positive sheep flocks were found in the Iranian province of Western Azarbaijan through a PCR-based test. The results of studies 1 and 2 showed that the amino acid substitution at position 35 of *TMEM154* could be a promising marker for breeding a lower number of serologically MV positive sheep in German and Iranian flocks. Concerning the epidemiology of SRLV infection in German sheep flocks, the range of infection varies from 10–100% and corresponds to positive flocks. Among Iranian sheep flocks, one-third of flocks (10 out of 30 flocks) were SRLV positive and had at least one sample that was SRLV positive. The range of infection within samples of individual flocks varied from 0–89%.

The initial aim of study 3 was the phylogenetic classification of German and Iranian SRLVs. Based on 54 SRLV *gag* sequences from German (n=48) and Iranian (n=6) sheep, the results revealed: SRLV subtypes A4, A5, A11, A16 and A21 (new) are found in German sheep, and, SRLV subtype A22 (new) are found in Iranian sheep. Notably, in study 3, further analyses provided additional conclusions: (1) It was shown that the genotype A likely has two ancestors; one is related to Turkish SRLVs, and the other is related to Iranian SRLVs; and (2) the transmission routes of some SRLVs are likely in line with the domestication pathways of sheep.

2.1 Small ruminant lentiviruses (SRLVs)

The lentivirus genus belongs to the Retroviridae family and infects a wide range of species. Lentiviruses include the human immunodeficiency viruses (HIV) in humans, the simians immunodeficiency viruses (SIV) in monkeys, the felines immunodeficiency viruses (FIV) in cats, the equine infectious anaemia virus (EIAV) in horses, the bovines immunodeficiency viruses (BIV) in cattle, the maedi-visna virus (MVV) in sheep, and the caprine arthritis encephalitis virus (CAEV) in goats (Blacklaws et al., 2004).

MVV-like and CAEV-like strains have been first described in sheep and goats, respectively, and considered strictly host-specific for a long time. However, there are nowadays several studies indicating that most strains can cross the species barrier (reviewed by Minardi da Cruz et al., 2013). Lentivirus infections of sheep and goats have recently been grouped as small ruminant lentiviruses (SRLVs) (Larruskain and Jugo, 2013).

Among lentiviruses, SRLV, EIAV, and BIV have the only tropism for the monocyte and macrophage cells, while the others infect lymphocytes as well (Larruskain and Jugo, 2013). Monocytes and lymphocytes make up 2–8% and 20–30% of white blood cells, respectively. Monocytes are found in blood and are bone marrow-derived leukocytes. Monocytes can differentiate into macrophages and dendritic cells, once absorbed into body tissues or extracellular fluid.

2.2 The pathologic characteristics of SRLV infection in sheep (maedi-visna)

SRLVs cause disease in sheep called maedi-visna (MV). The name maedi-visna is derived from two frequent clinical signs: pneumonia (maedi), and encephalitis (visna) (Blacklaws, 2012).

The primary outcome of SRLV infection in sheep is lung enlargement. An infected lung will grow to three times the average size, and present as rounded, rubbery and greyish-yellow (Gomez-Lucia et al., 2018). A secondary outcome of SRLV infection in sheep is mastitis. The signs of mammary disease are usually subclinical, indurative, chronic, and nonpainful. Mastitis can lead to economic losses due to a decline in milk production (Gomez-Lucia et al., 2018). A less frequent outcome of SRLV infection among sheep is arthritis, which generally affects the carpal and tarsal joints, and presents as lameness (Peterhans et al., 2004). The nervous form of SRLV infection has been observed in sheep flocks that were kept stable in Iceland, Spain, and the United Kingdom (Benavides et al., 2006). Although weight loss is an element of SRLV infection, unnatural thinness is less common (Peterhans et al., 2004). When considering SRLV infection, it is essential to emphasise that, usually, only a portion of infected animals develop clinical signs (Peterhans et al., 2004).

2.3 History of MV disease and two hypotheses regarding its emergence

Sigurdsson and colleagues discovered SRLV infection in sheep in Iceland (Sigurdsson et al., 1952). Visna virus was isolated from an infected sheep brain in 1957, kept frozen for several years. Maedi virus was isolated from the lung of a natural case of maedi in 1958. Both visna and maedi viruses were studied thoroughly in subsequent years and reported to be closely related, if not the same. While the Icelandic name of the disease, maedi-visna, has generally been accepted worldwide, ovine progressive pneumonia (OPP) and ovine lentivirus (OvLV) are more commonly used in the United States (Thormar, 2005).

Before the discovery of MV disease, the symptoms of it had already been detected in other countries, including South Africa, the United States, France, and the Netherlands under

different names. It was likely called Graaf-Reinet disease in South Africa in 1915, progressive sheep pneumonia or Montana sheep disease in the United States in 1923, la bouhite in France in 1942 and zwoegerziekte in the Netherlands in 1943 (Thormar, 2005). Maedi has been reported in several countries to date, including Iran (Norouzi et al., 2015), Iraq (Hamza and Özkan, 2017), Kenya, India, Kyrgyzstan, Canada, the United States and most European countries (Thormar, 2005). Australia and New Zealand are the only geographic areas where MV in sheep has never been observed (Thormar, 2005).

MV was first detected in East and West Germany in 1967 and 1970, respectively. When MV become widely recognised in Germany in 1969, research into the disease began. This research continued until the year 2000 at the Federal Research Centre in Tübingen, where four sheep breeds were studied: Texel, Finnsheep, Ile de France, and Merinoland. All breeds except Merinoland developed clinical signs of the disease (Straub, 2004). Since 2004 there has been no updated information relating to SRLV infection in Germany.

In Iran, SRLV infection was first diagnosed in sheep using histopathological methods in 2001 (Sayari and Soltanian, 2001). Following this first survey, SRLV infection was reported in different parts of Iran, using serological methods or PCR techniques. Relevant reports mostly cover small areas of the country at the provincial level. The prevalence rate varies in different provinces, and its range is 6–72% in infected flocks (Azizi et al., 2012; Sasani et al., 2013; Norouzi et al., 2015).

There are two hypotheses regarding the emergence of SRLVs: (1) It has been proposed that MV disease, due to SRLV infection, has appeared in Iceland through a flock of Karakul rams imported from Halle in Germany in 1933. That Karakul sheep flock was first brought to Germany from Astrakhan in Russia in 1901 and still existed in East Germany in 1970 (Thormar, 2005). Following MV discovery and its characterisation in Iceland, infections were detected in various countries although with differing prevalence (Pépin et al., 1998).

(2) The second hypothesis, which was proposed by Muz and colleagues (2013), indicated that the ancestors of all SRLVs, especially genotype A, are Turkish. The hypothesis suggests that the domestication process (10000 years ago) or a more recent transmission event during the Ottoman Empire (14th–19th centuries), resulted in the evolution of some SRLVs of genotype A (Muz et al., 2013).

2.4 Organisation and classification of SRLVs

SRLVs have a genetic organisation typical of lentiviruses: their genome is a dimer of RNA of positive-strand polarity, about 10 Kb in size, which is reverse transcribed into the chromosomal DNA (the provirus). Proviral DNA is flanked by a non-coding long terminal repeat (LTR) at both ends containing promoter elements that provide cis signals required for transcription of DNA. Between the LTRs, there are three structural genes, including *gag* (group-specific antigen), *pol* (polymerase), and *env* (envelope), and also several auxiliary genes. Gene *gag* encodes the capsid protein (CA), the matrix (MA), and the nucleocapsid (NC) proteins. *Pol* encodes the enzymes, namely protease (PR), reverse transcriptase (RT), deoxyuridine triphosphatase (dUTPase), and integrase (IN). Finally, *env* encodes the surface (SU) and the transmembrane (TM) glycoproteins. Both *gag* and *env* genes contain conserved regions. They are suitable candidates to be used for diagnosis by ELISA and PCR-based techniques. Auxiliary genes, including *tat*, *vif* (viral infectivity factor), and *rev* (regulator of virion protein expression) contain sequences that regulate viral replication (Pépin et al., 1998; Gomez-Lucia et al., 2018). Schematic representation of the organisation of the MVV genome is given in Fig 1.

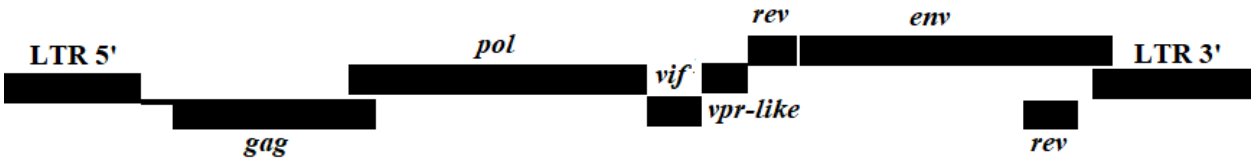


Figure 1. The proviral genome of MVV (modified from Gomez-Lucia et al., 2018).

SRLVs are quite heterogeneous. Already a nomenclature was proposed by Shah et al. (2004) for classification of SRLVs, based on DNA sequences of two long segments of *gag-pol* (1.8 kb) and *pol* (1.2 kb). SRLV genotypes differ about 25–37%, and SRLV subtypes vary about 15–27% (Shah et al., 2004). SRLVs are classified into four genotypes (A–C and E). Genotype A has 20 recognised subtypes to date (A1–A20), genotype B has three subtypes (B1–B3), and genotype E has two subtypes (E1 and E2) (Kuhar et al., 2013; Ramírez et al., 2013; Olech et al., 2018, 2019; Colitti et al., 2019). Genotype D has been found in Swiss and Spanish sheep but now reclassified as genotype A (Bertolotti et al., 2011). Although SRLV infection in small ruminants is widely distributed around the world, no knowledge is available on SRLV variants circulating in Germany and Iran. Distribution of SRLV subtypes of genotype A in different countries is shown in Table 1.

Table 1. SRLV subtypes A1–20 and their geographical distribution (Ramírez et al., 2013; Colitti et al., 2019; Olech et al., 2019).

	SRLV subtype																			
Country	A 1	A 2	A 3	A 4	A 5	A 6	A 7	A 8	A 9	A 10	A 11	A 12	A 13	A 14	A 15	A 16	A 17	A 18	A 19	A 20
Italy	x							x	x	x	x								x	x
Poland	x											x	x			x	x	x		
Switzerland	x		x	x	x	x	x													
Turkey		x	x		x				x		x									
Slovenia					x									x	x					
Finland	x	x																		
France	x					x														
Spain	x		x																	
Brazil	x																			
Canada		x																		
England	x																			
Iceland	x																			
Ireland	x																			
Netherlands	x																			
Norway	x																			
Portugal	x																			
South Africa	x																			
USA		x																		

2.5 The history of sheep domestication

Over the years, many researchers have shared their knowledge of animal domestication. Among them is Mrs Melinda A. Zeder, an American archaeologist, who revolutionised people's understanding of animal domestication. Zeder demonstrated that early domesticated sheep, which were morphologically wild but managed, appeared in the ancient Fertile Crescent region, including parts of Iran, Iraq, Turkey, Syria, and Jordan, approximately 10,000 to 8,000 years before present (BP). The domestication of goats started in the same region between 500 and 1,000 years earlier, at about 11,000 BP (Zeder, 1999, 2008; Zeder and Hesse, 2000). Later, host molecular genetic data combined with archaeological evidence

indicated that, during two domestication waves, sheep were distributed from the ancient Fertile Crescent to the West and the East (Chessa et al., 2009). Zeder demonstrated the first wave, and, later, Chessa and colleagues described the second wave (Zeder, 1999, 2008; Chessa et al., 2009). Recently, Zeder, in her book, showed that domestication was distributed from west Asia (Iran and Turkey) to Europe and Africa through three domestication pathways (Zeder, 2017). The three proposed domestication pathways are as follows: (1) the Danubian pathway (first pathway), which relates to the Danube river distribution, starting in Germany and passing through different European countries, eventually reaching the Black Sea (a geographical region of Turkey); (2) the northern Mediterranean route (second pathway), which starts in Turkey and comprises those countries geographically located above the Mediterranean Sea (e.g. southern Europe, including Italy), of which the Italian island of Sardinia and its history has been linked to the first wave of domestication; and (3) the southern Mediterranean pathway (third pathway), including those places located below the Mediterranean Sea, from Iran and Turkey to northern parts of Africa (Zeder, 2017).

2.6 Host-pathogen co-evolution in lentiviruses

Recent studies have demonstrated the usefulness of pathogens to elucidate their hosts' evolution across time and location (Gifford, 2012; Kaewthamasorn et al., 2018; Mühlemann et al., 2018; Otchere et al., 2018). Lentiviruses have a deep evolutionary history and have evolved alongside their mammalian hosts (Gifford, 2012). Previous reports have shown that investigating the phylogeny of SRLVs has the potential to enhance our knowledge of sheep and goat domestication (Reina et al., 2006, 2010; Bertolotti et al., 2011). The identification of SRLV subtype B3 in sheep/goats from Italy and Turkey as well as the finding that some bulk milk samples from Turkish sheep and goats were reactive against antigen derived from genotype E (a rare genotype found in Sardinia and other parts of Italy), supports the

hypothesis of migration of domesticated sheep from the Fertile Crescent into the Mediterranean Basin during the Neolithic age (Bertolotti et al., 2011). However, phylogenetic studies involving SRLV sequences from other regions of the Fertile Crescent (except Turkey), which would potentially enhance our knowledge of the domestication process within the domestication origin itself, have been absent until now. While few SRLV sequences from Jordan and Lebanon are available in the database, no SRLV sequence information has been published from Iran, Iraq, and Syria. Also, while previous studies related to Italy (located on the second pathway of domestication) and Turkey proposed the association of SRLVs with domestication process (Bertolotti et al., 2011; Muz et al., 2013), there is no report about SRLVs of Germany (located on the first domestication pathway).

2.7 The consequences of SRLV infection

The economic losses due to SRLV infection are significant. Milk production can be decreased by approximately 10% due to mastitis, which is often a feature of SRLV infection. Infection increases somatic cell counts in the milk, which may negatively impact its quality. Animals persistently infected with SRLVs may not meet the quality requirements as consumer demand increases for products from animals in excellent health (Peterhans et al., 2004). Low birth weight negatively affects development in both lambs and kids, thus lowering productivity, particularly in heavily infected flocks (Peterhans et al., 2004). SRLV infection inhibits weight gain in lambs and is caused by low-quality milk yield from infected ewes. Weight loss may range from 0.3 to 3.0 kg per infected lamb at weaning (Peterhans et al., 2004). Mortality rates due to SRLV infection depend on genetic background and environmental factors. During the epizootic of MV in Iceland, observations indicated that mortality might reach 20–30% in newly infected animals (Peterhans et al., 2004).

The effects of SRLV infection have a significant impact also on animal welfare due to pain and disability (Peterhans et al., 2004). The term "animal welfare" is used concerning SRLV infection aspects that directly affect animals' quality of life.

2.8 Virus transmission, diagnosis, and different strategies against SRLVs

The main transmission routes of SRLVs are vertical (via ingestion of infected colostrum or milk) and horizontal (via inhalation of respiratory secretions). Also, as SRLVs are present in all body fluids, venereal transmission (via mating) could be a risk factor (Peterhans et al., 2004).

Lentivirus infection in sheep is usually detected by serological or PCR tests. Serological tests include agar gel immunodiffusion (AGID), enzyme-linked immunosorbent assay (ELISA), Western blot, and radioimmunoprecipitation assay (RIPA). PCR tests include detecting the proviral DNA of the virus or isolation of the virus (RT-PCR) (Ramírez et al., 2013).

There is no gold standard for detecting SRLV infection. PCR tests with ELISA based tests can improve specificity and sensitivity compared with single ELISA or PCR (Extramiana et al., 2002; De Andres et al., 2005; Reina et al., 2006). So far, peripheral blood mononuclear cells are the primary source of target DNA. Milk, colostrum, semen, and synovial fluid give a lower sensitivity and are less reliable than blood as an SRLV DNA source for PCR (Reina et al., 2009).

There is no vaccine or drug treatments against SRLVs at present. Intervention strategies against SRLV infection are including diagnostic testing followed by culling SRLV positive animals and separation of offspring/s from the dam at birth to prevent vertical transmission. However, these strategies are neither cost-effective nor sustainable as SRLV-free flocks are still susceptible to infection if exposed to other infected sheep or goats. Another strategy is the breeding of sheep for lower SRLV infection susceptibility. Certain sheep breeds have

been associated consistently with higher (e.g., Columbia sheep) or lower (e.g., Rambouillet) SRLV infection. Also, SRLV susceptibility among different sheep breeds varies based on the determination of SRLV proviral concentration (White and Knowles, 2013).

2.9 Host genetic factors involved in SRLV-associated disease

SRLV induce both innate and adaptive (humoral and cellular) immune responses. Various research studies have revealed the presence of different loci that could influence resistance or susceptibility to SRLV infection, suggesting that variation in genetic factors might affect the outcome. Several host genetic factors that affect SRLV infection are shown below:

The major histocompatibility complex (*MHC*) region is a polymorphic multi-gene, has been implicated in SRLV infection. *MHC* genes located on chromosome 20 of sheep and chromosome 23 of goats. Genes of the *MHC* Class I and II encode those receptor glycoproteins that bind and present antigenic peptides to T cells and finally initiate the immune response (Herrmann-Hoesing et al., 2008; Larruskain et al., 2010).

Cytokines are signalling proteins that affect the immune system and initiate many processes. Several types of cytokines have been found that will regulate up or down in SRLV infection, although more data on cytokines is necessary. Therein, chemokine receptor type 5 (*CCR5*) is an example of how polymorphisms could affect SRLV infection. A four nucleotide deletion in this locus has been associated with an almost four-fold reduction in transcription in the US sheep (Larruskain and Jugo, 2013). Genes such as Toll-like receptors (*TLRs*) and their variation have been investigated against lentivirus infection in various species and sheep. *TLRs* modulate both the innate and the adaptive immune response mechanisms in response to pathogen-associated molecular patterns (PAMPs), including viral components.

In a study, a significant association was found between variants in *TLR7* and *TLR8* genes and SRLV infection (Mikula Jr and Mikula Sr, 2011; Larruskain and Jugo, 2013; Sarafidou et al., 2013; White and Knowles, 2013).

Further research is needed concerning additional innate immunity genes. For instance, Apolipoprotein B mRNA-editing enzyme type 3 (*APOBEC3*) proteins, which act against a wide variety of retroviruses, including HIV, have been studied in sheep and may be favourable candidates for SRLVs. A human *APOBEC3G* variant has been associated with rapid HIV-1 disease progression (Larruskain and Jugo, 2013).

Heaton et al. (2012) have evidenced that specific haplotypes of the gene encoding an ovine transmembrane protein 154 (*TMEM154*) are associated with the sheep's SRLV infection (Heaton et al., 2012). GWAS identified the *TMEM154* gene, and accordingly, sequencing methods detected a range of potentially functional mutations in this gene. Specifically, variants that contain glutamic acid (E) at position 35 and asparagine (N) at position 70 (variant 3), or isoleucine (I) at position 70 (variant 2), full-length version of the protein, are associated with increased susceptibility to the lentivirus, whereas variants that contain lysine (K) at position 35 or deletion mutants are associated with reduced susceptibility. Sheep homozygous for haplotype 1 were less susceptible to SRLV infection. However, sheep with at least one copy of either haplotype 2 or haplotype 3 were 69 times more susceptible. Besides, possibly, sheep homozygous for *TMEM154* haplotype 1 may also have lower proviral concentrations and lesion severity among infected sheep and would be less likely to spread the infection to other sheep. The three haplotypes observed at the highest frequencies were named haplotypes 1–3, and phylogenetic analysis showed that haplotype 3 was ancestral in ruminants (Heaton et al., 2012; Leymaster et al., 2013; White and Knowles, 2013). Although little is known about the functions of the *TMEM154*, among different

candidate genes, variants in the *TMEM154* gene have resulted in consistently replicated association for a validated genetic marker test (White and Knowles, 2013).

2.10 Aims and hypotheses

1) Testing the association of *TMEM154* and *CCR5* variants with SLRV antibody titers (ELISA test) in German sheep populations (studies 1 and 2).

2) Examing the association of *TMEM154* variants with SRLV provirus presence (PCR test) in Iranian sheep breeds (study 2).

3) Comparing SRLV susceptibility in 13 German and 8 Iranian sheep breeds based on the frequency of the *TMEM154* E allele (study 2).

4) Establishment of a PCR test for the diagnosis of Iranian SRLVs (study 2).

5) Providing general knowledge regarding the epidemiology of SRLV infection in Germany and Iran (studies 1 and 2).

6) Phylogenetic classification of German and Iranian SRLVs (study 3).

7) Investigating the evolution of German and Iranian SRLVs with those from other countries (study 3).

8) Assessing the relationship between the distribution of SRLVs and domestication pathways (study 3).

9) Checking whether the Icelandic SRLV subtype (A1) is found in German sheep or not (study 3).

Study 1: First survey on association of *TMEM154* and *CCR5* variants with serological maedi-visna status of sheep in German flocks

Molaee *et al. Vet Res (2018) 49:36*
https://doi.org/10.1186/s13567-018-0533-y

VETERINARY RESEARCH

RESEARCH ARTICLE

Open Access

First survey on association of *TMEM154* and *CCR5* variants with serological maedi-visna status of sheep in German flocks

Vahid Molaee[†], Marwa Eltanany and Gesine Lühken[*†]

Abstract

Maedi-visna, a disease caused by small ruminant lentiviruses (SRLVs), is present in sheep from many countries, also including Germany. An amino acid substitution (E/K) at position 35 of the transmembrane protein 154 (*TMEM154*) as well as a deletion in the chemokine (C-C motif) receptor type 5 gene (*CCR5*) were reported to be associated with the serological MV status and/or the SRLV provirus concentration in North American sheep populations. The aim of this study was to test if those two gene variants might be useful markers for MV susceptibility in Germany. For this purpose, more than 500 sheep from 17 serologically MV positive German sheep flocks with different breed backgrounds were genotyped applying PCR-based methods. Both, crosstab and non-parametric analyses showed significant associations of the amino acid substitution at position 35 of *TMEM154* with the serological MV status (cut-off-based classification) and the median MV ELISA S/P value in all samples and in two of the four analyzed breed subsets. The deletion in the *CCR5* promoter did not show a consistent association with serological MV status or median ELISA S/P value. It can be concluded that the amino acid substitution at position 35 of *TMEM154* is a promising marker for breeding towards a lower number of serologically MV positive sheep in German flocks, at least in flocks of the Texel breed, while this remains questionable for the deletion in the *CCR5* promoter. The findings of this study still need to be verified in additional sheep breeds.

Introduction

Small ruminant lentiviruses (SRLVs) belong to the family *Retroviridae* and cause diseases called maedi-visna (MV) in sheep and caprine arthritis encephalitis (CAE) in goats. Maedi-visna is widespread in sheep from many countries around the world. No European country except for Iceland can be considered to be free of SRLV infection [1]. The disease is also substantially spread among sheep flocks in Germany [2]. There is no cure for the chronic disease caused by SRLVs, which includes symptoms such as pneumonia, wasting, mastitis, arthritis and progressive paralysis [3]. A vaccine preventing SRLV infection has not been developed yet [4]. Production losses stem from lamb mortality, lower lamb weights and milk production from older infected ewes [5, 6], early culling [1] and export restrictions [7]. The significant economic losses have led to the development of control programs in Europe and elsewhere which commonly include separation of lambs from dams at birth to prevent virus transmission and test/cull methods. Although they can be successful [1], these methods are neither cost-effective [8] nor sustainable, as SRLV-free flocks are still susceptible to infection if exposed to other infected sheep or goats [9].

Due to these limitations of conventional strategies for the control of SRLV infections, a genetically based approach would be favorable. However, significant evidence for host genetic variation of resistance/susceptibility is a basic requirement for such a strategy [10]. Indeed, there is a confirmed genetic predisposition to resist SRLV infection at different levels among and within breeds [11–14].

*Correspondence: Gesine.Luehken@agrar.uni-giessen.de
[†]Vahid Molaee and Gesine Lühken contributed equally to this work
Department of Animal Breeding and Genetics, Justus Liebig University of Giessen, Ludwigstrasse 21, 35390 Giessen, Germany

The apparent differences in MV susceptibility between sheep breeds led to studies on possibly associated host genetic variation and finally to the detection of variants in several genes (e.g. *TMEM154*, *CCR5*, *MHC*, *ZNF389*, *TLRs* and *APOBEC3*) showing association with parameters of MV susceptibility [15–17]. Specifically, variants in the genes coding for transmembrane protein type 154 (*TMEM154*) and chemokine (C-C motif) receptor type 5 (*CCR5*) have been reported to be associated with the serological status and/or the provirus concentration of sheep in MV affected US sheep flocks [18, 19].

After the initial discovery of a possible involvement of *TMEM154* in the genetic control of MV susceptibility by genome-wide association studies, several polymorphic positions were identified in the coding region of this gene. Two *TMEM154* haplotypes, both carrying a nucleotide coding for glutamate (E) at amino acid position 35 of the *TMEM154* protein were associated with susceptibility to SRLV infection (as determined by the serological MV status of sheep), whereas a third haplotype, carrying a nucleotide coding for lysine (K) at position 35 was not [19]. This association was confirmed in additional cohorts in the same and in other studies [20, 21], but to our knowledge until now has not been tested in sheep populations outside of North America.

Approaching potential genetic factors for MV susceptibility from a different point of view, existing data in diverse species indicated that *CCR5* could be a candidate gene for resistance to various pathogens [16, 22–27]. Even more important in the context of this study is that humans carrying *CCR5* variants seem not to acquire human immunodeficiency virus (HIV) [28–30], the best-explored lentivirus. Based on this knowledge, White et al. [18] analyzed the *CCR5* variation in sheep and tested a 4-base promoter deletion for association with proviral levels of SRLV in sheep. Individuals carrying two copies of the mutant variant (deletion) had significantly reduced proviral levels [18].

Thus, the amino acid substitution at position 35 of *TMEM154* as well as the *CCR5* promoter deletion are promising candidates as selection tools to decrease MV susceptibility in sheep. However, once a genetic association with pathogen resistance has been detected in a population and/or country, the practical relevance of the identified genetic marker(s) cannot be generalized. In other populations and/or countries, variations in pathogen strains, host breeds, and environmental conditions may impair the effect of the identified variant(s). In this respect, the repeated testing in different animal sets is necessary and can confirm or disprove a proposed candidate gene [17]. Therefore, the aim of this study was to test if the amino acid substitution at position 35 of *TMEM154* as well as the *CCR5* promoter deletion could be useful markers for selection against MV susceptibility in the German sheep population. We report the results of the analysis of their association with the serological MV status and median MV ELISA S/P value of more than 500 sheep from 17 MV positive German sheep flocks with different breed backgrounds.

Materials and methods

Collection of animal samples

A total of 656 sheep from 23 German sheep flocks were sampled between 2014 and 2016 to evaluate the serological MV status. From the oldest ewes (minimum age was 4 years) of each flock whole blood samples were drawn from the jugular vein into 9 mL EDTA monovettes. In order to get a higher number of samples, in one flock (which was of special interest because it contained two pure breeds), samples were collected from all ewes with a minimum age of 3 years (few samples from even younger ewes) and also from rams older than 4 years. The sampled sheep flocks were located in the German states Schleswig-Holstein, Nordrhein-Westfalen, Mecklenburg-Vorpommern, Hessen and Baden-Württemberg. Only samples from flocks with serologically MV positive sheep were genotyped and included into association analyses, and divided into four sample sets (breed subsets) according to their breed background.

Serological testing for MV status

The plasma was separated by centrifugation at 3000 *g* for 10 min and stored at −20 °C. Plasma samples were shipped to the laboratory of the Animal Health Service of Thüringen (Jena, Germany) for serological testing with an enzyme-linked immunosorbent assay (ELISA) (IDEXX CAEV-MVV Total Ab ELISA, IDEXX GmbH, Ludwigsburg, Germany), according to the manufacturer's instructions. This ELISA is one of three ELISA tests officially approved for SRLV diagnosis in sheep and goats by the German licensing authority [Friedrich-Loeffler-Institute (FLI), Federal Research Institute for Animal Health, Island of Riems, Germany].

According to the guidelines of the used ELISA kit, the cut-off value is defined based on the corrected optical density (OD) at a wavelength of 450 nm ratio of sample to positive control (S/P):

$$S/P = 100 \times \frac{\text{mean OD 450 value of the sample} - \text{mean OD 450 value of the negative controls}}{\text{mean OD 450 value of the positive controls} - \text{mean OD 450 value of the negative controls}}$$

Samples were considered as serologically MV negative with an S/P value ≤ 110% and recorded as positive with an S/P value ≥ 120%. Suspicious results were in the range between 110 and 120% and excluded from further analyses.

Genotyping of *TMEM154* and *CCR5* variants

After plasma separation, residual blood was stored at −20 °C until extraction of genomic DNA using a modified salting out method [31].

The KASP technology (LGC, Hoddesdon, UK) was used for genotyping a nucleotide substitution in the coding region of *TMEM154* (*Ovis aries* chromosome 17, Oar_v4.0, NC_019474.2:g.4860407G>A), resulting in the substitution of the ancestral glutamic acid (E) with lysine (K) at position 35 of the mature protein. For this purpose, a common forward primer and two allele-specific primers were designed and synthesized by LGC (primer details are given in Additional file 1). Polymerase chain reactions (PCR) including the respective primers, the KASP master mix and 50 ng DNA were set up as recommended by LGC. PCR amplification was done in a Bio-Rad CFX96 thermal cycler (Bio-Rad Laboratories, München, Germany) under the following conditions: initial denaturation at 94 °C for 15 min, 10 touchdown cycles at 94 °C for 20 s, 61–55 °C (decreasing 0.6 °C per cycle) for 1 min, and 26 cycles at 94 °C for 20 s and 50 °C for 1 min, followed by a final cycle at 37 °C for 1 min. Fluorescence measurement and end point allelic discrimination were done in the same instrument and with the included software (CFX Manager 3.1).

Fragment-length analysis was used for determination of the presence or absence of a deletion in the promoter region of *CCR5* (*Ovis aries* chromosome 19, Oar_v4.0, NC_019476.2:g.52961717_52961714delAATG, minus strand). The reverse primer of the selected primer pair was fluorescence-labeled (primer details are given in Additional file 1). PCR reactions were performed in a final volume of 15 µL containing 20–70 ng of template DNA, 10 pmol of each primer, 2 mM dNTPs, 1× Go Taq Flexi PCR buffer and 0.5 units Go Taq Polymerase (Promega, Mannheim, Germany). The following PCR conditions were used: initial denaturation at 95 °C for 1.5 min, followed by 30 cycles at 95 °C for 15 s, 52 °C for 15 s and 72 °C for 15 s, and a final extension at 72 °C for 5 min. Fragment-length analysis of denaturated PCR products was done with an ABI 3130 automated sequencer and the software GeneMapper version 4.0 as recommended by the manufacturer (Applied Biosystems, Darmstadt, Germany).

Both genotyping methods were tested for their accuracy by a comparison of assay-derived genotypes with the results of direct Sanger sequencing. For this purpose, PCR products including the genotyped variants were amplified and sequenced (primer details are given in Additional file 1).

Statistical analyses

The SPSS program (version 23.0) for Windows (IBM SPSS Statistics, Armonk, NY: IBM Corp) was used for statistical analyses.

The independent segregation of alleles was assessed for Hardy–Weinberg equilibrium (HWE) by using Fisher's exact test. Differences in the distribution of allele and genotype frequencies between groups of serologically MV negative and positive samples were tested by Fisher's exact test (when expected values were < 5) or χ^2 test. Consistent with a dominant effect of the risk allele and a recessive effect of the protective allele, frequencies of genotypes with one or two copies of the putative risk allele (*TMEM154* E, or the wild type variant of the *CCR5* promoter) were combined.

The Shapiro–Wilk's test [32] was used to check whether ELISA S/P values were normally distributed. ELISA S/P values were described using median and interquartile range (IQR). Independent sample nonparametric analysis using median test was performed to compare ELISA S/P values of groups. For multiple comparisons, the threshold for significance was corrected using Bonferroni's correction.

The relative risk (RR) to be serologically MV positive (in a MV affected flock) was estimated for animals carrying one and/or two copies of the putative susceptible allele (risk factor) with the method of Altman [33] using the following equation:

$$RR = \frac{a/(a + b)}{c/(c + d)}$$

where *a* is the number of serologically MV positive individuals carrying the risk factor, *b* is the number of serologically MV negative individuals carrying the risk factor, *c* is the number of serologically MV positive individuals carrying no risk factor, and *d* is the number of serologically MV negative individuals carrying no risk factor.

Results

Serological MV status of sampled flocks and breed composition of MV positive flocks

A total of 656 samples from 23 flocks, originating from five German states, were sampled and serologically tested for MV status. Samples from 11 sheep with MV ELISA S/P values between 110 and 120% were excluded from all analyses. Details on state of origin, breed composition, sample numbers and percentage of serologically MV positive samples for each flock are given in Additional file 2. In six of these flocks, all tested samples ($n = 115$) were

serologically negative. The ELISA S/P values of samples from these six flocks, which were not included in further analyses, ranged from 0.43 to 72.73% with a median of 9.69% and interquartile range from 6.23 to 16.32%.

In contrast, in the other 17 sampled sheep flocks, a minimum of 10% up to 100% of the collected samples showed ELISA S/P values higher than 120%. These flocks were considered to be MV affected and assigned to four breed subsets according to the breed composition (details in Additional file 2). Subset 1 (TEX-x) included samples from 15 different flocks which all consisted of purebred and/or crossbred German Texel sheep (flocks 1–12 and 14–16a). Subsets 2–4 included samples from single flocks each, with purebred German Blackheaded Mutton (GBM) in subset 2, purebred and crossbred Merinoland sheep (MLS-x) in subset 3, and East Friesian milk and Lacaune sheep and the respective cross-breeds (EFM–LAC) in subset 4 (flocks 16b, 22, 23, respectively). The Texel and German Blackheaded Mutton sheep from flocks 16a (part of subset 1) and 16b (subset 2) were kept together.

In all 17 MV affected flocks, ELISA S/P values ranged from −6.34 to 109.65% in the group of serologically negative samples ($n=207$) and from 121.89 to 328.63% in the group of serologically positive samples ($n=323$). More descriptive MV-ELISA S/P data, including each breed subset, are given in Additional file 3. The Shapiro–Wilks test for normality showed a significant deviation of the ELISA S/P values from normal distribution in serologically negative and positive sheep in all samples and within breed subsets, except for the small group of serologically positive GBM (subset 2). In positive samples it was right-skewed (medians higher than means) and in negative samples it was left-skewed (medians lower than means).

Frequencies of putative risk/protective alleles and genotypes of *TMEM154* and *CCR5* in serologically MV positive and negative sheep

In all genotyped samples from 17 MV affected flocks, the putative protective allele (K) and the risk allele (E) at amino acid position 35 of *TMEM154* were observed at frequencies of 48 and 52%, respectively. Deviation from the Hardy–Weinberg equilibrium was only observed for breed subset 3 (MLS-x). In the breed subsets GMB, MLS-x and EFM–LAC the K allele was more frequent (63, 80 and 68%, respectively), whereas in the subset TEX-x it was less frequent (33%) than the E allele.

Comparing groups of serologically MV negative and positive sheep for the *TMEM154* variation, in all samples and in all four breed subsets, frequencies of the protective allele (K) and genotype (KK) were higher in serologically MV negative sheep than in serologically MV positive sheep (Table 1). Vice versa, frequencies of the risk allele (E) and the risk genotypes (EK and EE) were higher in MV positive sheep. These differences between groups of MV negative and positive sheep, either for comparison of the putative protective allele against the risk allele, or the protective genotype against the risk genotypes, were statistically significant in all samples. Concerning breed subsets, this association was significant in TEX-x and EFM–LAC, narrowly missed the significance threshold in MLS-x, and was not significant at all in the GMB subset (Table 1).

In all genotyped samples from 17 MV affected flocks, the putative protective *CCR5* promoter deletion was

Table 1 *TMEM154* E/K allele and genotype frequencies in serologically MV positive and negative sheep

Breed subset (n sheep)	MV status (n sheep)	*TMEM154* allele frequency (n chromosomes)		*P* value	*TMEM154* genotype frequency		*P* value
					(n sheep)	(n sheep)	
		K	E		KK	EK, EE	
All (527)	Negative (206)	0.65 (268)	0.35 (144)	< 0.001	0.48 (100)	0.52 (106)	< 0.001
	Positive (321)	0.36 (233)	0.64 (409)		0.15 (47)	0.85 (274)	
TEX-x (341)	Negative (95)	0.48 (92)	0.52 (98)	< 0.001	0.30 (29)	0.70 (66)	< 0.001
	Positive (246)	0.26 (130)	0.74 (362)		0.04 (9)	0.96 (237)	
GBM (39)	Negative (35)	0.66 (46)	0.34 (24)	0.140	0.40 (14)	0.60 (21)	0.277
	Positive (4)	0.37 (3)	0.63 (5)		0.00 (0)	1.00 (4)	
MLS-x (125)	Negative (62)	0.85 (105)	0.15 (19)	0.067	0.74 (46)	0.26 (16)	0.067
	Positive (63)	0.75 (95)	0.25 (31)		0.59 (37)	0.41 (26)	
EFM–LAC (22)	Negative (14)	0.89 (25)	0.11 (3)	0.001	0.79 (11)	0.21 (3)	0.006
	Positive (8)	0.31 (5)	0.69 (11)		0.13 (1)	0.87 (7)	

TEX-x: purebred and crossbred German Texel sheep, GBM: purebred German Blackheaded Mutton sheep, MLS-x: purebred and crossbred Merinoland sheep, EFM–LAC: East Friesian milk and Lacaune sheep and crosses of both breeds.

observed at a relatively low frequency of about 16%. Its frequency was even lower in the breed subsets EFM–LAC (10%) and TEX-x (11%), whereas it was higher in MLS-x (21%) and GBM (40%). No deviation from the Hardy–Weinberg equilibrium was found for *CCR5* genotype frequencies in all samples and in breed subsets.

In all sheep as well as in the breed subsets TEX-x and EFM–LAC, the putative protective allele (del), as well as the putative protective genotype (del/del) of the *CCR5* promoter occurred at higher frequencies in the MV negative compared to the MV positive sheep (Table 2). However, the opposite situation was observed in the breed subsets GBM and MLS-x, where this allele and genotype were observed more frequently in MV positive than MV negative sheep. The differences in allele frequencies between MV negative and positive samples were statistically significant in all samples, but not in any of the breed subsets. In all samples and in all breed subsets, no significant association with the serological MV status was found for the putative protective and risk genotypes of the *CCR5* promoter (Table 2).

Median MV ELISA S/P values of sheep with and without putative risk alleles of *TMEM154* and *CCR5*

For *TMEM154*, in all sheep and in all breed subsets, except for GBM, the median ELISA S/P values of sheep with risk allele (*TMEM154* genotypes EE and EK) were higher than the cut-off value (Figure 1, solid red line) and varied around 190%. In contrast, in all sheep and in all breed subsets, the median ELISA S/P values of sheep without risk allele (carrying the genotype KK) were lower than the cut-off value (Figure 1). In all sheep and in subsets TEX-x and EFM–LAC, the median ELISA S/Ps of sheep with the protective genotype (KK) were lower compared to the putative risk genotypes (EE or EK). In detail, in all samples and in TEX-x and EFM–LAC breed subsets, the median of S/P values from sheep with the protective genotype (KK) were 13.76, 21.69 and 25.11 fold lower than those of sheep with risk genotypes (EK or EE). These differences were not significant for the breed subsets GBM and only approached statistical significance ($P=0.077$) in MLS-x.

The median ELISA S/P value of sheep from the 6 MV negative flocks is indicated in Figure 1 (black dotted line). Interestingly, this median ELISA S/P value was not significantly different (pairwise comparisons revealed $P>0.05$) from those of sheep from the subsets GBM, EFM–LAC and TEX-x, carrying no *TMEM154* risk allele. In contrast, a noticeably high median ELISA S/P value (84.67%) was observed in sheep from the subset MLX-x with this genotype (KK). It differed significantly (P values of pairwise comparisons ranged from 0.023 to <0.001) from those of sheep with the same genotype from other subsets as well as from those of sheep from negative flocks.

Concerning sheep with and without *CCR5* risk allele (promoter deletion), the differences in median ELISA S/P values in all sheep and in breed subsets did not reach statistical significance in any group (Additional file 4).

Relative risk to be serologically positive for sheep with and without putative risk alleles of *TMEM154* and *CCR5*

The relative risk to be serologically MV positive was calculated for sheep carrying one or two copies of the putative risk allele (*TMEM154*: E at position 35, *CCR5*: wild type promoter sequence), and compared to that of sheep carrying no risk allele. In the breed subgroups GBM and

Table 2 *CCR5* promoter variant (wild type/deletion) allele and genotype frequencies in serologically MV positive and negative sheep

Breed subset (n sheep)	MV status (n sheep)	*CCR5* allele frequency (n chromosomes)		*P* value	*CCR5* genotype frequency		*P* value
		del	wt		(n sheep) del/del	(n sheep) del/wt, wt/wt	
All (521)	Negative (206)	0.19 (77)	0.81 (335)	0.021	0.04 (8)	0.96 (198)	0.519
	Positive (315)	0.13 (85)	0.87 (545)		0.03 (9)	0.97 (306)	
TEX-x (337)	Negative (96)	0.14 (26)	0.86 (166)	0.177	0.02 (2)	0.98 (94)	0.321
	Positive (241)	0.10 (49)	0.90 (433)		0.01 (2)	0.99 (239)	
GBM (39)	Negative (35)	0.37 (26)	0.63 (44)	0.253	0.11 (4)	0.89 (31)	0.321
	Positive (4)	0.63 (5)	0.37 (3)		0.25 (1)	0.75 (3)	
MLS-x (125)	Negative (62)	0.17 (21)	0.83 (103)	0.135	0.02 (1)	0.98 (61)	0.114
	Positive (63)	0.25 (31)	0.75 (95)		0.10 (6)	0.90 (57)	
EFM–LAC (20)	Negative (13)	0.15 (4)	0.85 (22)	0.278	0.08 (1)	0.92 (12)	1.000
	positive (7)	0.00 (0)	1.00 (14)		0.00 (0)	1.00 (7)	

del: deletion, wt: wild type, TEX-x: purebred and crossbred German Texel sheep, GBM: purebred German Blackheaded Mutton sheep, MLS-x: purebred and crossbred Merinoland sheep, EFM–LAC: East Friesian milk and Lacaune sheep and crosses of both breeds.

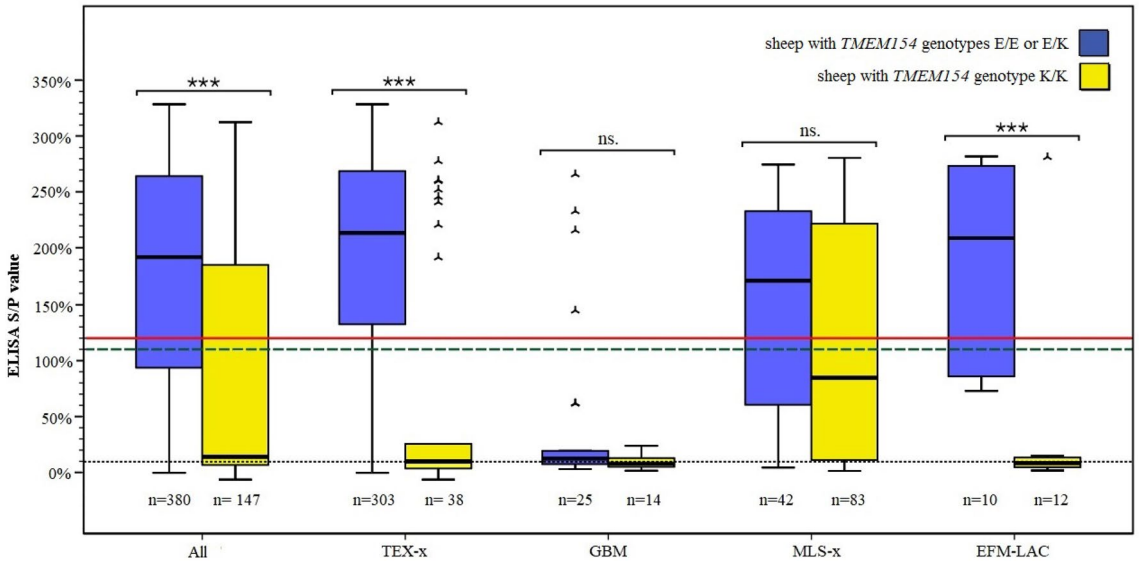


Figure 1 Box plots depicting MV ELISA S/P values in sheep carrying genotypes with (blue) and without (yellow) the putative *TMEM154* risk allele (E), in all sheep and in breed subsets. P values are resulting from nonparametric analyses comparing median MV ELISA S/P values of groups. ***$P < 0.001$, *ns* not significant ($P > 0.05$). The black dotted line indicates the median ELISA S/P value of sheep from six serologically MV negative flocks (about 9%). Sheep with ELISA S/P values below 110% (green dashed horizontal line) were considered serologically MV negative. Sheep with ELISA S/P values over 120% (red solid line) were considered serologically MV positive. TEX-x: purebred and crossbred German Texel sheep, GBM purebred German Blackheaded Mutton sheep, MLS-x: purebred and crossbred Merinoland sheep, EFM–LAC: East Friesian milk and Lacaune sheep and crosses of both breeds.

Table 3 Relative risk of infection in sheep with one or two copies of the *TMEM154* E allele

Breed subset	Parameters	*TMEM154* genotypes EE, EK vs. KK
All	Relative risk	2.255
	95% CI	1.767–2.878
	P value	< 0.001
TEX-x	Relative risk	3.302
	95% CI	1.860–5.862
	P value	< 0.001
MLS-x	Relative risk	1.389
	95% CI	0.991–1.945
	P value	0.056

CI: confidence interval, TEX-x: purebred and crossbred German Texel sheep, MLS-x: purebred and crossbred Merinoland sheep.

EFM–LAC, zero observations in serologically MV positive or negative groups produced large confidence intervals. Therefore, this analysis was only done in all sheep and for the two larger breed subsets, TEX-x and MLS-x (Table 3). In all samples and in the breed subgroup TEX-x, sheep carrying at least a single copy of the *TMEM154* risk allele had a significantly higher risk to be serologically MV positive than sheep with the KK genotype (relative risk of 2.26 and 3.30, respectively). Within the breed subset MLS-x, the relative risk to be serologically MV positive for sheep carrying one or two copies of the *TMEM154* risk allele was only 1.38. Moreover, the *P* value narrowly missed the significance threshold in this breed subset (Table 3).

A significant difference in the relative risk to be serologically MV positive for sheep with one or two copies of the *CCR5* promoter wild type allele, compared to sheep without this putative risk factor, was only observed for the breed subset MLS-x. However, in this subset, the risk for sheep to be serologically positive with one or two copies of the putative risk allele was about half (0.56-fold) of that of sheep without this allele (Additional file 5).

Discussion

This study aimed at evaluating the association of sequence variants in two candidate genes for MV susceptibility (*TMEM154* and *CCR5*) with serological MV infection status (cut-off-based classification) as well as with ELISA values (S/P ratio) in German sheep flocks. For this purpose, samples were collected from sheep flocks already known, or suspected to be MV affected, without considering their breed background. It was remarkable but not unexpected that the great majority (15 out of 17) of the MV positive flocks contained

purebred and/or crossbred sheep of the breed Texel (TEX-x), allowing clustering of these samples into the largest breed subset for following analyses. This high proportion of MV positive TEX-x flocks among the sampled flocks may be explained on the one hand by the fact that Texel is the dominating breed in commercial sheep flocks in Northern Germany, where most flocks for this study were sampled. On the other hand, it should also be taken into consideration that the Texel breed is known for high MV susceptibility, based on both seroconversion rates and number of sheep showing clinical signs in an infected flock [13]. In the past it was suspected that this cosmopolitan breed might be responsible for the introduction of MV into some previously MV-free countries, e.g. Great Britain [11, 13]. One of the other breed subsets of this study included sheep from two milk sheep breeds, East Friesian milk and Lacaune, and their crosses (EFM–LAC). In Germany, the breeds Texel and East Friesian milk are targeted in regional or federal MV monitoring and eradication programs (e.g., via serological testing and culling). Also in other countries, East Friesian milk sheep were observed to be higher susceptible to MV than other breeds [12]. In a German study on MV seropositive sheep of the breeds Texel, Finnsheep, Ile the France and Merinoland, sheep of all breeds except for Merinoland developed clinical signs of disease [13].

In all samples from 17 MV affected German sheep flocks, both crosstab and non-parametric analyses showed a significant association between the amino acid substitution at position 35 of *TMEM154* and both the serological MV status (cut-off-based classification) and the median ELISA S/P value. More precisely, the *TMEM154* genotype KK was associated with a lower number of serologically MV positive sheep and with a lower median ELISA S/P value, whereas for sheep with one or two copies of the E allele, the situation was reversed. The same was observed for the breed subsets, partly with an even more explicit effect of the genotype KK. However, this was not statistically significant in the breed subsets GBM and MLS-x. For the GBM subset, the missing significant association can be explained by the very low number of serologically positive sheep (four out of 39 sheep). It is of interest that these purebred GBM sheep were kept together with purebred TEX sheep (Additional file 2, flocks 16b and 16a, respectively). Among the TEX sheep of this flock, a noticeably higher proportion (65%) was serologically MV positive. As sheep of similar ages were compared and kept in the same, relatively small flock, different environmental factors can be neglected. Hence, the different genetics of the two breeds, e.g., different frequencies of protective or risk alleles, is the most likely reason for the observed difference in the proportion of serologically MV positive sheep. In fact, the frequency of sheep carrying the *TMEM154* genotype KK was 40% among the GBM and 0% among the TEX sheep of this flock.

In the breed subset MLS-x, the association of *TMEM154* E/K with serological MV status and median ELISA S/P value was not significant either but approached the significance threshold. A remarkably high frequency of MV seropositive sheep with the genotype KK was observed in this breed subset (47%), compared to 24% in TEX-x, 8% in EFM–LAC and 0% in GBM. This was also illustrated by a notably higher median ELISA S/P value in sheep with the KK genotype of this subset compared with the other breed subsets as well as with the median ELISA value of MV negative flocks (Figure 1). It has been shown in other studies that the KK genotype is not fully protective. In a previous study [19], about 26% of sheep carrying two K alleles were seropositive. Several factors may influence the proportion of seropositive sheep with this genotype in a MV positive flock, such as viral dose, route of infection, additional host risk factors (e.g., animal crowding, other breed-specific genetic factors), presence of a distinct and/or more than one SRLV strain(s) or subtype(s), and coinfection with another disease at the same time [19, 20, 34]. Sider et al. [34] reported different odds of infection for sheep carrying the same *TMEM154* E35K genotypes, but infected with different SRLV genotypes, postulating that not only the host but also the SRLV genotype affects the relative risk of infection in sheep. Hence, an (additional) infection with an SRLV strain or subtype that is different from the virus strain(s) or subtype(s) circulating in the other analyzed flocks could be the reason for the higher percentage of seropositive sheep with the KK genotype in the MLS-x flock compared to the other subsets. Until now, no information has been available on virus subtypes or strains circulating in the German sheep population. In a follow-up study, the samples collected in this study, together with additional samples from other German regions, should be used for genotyping and characterization of virus strains. On the one hand, this would make it possible to test for a virus subtype-dependent association of host gene variants with susceptibility or resistance to infection. On the other hand, this is a prerequisite to establish more precise tools for phenotyping the MV infection status of the German sheep population.

In all samples, sheep with one or two copies of the E allele of *TMEM154* had a relative risk of 2.26 to be serologically MV positive. This is comparable to another study [19] in which the relative risk of infection for sheep carrying one or two copies of the *TMEM154* risk haplotypes (haplotypes including E at position 35) varied from 1.27 to 5.30 between cohorts (populations with different

breed background, age at sampling and seroprevalence) with an overall relative risk of 2.85.

In this study, the deletion in the *CCR5* promotor did not show a consistent association with serological MV status or median ELISA S/P values. The main reason might be the low number of sheep carrying two copies of the *CCR5* promoter deletion (only 17 out of 521 genotyped sheep). For sheep carrying one or two copies or the *CCR5* promoter wild type, a significant difference in the relative risk to be serologically MV positive compared to sheep with the promoter deletion was only found in the breed subgroup MLS-x. Unexpectedly, in this subset, the relative risk was higher for sheep carrying two copies of the putative protective allele (deletion) than for sheep carrying one or two copies of the putative risk allele (wild type). Hence, the direction of the association was opposed to that observed in a US sheep flock originating from Idaho [18], but consistent with results in a US sheep flock from Iowa [20]. It should be reconsidered that in the study of White et al. [18], a different phenotype was used for association analysis with *CCR5* (proviral load). It is possible that different susceptibility phenotypes (here antibody response and provirus load) are controlled in part or completely by different host genes. At this point it should be mentioned that, up to our knowledge, the amino acid mutation at position 35 of *TMEM154* is the only genetic marker displaying association with both the serological status as well as the proviral level regarding MV infection in sheep [19, 20].

It is also possible that the deletion in the *CCR5* promoter is not directly influencing the analyzed phenotype(s), but is linked with an unknown causal genetic variant. In such a case, the allele which is linked with the causal protective or risk allele can vary between breeds and populations.

The present study was based on the collection of samples in commercial German sheep flocks which were tested for MV infection status by a common and approved serological method. The median ELISA S/P value of samples from 6 MV negative flocks was 9.69% with a non-conservative approximate 95% confidence interval (95% CI) of 1.78–52.30%. Hence, the recommended cut-off value for the used ELISA (≤ 110 negative, ≥ 120 positive) is approximately twofold the upper limit of the confidence interval of the negative flocks. Therefore, it is possible that some sheep which were classified as serologically MV negative or doubtful in fact were infected.

A further limitation of this study may be the well-known high genetic variability of the small ruminant lentivirus, which is promoted by high mutation and recombination rates during viral replication [35]. Therefore, the diagnostic power of a MV ELISA can be limited by the virus subtype(s) or strain(s) present in a flock. The MV ELISA used in the present study is one of three ELISAs which are approved for determination of SRLV infection in sheep and goats by the German licensing authority. In a comparison test including these three and also other ELISAs used in other countries, the proportion of samples from sheep with clinical MV symptoms detected as serologically positive varied between the different ELISAs (Günter Kotterba, German Friedrich-Loeffler-Institut (FLI), Federal Research Institute for Animal Health, Island of Riems, Germany, personal communication). This may be due to differences in the sensitivity of ELISAs for certain virus strains or genotypes. For instance, correct identification of SRLV strain E, which is present in Italy and highly divergent from other SRLV strains, strain-specific antigens had to be developed and tested [36]. In a study on SRLVs in Jordan, where no information on present SRLV genotypes was available, the maximum of positive samples was found using three different ELISA kits [37].

It can be concluded that the amino acid substitution at position 35 of *TMEM154* is a promising marker for breeding towards a lower MV susceptibility (in terms of a lower number of serologically positive sheep) in Germany, at least in flocks based on the Texel breed, while this remains questionable for the deletion in the *CCR5* promoter. The presence of the protective *TMEM154* variant in diverse sheep breeds all over the world indicates that it is an old mutation and carried in a most likely quite short haplotype. Therefore, the risk of associated negative effects on other selection traits will be low. However, this has to be monitored carefully during selection.

In consecutive analyses, the findings of this study should be verified and possibly enlarged by including a higher number of MV affected flocks of other breeds and employing additional phenotyping tools, e.g., other MV ELISA kits and methods for the measurement of provirus load. In particular, a follow-up study employing genome-wide association analysis in a higher number of samples from MV affected Merinoland sheep might be able to identify additional genetic variants with a breed-specific impact on MV susceptibility.

Additional files

Additional file 1. Primers used for genotyping and sequencing.

Additional file 2. Origin (German state), breed composition, percentage of serologically MV-positive samples and classification into breed subset for each sampled sheep flock.

Additional file 3. Numbers of serologically MV negative and positive sheep and description of MV ELISA S/P values (%) in all 17 MV affected flocks and within breed subsets.

Additional file 4. Box plots depicting MV ELISA S/P values in sheep carrying genotypes with (blue) and without (yellow) the putative *CCR5* promoter risk allele (wt), in all sheep and in breed subsets. *P* values are resulting from nonparametric analyses comparing median MV ELISA S/P values of groups. ns: not significant ($P > 0.05$). The black dotted line indicates the median ELISA S/P value of sheep from six serologically MV negative flocks (about 9%). Sheep with ELISA S/P values below 110% (green dashed horizontal line) were considered serologically MV negative. Sheep with ELISA S/P values over 120% (red solid line) were considered serologically MV positive. wt: wild type; del: deletion; TEX-x: purebred and crossbred German Texel sheep; GBM: purebred German Blackheaded Mutton sheep; MLS-x: purebred and crossbred Merinoland sheep; EFM-LAC: East Friesian Milk and Lacaune sheep and crosses of both breeds.

Additional file 5. Relative risk of infection in sheep with one or two copies of the *CCR5* promoter deletion.

Abbreviations

CCR5: chemokine (C-C motif) receptor type 5; EFM: East Friesian milk; ELISA: enzyme-linked immunosorbent assay; GBM: German Blackheaded Mutton sheep; LAC: Lacaune sheep; MLS: Merinoland sheep; MV: maedi-visna; OD: optical density; PCR: polymerase chain reaction; S/P: ratio between mean OD of sample and mean OD of positive control; SRLV: small ruminant lentiviruses; TEX: German Texel sheep; *TMEM154*: transmembrane protein 154.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

VM: contributed to genotyping of samples, carried out the statistical analyses and wrote significant parts of the manuscript. ME: contributed to establishment of genotyping methods, genotyping of samples and contributed to data analysis. GL: designed the experiment, contributed samples, contributed to establishment of genotyping methods, contributed to data analysis and wrote significant parts of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank Carina Crispens and Stephanie Steitz for their excellent technical assistance. We are grateful to sheep farmers and veterinarians for getting access to sheep samples and data.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its Additional files.

Ethics approval and consent to participate

Collection of blood samples from sheep was done by trained veterinarians (mostly from animal health services). The purpose of sampling was to test the serological MV status of MV suspected flocks in order to decide on subsequent veterinary measures and not for the purpose of this study. According to the German animal protection law, this approach obviated the need for an explicit ethics committee approval.

Funding

No funding has been received for the conduct of this study and preparation of this manuscript.

Publisher's Note

Received: 30 December 2017 Accepted: 3 April 2018
Published online: 19 April 2018

References

1. Peterhans E, Greenland T, Badiola J, Harkiss G, Bertoni G, Amorena B, Eliaszewicz M, Juste RA, Krassnig R, Lafont JP, Lenihan P, Petursson G, Pritchard G, Thorley J, Vitu C, Mornex JF, Pepin M (2004) Routes of transmission and consequences of small ruminant lentiviruses (SRLVs) infection and eradication schemes. Vet Res 35:257–274
2. Hüttner K, Seelmann M, Feldhusen F (2010) Prevalence and risk factors for Maedi-Visna in sheep farms in Mecklenburg-Western-Pomerania. Berl Munch Tierarztl Wochenschr 123:463–467
3. Leroux C, Cruz JC, Mornex JF (2010) SRLVs: a genetic continuum of lentiviral species in sheep and goats with cumulative evidence of cross species transmission. Curr HIV Res 8:94–100
4. Blacklaws BA (2012) Small ruminant lentiviruses: immunopathogenesis of visna-maedi and caprine arthritis and encephalitis virus. Comp Immunol Microbiol Infect Dis 35:259–269
5. Keen JE, Hungerford LL, Littledike ET, Wittum TE, Kwang J (1997) Effect of ewe ovine lentivirus infection on ewe and lamb productivity. Prev Vet Med 30:155–169
6. Arsenault J, Dubreuil P, Girard C, Simard C, Belanger D (2003) Maedi-visna impact on productivity in Quebec sheep flocks (Canada). Prev Vet Med 59:125–137
7. Reina R, Berriatua E, Lujan L, Juste R, Sanchez A, de Andres D, Amorena B (2009) Prevention strategies against small ruminant lentiviruses: an update. Vet J 182:31–37
8. Houwers DJ (1990) Economic importance, epidemiology and control. In: Petursson G, Hoff-Jorgensen R (eds) Maedi-visna and related diseases. Kluwer Academic Press, Boston, pp 83–117
9. Gjerset B, Rimstad E, Teige J, Soetaert K, Jonassen CM (2009) Impact of natural sheep-goat transmission on detection and control of small ruminant lentivirus group C infections. Vet Microbiol 135:231–238
10. Davies G, Genini S, Bishop SC, Giuffra E (2009) An assessment of opportunities to dissect host genetic variation in resistance to infectious diseases in livestock. Animal 3:415–436
11. Keen JE, Hungerford LL, Wittum TE, Kwang J, Littledike ET (1997) Risk factors for seroprevalence of ovine lentivirus in breeding ewe flocks in Nebraska, USA. Prev Vet Med 30:81–94
12. Schaller P, Vogt HR, Strasser M, Nettleton PF, Peterhans E, Zanoni R (2000) Seroprevalence of maedi-visna and border disease in Switzerland. Schweiz Arch Tierheilkd 142:145–153 (in German)
13. Straub OC (2004) Maedi-Visna virus infection in sheep. History and present knowledge. Comp Immunol Microbiol Infect Dis 27:1–5
14. Herrmann-Hoesing LM, White SN, Mousel MR, Lewis GS, Knowles DP (2008) Ovine progressive pneumonia provirus levels associate with breed and Ovar-DRB1. Immunogenetics 60:749–758
15. Stonos N, Wootton SK, Karrow N (2014) Immunogenetics of small ruminant lentiviral infections. Viruses 6:3311–3333
16. Larruskain A, Jugo BM (2013) Retroviral infections in sheep and goats: small ruminant lentiviruses and host interaction. Viruses 5:2043–2061
17. White SN, Knowles DP (2013) Expanding possibilities for intervention against small ruminant lentiviruses through genetic marker-assisted selective breeding. Viruses 5:1466–1499
18. White SN, Mousel MR, Reynolds JO, Lewis GS, Herrmann-Hoesing LM (2009) Common promoter deletion is associated with 3.9-fold differential transcription of ovine *CCR5* and reduced proviral level of ovine progressive pneumonia virus. Anim Genet 40:583–589
19. Heaton MP, Clawson ML, Chitko-Mckown CG, Leymaster KA, Smith TP, Harhay GP, White SN, Herrmann-Hoesing LM, Mousel MR, Lewis GS, Kalbfleisch TS, Keen JE, Laegreid WW (2012) Reduced lentivirus susceptibility in sheep with *TMEM154* mutations. PLoS Genet 8:e1002467
20. Alshanbari FA, Mousel MR, Reynolds JO, Herrmann-Hoesing LM, Highland MA, Lewis GS, White SN (2014) Mutations in *Ovis aries TMEM154* are

associated with lower small ruminant lentivirus proviral concentration in one sheep flock. Anim Genet 45:565–571
21. Leymaster KA, Chitko-McKown CG, Clawson ML, Harhay GP, Heaton MP (2013) Effects of *TMEM154* haplotypes 1 and 3 on susceptibility to ovine progressive pneumonia virus following natural exposure in sheep. J Anim Sci 91:5114–5121
22. Connor RI, Sheridan KE, Ceradini D, Choe S, Landau NR (1997) Change in coreceptor use correlates with disease progression in HIV-1–infected individuals. J Exp Med 185:621–628
23. Dawson TC, Beck MA, Kuziel WA, Henderson F, Maeda N (2000) Contrasting effects of CCR5 and CCR2 deficiency in the pulmonary inflammatory response to influenza A virus. Am J Pathol 156:1951–1959
24. Hull J, Rowlands K, Lockhart E, Moore C, Sharland M, Kwiatkowski D (2003) Variants of the chemokine receptor CCR5 are associated with severe bronchiolitis caused by respiratory syncytial virus. J Infect Dis 188:904–907
25. Kaslow RA, Dorak T, Tang JJ (2005) Influence of host genetic variation on susceptibility to HIV type 1 infection. J Infect Dis 191(Suppl 1):S68–S77
26. Thio CL, Astemborski J, Bashirova A, Mosbruger T, Greer S, Witt MD, Goedert JJ, Hilgartner M, Majeske A, O'Brien SJ, Thomas DL, Carrington M (2007) Genetic protection against hepatitis B virus conferred by *CCR5Delta32*: evidence that CCR5 contributes to viral persistence. J Virol 81:441–445
27. Lim JK, Louie CY, Glaser C, Jean C, Johnson B, Johnson H, McDermott DH, Murphy PM (2008) Genetic deficiency of chemokine receptor CCR5 is a strong risk factor for symptomatic West Nile virus infection: a meta-analysis of 4 cohorts in the US epidemic. J Infect Dis 197:262–265
28. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R, Goedert JJ, Buchbinder SP, Vittinghoff E, Gomperts E, Donfield S, Vlahov D, Kaslow R, Saah A, Rinaldo C, Detels R, O'Brien SJ (1996) Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the *CKR5* structural gene. Science 273:1856–1862
29. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, MacDonald ME, Stuhlmann H, Koup RA, Landau NR (1996) Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell 86:367–377
30. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, Saragosti S, Lapoumeroulie C, Cognaux J, Forceille C, Muyldermans G, Verhofstede C, Burtonboy G, Georges M, Imai T, Rana S, Yi Y, Smyth RJ, Collman RG, Doms RW, Vassart G, Parmentier M (1996) Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. Nature 382:722–725
31. Montgomery GW, Sise JA (1990) Extraction of DNA from sheep white blood-cells. N Z J Agric Res 33:437–441
32. Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52:591–611
33. Altman DG (1990) Practical statistics for medical research. Chapman and Hall, London
34. Sider LH, Heaton MP, Chitko-McKown CG, Harhay GP, Smith TP, Leymaster KA, Laegreid WW, Clawson ML (2013) Small ruminant lentivirus genetic subgroups associate with sheep *TMEM154* genotypes. Vet Res 44:64
35. Minardi da Cruz JC, Singh DK, Lamara A, Chebloune Y (2013) Small ruminant lentiviruses (SRLVs) break the species barrier to acquire new host range. Viruses 5:1867–1884
36. Grego E, Bertolotti L, Quasso A, Profiti M, Lacerenza D, Muz D, Rosati S (2007) Genetic characterization of small ruminant lentivirus in Italian mixed flocks: evidence for a novel genotype circulating in a local goat population. J Gen Virol 88:3423–3427
37. Tolari F, Al-Ramadneh W, Mazzei M, Carrozza ML, Forzan M, Bandecchi P, Grego E, Rosati S (2013) Small ruminant lentiviruses in Jordan: evaluation of sheep and goat serological response using recombinant and peptide antigens. Trop Anim Health Prod 45:1335–1340

Gene	primer sequences	product size	purpose
TMEM154	5'- CCACAGGAGAGGAGRACACA-3' (forward) 5'-GGGCACGTCTCCTGACAGTT**T**-3' (reverse, FAM-labeled, K allele) 5'-GGCACGTCTCCTGACAGTT**C**-3' (reverse, HEX-labeled, E allele)	40/41 bp	determination of nucleotide substitution NC_019474.2:g.4860407G>A, resulting in amino acid substitution E35K, using KASP technology (LGC, Hoddesdon, UK)
TMEM154	5'-GCTAGACACTGCCAAGCTTC-3' (forward) 5'-TGTCACTGAAACAAGTCATCACT-3' (reverse)	788 bp	amplification and sequencing for verification of genotyping results
CCR5	5'-CCCCATTGATAAGCCCTACA-3' (forward) 5'-CACCCAACTACCCAAATGGT-3' (reverse, FAM-labeled or unlabeled)	160/156 bp	determination of promotor region deletion NC_019476.2:g.52961717_52961714delAATG (minus strand); amplification and sequencing for verification of genotyping results

Bold letters: nucleotide positions leading to allele specific binding of primers.

Flock no.	state	Breed background of flocks	sampled sheep (n)[1]	serologically MV positive (%)	breed subset
1	SH	TEX-x	33	60.60	1
2	SH	TEX-x	16	68.75	1
3	SH	TEX-x	26	57.69	1
4	SH	TEX-x	41	73.17	1
5	SH	TEX-x	11	100.00	1
6	SH	TEX-x	30	76.67	1
7	SH	TEX-x	21	66.66	1
8	SH	TEX-x	22	77.27	1
9	SH	TEX-x	30	83.33	1
10	SH	TEX-x	11	63.64	1
11	SH	TEX	19	78.95	1
12	SH	TEX	5	80.00	1
13	SH	EFM	6	0.00	-
14	NRW	TEX	21	80.95	1
15	NRW	TEX-x	15	73.33	1
16.a	NRW	TEX	43	65.11	1
16.b	NRW	GBM	39	10.26	2
17	NRW	TEX, EFM	17	0.00	-
18	NRW	RHO	25	0.00	-
19	NRW	Mix of several breeds	17	0.00	-
20	NRW	SUF	32	0.00	-
21	MVP	RPL	18	0.00	-
22	HE	MLS-x	125	50.40	3
23	BW	EFM, LAC, EFM-LAC	22	36.36	4

[1] without sheep with suspicious ELISA S/P results (n = 11).

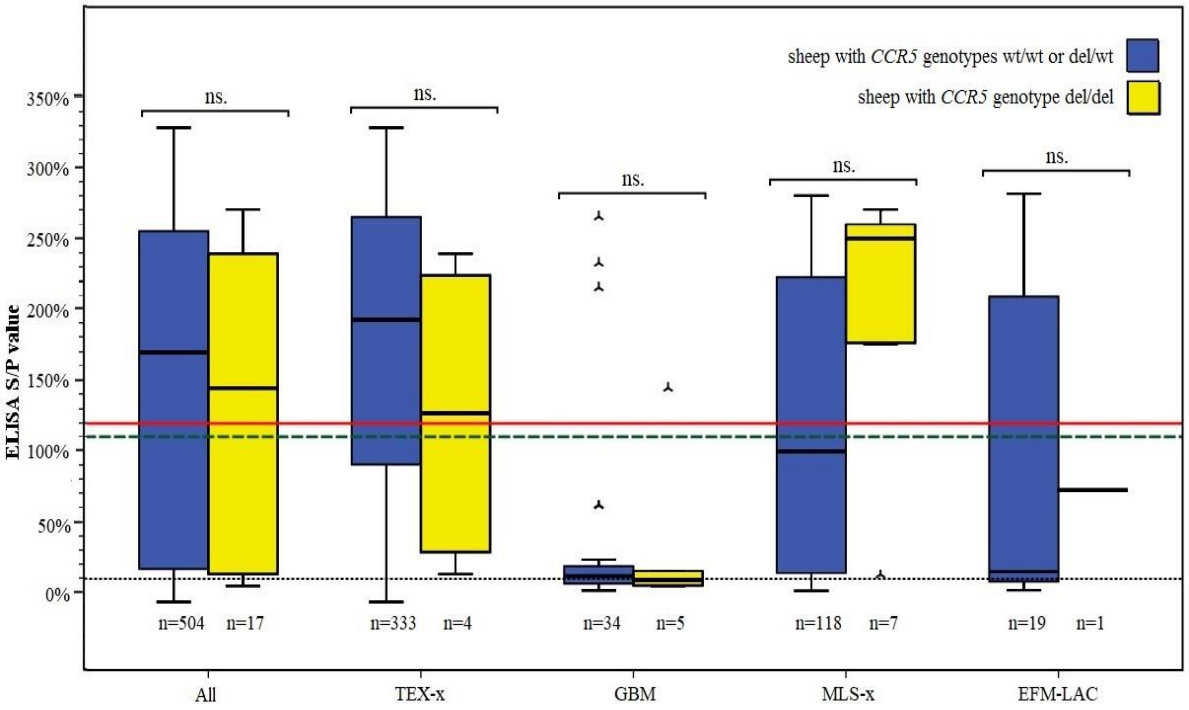
Abbreviation of German states: SH (Schleswig-Holstein), NRW (Nordrhein-Westfalen), MVP (Mecklenburg-Vorpommern), HE (Hessen), BW (Baden-Württemberg).

Abbreviation of breeds: -x (not all sheep were purebred), TEX (German Texel), EFM (East Friesian Milk sheep), GMB (German Blackheaded Mutton), RHO (Rhoen sheep), SUF (Suffolk), RPL (Rough-coated Pomeranian Landrace), MLS-x (Merinoland sheep), LAC (Lacaune).

Breed subsets: 1 = purebred and crossbred German Texel (TEX-x), 2 = purebred German Blackheaded Mutton (GBM), 3 = purebred and crossbred Merinoland sheep (MLS-x), 4 = East Friesian Milk and Lacaune sheep and crosses of both breeds (EFM-LAC).

Breed subset	MV status (n)	MV ELISA S/P value range min, max	median MV ELISA S/P values	quartile of MV ELISA S/P values Q1, Q3	*p* value (test of normal distribution)
All	negative (207)	-6.34, 109.65	12.95	5.56, 37.37	<0.001
	positive (323)	121.89, 328.63	245.89	182.27, 270.72	< 0.001
TEX-x	negative (96)	-6.34, 109.65	13.67	5.44, 42.07	< 0.001
	positive (248)	123.11, 328.63	246.70	183.07, 273.29	< 0.001
GMB	negative (35)	1.61, 61.92	9.14	5.19, 13.60	< 0.001
	positive (4)	144.65, 266.04	224.51	162.48, 257.80	0.681
MLS-x	negative (62)	1.45, 104.05	14.14	6.47, 59.03	< 0.001
	positive (63)	122.10, 280.70	233.17	176.20, 263.39	< 0.001
EFM-LAC	negative (14)	1.80, 85.84	10.25	4.98, 29.42	< 0.001
	positive (8)	121.89, 282.01	269.10	183.38, 280.92	0.002

TEX-x: purebred and crossbred German Texel sheep; GBM: purebred German Blackheaded Mutton sheep; MLS-x: purebred and crossbred Merinoland sheep; EFM-LAC: East Friesian Milk and Lacaune sheep and crosses of both breeds.


sheep with *CCR5* genotypes wt/wt or del/wt
sheep with *CCR5* genotype del/del
ELISA S/P value
350%
300%
250%
200%
150%
100%
50%
0%
ns.
ns.
ns.
ns.
ns.
n=504 n=17
n=333 n=4
n=34 n=5
n=118 n=7
n=19 n=1
All
TEX-x
GBM
MLS-x
EFM-LAC

Breed subset	parameters	*CCR5* genotypes wt/ wt, del/wt vs. del/del
All	relative risk	1.146
	95% CI	0.728 - 1.805
	p value	0.554
TEX-x	relative risk	1.435
	95% CI	0.537 - 3.833
	p value	0.471
MLS-x	relative risk	0.563
	95% CI	0.395 - 0.804
	p value	0.001

del: deletion; wt: wild type; CI: confidence interval; TEX-x: purebred and crossbred German Texel sheep; MLS-x: purebred and crossbred Merinoland sheep.

Study 2: Lentivirus susceptibility in Iranian and German sheep assessed by determination of *TMEM154* E35K

animals

MDPI

Article

Lentivirus Susceptibility in Iranian and German Sheep Assessed by Determination of *TMEM154* E35K

Vahid Molaee [1], Vahid Otarod [2], Darab Abdollahi [3] and Gesine Lühken [1,*]

[1] Department of Animal Breeding and Genetics, Justus Liebig University Giessen, Ludwigstraße 21, 35390 Gießen, Germany; Vahid.Molaee@agrar.uni-giessen.de
[2] Quarantine and Biosafety Directorate General, Iran Veterinary Organization (IVO), Vali Asr Avenue, Seyd Jamaledin Asad Abadi Street, 6349 Tehran, Iran; votarod@hotmail.com
[3] Bureau of Animal Health and Disease Management, Iran Veterinary Organization (IVO), Vali Asr Avenue, Seyd Jamaledin Asad Abadi Street, 6349 Tehran, Iran; d.abdollahi@ivo.ir
* Correspondence: Gesine.Luehken@agrar.uni-giessen.de

Received: 21 July 2019; Accepted: 11 September 2019; Published: 15 September 2019

Simple Summary: There are no data on the effect of the transmembrane protein 154 (TMEM154) E35K variants on susceptibility to small ruminant lentivirus (SRLV) infection in Iranian sheep breeds, and only limited data for German sheep flocks. This study aimed at investigating the association of *TMEM154* variants and SRLV infection status in Iranian and German sheep flocks and breeds. Three out of the four analyzed sheep flocks/breeds showed a significant association between *TMEM154* variants and SRLV prevalence. A complementary analysis was carried out based on regression analysis to test the relationship between frequency of the *TMEM154* E allele and SRLV prevalence in different flocks/breeds. Results showed that the *TMEM154* E allele frequencies could be useful for predicting genetic susceptibility to SRLV infection in a sheep flock or breed. Finally, the genetic susceptibility of different Iranian and German sheep breeds was compared based on the frequency of the *TMEM154* E allele.

Abstract: Small ruminant lentiviruses (SRLVs) cause maedi-visna disease in sheep and are prevalent in Iran and Germany. The association of the transmembrane protein 154 (TMEM154) variants with SRLV infection has been previously identified by a genome-wide association (GWAS) approach and subsequent analyses, and validated in some US, German, and Turkish sheep flocks. We aimed at evaluating these findings for the first time in Iranian, and in some more German sheep flocks/breeds. Also, we aimed at comparing the SRLV susceptibility in Iranian and German sheep based on the frequency of the *TMEM154* E35 allele. About 800 blood samples were collected from 21 Iranian and German sheep flocks/breeds for different purposes: (1) The association of *TMEM154* E35K with SRLV infection status was tested in four sheep breeds and found to be significant in Kermani, Merinoland, and Brown Hair. (2) The usefulness of the *TMEM154* E35 frequency for predicting SRLV susceptibility was evaluated by regression analysis, combining data from this study and some already published data. Results showed a significant association between E35 frequency and SRLV prevalence. (3) SRLV susceptibility was compared based on E35 frequency in Iranian and German sheep. Altogether, findings of this study provide valuable information on SRLV susceptibility, using *TMEM154* E35, in Iranian and German sheep.

Keywords: maedi-visna; small ruminant lentivirus; sheep; susceptibility; transmembrane protein; TMEM154

1. Introduction

Lentiviruses can cause maedi-visna (MV) disease in sheep and caprine arthritis encephalitis disease (CAE) in goats. Since these viruses infect both host species, they are generally called small ruminant lentiviruses (SRLV). Leading to progressive pneumonia, early culling, lower lamb weight, and lamb mortality of infected ewes, SRLVs cause high economic losses [1].

There is not any vaccine or cure for SRLV infection in sheep and goats. Eradication programs have already been established in many countries by culling infected animals and feeding kids with virus-free colostrum [2]. This strategy comes with costs because of cull replacements and the maintenance of SRLV-free animals [3]. A prerequisite for this eradication strategy is establishing a reliable tool for identifying infected animals [4]. Currently, there are several different laboratory techniques for the diagnosis of SRLV infection. These include serological methods, such as enzyme-linked immune assay (ELISA) and agar gel immunodiffusion (AGID), and polymerase chain reaction (PCR)-based techniques, such as nested and real-time PCR [2,5]. There is no gold standard in virology to diagnose SRLVs, and especially different ELISAs are still under debate. As reviewed by Herrmann-Hoesing [6], PCR tests compared to indirect ELISA tests have 92% to 95% positive concordance and 87% to 88% negative concordance. The genome of SRLVs contains two non-coding genes of long terminal repeats (LTR) and six coding genes encoding for structural proteins of *gag*, *pol*, and *env* and non-structural proteins of *tat*, *rev*, and *vif*. The *gag* gene includes matrix (MA), capsid (CA), and nucleocapsid (NC) domains, and the *env* gene contains surface (SU) and transmembrane (TM) proteins. A part of the *gag* gene (CA) and a part of the *env* gene (TM) are conserved among SRLVs, which make them suitable for designing immunodominant epitopes for serological purposes [7] and primers for both conventional and quantitative PCR [8]. Therefore, the major segments of *gag* (CA) and *env* (TM) in SRLVs could be targeted for screening SRLVs in sheep and goat flocks.

Another preventive intervention is to breed sheep for lower SRLV infection susceptibility [3]. There is much evidence that the susceptibility to SRLV infection varies among different sheep breeds [9,10]. Recently, various sheep genes, such as *MHC, CCR5, TLRs,* transmembrane protein 154 (*TMEM154*), *APOBEC3,* and *ZNF389*, were studied and associated with SRLV infection [3,11]. The association of *TMEM154* variants with SRLV infection has been identified by a genome-wide association (GWAS) approach and subsequent analysis, and validated in some North American sheep populations as well as recently in German and Turkish sheep flocks [12–15]. According to this, the susceptibility of sheep breeds to SRLV infection was assessed based on the frequency of *TMEM154* haplotypes [12,13,15] and alleles [14]. Twelve haplotypes were identified based on sequence variations in the *TMEM154* coding region. Haplotypes 2 and 3 (ancestral) carry a nucleotide (g) coding for glutamic acid (E) at position 35, whereas this nucleotide is substituted by a nucleotide (a) in haplotype 1 coding for lysine (K) [12]. In a study by Heaton et al. [16], using 50K single nucleotide polymorphism (SNP) chip data, the susceptibility of 74 globally distributed sheep breeds to SRLV infection was predicted based on the frequency of allele "c" which is in linkage disequilibrium (LD) with allele E at position 35 of *TMEM154* [16]. However, at present, the susceptibility status of certain sheep breeds cannot be predicted only based on the knowledge of the frequencies of *TMEM154* variants (alleles or haplotypes) or linked markers. In a German study, among the sheep breeds tested for association of *TMEM154* E/K with SRLV susceptibility, the results for Merinoland sheep and crossbreds of this breed indicated a possible deviation from the findings in other breeds, as a relatively high number of animals with the genotype KK were found to be serologically SRLV positive [14].

Iran is one of the main producers of small ruminants in the world. According to the large-scale surveys of the Statistical Center of Iran (SCI) in winter 2011, the population of small ruminants of Iran was composed of 50.2 million heads of sheep and 22.1 million heads of goats. The sheep population of Iran comprises 27 different sheep breeds, which were classified based on physical appearances such as size, shape, color, horn status, and tail type [17]. In Iran, SRLV infection was first diagnosed in the southwest of Iran (Khuzestan Province) using histopathological methods [18]. Later, many reports

restricted to local areas of Iran, either using serological methods or PCR techniques, have shown a variable level (2.2% to 34.5%) of SRLV prevalence in sheep [19–22].

A prerequisite for the practical usage of a genetic marker is to validate the observed association with the trait of interest in other populations and countries. In the US studies [12,13] as well as a Turkish study [15], the association between *TMEM154* variants and SRLV infection status was evaluated based on frequencies of *TMEM154* haplotypes. However, we decided to select the single E allele at position 35 as a surrogate for the two most common highly-susceptible *TMEM154* haplotypes (i.e., "2" and "3"), and K35 as a surrogate for the most common *TMEM154* haplotype with reduced susceptibility (i.e., "1"). Together, these three haplotypes are present in about 98% of US sheep [12] and 86% of Turkish sheep [15]. Investigating the association of *TMEM154* E35K allele/genotype frequencies with SRLV susceptibility in Iranian sheep breeds was the main goal of the present study. A second aim was to validate previous observations in the German sheep population by including further sheep breeds and an additional flock of the Merinoland sheep breed from another region. Finally, we aimed at comparing SRLV susceptibility in different sheep breeds from Iran and Germany based on E or c allele frequencies.

2. Materials and Methods

2.1. Animals and Blood Collection

For this study, three sets of samples from different sheep breeds of Iran (set 1) and Germany (sets 2 and 3) were collected.

The first set of blood samples (set 1, $n = 365$), originating from 30 flocks located in six provinces of Iran, were collected between 2015 and 2016. All sampled sheep were between four and seven years old and purebreds from the following sheep breeds: Makouee, Qezel, Moghani, Bakhtiari, Kaboudeh, Kermani, Balouchi, and Karakul (details on numbers of flocks, sheep per breed, and province are given in Table 1). There were no clinical signs of SRLV infection at the time of sampling and flocks were without any history of SRLV infection. No animal transfer happened between the sampled flocks. Therefore, all flocks were epidemiologically unrelated to each other. No other species than sheep were kept in each flock.

The second set of samples (set 2, $n = 127$) was collected from four SRLV positive sheep flocks located in four provinces of Germany. Set 2 included samples from breeds Texel, Merinoland, Brown Hair, and Zackel (Table 1).

The third set of samples (set 3) included 302 samples from 11 pure sheep breeds of Germany which were collected during the years 2003–2018 for routine breeding requirements (e.g., parentage testing, prion protein genotyping) (Appendix A Table A1). As this set of samples was used to estimate *TMEM154* E/K allele frequencies within breeds, not more than three sheep originated from the same flock in order to minimize possible relationships between animals. There were no epidemiological data about SRLV infection status available for this set of samples.

All blood samples (10 mL each) were drawn by puncture of the jugular vein and collected into ethylene-diamine-tetra-acetic acid (EDTA) tubes for DNA isolation.

2.2. Ethics Statements

Blood samples were collected by veterinarians. Independent from this study, the original purpose of sampling was to test the MV status of sheep flocks (sample sets 1 and 2) or the scrapie resistance status of sheep (sample set 3) in order to decide on subsequent veterinary or breeding measures. According to the German Animal Welfare Act (released on 18th of May 2006, last changes on 17th of December 2018), regulations for animal protection, this origin of samples obviates the need for an explicit ethics committee approval. The Iranian Veterinary Organization (IVO) was directly responsible for gathering Iranian sheep samples as a part of routine examinations for care and control of animal diseases (permission released on 20th of December 2014, tracking number 93/22/70521). The ethics

responsibility for the collection of Iranian sheep samples was in accordance with the legal requirements of that national authority.

Table 1. Small ruminant lentivirus (SRLV) infection status in sampled sheep flocks of Iran and Germany.

Flock No.	Country	Province	Breed	Sampled Sheep (*n*)	Positive [1]
1	Iran	Western Azarbaijan	Makouee	9	1 (11%)
2	Iran	Western Azarbaijan	Makouee	7	0 (0%)
3	Iran	Western Azarbaijan	Makouee	5	0 (0%)
4	Iran	Western Azarbaijan	Makouee	7	1 (14%)
5	Iran	Western Azarbaijan	Makouee	9	1 (11%)
6	Iran	Western Azarbaijan	Makouee	10	0 (0%)
7	Iran	Western Azarbaijan	Makouee	10	1 (10%)
8	Iran	Western Azarbaijan	Makouee	4	0 (0%)
9	Iran	Western Azarbaijan	Makouee	13	0 (0%)
10	Iran	Western Azarbaijan	Qezel	6	0 (0%)
11	Iran	Western Azarbaijan	Qezel	20	1 (5%)
12	Iran	Ardebil	Moghani	21	0 (0%)
13	Iran	Chaharmahal-Va-Bakhtiari	Bakhtiari	18	16 (89%)
14	Iran	Chaharmahal-Va-Bakhtiari	Bakhtiari	15	10 (67%)
15	Iran	Chaharmahal-Va-Bakhtiari	Bakhtiari	17	15 (88%)
16	Iran	Fars	Kaboudeh	10	0 (0%)
17	Iran	Fars	Kaboudeh	21	0 (0%)
18	Iran	Fars	Kaboudeh	12	0 (0%)
19	Iran	Fars	Kaboudeh	8	0 (0%)
20	Iran	Kerman	Kermani	14	3 (21%)
21	Iran	Kerman	Kermani	22	12 (55%)
22	Iran	Kerman	Kermani	10	0 (0%)
23	Iran	Khorasan Razavi	Balouchi	10	0 (0%)
24	Iran	Khorasan Razavi	Balouchi	2	0 (0%)
25	Iran	Khorasan Razavi	Balouchi	10	0 (0%)
26	Iran	Khorasan Razavi	Balouchi	23	0 (0%)
27	Iran	Khorasan Razavi	Karakul	12	0 (0%)
28	Iran	Khorasan Razavi	Karakul	11	0 (0%)
29	Iran	Khorasan Razavi	Karakul	9	0 (0%)
30	Iran	Khorasan Razavi	Karakul	20	0 (0%)
31	Germany	Schleswig-Holstein	Texel	39	39 (100%)
32	Germany	Bayern	Merinoland	49	31 (63%)
33	Germany	Nordrhein-Westfalen	Brown Hair	25	12 (48%)
34	Germany	Baden-Württemberg	Zackel	14	8 (57%)

[1] SRLV positive samples in Iranian and German flocks were detected by PCR test (*env*-SU/TM) or serology (ELISA technique), respectively.

2.3. DNA Extraction

Genomic DNA was obtained from peripheral blood leukocytes by using a DNA extraction kit (MBST, Tehran, Iran) for Iranian sheep samples in accordance to the manufacturer's protocol and for German sheep samples by a modified salting-out method [23].

2.4. Quantity and Quality Control of Extracted DNA

The concentration (quantity) and OD260/280 ratio (quality) of the extracted DNA were measured with the Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Based on these results, the concentration of all DNA samples was standardized to the same value. Additionally, the integrity of the extracted DNA was determined by amplifying *SRY* and *AMLX* genes usually used for sex determination [24] (details on primers are given in Appendix A Table A2). The clear detection of PCR products displaying the male or female sex was a prerequisite for using a DNA sample in further analyses.

2.5. PCR Amplification for Detection of SRLVs in Iranian Sheep Samples (Set 1)

Due to animal health requirements, serum samples of Iranian sheep breeds could not be shipped to Germany, where all laboratory analyses, except DNA extraction from Iranian samples, were carried out. Hence, SRLV infection status of Iranian sheep was determined by PCR test. For this purpose, a semi-nested PCR was designed corresponding to the *env* SU/TM fragment, by selecting the conserved regions with the highest sequence homology between different SRLV genotypes (A–C and E). Detailed information on primers is given in Table A2. To amplify a first PCR product with a size of ~1 kb, the forward (*env*-SU-F1) and the reverse primer (*env*-TM-R1) were used. For the second PCR product (default PCR), an internal reverse primer (*env*-TM-R2) was used with the forward primer of the first PCR (*env*-SU-F1), providing a ~0.4 kb fragment (Appendix A Figure A1, Table A2). The first PCR was carried out in a total volume of 15.0 µL, containing 150 ng template DNA, 1× Go Taq Flexi PCR buffer (Promega, Mannheim, Germany), 0.2 mM dNTPs, 1.5 mM $MgCl_2$, 9 pmol of each primer (*env*-SU-F1 and *env*-TM-R1), and 0.6 units Go Taq-polymerase (Promega, Mannheim, Germany). The conditions for PCR were 94 °C for 1.5 min, followed by 35 cycles at 94 °C for 45 s, 55 °C for 1 min, 72 °C for 1 min, and a final extension at 72 °C for 5 min. The second PCR (default PCR) was carried out in a total volume of 50.0 µL, containing 3.0 µL of the first PCR product, 1× Go Taq Flexi PCR buffer (Promega, Mannheim, Germany), 0.2 mM dNTPs, 1.5 mM $MgCl_2$, 30 pmol of each primer (*env*-SU-F1 and *env*-TM-R2), and 2 units Go Taq-polymerase. The conditions for the second PCR (default PCR) were 94 °C for 1.5 min, followed by 35 cycles at 94 °C for 20 s, 55 °C for 30 s, 72 °C for 45 s, and a final extension at 72 °C for 5 min.

In order to validate the PCR results (default PCR), the second PCR was also carried out with an alternative primer pair, now using a new forward primer (*env*-TM-F2) together with the *env*-TM-R1 primer, resulting in a ~0.8 kb fragment (Figure A1, Table A2). The second PCR (alternative) was carried out in a total volume of 25.0 µL, containing 1.5 µL of the first PCR product, 1× Go Taq Flexi PCR buffer (Promega, Mannheim, Germany), 0.2 mM dNTPs, 1.5 mM $MgCl_2$, 12.5 pmol of each primer (*env*-TM-F2 and *env*-TM-R1), and 1 unit Go Taq-polymerase. The condition for the second PCR (alternative) was 94 °C for 1.5 min, followed by 40 cycles at 94 °C for 30 s, 58 °C for 45 s, 72 °C for 50 s, and a final extension at 72 °C for 10 min.

For each amplification, a positive control (DNA from German sheep serologically diagnosed to be SRLV positive) [14] and a negative control (no DNA template) were run in parallel with all samples. All PCR products were analyzed by 1.7% agarose gel electrophoresis and visualized with ethidium bromide staining.

2.6. Serological Test for Detection of SRLVs in German Sheep Samples (Set 2)

A serological test with ELISA (IDEXX CAEV-MVV Total Ab ELISA, IDEXX GmbH, Ludwigsburg, Germany) as described previously [14] was carried out to determine the SRLV infection status in German sheep flocks (set 2). The used ELISA is an indirect ELISA designed to detect an immunogenic peptide of the transmembrane protein (*env* gene, TM) and the recombinant P28 protein which enters into the composition of the viral capsid (*gag* gene) (POURQUIER ELISA Maedi-Visna/CAEV Serum Verification, version P00302/07).

2.7. TMEM154 Genotyping

A total of 794 sheep samples from sets 1–3 were genotyped for *TMEM154* E35K by allele-specific PCR method. Details of the established method for genotyping of *TMEM154* E35K were described elsewhere [14].

2.8. Data Analysis

The SPSS program (version 25.0) for Windows (IBM SPSS Statistics, IBM Corp, Armonk, NY, USA) was used for statistical analyses. Differences in the distribution of genotype frequencies between sets

of SRLV negative and positive samples were tested by chi-square or Fisher's exact test (when expected values were < 5). The relative risk analyses were conducted based on the method of Altman [25].

Association analyses using chi-square/Fisher's exact test were done only for those sheep flocks/breeds that complied with two conditions: (1) They had an SRLV prevalence between 10% and 90%, and (2) They included carriers of protective (KK) as well as of non-protective (EK or EE) genotypes based on a protocol by Clarke et al. [26].

For regression analysis, the average allele frequencies of the *TMEM154 E* allele in SRLV affected flocks of the same breed (or dominated by the same breed) were plotted against the average SRLV infection in the flocks of these breeds/breed mixes. This meta-analysis includes data generated in the present study (four sheep breeds from Iran (set 1, only positive flocks) and four sheep breeds from Germany (set 2), a previous German study (four sheep breeds/breed mixes) [14] and a previous US study (eight sheep breeds) [12].

For comparing SRLV susceptibility between different sheep breeds of Iran and Germany, in addition to E allele frequencies, c allele frequencies of SNP no. 5388531 (OVAR 17) located in the intronic region of *TMEM154* were extracted from 50K SNP chip data (International Sheep Genomics Consortium, ISGC, https://www.sheephapmap.org). Information on *TMEM154* E35 allele frequencies (combined frequencies of haplotypes 2 and 3) in three US sheep [12] were also used for comparison purposes.

3. Results

3.1. SRLV Infection Status of Iranian and German Sheep Flocks (Sets 1 and 2)

A proportion of 61 out of 365 (16.7%) Iranian sheep samples were determined to be SRLV positive by PCR test. It should be noted that changing primers and arrangement of the second PCR (alternative instead of default PCR) did not result in a different PCR status (positive or negative) of samples. In 10 out of 30 sampled Iranian sheep flocks (33%), at least one SRLV positive sheep was identified by PCR (Table 1). The SRLV prevalence ranged from 5% (flock no. 11) to 89% (flock no. 13) in single flocks. Samples with SRLV positive status were identified in the four sheep breeds Bakhtiari (all three sampled flocks), Kermani (two out of three flocks), Makouee (four out of nine flocks), and Qezel (one out of two flocks).

In the German sample set (set 2), 90 out of 127 (71%) samples were tested SRLV positive by ELISA diagnosis. The SRLV prevalence in the German sheep breeds ranged from 48% (Brown Hair, flock no. 33) to 100% (Texel, flock no. 31). Details on SRLV infection status within Iranian and German flocks are given in Table 1.

3.2. TMEM154 Genotyping and Association Analyses (Chi-Square and Fisher's Exact Test)

All Iranian and German sheep samples were successfully genotyped for the amino acid substitution at position 35 of *TMEM154* (sets 1–3).

Association analyses were limited to two sheep breeds from Iran (Kermani and Makouee) and two sheep breeds from Germany (Merinoland sheep and Brown Hair) (Table 2). A significant association ($p < 0.05$) between *TMEM154* genotypes at position 35 and SRLV status (positive vs. negative) was found in three of these analyzed sheep breeds, but not in Makouee from Iran.

In the four analyzed flocks/breeds, the relative risk to be SRLV positive for sheep carrying one or two E alleles compared to those without E allele ranged from 0.48 to 20 (Table 2). In the Kermani sheep breed, the risk to be SRLV positive with E allele was almost half (0.48) compared to sheep with KK. In the German Brown Hair sheep breed, a noticeable high relative risk (20.0) was found for sheep with one or two copies of risk allele E compared to sheep without risk allele (KK). In this sheep breed, all sheep with the protective genotype (KK) were SRLV negative.

TMEM154 genotyping results of sheep flocks/breeds excluded from association analyses for different reasons (less than 10% or more than 90% SRLV positive sheep and/or non-adequate balance of genotype frequencies) are shown in Appendix A Table A3.

Table 2. *TMEM154* E35K genotype frequencies in four SRLV positive sheep breeds of Iran (IRN) and Germany (GER) with results from chi-square/Fisher's exact test and relative risk (RR) analyses.

Breed Subset	MV Status	*TMEM154* Genotype Frequencies			*p* Value	RR	95% CI	*p* Value
(*n* Sheep)	(*n* Sheep)	KK	EK	EE				
Makouee, IRN (35)	neg (31) pos (4)	0.71 (22) 0.50 (2)	0.29 (9) 0.25 (1)	0.00 (0) 0.25 (1)	0.118	2.18	0.35 to 13.53	0.402
Kermani, IRN (36)	neg (21) pos (15)	0.09 (2) 0.33 (5)	0.24 (5) 0.00 (0)	0.67 (14) 0.67 (10)	0.047	0.48	0.24 to 0.96	0.038
Merinoland, GER (49)	neg (18) pos (31)	1.00 (18) 0.74 (23)	0.00 (0) 0.26 (8)	0.00 (0) 0.00 (0)	0.0197	1.782	1.359 to 2.337	<0001
Brown Hair, GER (25)	neg (13) pos (12)	0.85 (11) 0.00 (0)	0.15 (2) 0.83 (10)	0.00 (0) 0.17 (2)	0.00002	20	1.313 to 304.494	0.0311

3.3. Regression Analysis Based on a Combination of Previously Published Data and Those of the Present Study

A regression analysis was conducted with all available data from different sheep flocks/breeds of Iran, Germany, and data from North America. The regression line indicates a significant relationship ($p < 0.001$) between changes in the frequency of the allele E and changes in SRLV prevalence (Figure 1). The R-squared measure and standard error of estimate (SSE) were 0.465 and 19.70%, respectively.

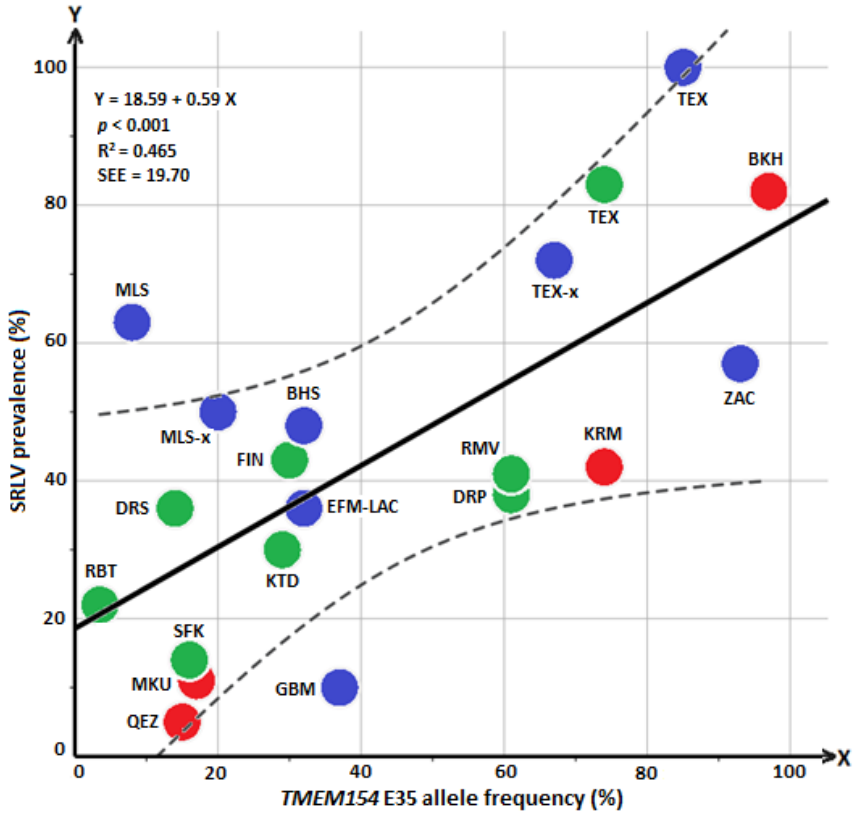


Figure 1. Relationship between *TMEM154* E allele frequency (%) and SRLV prevalence (%) in 20 different breeds/flocks from Iran, Germany, and the USA is shown by regression line (black). The dashed lines indicate 99% confidence interval. Sheep breeds/flocks of Iran, Germany, and the USA were labeled with red, blue and green solid colors, respectively. Breed names were abbreviated for Iran (Bakhtiari, BKH; Makouee, MKU; Kermani, KRM; Qezel, QEZ), Germany (present study: German Texel, TEX; Merinoland sheep, MLS; Brown Hair sheep, BHS; Zackel, ZAC; published by Molaee et al. [14]: purebred and crossbred German Texel sheep, TEX-x; German Blackheaded Mutton, GBM; purebred and crossbred Merinoland, MLS-x; East Friesian milk and Lacaune, EFM–LAC) and USA (published by Heaton et al. [12]: Dorset, DRS; Dorper, DRP; Finnsheep, FIN; Kathahdin, KTD; Rambouillet, RBT; Romanov, RMV; Suffolk, SFK; Texel, TEX). SEE, standard error of estimate.

3.4. SRLV Susceptibility in Sheep Breeds of Iran and Germany Based on the Frequency of TMEM154 Alleles E and c

Frequency of the E allele at position 35 of *TMEM154* found in different sheep breeds of Iran (n = 8) and Germany (n = 13) (sets 1–3), as well as the frequency of the intronic c allele (n = 7), are shown in Figure 2. For some sheep breeds, frequency information on both alleles (E and c) was available (n = 4). Information on E allele frequency (haplotypes 2 and 3) of three US sheep breeds with a known SRLV susceptibility status [12] was also included in the diagram as a criterion for comparing SRLV susceptibility of different sheep breeds. E/c allele frequencies ranged from 2% (German Grey Heath, Germany) to 100% (Kaboudeh, Iran) with a median of 56% and interquartile range from 15% (comparable with E allele frequencies of the US Suffolk) to 88% (comparable with E allele frequencies of the US Texel). In most of the analyzed Iranian sheep breeds, the observed E/c allele frequencies were relatively high. In five of them, the E/c allele frequencies were higher than or equal to the third quartile value (≥88%), in three of them, the E/c allele frequencies were lower than the first quartile value. In the breed Qezel, E and c allele frequencies were not congruent. In the German data set, E/c allele frequencies mostly ranged between quartile 1 and quartile 3 values (15–88%). In five out of the 15 analyzed breeds, E/c allele frequencies were lower than first quartile value (≤15%), whereas, in two breeds, E allele frequencies were higher than the third quartile value.

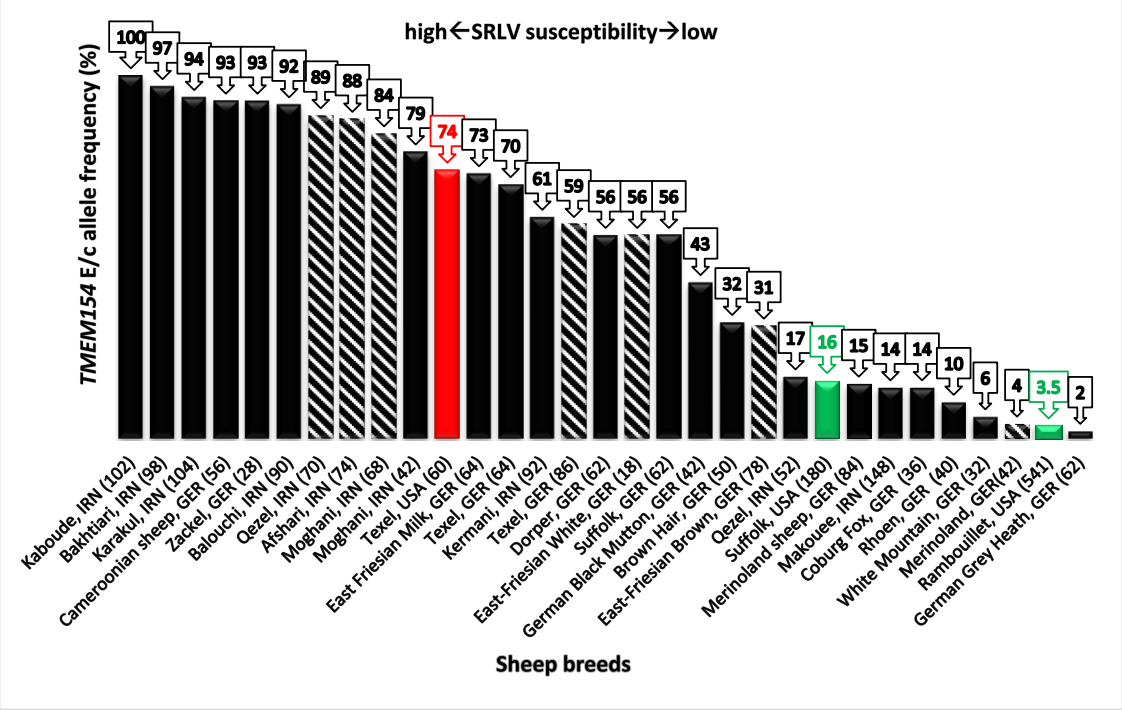


Figure 2. The degree of susceptibility to SRLV infection in nine sheep breeds from Iran (IRN) and fifteen sheep breeds from Germany (GER) is shown based on the frequency of the *TMEM154* E allele (solid bars, present study) or the c allele of SNP OVAR17-5388531 (shaded bars). The c allele is in linkage disequilibrium with the E allele. The c allele frequencies derived from 50K SNP chip data (International Sheep Genomics Consortium, ISGC). Colored solid bars indicate certain US sheep breeds with known SRLV susceptibility based on epidemiological data [12] (red = high susceptibility, green = low susceptibility).

4. Discussion

This study is the first to assess the usefulness of an amino acid substitution (E/K) in *TMEM154* for selection against SRLV susceptibility in sheep breeds of Iran. Moreover, it enlarges existing data for the German sheep population. SRLV positive samples were found in flocks of eight different sheep

breeds, and in four of them, the association of *TMEM154* E35K with SRLV status of sheep was tested applying the chi-square/Fisher's exact test. A significant association was found in flocks of the breeds Kermani, Merinoland and Brown Hair. Data from flocks of the other four breeds were not suitable for this method of association testing. To provide adequate power in testing for statistical significance it is necessary to find flocks with a "moderate" level of infection (e.g., not all or most sheep should be positive or negative) and a balanced ratio of genetically susceptible and resistant sheep. This is why in a study by Leymaster et al. [13] ewes which were infected and had *TMEM154* diplotype "1,3" were mated with "1,3" rams so that lambs with *TMEM154* diplotypes "1,1", "1,3" and "3,3" were produced in a 1:2:1 ratio in an environment with maximum pathogen exposure. The outcome was nearly 100% infected lambs with either one or two copies of haplotype "3" (EK/EE) and only 10% infected with haplotype "1,1" (KK) after five years. Of course, such optimal conditions are hard to find in field studies. For example, in the present study association testing for the Qezel flock with only 5% infected sheep was prevented because the SRLV prevalence was too low. We also had to withhold the Bakhtiari, German Texel and Zackel flocks/breeds from conventional association analyses because of the absence of the KK genotype. However, in these three breeds, a high SRLV prevalence came along with a high E allele frequency. Accordingly, a low E frequency was related to a low SRLV prevalence in Qezel. This positive correlation between E allele frequency and SRLV prevalence could be demonstrated by regression analysis, including data of sheep flocks/breeds from Iran, Germany, and the USA (Figure 1). Regarding the regression line, an SRLV prevalence of about 30% is expected in infected flocks with an E frequency of 20%, while an SRLV prevalence of about 60% is expected in sheep flocks with an E frequency of 70% (Figure 1). Regression analysis was used for two reasons: (1) independently from any prerequisite for carrying out a conventional association analysis, the regression analysis was helpful to visualize the relationship between "SRLV prevalence" and "E allele frequency" in all eight positive breeds of this study; (2) a general agreement between frequency of risk allele/genotypes of breeds and SRLV infection status was already observed by Heaton et al. [12]. Thus, we aimed to examine these relationships with a combined data set from sheep flocks of different countries (Iran, Germany, and the US), using regression analysis. Regression analysis showed a moderate R-square ($R^2 > 0.6$, strong accuracies; $0.4 < R^2 < 0.6$, moderate accuracies). For instance, an SRLV prevalence of about 62% would be predicted for flocks/breeds with an E allele frequency of about 75% (Figure 1; "KRM" and "TEX"), with a range from 42 to 82% (SEE ± 20%). With regard to prediction accuracy, a moderate value is common for most of the traits which are controlled by multiple genes, including susceptibility to disease [3]. Other factors such as effect of the different SRLV subgroups [27], other possible gene regions [3,11], and management also may prevent a precise prediction of the susceptibility to SRLV infection based on the *TMEM154* E allele frequency.

The relative risk levels estimated for sheep carrying one or two E alleles in the four analyzed flocks/breeds in this study (Kermani, RR = 0.48 to Makouee, RR = 2.18) were, with one exception (Brown Hair, RR = 20), comparable with previous studies in German [14] and US sheep flocks [12]. White and Knowles [3] explained that a ~2.5-fold relative risk for *TMEM154* haplotypes 2 and 3 was quite large for a single gene. The estimated relative risk for Merinoland sheep carrying one or two copies of the E allele was in the same range (95% Cl = 1.36 to 2.34) as previously estimated for this breed with another set of samples from another German region [14]. For the analyzed flock of Brown Hair sheep, an outstanding high risk (RR = 20) for sheep carrying one or two copies of the E allele was obtained with a high confidence interval because of zero denominators (absence of positive sheep with KK genotype).

In this study, SRLV susceptibility of sheep breeds of Iran and Germany was compared based on frequencies of the *TMEM154* E and/or c allele (the latter located in an intronic region of *TMEM154*). For the breeds Moghani, Texel, Merinoland, and Qezel, frequency information was available for both the E and the c allele. The frequencies of E and c alleles matched for all these breeds but not for Qezel (17 vs. 89%). Heaton and colleagues [12,16] showed a strong LD (p-value = 3.19×10^{-9}, $r^2 = 0.98$) between the causative allele (E) and the c allele in an intron of *TMEM154* in US sheep. The question is,

in the light of the possible age of the *TMEM154* variants, which most likely occurred before modern breed separation, if the used SNP marker from the ovine SNP array is a good predictor for the close coding variant. In this study, with the example of the Iranian Qezel breed, we were able to show that data on the frequency of the c allele, which are available for a high number of international sheep breeds, can result in a wrong assessment of the SRLV susceptibility of a breed. Hence, the knowledge of the *TMEM154* E35 frequency is a crucial factor for the assessment of SRLV susceptibility of a single breed. However, another explanation may be related to the relatively low sample size for this breed in the present study and also for the generation of the ovine SNP chip data (26 and 35 sheep, respectively), maybe leading to an inaccurate estimation of the E and c allele frequencies for this breed. Although a low E allele frequency is consistent with the low SRLV prevalence (5%) in the Qezel flock, further sampling from additional flocks is necessary to provide a more accurate estimation of the SRLV susceptibility of this breed.

In Germany, the three sheep breeds Texel, East Friesian Milk, and Cameroon are historically known to be highly susceptible to SRLV infection. In contrast, the three sheep breeds German Grey Heath, Merinoland, and Suffolk are commonly not affected by this disease. The *TMEM154* E allele frequencies of these breeds are consistent with these observations. Additionally, in the Merinoland sheep breed, the relatively low E allele frequency seems to reflect the "general" low susceptibility of this breed [28]. For this reason, it was quite surprising to repeatedly find a high number of KK sheep (55% in this study; 44% in a previous study [14]) to be serologically SRLV positive in flocks of the Merinoland sheep breed. A follow-up GWAS study could be useful for detecting potential additional variants/markers related to SRLV infection status in this breed.

The SRLV infection status observed in Iranian sheep flocks (Table 1) was consistent with earlier studies. Nevertheless, we did not find any positive samples in the province of Khorasan Razavi that belonged to the northeast of Iran (Table 1, flock nos. 23–30), whereas serological studies have shown about 20% [22] and 35% [20] seropositivity in sheep of that region. It could be possible that our sampling missed infected flocks in this province. Furthermore, unknown mutations in the provirus sequences could have created primer mismatches and consequently prevented PCR amplification. However, this will only be an explanation if other SRLV strains are present in this region. There has been no report of any SRLV infection in the Fars Province, and we also did not find positive samples in this area. In this study, screening for the presence of SRLV infection was done for the first time from the provinces Western Azarbaijan and Ardebil. SRLV positive samples were found in the province of Western Azarbaijan, which is at the border of Turkey. In a follow-up study, SRLV sequences from Iran should be compared with those already observed in Turkey [29] in order to gain a deeper insight into the variability of SRLVs of the entire region. This could also help to develop more precise tools for diagnosis and prevention of SRLV infection in sheep.

4.1. Limitation of the Study

In other studies, it has been shown that rare E35-containing haplotypes also can confer recessively a lower susceptibility to SRLV (i.e., haplotype "4" together with haplotype "16" according to the study of Yaman et al. [15]). Therefore, the findings of the present study should not only be verified with more samples from additional flocks and breeds, but also with *TMEM154* haplotype data.

5. Conclusions

In this study, a significant association was found between *TMEM154* E35K variants and SRLV infection status in three out of four analyzed sheep breeds. Although our results showed a significant association in the flock of Merinoland sheep, a high number of KK sheep (carrier of the protective genotype) were SRLV positive. A follow-up GWAS study could be useful for detecting potential additional variants/markers related to SRLV infection status in Merinoland sheep.

An additional point of this study was testing the usefulness of the frequency of the *TMEM154* E allele for prediction of SRLV susceptibility in sheep flocks or breeds. Regression analysis showed that

the SRLV susceptibility can be predicted based on the frequency of allele E with a variation of about plus/minus 20%.

Additionally, we found, for the first time and by a PCR-based test, SRLV positive sheep flocks in the Iranian province of Western Azarbaijan. Characterization of SRLVs in this province (located on Iran's border with Turkey), as well as other regions of Iran, will provide insight into the variability of SRLVs of the entire region.

Author Contributions: Conceptualization, G.L., V.M. and V.O.; Formal Analysis, V.M.; Investigation, V.M., G.L. and D.A.; Resources, G.L., V.O. and D.A.; Writing—Original Draft Preparation, V.M. and G.L.; Writing—Review & Editing, V.M., G.L. and V.O.; Visualization, V.M. and G.L.; Supervision, G.L.; Project Administration, G.L.

Funding: This research received no external funding.

Acknowledgments: We thank Carina Crispens and Stephanie Steitz for their excellent technical support. We also thank Ute Richter for taking the time for language editing of the manuscript. We would like to sincerely thank sheep farmers and veterinarians for getting access to sheep samples.

Conflicts of Interest: The authors declare that they have no conflict of interest.

Appendix A

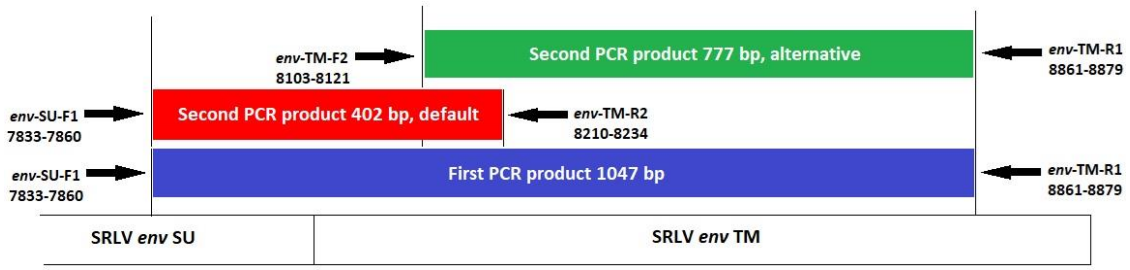


Figure A1. Primer localization in the *env* SU/TM region of small ruminant lentiviruses. Position of primers is according to SRLV strain K1514 (accession number: M60609). SU, surface protein; TM, Transmembrane protein.

Table A1. Samples collected from German breeding flocks (set 3).

Breed No.	Breed	Sheep (*n*)
1	German Grey Heath	31
2	White Moutain	16
3	Rhoen	20
4	Coburg Fox	18
5	Merinoland	42
6	German Black Mutton	21
7	Dorper	31
8	Suffolk	31
9	Texel	32
10	East Friesian Milk	32
11	Cameroonian sheep	28
Total		302

Table A2. Details on PCR amplifications done in this study.

Gene	Primer Sequences (5′ → 3′)	Purpose of PCR	Product Size	Reference
TMEM154	CCACAGGAGAGGAGRACACA (forward) GGGCACGTCTCCTGACAGTTT (reverse, FAM-labeled, matching with K allele) GGCACGTCTCCTGACAGTTC (reverse, HEX-labeled, matching with E allele)	genotyping *TMEM154* E35K	40/41 bp	[14]
TMEM154	GCTAGACACTGCCAAGCTTC (forward) TGTCACTGAAACAAGTCATCACT (reverse)	sequencing *TMEM154* region around E35K	788 bp	[14]
env Surf-Trans	TAATAARRGTRAGAGCTTACACATATGG (*env*-SU-F1) CCTGACAGTCCACCCTTTC (env-TM-R1)	diagnosis of SRLV infection, first PCR	1047 bp	this study
env Surf-Trans	TAATAARRGTRAGAGCTTACACATATGG (*env*-SU-F1) TTACACAGWAGTGTTGATAATGCCA (*env*-TM-R2)	diagnosis of SRLV infection, second PCR (default)	402 bp	this study
env Surf-Trans	ATGCCATGGTACAGCATGT (env-TM-F2) CCTGACAGTCCACCCTTTC (*env*-TM-R1)	diagnosis of SRLV infection, second PCR (alternative)	777 bp	this study
AMELX (*Capra hircus*)	CAGTAGCTCCAGCTCCAGCT (*Aml*.X-F1) GTGCATCCCTTCATTGGC (*Aml*.X-R1)	assessment of DNA quality	300 bp	[24]
SRY (*Capra hircus*)	ATGAATAGAACGGTGCAATCG (*Sry*-F) GAAGAGGTTTTCCCAAAGGC (*Sry*-R)	assessment of DNA quality	116 bp	[24]

Table A3. Genotyping results of *TMEM154* E35K variants in four positive sheep flocks of Iran (IRN) and Germany (GER).

Breed Subset	MV Status	*TMEM154* Genotype Frequencies		
(*n* Sheep)	(*n* Sheep)	KK	EK	EE
Bakhtiari, IRN (50)	neg (9)	0.00 (0)	0.11 (1)	0.89 (8)
	pos (41)	0.00 (0)	0.05 (2)	0.95 (39)
Qezel, IRN (20)	neg (19)	0.74 (14)	0.21 (4)	0.05 (1)
	pos (1)	1.00 (1)	0.00 (0)	0.00 (0)
Texel, GER (39)	neg (0)	0.00 (0)	0.00 (0)	0.00 (0)
	pos (39)	0.00 (0)	0.31 (12)	0.69 (27)
Zackel, GER (14)	neg (6)	0.00 (0)	0.17 (1)	0.83 (5)
	pos (8)	0.00 (0)	0.12 (1)	0.88 (7)

References

1. Peterhans, E.; Greenland, T.; Badiola, J.; Harkiss, G.; Bertoni, G.; Amorena, B.; Eliaszewicz, M.; Juste, R.A.; Kraßnig, R.; Lafont, J.-P.; et al. Routes of transmission and consequences of small ruminant lentiviruses (SRLVs) infection and eradication schemes. *Vet. Res.* **2004**, *35*, 257–274. [CrossRef] [PubMed]
2. Reina, R.; Berriatua, E.; Luján, L.; Juste, R.; Sánchez, A.; de Andrés, D.; Amorena, B. Prevention strategies against small ruminant lentiviruses: An update. *Vet. J.* **2009**, *182*, 31–37. [CrossRef] [PubMed]
3. White, S.; Knowles, D. Expanding possibilities for intervention against small ruminant lentiviruses through genetic marker-assisted selective breeding. *Viruses* **2013**, *5*, 1466–1499. [CrossRef] [PubMed]
4. Pépin, M.; Vitu, C.; Russo, P.; Mornex, J.-F.; Peterhans, E. Maedi-visna virus infection in sheep: A review. *Vet. Res.* **1998**, *29*, 341–367. [PubMed]
5. Ramírez, H.; Reina, R.; Amorena, B.; de Andrés, D.; Martínez, H.A. Small ruminant lentiviruses: Genetic variability, tropism and diagnosis. *Viruses* **2013**, *5*, 1175–1207. [CrossRef] [PubMed]
6. Herrmann-Hoesing, L.M. Diagnostic assays used to control small ruminant lentiviruses. *J. Vet. Diagn. Investig.* **2010**, *22*, 843–855. [CrossRef] [PubMed]

7. Rosati, S.; Profiti, M.; Lorenzetti, R.; Bandecchi, P.; Mannelli, A.; Ortoffi, M.; Tolari, F.; Ciabatti, I.M. Development of recombinant capsid antigen/transmembrane epitope fusion proteins for serological diagnosis of animal lentivirus infections. *J. Virol. Methods* **2004**, *121*, 73–78. [CrossRef]
8. Herrmann-Hoesing, L.M.; White, S.N.; Lewis, G.S.; Mousel, M.R.; Knowles, D.P. Development and validation of an ovine progressive pneumonia virus quantitative PCR. *Clin. Vaccine Immunol.* **2007**, *14*, 1274–1278. [CrossRef]
9. Keen, I.E.; Hungerford, L.L.; Wittum, T.E.; Kwang, J.; Littledike, E.T. Rick factors for seroprevalence of ovine lentivirus in breeding ewe flocks in Nebraska, USA. *Prev. Vet. Med.* **1997**, *30*, 81–94. [CrossRef]
10. Gates, N.L.; Winward, L.D.; Gorham, J.R.; Shen, D.T. Serologic survey of prevalence of ovine progressive pneumonia in Idaho range sheep. *J. Am. Vet. Med. Assoc.* **1978**, *173*, 1575–1577.
11. Larruskain, A.; Jugo, B. Retroviral infections in sheep and goats: Small ruminant lentiviruses and host interaction. *Viruses* **2013**, *5*, 2043–2061. [CrossRef] [PubMed]
12. Heaton, M.P.; Clawson, M.L.; Chitko-Mckown, C.G.; Leymaster, K.A.; Smith, T.P.L.; Harhay, G.P.; White, S.N.; Herrmann-Hoesing, L.M.; Mousel, M.R.; Lewis, G.S.; et al. Reduced lentivirus susceptibility in sheep with TMEM154 mutations. *PLoS Genet.* **2012**, *8*, e1002467. [CrossRef] [PubMed]
13. Leymaster, K.A.; Chitko-McKown, C.G.; Clawson, M.L.; Harhay, G.P.; Heaton, M.P. Effects of TMEM154 haplotypes 1 and 3 on susceptibility to ovine progressive pneumonia virus following natural exposure in sheep. *J. Anim. Sci.* **2013**, *91*, 5114–5121. [CrossRef] [PubMed]
14. Molaee, V.; Eltanany, M.; Lühken, G. First survey on association of TMEM154 and CCR5 variants with serological maedi-visna status of sheep in German flocks. *Vet. Res.* **2018**, *49*, 36. [CrossRef] [PubMed]
15. Yaman, Y.; Keleş, M.; Aymaz, R.; Sevim, S.; Sezenler, T.; Önaldı, A.T.; Kaptan, C.; Başkurt, A.; Koncagül, S.; Öner, Y.; et al. Association of TMEM154 variants with visna/maedi virus infection in Turkish sheep. *Small Rumin. Res.* **2019**, *177*, 61–67. [CrossRef]
16. Heaton, M.P.; Kalbfleisch, T.S.; Petrik, D.T.; Simpson, B.; Kijas, J.W.; Clawson, M.L.; Chitko-McKown, C.G.; Harhay, G.P.; Leymaster, K.A. Genetic testing for TMEM154 mutations associated with lentivirus susceptibility in sheep. *PLoS ONE* **2013**, *8*, 55490. [CrossRef]
17. Tavakolian, J. *An introduction to genetic resources of native farm animals in Iran*; Animal Science Genetic Research Institute Press: Tehran, Iran, 2000. (In Persian)
18. Sayari, M.; Soltanian, S. Serological and pathological study of Maedi-like diseases in mammary glands of sheep of Ahwaz region. *Pajouhesh Sazandegi* **2001**, 2–7. (In Persian)
19. Azizi, S.; Tajbakhsh, E.; Fathi, F.; Oryan, A.; Momtaz, H.; Goodarzi, M. Maedi in slaughtered sheep: A pathology and polymerase chain reaction study in southwestern Iran. *Trop. Anim. Health Prod.* **2012**, *44*, 113–118. [CrossRef] [PubMed]
20. Norouzi, B.; Taghavi-Razavizadeh, A.; Azizzadeh, M.; Mayameei, A.; Najar-Nezhad-Mashhadi, V. Serological study of small ruminant lentiviruses in sheep population of Khorasan-e-Razavi province in Iran. *Vet. Res. Forum* **2015**, *6*, 245–249.
21. Sakhaee, E.; Khalili, M. Serological study of Maedi-Visna virus among sheep flocks in Kerman province of Iran. *Iran. J. Virol.* **2010**, *4*, 29–33. [CrossRef]
22. Sasani, F.; Javanbakht, J.; Hemmatzadeh, F.; Moghadam, M.R.; Hassan, M.A.M. Evaluation of histopathological on maedi disease with serological confirmation in North-East of Iran. *Res. J. Infect. Dis.* **2013**, *1*, 5. [CrossRef]
23. Montgomery, G.W.; Sise, J.A. Extraction of DNA from sheep white blood cells. *New Zeal. J. Agric. Res.* **1990**, *33*, 437–441. [CrossRef]
24. Phua, A.C.Y.; Abdullah, R.B.; Mohamed, Z. A PCR-based sex determination method for possible application in caprine gender selection by simultaneous amplification of the Sry and Aml-X genes. *J. Reprod. Dev.* **2003**, *49*, 307–311. [CrossRef] [PubMed]
25. Altman, D.G. *Practical Statistics for Medical Research*; Chapman and Hall/CRC: London, UK, 1990; ISBN 9780412276309.
26. Clarke, G.M.; Anderson, C.A.; Pettersson, F.H.; Cardon, L.R.; Morris, A.P.; Zondervan, K.T. Basic statistical analysis in genetic case-control studies. *Nat. Protoc.* **2011**, *6*, 121–133. [CrossRef] [PubMed]
27. Sider, L.H.; Heaton, M.P.; Chitko-McKown, C.G.; Harhay, G.P.; Smith, T.P.; Leymaster, K.A.; Laegreid, W.W.; Clawson, M.L. Small ruminant lentivirus genetic subgroups associate with sheep TMEM154 genotypes. *Vet. Res.* **2013**, *44*, 64. [CrossRef] [PubMed]

28. Straub, O.C. Maedi–Visna virus infection in sheep. History and present knowledge. *Comp. Immunol. Microbiol. Infect. Dis.* **2004**, *27*, 1–5. [CrossRef]
29. Muz, D.; Oğuzoğlu, T.Ç.; Rosati, S.; Reina, R.; Bertolotti, L.; Burgu, I. First molecular characterization of visna/maedi viruses from naturally infected sheep in Turkey. *Arch. Virol.* **2013**, *158*, 559–570. [CrossRef] [PubMed]

Study 3: Phylogenetic analysis of small ruminant lentiviruses in Germany and Iran suggests their expansion with domestic sheep

SCIENTIFIC REPORTS
nature research

OPEN

Phylogenetic analysis of small ruminant lentiviruses in Germany and Iran suggests their expansion with domestic sheep

Vahid Molaee[1*], Moira Bazzucchi[2], Gian Mario De Mia[2], Vahid Otarod[3], Darab Abdollahi[4], Sergio Rosati[5] & Gesine Lühken[1]

Small ruminant lentiviruses (SRLVs) are found in sheep in Germany and Iran. SRLVs have been classified into four genotypes: A–C and E. Genotype A has been subdivided into 20 subtypes. Previous studies suggested that, first, the ancestors of genotype A are those SRLVs found in Turkey, second, the evolution of SRLVs is related to the domestication process, and, third, SRLV infection was first observed in sheep in Iceland and the source of that infection was a flock imported from Germany. This study generated, for the first time, partial SRLV sequence data from German and Iranian sheep, enhancing our knowledge of the genetic and evolutionary relationships of SRLVs, and their associations with the domestication process. Based on 54 SRLV sequences from German and Iranian sheep, our results reveal: (1) SRLV subtypes A4, A5, A11, A16 and A21 (new) are found in German sheep and A22 (new) in Iranian sheep. (2) Genotype A has potentially an additional ancestor (A22), found in Iran, Lebanon and Jordan. (3) Subtype A22 is likely an old version of SRLVs. (4) The transmission routes of some SRLVs are compatible with domestication pathways. (5) This study found no evidence of Icelandic subtype A1 in German sheep.

Small ruminant lentiviruses (SRLVs), which comprise maedi-visna virus (MVV) and caprine arthritis encephalitis virus (CAEV), belong to the genus Lentivirus and the family Retroviridae. SRLVs can cause progressive multisystem disease in sheep involving lungs, joints, mammary gland and the central nervous system[1]. There is no cure or vaccine available against SRLV infection. SRLV-related diseases are distributed worldwide among sheep and goats, resulting in considerable economic losses[2].

Like in other lentiviruses, the SRLV genome includes three structural genes, coding for the group-specific antigens (*gag*), the polymerase (*pol*) and the envelope (*env*). The *gag* gene encodes the matrix (MA) protein (p17), capsid (CA) protein (p25) and nucleocapsid (NC) protein (p14)[3]. Both *gag* and *pol* genes are relatively conserved, and phylogenetic analyses of SRLVs have been established based on these two genes[4].

SRLV isolates can be classified into four genotypes, A–C and E[4–6]. Genotypes A and B are widespread and refer to MVV-like and CAEV-like viruses, respectively. MVV-like and CAEV-like strains have been first described in sheep and goats, respectively, and considered strictly host-specific for a long time. However, there are nowadays several studies indicating that most strains can cross the species barrier (reviewed by Minardi da Cruz *et al.*[7]). Genotype A is the most heterogeneous group and has so far been subdivided into 20 subtypes, A1 to A20[4,8–15]. Two recently published studies have to be noted, one by Olech *et al.*[13] that defines SRLV subtype A18, and the other by Colitti *et al.*[15] that defines SRLV subtypes A18 and A19. In the present study, the SRLVs found by Colitti *et al.*[15] are renamed from 'A18' to 'A19' and from 'A19' to 'A20'. Genotype B contains three subtypes, B1 to B3[4,16]. SRLVs restricted to certain geographical areas have been assigned to other genotypes: genotype C is divided

[1]Institute of Animal Breeding and Genetics, Justus Liebig University Giessen (JLU), Ludwigstraße 21, 35390, Gießen, Germany. [2]Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche Togo Rosati (IZSUM), Via G. Salvemini 1, 06126, Perugia, Italy. [3]Quarantine and Biosafety Directorate General, Iran Veterinary Organization (IVO), Vali Asr Avenue, Seyed Jamaledin Asad Abadi Street, 6349, Tehran, Iran. [4]Bureau of Animal Health and Disease Management, Iran Veterinary Organization (IVO), Vali Asr Avenue, Seyed Jamaledin Asad Abadi Street, 6349, Tehran, Iran. [5]Department of Veterinary Science, University of Turin (UNITO), Largo Paolo Braccini 2, 10095, Grugliasco, Torino, Italy. *email: Vahid.Molaee@agrar.uni-giessen.de

into two subtypes and refers to Norwegian isolates[15,17,18], genotype D was found in few isolates originating from Switzerland and Spain, but they are now re-classified as genotype A[4,6,19]; genotype E comprises subtypes E1 and E2 and was isolated in Italy[9,20].

Lentiviruses have a deep evolutionary history and have evolved alongside their mammalian hosts[21–25]. During the process of evolution, exogenous retroviruses (e.g. Jaagsiekte sheep retrovirus; a retrovirus that has many similarities with SRLVs and imposes pulmonary adenocarcinoma disease in sheep) have inserted into the germline of the infected host, leading to endogenous retroviruses (ERVs)[26]. In a study by Chessa *et al.*[27], the presence of six variants of endogenous Jaagsiekte retrovirus (enJSRVs) was examined in 65 global domestic sheep, to investigate the history of sheep domestication. During the first wave of domestication, early domesticated sheep, which were morphologically wild but managed, appeared in the ancient Fertile Crescent region, including parts of Iran, Iraq, Turkey, Syria and Jordan, approximately 10,000 to 8,000 years before present (YBP)[28,29]. The domestication of goats started in the same region between 500 and 1,000 years earlier at about 11,000 YBP[28,30]. During the second wave of domestication, sheep with 'modern' features, typical of present-day breeds (e.g. with woolly fleeces and polled), appeared in West Asia and other areas in the world at approximately 6,000 YBP[27]. Host molecular genetic data combined with archaeological evidence indicates that sheep were distributed from the Fertile Crescent to the West and East likely during both waves of domestication[27,31,32]. For instance, during the first wave of domestication, the European mouflon migrated through the Mediterranean Basin to the islands of Corsica and Sardinia at around 7,000 YBP[21,27,32,33]. During the second wave, sheep with improved production traits were introduced from the East into Europe at the beginning of the 5th millennium YBP[34].

In parallel with researches on host evolution, recent studies have demonstrated the usefulness of pathogens to elucidate the evolution of their hosts across time and location[21,35–37]. In this respect, besides enJSRVs[27], investigating the phylogeny of SRLVs has the potential to enhance our knowledge of sheep and goat domestication[16,20,21]. The identification of SRLV subtype B3 in sheep/goats from Italy and Turkey as well as the finding that some bulk milk samples from Turkish sheep and goats were reactive against antigen derived from genotype E (a genotype found in Sardinia and other parts of Italy), supports the hypothesis of migration of domesticated sheep from the Fertile Crescent into the Mediterranean Basin during the Neolithic age[16]. However, phylogenetic studies involving SRLV sequences from other regions of the Fertile Crescent (except Turkey), which would potentially enhance our knowledge of the domestication process within the domestication origin itself, have been absent until now. While few SRLV sequences from Jordan and Lebanon are available in the database, no SRLV sequence information has been published from Iran, Iraq and Syria.

Historically, it has been suggested that maedi-visna was introduced to Iceland through German sheep[2,6,38–40]. In the early 1930s, a set of Karakul sheep (n = 20) was imported from Halle in Germany to Iceland. After several months, signs of maedi-visna were observed in some Icelandic sheep flocks, which hosted German Karakul rams[38]. The German Karakul flock originated from Astrakhan (Russia)[41]. The Icelandic SRLV strain (subtype A1) was first characterised by Sigurdsson *et al.*[38] about 15 years after the observation of the first signs of maedi-visna in the Icelandic sheep flocks. Subtype A1 had already been detected in many countries[6]. However, German SRLVs have not yet been characterised.

In Iran, SRLV infection was first diagnosed in sheep using histopathological methods in 2001[42]. Following this first survey, SRLV infection was reported in different parts of Iran, using serological methods or PCR techniques[43–46].

Currently, researchers have suggested three domestication pathways, from West Asia (Iran and Turkey) to Europe and Africa[32,47]. Germany is located on the end terminal of the Danubian pathway, Italy is located on the northern Mediterranean pathway and Morocco is located on the southern Mediterranean pathway[47]. In this respect, Iran belongs to the ancient Fertile Crescent region, where the initial domestication of sheep and goats occurred[28–30]. Also, the geographical position of Germany for investigating domestication events is important, as it has long been connected with the ancient Fertile Crescent via the Danube River[32,47]. This study aimed at generating, for the first time, SRLV sequence data from German and Iranian sheep, for which no such data has been available until now. The results were expected to enhance our knowledge of the genetic and evolutionary relationships of German and Iranian SRLVs, as compared to those from other countries, as well as their associations with the domestication process.

Results

SRLV sequences from German and Iranian sheep flocks. All DNA samples (n = 54), were successfully amplified using *gag-pol* primers. None of the sequences showed evidence of recombination. Information on *gag-pol* SRLV sequences collected from different sheep flocks in Germany and Iran, and the accession numbers are given in Table 1. Notably, due to unmatched sequence data for the alignment (Supplementary Fig. S1 online), genetic and phylogenetic analyses on *gag-pol* sequences correspond to a part of the *gag* gene. Based on initial analyses, 23 out of the 54 *gag* sequences were selected for phylogenetic and genetic analyses. Of the 23 selected sequences, 17 *gag* sequences belonged to 13 German sheep flocks and 6 corresponded to 6 Iranian sheep flocks.

Phylogenetic analysis of SRLVs based on *gag* fragment. The constructed phylogenetic tree with *gag* sequences is shown in Fig. 1. According to the phylogenetic analysis, the *gag* sequences from Germany (n = 17), which were distributed in five distinct clusters, were strongly related to genotype A. The tree showed that these sequences were affiliated to subtypes A4 (flock 12), A5 (flocks 2–4), A11 (flocks 8, 10 and 13), A16 (flock 1) and potentially to a new subtype, which could be tentatively named A21 (flocks 3 and 5–11). Not all of these clusters were supported with high bootstrap values. The six Iranian SRLV sequences belonged to genotype A with a bootstrap value of 96%. Furthermore, the Iranian SRLV sequences clustered together with the Jordanian (KT898826 and KT921318) and Lebanese (KU170760) SRLV prototypes. This cluster was tentatively named A22.

sample no.	flock no.	country	German state/Iranian province	strain	accession number	proposed subtype
1	1	Germany	Baden Württemberg	BW1	MN233104	A16
2	2	Germany	Bayern	BY1	MN233105	A5
3	3	Germany	Hessen	HE1	MN233106	A21
4	3	Germany	Hessen	HE2	MN233107	A5
5	4	Germany	Nordrhein-Westfalen	NW1	MN233108	A5
6	5	Germany	Nordrhein-Westfalen	NW2	MN233109	
7	5	Germany	Nordrhein-Westfalen	NW3	MN233110	
8	5	Germany	Nordrhein-Westfalen	NW4	MN233111	A21
9	5	Germany	Nordrhein-Westfalen	NW5	MN233112	
10	5	Germany	Nordrhein-Westfalen	NW6	MN233113	
11	6	Germany	Schleswig-Holstein	SH1	MN233114	
12	6	Germany	Schleswig-Holstein	SH2	MN233115	
13	6	Germany	Schleswig-Holstein	SH3	MN233116	A21
14	7	Germany	Schleswig-Holstein	SH4	MN233117	
15	7	Germany	Schleswig-Holstein	SH5	MN233118	A21
16	8	Germany	Schleswig-Holstein	SH6	MN233119	
17	8	Germany	Schleswig-Holstein	SH7	MN233120	A21
18	8	Germany	Schleswig-Holstein	SH8	MN233121	
19	8	Germany	Schleswig-Holstein	SH9	MN233122	
20	8	Germany	Schleswig-Holstein	SH10	MN233123	
21	8	Germany	Schleswig-Holstein	SH11	MN233124	A11
22	8	Germany	Schleswig-Holstein	SH12	MN233125	
23	8	Germany	Schleswig-Holstein	SH13	MN233126	
24	9	Germany	Schleswig-Holstein	SH14	MN233127	
25	9	Germany	Schleswig-Holstein	SH15	MN233128	A21
26	9	Germany	Schleswig-Holstein	SH16	MN233129	
27	9	Germany	Schleswig-Holstein	SH17	MN233130	
28	9	Germany	Schleswig-Holstein	SH18	MN233131	
29	9	Germany	Schleswig-Holstein	SH19	MN233132	A21
30	9	Germany	Schleswig-Holstein	SH20	MN233133	
31	9	Germany	Schleswig-Holstein	SH21	MN233134	
32	9	Germany	Schleswig-Holstein	SH22	MN233135	
33	9	Germany	Schleswig-Holstein	SH23	MN233136	
34	10	Germany	Schleswig-Holstein	SH24	MN233137	
35	10	Germany	Schleswig-Holstein	SH25	MN233138	
36	10	Germany	Schleswig-Holstein	SH26	MN233139	
37	10	Germany	Schleswig-Holstein	SH27	MN233140	
38	10	Germany	Schleswig-Holstein	SH28	MN233141	A21
39	10	Germany	Schleswig-Holstein	SH29	MN233142	
40	10	Germany	Schleswig-Holstein	SH30	MN233143	A11
41	11	Germany	Schleswig-Holstein	SH31	MN233144	
42	11	Germany	Schleswig-Holstein	SH32	MN233145	A21
43	11	Germany	Schleswig-Holstein	SH33	MN233146	
44	11	Germany	Schleswig-Holstein	SH34	MN233147	
45	12	Germany	Schleswig-Holstein	SH35	MN233148	A4
46	13	Germany	Schleswig-Holstein	SH36	MN233149	
47	13	Germany	Schleswig-Holstein	SH37	MN233150	
48	13	Germany	Schleswig-Holstein	SH38	MN233151	A11
49	14	Iran	Chaharmahal-Va-Bakhtiari	BKH1	MK098477	A22
50	15	Iran	Chaharmahal-Va-Bakhtiari	BKH2	MK098478	A22
51	16	Iran	Kerman	KRM1	MK098479	A22
52	17	Iran	Kerman	KRM2	MK098480	A22
53	18	Iran	Western Azarbaijan	MKU1	MK098481	A22
54	19	Iran	Western Azarbaijan	MKU2	MK098482	A22

Table 1. Information on *gag-pol* SRLV sequences collected from different sheep flocks in Germany and Iran.

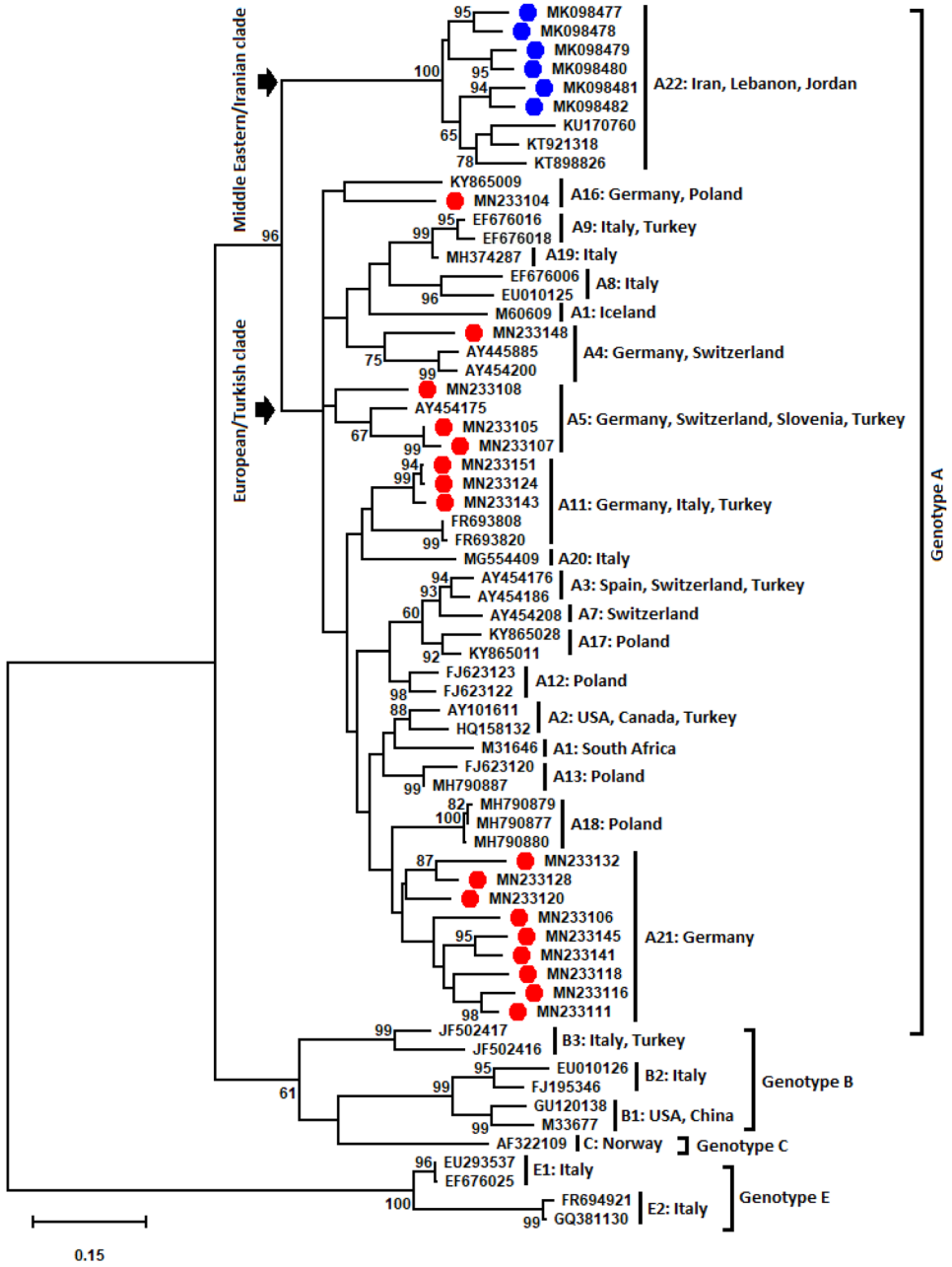


Figure 1. Phylogenetic tree indicates SRLV subtypes and their differing geographical distribution (country). The analysis was performed using the Maximum Likelihood method and was based on the Tamura-Nei model[61]. The analysis involved a total of 399 bp from 65 nucleotide sequences: 17 German *gag* sequences (labelled by solid red circles), six Iranian sequences (labelled by solid blue circles), two Jordanian SRLV sequences (KT898826 and KT921318), a Lebanese SRLV sequence (KU170760) and 39 reference SRLV sequences originating from different geographical areas (retrieved from the GenBank database). The position of the analysed fragments was related to coding regions of the capsid gene (p25) and the nucleocapsid gene (p14) located on the *gag* fragment (nucleotide: 1114–1506; numbering according to prototype strain K1514[62]. The difference in the evolutionary rate among sites was considered using the discrete gamma distribution (+G parameter = 0.5202) and the invariable sites (+I = 21.68% sites). The numbers on the nodes indicate the percentage of bootstrap values obtained from 1,000 replicates. Bootstrap values of less than 60% were excluded from nodes. The branch lengths show the number of substitutions per site. Evolutionary analyses were performed in MEGA7[60]. The two suggestive clades of genotype A, "European/Turkish clade" and "Middle Eastern/Iranian clade" are shown with arrows.

Genetic distance analyses of German and Iranian SRLV sequences based on *gag* gene. As all German and Iranian *gag* sequences tend to be genotype A, a cut-off value was determined for assigning a new subtype within genotype A. The mean genetic distances between available *gag* reference data (A1–A20 except

sequence	country	A1	A2	A3	A4	A5	A7	A8	A9	A11	A12	A13	A16	A17	A18	A19	A20	Mean A
MN233148	Germany	17.43	16.28	14.89	13.23	12.98	15.27	16.54	16.69	16.79	15.78	15.39	17.05	14.41	16.92	14.50	17.30	15.72
MN233105	Germany	16.41	15.01	13.23	16.03	9.41	13.49	17.18	13.50	15.78	13.61	13.61	16.28	11.35	14.12	12.21	13.74	14.06
MN233108	Germany	16.16	14.50	14.89	15.65	12.21	14.50	18.07	15.92	13.23	14.89	14.63	15.01	15.69	14.08	16.03	15.27	15.05
MN233107	Germany	16.28	15.52	12.47	16.79	9.92	12.72	17.94	13.76	15.78	13.61	13.10	16.79	11.10	13.10	12.47	14.76	14.13
MN233143	Germany	14.50	11.96	14.12	16.92	14.25	15.52	16.92	15.29	12.98	12.98	13.49	17.56	14.16	12.47	14.76	12.98	14.43
MN233124	Germany	14.38	11.20	14.25	16.92	13.99	15.52	17.05	14.27	12.47	12.21	13.49	16.03	13.90	12.72	14.25	13.49	14.13
MN233151	Germany	14.38	10.69	14.25	16.41	13.49	15.52	17.05	14.27	12.47	12.21	13.49	16.03	13.39	12.72	13.74	13.49	13.98
MN233104	Germany	15.78	16.28	17.18	16.54	16.03	16.79	19.47	18.73	17.18	15.14	15.27	14.25	14.29	15.52	17.30	15.78	16.35
MN233116	Germany	16.67	15.52	15.39	18.70	15.78	15.27	18.83	16.82	14.50	14.63	14.89	17.56	15.43	14.63	16.54	16.28	16.09
MN233111	Germany	15.52	13.87	14.12	19.85	13.74	15.01	18.07	15.80	13.49	13.87	13.10	15.78	14.80	14.25	13.99	15.52	15.05
MN233118	Germany	15.65	13.49	14.89	17.94	13.99	15.01	17.94	15.54	15.52	13.74	12.98	17.05	13.78	13.10	15.01	16.79	15.15
MN233132	Germany	18.32	15.14	16.16	17.43	15.52	16.54	20.48	18.47	15.65	15.27	15.52	18.32	15.94	16.03	16.79	17.30	16.81
MN233120	Germany	14.76	12.47	15.01	15.78	14.25	14.50	16.92	15.41	12.85	12.47	10.81	14.50	13.65	11.70	14.25	13.74	13.94
MN233145	Germany	16.16	14.38	11.83	16.79	14.50	11.96	16.92	15.16	13.49	11.07	14.38	16.28	13.90	16.28	14.76	16.28	14.63
MN233128	Germany	16.67	12.98	14.38	15.39	13.23	14.50	18.32	16.94	16.16	13.10	11.45	16.28	14.41	11.79	14.25	17.05	14.80
MN233141	Germany	16.92	13.74	13.49	17.18	13.99	13.23	18.58	16.31	14.76	14.50	14.50	17.30	13.78	13.74	15.27	15.27	15.16
MN233106	Germany	15.14	13.61	13.99	16.79	13.99	14.76	16.28	14.90	15.01	14.76	14.12	16.28	12.76	12.60	13.74	16.28	14.69
MK098477	Iran	19.34	18.32	17.81	18.32	15.01	17.81	19.72	19.24	19.59	18.07	19.21	18.83	16.58	18.15	18.32	18.83	18.32
MK098478	Iran	19.97	18.58	18.70	18.96	15.78	19.08	20.48	19.49	18.19	17.68	20.23	20.10	18.11	18.24	19.08	19.59	18.89
MK098479	Iran	20.10	18.07	15.65	19.72	15.27	17.05	18.58	16.69	18.07	17.30	20.23	19.08	17.47	19.93	16.28	16.54	17.88
MK098480	Iran	19.85	19.21	15.65	19.59	15.78	16.28	17.18	16.82	18.83	17.94	19.59	19.85	16.45	19.17	16.79	16.54	17.85
MK098481	Iran	19.21	19.59	17.94	19.21	16.03	19.34	20.74	18.85	17.94	17.05	18.96	19.34	18.37	18.83	18.32	17.30	18.56
MK098482	Iran	18.58	18.96	16.92	19.47	15.01	18.58	19.34	17.32	17.81	17.30	19.08	19.85	17.35	19.76	18.07	18.32	18.23

Table 2. Estimates of evolutionary divergences of German and Iranian SRLV sequences compared to different subtypes of genotype A based on the *gag* fragment (nucleotide: 1114–1506; numbering according to prototype strain K1514, Staskus *et al.*[62]). For each reference subtype, up to three strains were selected and the evolutionary divergences calculated based on the mean divergence of each set of subtypes and representative sequences of this study. 17 German and six Iranian SRLV sequences were compared with subtypes A1 (M60609 and M31646), A2 (AY101611 and HQ158132), A3 (AY454186 and AY454176), A4 (AY454200 and AY445885), A5 (AY454175), A7 (AY454208), A8 (EU010125 and EF676006), A9 (EF676018 and EF676016), A11 (FR693820 and FR693808), A12 (FJ623122 and FJ623123), A13 (MH790877and FJ623120), A16 (KY865009), A17 (KY865011 and KY865028), A18 (MH790877, MH790879 and MH790880), A19 (MH374287) and A20 (MG554409).

for A6, A10, A14 and A15) was 14.67% (SE = 0.19%, 95%CI = 14.29–15.05%). The value of 14.67% was the used cut-off value for defining a new subtype within genotype A. German *gag* sequences of the proposed subtype A21 (n = 9), had mean sequence similarity of 11.62% (range = 0.06–15.01%). The following mean genetic distances between the new subtype (A21) and different genotypes were observed: A: 15.15%, range = 13.94–16.81%; B: 23.11%, range = 22.01–24.17%; C: 23.61%, range = 21.88–25.95%; E: 28.50%, range = 27.42–29.46% (Table 2 and Supplementary Table S1 online). Based on phylogeny, the most similar subtype with cluster A21 was A18. In two out of nine *gag* sequences of cluster A21 (MN233132 and MN132145), genetic distances compared to A18 were higher than the cut-off value. Observation of these two sequences in cluster A21 that genetically differed from A18 (>14.67), fulfils the criterion of proposing a new subtype for a cluster in SRLVs[4]. Thus, all nine sequences of cluster A21 constitute subtype A21 (new). In other German *gag* sequences, the genetic distances with the most similar subtypes (corresponding to subtypes A4, A5, A11 and A16 based on phylogeny) were lower than the cut-off value (<14.67). Therefore, other German *gag* sequences did not constitute new subtypes and joined the reference subtypes A4 (MN233148), A5 (MN233105, MN233107 and MN233108), A11 (MN233124, MN233143 and MN233151) and A16 (MN233104). Notably, according to phylogeny (Fig. 1), the most similar subtypes with the three sequences of MN233143, MN233124 and MN233151 were A11 and A20 (both SRLV subtypes from Italy). Based on genetic distance analysis these three sequences were more similar with A11 than with A20. More details are shown in Table 2 and Supplementary Table S1 online.

The *gag* sequences of six Iranian sheep constituted a cluster (subtype A22) within genotype A. The mean genetic similarity between Iranian SRLVs was 10.62% and varied from 5.34% to 12.72%. The following mean genetic distances between Iranian SRLV sequences (A22) and different genotypes were observed: A: 18.29%, range = 17.85–18.89%; B: 23.45%, range = 22.31–24.77%; C: 24.13%, range = 22.65–25.45%; E: 28.73%, range = 27.87–29.84% (Table 2 and Supplementary Table S1 online). The genetic divergences between six Iranian, two Jordanian (KT898826 and KT921318) and one Lebanese (KU170760) *gag* sequences were lower than the cut-off value (<14.67). Therefore, all these nine *gag* sequences together constitute subtype A22. More details are shown in Table 2 and Supplementary Table S1 online.

Amino acid (aa) sequence analysis based on the *gag-pol* fragment. The alignment of 65 *gag-pol* aa sequences (up to 214 aa) is shown in Fig. 2 and Supplementary Fig. S2 online. The reference *gag-pol* sequence in both Fig. 2 and Supplementary Fig. S2 online is the Iranian SRLV strain of BKH1 (accession number: MK098477). Except for the reference *gag-pol* sequence (MK098477), the *gag-pol* aa sequences in Fig. 2 (n = 32) and Supplementary Fig. S2 online (n = 32) are not identical. Therefore, the sum of *gag-pol* aa sequences that have been aligned with MK098477 is 64, and altogether a total of 65 *gag-pol* aa sequences is shown in Fig. 2 and Supplementary Fig. S2 online.

Both German and Iranian SRLV sequences share similar epitope patterns with other sequences of genotype A. Also, they did not share the double glycine "GG" motif[48], as other sequences of genotype A (Fig. 2). When comparing amino acid sequence substitutions at epitopes 2, major homology region (MHR) and epitopes 3, few alterations were observed mostly in the middle part of the conserved SRLV domains found in Germany and Iran.

All the Iranian SRLVs, as well as the Jordanian (KT898826 and KT921318) and Lebanese (KU170760) strains, contained an insertion at position 173, which was not seen in any other small ruminant lentiviruses (Fig. 2 and Supplementary Fig. S2 online). A second insertion at position 172 was found exclusively in two SRLV *gag-pol* sequences in sheep flocks from the Iranian province of Chaharmahal-Va-Bakhtiari. Interestingly, these insertions (positions 172 or 173) are found exclusively in SRLV subtype A22 and as well in other lentiviruses, including bovine immunodeficiency virus (BIV), human immunodeficiency virus type 1 (HIV1), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV), equine infectious anemia virus (EIAV) and human immunodeficiency virus type 2 (HIV2) (Supplementary Fig. S3 online).

Geographical distribution of SRLV subtypes found in this study. SRLV subtypes of A4, A5, A11, A16, A21 and A22 were observed in this study. Except for A21, the other SRLVs have also been found in other countries. The qualitative aspects of the phylogeography were presented in Fig. 3 by showing the geographic dispersal of pairwise SRLV subtypes of A5, A11 and A22. Transmission routes of SRLV subtypes A5 (solid red line), A11 (solid black line) and A22 (solid green line) follow the Danubian pathway, northern Mediterranean pathway and the ancient Fertile Crescent, respectively. As subtype A4 was only determined in Germany and Switzerland[4], and subtype A16 was only observed in Germany and Poland[12], for simplicity, their transmission routes were not shown in Fig. 3.

Discussion

Maedi-visna disease, already distributed among the sheep flocks of Germany[41,49], was first detected in the German East (1967) and then in the West (1972)[41]. It was reported in southwestern Iran in 2001[42] and later in other parts of the country[43–45]. In this study, for the first time, *gag* sequences of German and Iranian SRLVs were generated and analysed, using samples from five German states and three Iranian provinces. Our analyses identified a noticeable variation in German SRLV sequences that was absent in the Iranian ones. In the 13 studied German flocks, SRLV subtypes A4, A5, A11, A16 and A21 were found, with co-infection being identified in three flocks: flock 3 (A5 and A21), flocks 8 and 10 (A11 and A21) (Table 1). In the six Iranian flocks, only subtype A22 was found.

In earlier studies, migration of at least some SRLVs was related to the old times[16,50]. For example, subtype B3 and genotype E, which are rare, were found in Turkey and Italy (Sardinia and other regions)[16]. Turkey is one of the centres of sheep and goat domestication[28,32], and the early agriculture of Sardinia is related to the Neolithic period[21,27,32,33]. Later, observations of various SRLV subtypes of genotype A (A2, A3, A5, A9 and A11) in sheep flocks of Turkey let Muz *et al.*[50] propose that the ancestors of all SRLVs, especially genotype A, are Turkish SRLVs, and the evolution of some SRLVs is related to the domestication process or a more recent transmission pathway during the Ottoman Empire (about 14th–19th centuries). Our results are consistent with this scenario, which relates the evolution of SRLVs to the domestication process[16,50].

So far, SRLV subtype A22 has been observed in sheep in Iran, Lebanon and Jordan but not in any European countries, including Turkey. The absence of A22 among small ruminants of Turkey and Europe is likely related to their hosts' lineages. Several reports have shown that domestic sheep have different wild ancestors. Already, five lineages (A–E) have been identified in modern domestic sheep through global studies of mitochondrial DNA[51–55]. There is a relatively low percentage of lineages C–E, whilst the two most frequent lineages in domestic sheep are A and B, as identified by Hiendleder and colleagues[51,52,55]. These two lineages could be linked, but not completely, with modern fat- and thin-tailed where the first one is mostly found in Iran and Eastern Asia and the other one in Europe (including Turkey)[56]. Possibly, this could be a result of geographical separation of sheep lineages in the past. Accordingly, a long-term genetic differentiation took place between SRLV subtype A22, found in Iran, Lebanon and Jordan and other SRLVs of genotype A, found in Europe. Therefore, A22 could be suggested as an additional ancestor for genotype A of SRLVs (see Fig. 1; European/Turkish clade vs. Middle Eastern/Iranian clade).

In this study, we realised that the relationship between transmission routes of some SRLVs (A5, A11 and A22, this study; A3, A9, B3 and E reviewed by Ramírez *et al.*[6]) are compatible with domestication pathways of sheep. Three domestication pathways have been proposed from the Near East (Iran and Turkey) to the West and Africa[32,47]. The first pathway follows the Danube River from the Near East (Turkey) to Germany[32,47]. An excellent example of this distribution is the observation of SRLV subtype A5 in Turkey[50] (start terminal), Slovenia[10], Switzerland[4] and Germany (end terminal). This subtype's existence in countries lying on the route between Turkey and Germany (Fig. 3, red line) shows that the distribution is targeted. A further hint for the existence of this pathway is SRLV subtype A3, which has been identified in Turkey, Switzerland and Spain (reviewed by Ramírez *et al.*[6]). The second domestication pathway runs from the Near East via the Mediterranean route to southern Europe[32,47], including Italy. SRLV subtype A11 hints at this distribution. It has been detected in Turkey[50] and Italy[8] and in the current study also in Germany (Fig. 3, black line). Other clues for the second domestication pathway could be related to the previously observed subtypes A9, B3 and genotype E in Turkey and Italy[8,16,50].

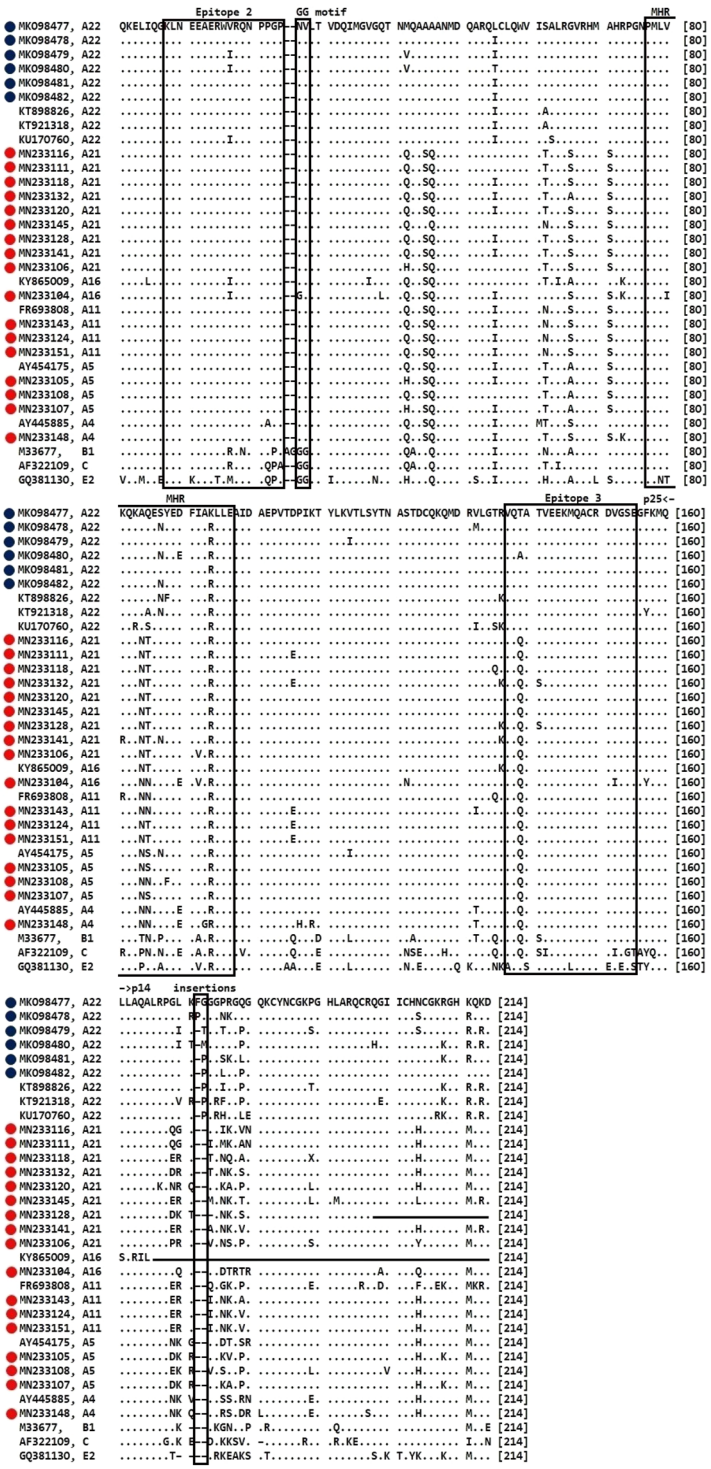


Figure 2. Amino acid sequence alignments of the SRLV *gag-pol* fragment. The reference sequence is the SRLV sequence of strain BKH1 (accession number: MK098477) from the Iranian province of Chaharmahal-Va-Bakhtiari. For this amino acid comparison, 6 Iranian (labelled by solid blue circles) and 17 German sequences (labelled by solid red circles) were compared with different SRLV subtypes/genotypes. Immunodominant epitopes 2 and 3, the major homology region (MHR), the double glycine motif (GG) and insertions (at position 172 or 173) are delineated with boxes. The major core protein (p25) and nucleic acid-binding protein (p14) are separated with left and right arrows (p25 ≤ position 160; p14 ≥ position 161). Dashes and dots indicate deletions and identical residues, respectively.

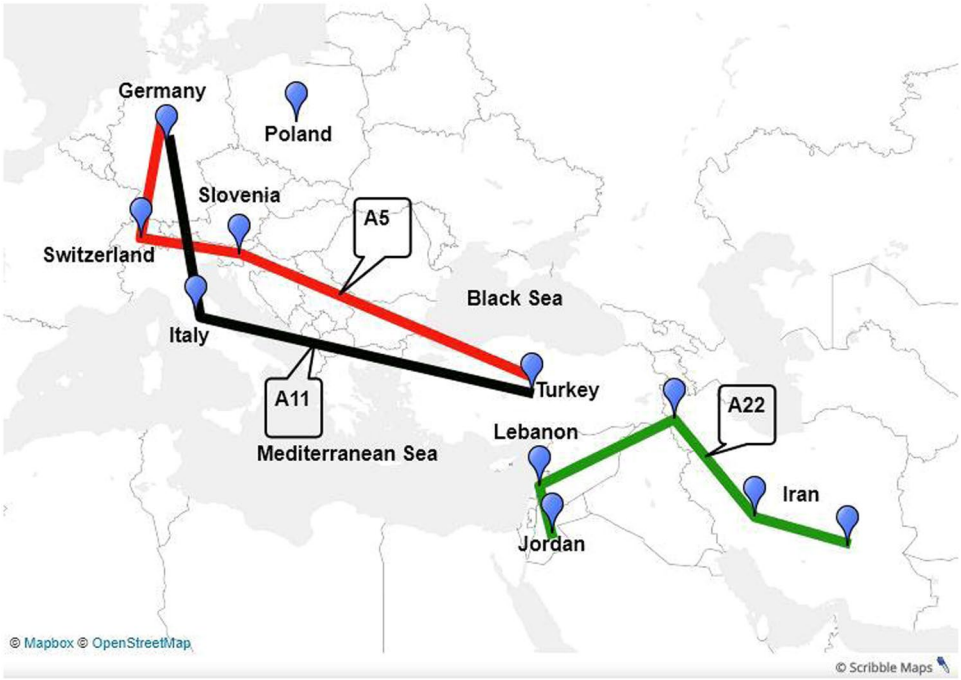


Figure 3. Putative transmission routes of SRLV subtypes A5, A11 and A22 based on their geographical distribution. The Danubian domestication pathway, corresponding to the Danube River, rises in Germany, flows through different European countries, and finally reaches the Black Sea. The northern Mediterranean pathway comprises Turkey and southern parts of Europe, including Italy. The ancient Fertile Crescent region incorporates Iran, Iraq, Turkey, Syria, Lebanon and Jordan (based on Harlan and Zohary[57]). The map was created using online Scribble Maps (https://www.scribblemaps.com/). For more information concerning printing permissions and inclusion of logos please visit https://help.scribblemaps.com/knowledgebase/articles/878916-printing-permissions-book-offline-etc. The online Scribble Maps used the databases of "© OpenStreetMap", "© Mapbox" and "Improve this map". The OpenStreetMap provides free geographic data and is available under the Open Database Licence (https://www.openstreetmap.org/copyright). To learn more about the Mapbox, visit https://www.mapbox.com/about/maps/. Further information regarding the use of "Improve this map" is available at https://www.mapbox.com/map-feedback/. The URL for linking to the map in Fig. 3 is https://www.scribblemaps.com/maps/view/Figure_3_/XUxAbALsVL.

The third proposed domestication pathway runs from the Near East via the southern Mediterranean route to the northern parts of Africa[32,47]. A good support of this distribution is shown by the observations of A22 in Iran, Lebanon and Jordan; these observations are also compatible with the location of the ancient Fertile Crescent[57] (Fig. 3, green line), where domestication originated. Therefore, our data reflect an association between the distribution of domestication and the transmission of some SRLVs. However, the distribution of some SRLV subtypes may not be related to antiquity. For example, subtype A16 was found only in Germany and Poland[12]. Subtype A21, the most common subtype in the North of Germany, is genetically close to subtype A18[13] in Poland. A study by Olech *et al.*[13] mentioned an epidemiological linkage between the small ruminants of Germany and Poland after the Second World War, which is supported by our results.

Amino acid insertions 172 or 173 were observed in SRLV subtype A22 but were not found in any other subtypes. Notably, we observed these insertions to be typical for subtype A22 and other lentiviruses but not for all other SRLVs. As these insertions are in a variable area of the *gag-pol* region (Supplementary Fig. S3 online), bioinformatics analyses may not show how these insertions have evolved in different lentiviruses. Observation of these insertions, however, confirms evolutionary processes within and between different lentiviruses. Furthermore, the hosts of the Jordanian isolates were sheep of the Awassi breed (M. Mazzei, University of Pisa, Italy, personal communication), which is a fat-tailed sheep with a history related to the second wave of domestication (about 5,000 years ago)[58]. Likewise, the Iranian samples came from three provinces close to the Zagros Mountains that are also known as primary centres of goat and sheep domestication[29,30]. The common presence of a unique variant of SRLVs (insertion at positions 172 or 173) among sheep from the countries of the ancient Fertile Crescent suggests that subtype A22 may represent an ancient precursor of modern SRLVs.

The detection of a single SRLV A subtype in Iranian, Jordanian and Lebanese sheep also suggests that these sheep have lived in isolation for centuries. The restricted variability points to this SRLV subtype as a potential marker to trace sheep domestication pathways. In contrast, the situation in Europe may be much more complex and the influence of trading overwhelming, confounding the potential link of SRLV genotypes to the domestication pathways, such as the Danubian or northern Mediterranean pathways. In this respect, Chessa and colleagues[27] showed that tracking infection and endogenization of the Jaagsiekte sheep retrovirus in different sheep breeds permits the reconstruction of the domestication history of sheep. Unlike SRLVs, Jaagsiekte sheep

retrovirus introduces a stable marker in the genome of the infected animals and may be more suitable to follow the domestication pathways of the species under study. Therefore, further studies on sheep genome variations using endogenous retroviruses may provide a more precise picture of the domestication pathways in Europe.

Earlier reports mentioned that maedi-visna (subtype A1) arrived in Iceland with Karakul sheep imported from Germany. Straub[41] showed that this Karakul flock existed in Germany until 1970 but that the sheep of this flock had never shown any signs of maedi-visna. Currently, there is a meager number of Karakul sheep in Germany (maybe a single breeding flock), and this flock is not known to be SRLV positive. However, SRLVs are not strictly breed-specific. They can circulate from one region to another and as well from one breed to another. In this study, we found no evidence for the presence of the Icelandic subtype (A1) among 48 *gag* sequences in German sheep. However, it is possible that there are additional subtypes or genotypes of SRLVs present in German sheep flocks that have not yet been identified, including A1. More work is therefore needed to firmly rule out the hypothesis of distribution of subtype A1 via Karakul sheep from Germany to Iceland.

Shah *et al.*[4] developed a systematic taxonomic classification for SRLVs, based on 1.8-kb *gag-pol* sequences. At that time and later, many strains have been included in the list of new SRLVs when the characteristic of a new strain was distinct enough from other strains. However, in different research projects, different fragments of *gag-pol* were sequenced. The available sequence data for the SRLV subtypes A12, A13, A14, A16, A17 and A18 were not fully matched to the sequence data of other SRLV subtypes of genotype A. As a result, the alignment and classification of SRLVs were limited to only 0.4-kb of *gag* gene (Supplementary Fig. S1 online). In this study, we used the same analysis for defining the cut-off value as it was used by Shah *et al.*[4]. As the 0.4-kb *gag* fragment we used in this study is more conserved than the 1.8-kb *gag-pol* fragment used by Shah *et al.*[4], we adapted the cut-off value (14.67% instead 15%) in order to get an accurate classification.

The principal limitation of this study is the restricted number of sequences obtained. 48 samples were collected in Germany, 38 of which came from Schleswig-Holstein, 6 from Nordrhein-Westfalen, and the remaining 4 from Baden-Württemberg, Hessen and Bayern. It is highly unlikely that these sequences are representatives of the SRLV strains circulating in this country. The same applies to Iran, with only six sequences. The reason for unbalanced sequencing results from German flocks is the identification of more positive flocks in the North of Germany[49]. Although Germany is not free of SRLVs, there are not many infected flocks, especially of low susceptible breeds[46]. German Texel is a sheep breed susceptible to SRLV infection and it is mostly found in the state of Schleswig-Holstein (North of Germany)[49]. We found less/no positive flocks in other German states. The six Iranian SRLV sequences were collected in a previous study by collecting samples from 30 sheep flocks in six Iranian provinces. We found SRLV positive flocks only in three provinces of Western Azarbaijan, Kerman and Chaharmahal-Va-Bakhtiari[46]. Similar to the situation in Germany, the distribution of SRLV positive samples in these three Iranian provinces was not balanced. Finally, we decided to select two samples per each province. Therefore, performing follow up studies on characterisation of SRLVs in sheep flocks from additional regions in Germany and Iran is necessary to provide a more in-depth insight into the variability of SRLVs circulating in both countries.

Materials and Methods

Samples. For characterisation of German and Iranian SRLVs, a total of 54 DNA samples was selected from SRLV positive sheep flocks in Germany (13 flocks, n = 48 samples) and Iran (6 flocks, n = 6 samples)[49]. The SRLV infection status of German and Iranian sheep flocks were determined in earlier studies based on serological (IDEXX CAEV-MVV Total Ab ELISA, Ludwigsburg, Germany)[49] and PCR (a semi-nested PCR on *env* gene)[46] tests, respectively. Origin of flocks and the number of sequenced samples from each flock in this study are shown in Table 1.

Ethical approval. No live animals were used for this study. All blood samples were collected by trained veterinarians exclusively as part of a routine clinical examination and the residuals of the samples would have otherwise been discarded. According to the German regulations for animal protection, the German Animal Welfare Act (released on 8/5/2006, modified on 17/12/ 2018), this origin of samples obviates the need for an explicit ethics committee approval. Iran Veterinary Organization (IVO), a national authority, participated in collecting Iranian sheep samples (permission issued on 20/12/2014, No: 93/22/70521). The ethical responsibility for Iranian animal welfare was carried out in accordance with the legal regulations of the IVO.

PCR amplification and sequencing of SRLV proviral DNA. To characterise the SRLVs found in Germany and Iran, genomic DNA of sampled sheep was analysed using a nested PCR targeting the *gag-pol* region of the provirus as described and utilised elsewhere[8,9,50].

For sequencing purposes, PCR products of the second PCR were purified and directly sequenced with either ABI PRISM 3130 Genetic Analyzer or by LGC Genomics GmbH, Germany (https://shop.lgcgenomics.com). To control PCR errors, both the sense and antisense strands were sequenced performing three independent reactions for each sample.

Analyses of SRLV sequence data. The obtained SRLV provirus nucleotide sequences were trimmed and analysed using the ChromasPro software version 2.1.6 (Technelysium Pty Ltd, Tewantin, Australia). Recombination analyses were carried out using RDP4 package[59]. Multiple alignments of the nucleotide sequences and the deduced amino acid sequences were generated with Muscle-built in Mega version 7.0.26[60].

For simplicity, only one or two sequences (Table 1) of each flock were selected and used for all analyses. The available sequences of the reference SRLV strains of genotypes A–C and E, and the SRLV strains most homologues to German and Iranian SRLVs (using the Basic Local Alignment Search Tool, BLAST) were included in the analysis. The sequence data of SRLV subtypes A6, A10, A15 were not available (a pol fragment defined these

subtypes) for the analyses. Additionally, A14 had to be excluded from analyses because of the shortness of the relevant sequence part. All analyses were carried out based on 65 aligned sequences from various sources (Germany 17, Iran 6, Jordan 2, Lebanon 1 and 39 from reference SRLV subtypes).

The evolutionary relationships of German and Iranian SRLVs with other published sequences were investigated by constructing the phylogenetic tree from alignments of the *gag* region. The best-fitting nucleotide substitution model for phylogenetic analysis was the Tamura-Nei (TN93)[61] model, with the gamma distribution (G) and invariant sites (I), and it was estimated using MEGA version 7.0.26[60]. The phylogeny was estimated using the maximum likelihood (ML) method. The percentage of bootstrap values were obtained based on 1,000 repetitions.

Genetic divergences between sequences were estimated with the *p*-distance model applying the gamma distribution parameter using Mega version 7.0.26[60]. Descriptive statistics of genetic distances were done with the SPSS program version 25.0 for Windows (IBM SPSS Statistics, Armonk, NY: IBM Corp). A cut-off value for assigning new subtypes within genotype A was determined according to the *gag* fragment (399 bp). This was done by calculating the mean genetic distances among SRLV subtypes of A1–A20 (except for A6, A10, A14 and A15).

The deducted *gag-pol* amino acid sequences (214 aa) of 6 Iranian and 17 German *gag-pol* sequences were aligned with all available sequence data. The purposes of this alignment were: (1) Checking amino acid variations of epitopes 2 and 3, especially the N-terminus and C-terminus, as such substitutions may affect the sensitivity of ELISA. (2) Checking amino acid variations, with regard to epitopes 2 and 3, and other conserved motifs (GG and MHR), to be sure whether our sequence data have been genotyped correctly. Previous studies have shown some specific differences between different genotypes of A-C and E[9,15]. (3) Finding potential amino acid variation/s in our sequence data that presumably is linked to other lentiviruses of the family Retroviridae.

Data availability

All data generated and analysed in this study are included in this published article and its supplementary information files.

Received: 22 August 2019; Accepted: 21 January 2020;
Published online: 10 February 2020

References

1. Narayan, O. *et al*. The lentiviruses of sheep and goats in *The Retroviridae 2* (ed. Levy, J.A.) 229–255 (Springer, Boston, MA, 1993).
2. Peterhans, E. *et al*. Routes of transmission and consequences of small ruminant lentiviruses (SRLVs) infection and eradication schemes. *Vet. Res.* **35**, 257–274, https://doi.org/10.1051/vetres:2004014 (2004).
3. Pépin, M., Vitu, C., Russo, P., Mornex, J.-F. & Peterhans, E. Maedi-visna virus infection in sheep: a review. *Vet. Res.* **29**, 341–367, https://hal.archives-ouvertes.fr/hal-00902532 (1998).
4. Shah, C. *et al*. Phylogenetic analysis and reclassification of caprine and ovine lentiviruses based on 104 new isolates: evidence for regular sheep-to-goat transmission and worldwide propagation through livestock trade. *Virology* **319**, 12–26, https://doi.org/10.1016/j.virol.2003.09.047 (2004).
5. Reina, R. *et al*. Prevention strategies against small ruminant lentiviruses: An update. *Vet. J.* **182**, 31–37, https://doi.org/10.1016/j.tvjl.2008.05.008 (2009).
6. Ramírez, H., Reina, R., Amorena, B., de Andrés, D. & Martínez, H. A. Small ruminant Lentiviruses: Genetic variability, tropism and diagnosis. *Viruses* **5**, 1175–1207, https://doi.org/10.3390/v5041175 (2013).
7. Minardi da Cruz, J. S., Singh, D., Lamara, A. & Chebloune, Y. Small ruminant lentiviruses (SRLVs) break the species barrier to acquire new host range. *Viruses* **5**, 1867–1884, https://doi.org/10.3390/v5071867 (2013).
8. Giammarioli, M. *et al*. Phylogenetic analysis of small ruminant lentivirus (SRLV) in Italian flocks reveals the existence of novel genetic subtypes. *Virus Genes* **43**, 380–384, https://doi.org/10.1007/s11262-011-0653-1 (2011).
9. Grego, E. *et al*. Genetic characterization of small ruminant lentivirus in Italian mixed flocks: evidence for a novel genotype circulating in a local goat population. *J. Gen. Virol.* **88**, 3423–3427, https://doi.org/10.1099/vir.0.83292-0 (2007).
10. Kuhar, U., Barlič-Maganja, D. & Grom, J. Phylogenetic analysis of small ruminant lentiviruses detected in Slovenia. *Vet. Microbiol.* **162**, 201–206, https://doi.org/10.1016/j.vetmic.2012.08.024 (2013).
11. Olech, M., Rachid, A., Croisé, B., Kuźmak, J. & Valas, S. Genetic and antigenic characterization of small ruminant lentiviruses circulating in Poland. *Virus Res.* **163**, 528–536, https://doi.org/10.1016/j.virusres.2011.11.019 (2012).
12. Olech, M., Valas, S. & Kuźmak, J. Epidemiological survey in single-species flocks from Poland reveals expanded genetic and antigenic diversity of small ruminant lentiviruses. *PLoS One* **13**, e0193892, https://doi.org/10.1371/journal.pone.0193892 (2018).
13. Olech, M., Murawski, M. & Kuźmak, J. Molecular analysis of small-ruminant lentiviruses in Polish flocks reveals the existence of a novel subtype in sheep. *Arch. Virol.* **164**, 1193–1198, https://doi.org/10.1007/s00705-019-04161-9 (2019).
14. Pisoni, G. *et al*. Genetic analysis of small ruminant lentiviruses following lactogenic transmission. *Virology* **407**, 91–99, https://doi.org/10.1016/j.virol.2010.08.004 (2010).
15. Colitti, B. *et al*. A new approach for Small Ruminant Lentivirus full genome characterization revealed the circulation of divergent strains. *PLoS One* **14**, e0212585, https://doi.org/10.1371/journal.pone.0212585 (2019).
16. Bertolotti, L. *et al*. Characterization of new small ruminant lentivirus subtype B3 suggests animal trade within the Mediterranean Basin. *J. Gen. Virol.* **92**, 1923–1929, https://doi.org/10.1099/vir.0.032334-0 (2011).
17. Gjerset, B., Storset, A. K. & Rimstad, E. Genetic diversity of small-ruminant lentiviruses: Characterization of Norwegian isolates of Caprine arthritis encephalitis virus. *J. Gen. Virol.* **87**, 573–580, https://doi.org/10.1099/vir.0.81201-0 (2006).
18. Gjerset, B., Rimstad, E., Teige, J., Soetaert, K. & Jonassen, C. M. Impact of natural sheep-goat transmission on detection and control of small ruminant lentivirus group C infections. *Vet. Microbiol.* **135**, 231–238, https://doi.org/10.1016/j.vetmic.2008.09.069 (2009).
19. Reina, R. *et al*. Molecular characterization and phylogenetic study of Maedi Visna and Caprine Arthritis Encephalitis viral sequences in sheep and goats from Spain. *Virus Res.* **121**, 189–198, https://doi.org/10.1016/j.virusres.2006.05.011 (2006).
20. Reina, R. *et al*. Small ruminant lentivirus genotype E is widespread in Sarda goat. *Vet. Microbiol.* **144**, 24–31, https://doi.org/10.1016/j.vetmic.2009.12.020 (2010).
21. Gifford, R. J. Viral evolution in deep time: lentiviruses and mammals. *Trends Genet.* **28**, 89–100, https://doi.org/10.1016/j.tig.2011.11.003 (2012).
22. Zhang, C., De Silva, S., Wang, J.-H. & Wu, L. Co-evolution of primate SAMHD1 and lentivirus Vpx leads to the loss of the vpx gene in HIV-1 ancestor. *PLoS One* **7**, e37477, https://doi.org/10.1371/journal.pone.0037477 (2012).
23. Yamada, E. *et al*. A naturally occurring bovine APOBEC3 confers resistance to bovine lentiviruses: implication for the co-evolution of bovids and their lentiviruses. *Sci. Rep.* **6**, 33988, https://doi.org/10.1038/srep33988 (2016).

24. Takeuchi, J. S. *et al.* Coevolutionary dynamics between tribe Cercopithecini tetherins and their lentiviruses. *Sci. Rep.* **5**, 16021, https://doi.org/10.1038/srep16021 (2015).
25. Worobey, M. *et al.* Island biogeography reveals the deep history of SIV. *Science* **329**, 1487, https://doi.org/10.1126/science.1193550 (2010).
26. Garcia-Etxebarria, K., Sistiaga-Poveda, M. & Marina Jugo, B. Endogenous retroviruses in domestic animals. *Curr Genomics* **15**, 256–265 (2014).
27. Chessa, B. *et al.* Revealing the history of sheep domestication using retrovirus integrations. *Science* **324**, 532–536, https://doi.org/10.1126/science.1170587 (2009).
28. Zeder, M. A. Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proc. Natl. Acad. Sci.* **105**, 11597–11604, https://doi.org/10.1073/pnas.0801317105 (2008).
29. Zeder, M. A. Animal domestication in the Zagros: a review of past and current research. *Paléorient* **25**, 11–25, http://www.jstor.org/stable/41496540 (1999).
30. Zeder, M. A. & Hesse, B. The initial domestication of goats (Capra hircus) in the Zagros Mountains 10,000 years ago. *Science* **287**, 2254–2257, https://doi.org/10.1126/science.287.5461.2254 (2000).
31. Pereira, F. & Amorim, A. Origin and spread of goat pastoralism. *e LS;* https://doi.org/10.1002/9780470015902.a0022864 (2001).
32. Zeder, M. A. Out of the Fertile Crescent: The dispersal of domestic livestock through Europe and Africa in *Hum. Dispersal Species Mov. From Prehistory to* Present (ed. Boivin, N., Crassard, R. & Petraglia, M.) 261–296 (Cambridge University Press, 2017)
33. Vigne, J. D. Zooarchaeology and the biogeographical history of the mammals of Corsica and Sardinia since the last ice age. *Mamm. Rev.* **22**, 87–96, https://doi.org/10.1111/j.1365-2907.1992.tb00124.x (1992).
34. Greenfield, H. J. *et al.* The origins of milk and wool production in the Old World: a zooarchaeological perspective from the Central Balkans [and comments]. *Curr. Anthropol.* **29**, 573–593 (1988).
35. Kaewthamasorn, M. *et al.* Genetic homogeneity of goat malaria parasites in Asia and Africa suggests their expansion with domestic goat host. *Sci. Rep.* **8**, 5827, https://doi.org/10.1038/s41598-018-24048-0 (2018).
36. Mühlemann, B. *et al.* Ancient human parvovirus B19 in Eurasia reveals its long-term association with humans. *Proc. Natl. Acad. Sci.* **115**, 7557–7562, https://doi.org/10.1073/pnas.1804921115 (2018).
37. Otchere, I. D. *et al.* Comparative genomics of Mycobacterium africanum Lineage 5 and Lineage 6 from Ghana suggests distinct ecological niches. *Sci. Rep.* **8**, 11269, https://doi.org/10.1038/s41598-018-29620-2 (2018).
38. Sigurdsson, B., Grímsson, H. & Pálsson, P. A. Maedi, a chronic, progressive infection of sheep's lungs. *J. Infect. Dis.* **90**, 233–241 (1952).
39. Blacklaws, B. A. *et al.* Transmission of small ruminant lentiviruses. *Vet. Microbiol.* **101**, 199–208, https://doi.org/10.1016/j.vetmic.2004.04.006 (2004).
40. Thormar, H. The origin of lentivirus research: Maedi-visna virus. *Curr. HIV Res.* **11**, 2–9, https://doi.org/10.2174/157016213804999212 (2013).
41. Straub, O. C. Maedi–Visna virus infection in sheep. History and present knowledge. *Comp. Immunol. Microbiol. Infect. Dis.* **27**, 1–5, https://doi.org/10.1016/S0147-9571(02)00078-4 (2004).
42. Sayari, M. & Soltanian, S. Serological and pathological study of Maedi-like diseases in mammary glands of sheep of Ahwaz region. *Pajouhesh Sazandegi (in Persian)*, 2–7 (2001).
43. Azizi, S. *et al.* Maedi in slaughtered sheep: A pathology and polymerase chain reaction study in southwestern Iran. *Trop. Anim. Health Prod.* **44**, 113–118, https://doi.org/10.1007/s11250-011-9896-z (2012).
44. Norouzi, B., Razavizadeh, A. T., Azizzadeh, M., Mayameei, A. & Mashhadi, V. N. N. Serological study of small ruminant lentiviruses in sheep population of Khorasan-e-Razavi province in Iran. *Veterinary Research Forum*, **6**, 245–249, PMC4611980 (2015).
45. Sasani, F., Javanbakht, J., Hemmatzadeh, F., Moghadam, M. R. & Hassan, M. A. M. Evaluation of histopathological on maedi disease with serological confirmation in North-East of Iran. *Res. J. Infect. Dis.* **1**, 5, https://doi.org/10.7243/2052-5958-1-5 (2013).
46. Molaee, V., Otarod, V., Abdollahi, D. & Lühken, G. Lentivirus susceptibility in Iranian and German sheep assessed by determination of TMEM154 E35K. *Animals.* **9**, 685, https://doi.org/10.3390/ani9090685 (2019).
47. Alberto, F. J. *et al.* Convergent genomic signatures of domestication in sheep and goats. *Nat. Commun.* **9**, 813, https://doi.org/10.1038/s41467-018-03206-y (2018).
48. L'Homme, Y. *et al.* Molecular characterization and phylogenetic analysis of small ruminant lentiviruses isolated from Canadian sheep and goats. *Virol. J.* **8**, 271, https://doi.org/10.1186/1743-422X-8-271 (2011).
49. Molaee, V., Eltanany, M. & Lühken, G. First survey on association of TMEM154 and CCR5 variants with serological maedi-visna status of sheep in German flocks. *Vet. Res.* **49**, 36, https://doi.org/10.1186/s13567-018-0533-y (2018).
50. Muz, D. *et al.* First molecular characterization of visna/maedi viruses from naturally infected sheep in Turkey. *Arch. Virol.* **158**, 559–570, https://doi.org/10.1007/s00705-012-1518-1 (2013).
51. Hiendleder, S., Mainz, K., Plante, Y. & Lewalski, H. Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different ancestral maternal sources: no evidence for contributions from urial and argali sheep. *J. Hered.* **89**, 113–120, https://doi.org/10.1093/jhered/89.2.113 (1998).
52. Hiendleder, S., Kaupe, B., Wassmuth, R. & Janke, A. Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proc. R. Soc. London. Ser. B Biol. Sci.* **269**, 893–904, https://doi.org/10.1098/rspb.2002.1975 (2002).
53. Pedrosa, S. *et al.* Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. *Proc. R. Soc. B Biol. Sci.* **272**, 2211–2217, https://doi.org/10.1098/rspb.2005.3204 (2005).
54. Meadows, J. R. S., Cemal, I., Karaca, O., Gootwine, E. & Kijas, J. W. Five ovine mitochondrial lineages identified from sheep breeds of the near East. *Genetics* **175**, 1371–1379, https://doi.org/10.1534/genetics.106.068353 (2007).
55. Meadows, J. R. S., Hiendleder, S. & Kijas, J. W. Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. *Heredity (Edinb).* **106**, 700, https://doi.org/10.1038/hdy.2010.122 (2011).
56. Bruford, M. W., Bradley, D. G. & Luikart, G. DNA markers reveal the complexity of livestock domestication. *Nat. Rev. Genet.* **4**, 900, https://doi.org/10.1038/nrg1203 (2003).
57. Harlan, J. R. & Zohary, D. Distribution of wild wheats and barley. *Science.* **153**, 1074–1080, https://doi.org/10.1126/science.153.3740.1074 (1966).
58. Sunderman, F. & Johns, M. Awassi fat tails: a chance for premium exports. *J. Dep. Agric. West. Aust. Ser. 4* **35**, 99–105 (1994).
59. Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.*, **1**; https://doi.org/10.1093/ve/vev003 (2015).
60. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874, https://doi.org/10.1093/molbev/msw054 (2016).
61. Tamura, K. & Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512–526, https://doi.org/10.1093/oxfordjournals.molbev.a040023 (1993).
62. Staskus, K. A. *et al.* Isolation of replication-competent molecular clones of visna virus. *Virology* **181**, 228–240, https://doi.org/10.1016/0042-6822(91)90488-W (1991).

Acknowledgements
We want to thank Renate Latzke-Reinhardt and Lena Vogel from the Institute of Animal Breeding and Genetics, Justus Liebig University Giessen, Germany, for their excellent technical assistance. We also want to thank technical assistants of the Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche Togo Rosati, Perugia, Italy, who participated in our study. We are grateful to thank those sheep farmers who shared samples for this study.

Author contributions
G.L. designed the research; G.L., V.O. and D.A. contributed to the collection of samples; V.M., M.B., G.M.D.M., G.L. and D.A. performed experiments; V.M. analysed data; V.M. generated figures; V.M. wrote the manuscript; G.L., S.R., M.B., G.M.D.M. and V.O. reviewed the manuscript; G.L. Supervision; All authors contributed to the final approval of the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-58990-9.

Correspondence and requests for materials should be addressed to V.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Phylogenetic analysis of small ruminant lentiviruses in Germany and Iran suggests their expansion with domestic sheep

Vahid Molaee[1,*], Moira Bazzucchi[2], Gian Mario De Mia[2], Vahid Otarod[3], Darab Abdollahi[4], Sergio Rosati[5] & Gesine Lühken[1]

[1]Institute of Animal Breeding and Genetics, Justus Liebig University Giessen (JLU), Ludwigstraße 21, 35390 Gießen, Germany

[2]Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche Togo Rosati (IZSUM), Via G. Salvemini 1, 06126 Perugia, Italy

[3]Quarantine and Biosafety Directorate General, Iran Veterinary Organization (IVO), Vali Asr Avenue, Seyd Jamaledin Asad Abadi Street, 6349 Tehran, Iran

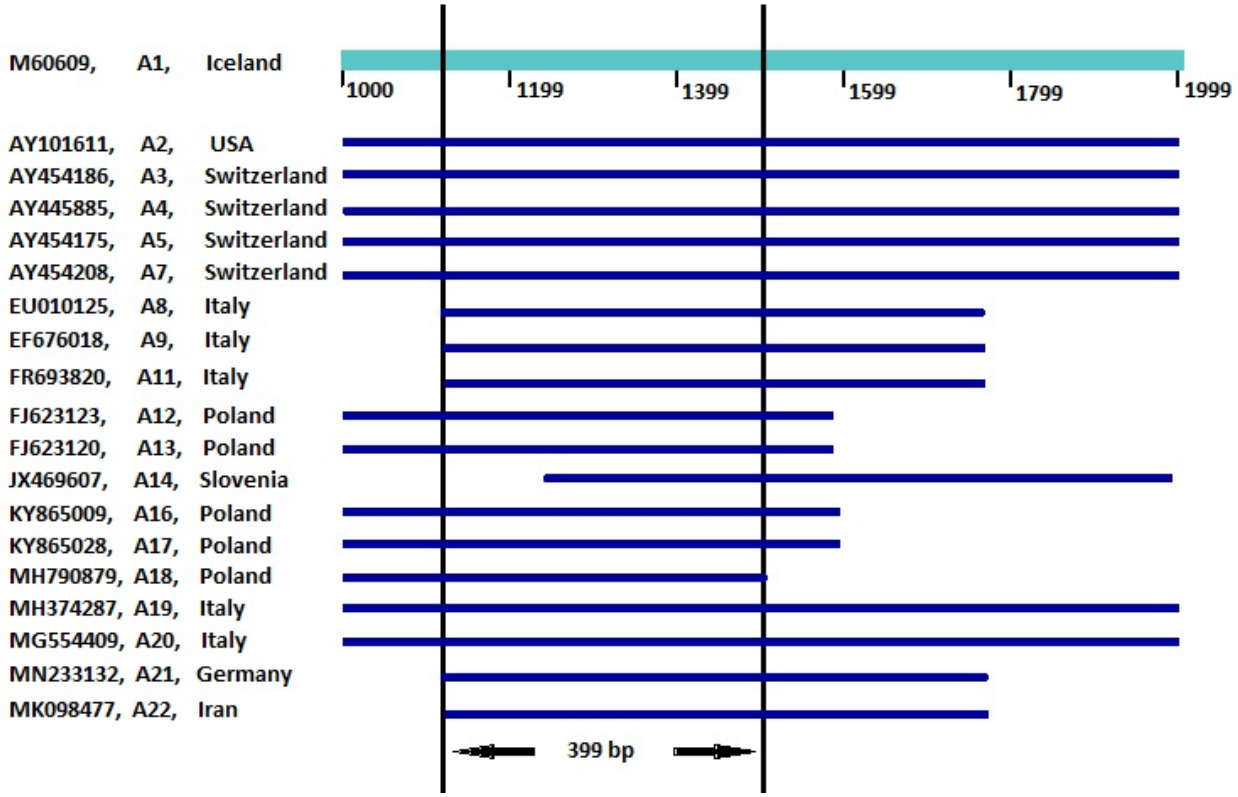
[4]Bureau of Animal Health and Disease Management, Iran Veterinary Organization (IVO), Vali Asr Avenue, Seyd Jamaledin Asad Abadi Street, 6349 Tehran, Iran

[5]Department of Veterinary Science, University of Turin (UNITO), Largo Paolo Braccini 2, 10095 Grugliasco, Torino, Italy

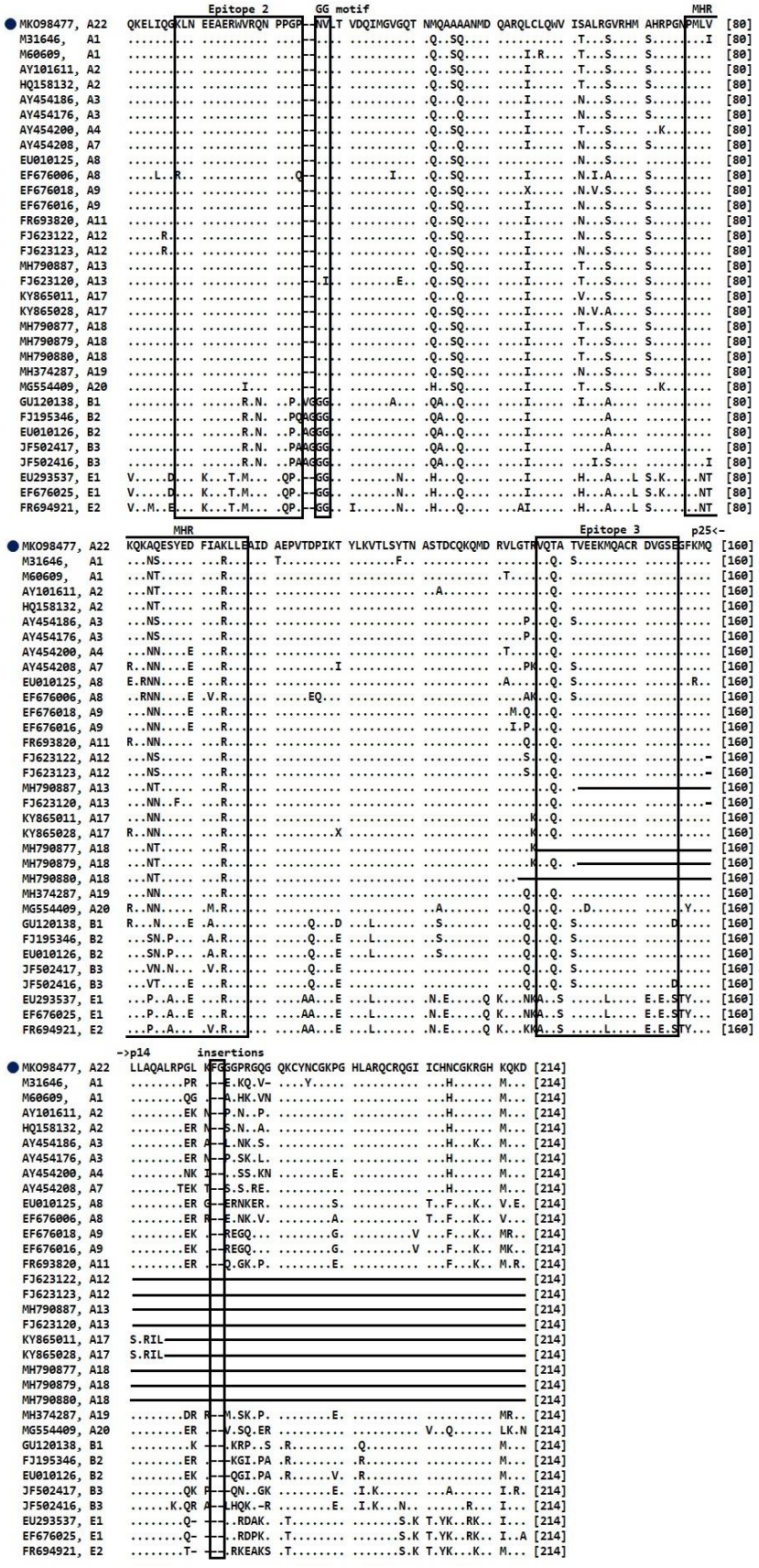
Supplementary Table S1. Estimates of evolutionary divergences of German and Iranian SRLV sequences compared to different subtypes of genotype B, C and E based on the *gag* fragment (nucleotide: 1114–1506; numbering according to prototype strain K1514[62]).

sequence	country	B1	B2	B3	mean B	C	E1	E2	mean E
MN233148	Germany	24.17	23.16	22.52	23.28	23.92	28.63	31.68	30.16
MN233105	Germany	23.16	22.52	22.01	22.56	23.16	27.23	28.5	27.87
MN233108	Germany	23.92	22.52	22.65	23.03	21.88	28.12	29.13	28.63
MN233107	Germany	22.90	22.26	20.99	22.05	23.16	26.72	28.24	27.48
MN233143	Germany	22.39	21.37	20.10	21.29	23.66	27.99	27.10	27.55
MN233124	Germany	21.63	21.63	19.97	21.08	24.43	26.59	26.84	26.72
MN233151	Germany	21.63	21.63	19.97	21.08	23.92	26.59	26.34	26.47
MN233104	Germany	23.54	19.47	20.99	21.33	20.87	25.70	27.35	26.53
MN233116	Germany	25.57	24.30	21.77	23.88	23.92	28.63	30.28	29.46
MN233111	Germany	23.54	24.68	21.25	23.16	24.68	28.63	30.28	29.46
MN233118	Germany	24.17	24.43	21.12	23.24	25.45	27.23	27.61	27.42
MN233132	Germany	25.19	25.70	21.63	24.17	25.95	26.59	29.39	27.99
MN233120	Germany	22.14	23.16	20.74	22.01	21.88	26.97	29.64	28.31
MN233145	Germany	23.92	22.39	21.50	22.60	22.65	26.59	28.63	27.61
MN233128	Germany	23.92	22.52	22.26	22.90	22.39	26.84	28.88	27.86
MN233141	Germany	24.30	24.68	22.26	23.75	22.90	28.12	30.41	29.27
MN233106	Germany	22.52	22.77	21.63	22.31	22.65	29.01	29.26	29.14
MK098477	Iran	25.45	24.43	21.50	23.79	24.68	27.23	28.88	28.06
MK098478	Iran	25.83	24.94	23.54	24.77	25.45	29.01	30.66	29.84
MK098479	Iran	23.92	22.90	20.10	22.31	22.65	27.35	30.28	28.82
MK098480	Iran	24.30	23.92	19.21	22.48	23.66	27.99	31.04	29.52
MK098481	Iran	25.19	25.19	21.37	23.92	23.66	25.83	29.90	27.87
MK098482	Iran	24.05	24.05	22.26	23.45	24.68	27.10	29.39	28.25

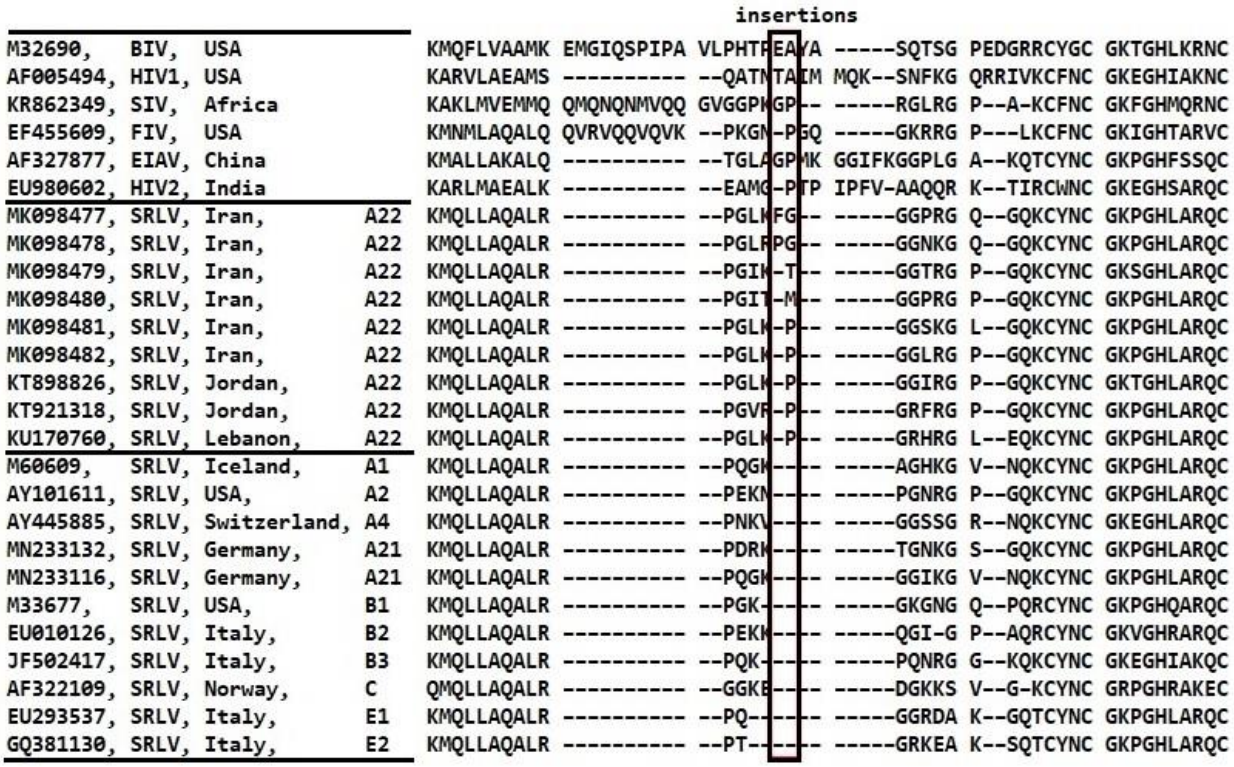
For each reference subtype up to two strains were selected and the evolutionary divergences calculated based on the mean divergence of each set of subtypes and representative sequences of this study. 17 German and six Iranian SRLV sequences were compared with subtypes/genotypes B1 (M33677 and GU120138), B2 (FJ195346 and EU010126), B3 (JF502417 and JF502416), C (AF322109), E1 (EU293537 and EF676025), and E2 (FR694921 and GQ381130).



Supplementary Figure S1. Status of SRLV *gag-pol* fragment sequence data derived from previous studies and from the present study. The SRLV subtype A14 had to be excluded from analyses because of the shortness of the relevant sequence part. The sequence data available for the SRLV subtypes A12, A13, A16, A17 and A18 were not match with the sequence data of other SRLV subtypes of genotype A.


Epitope 2
GG motif
MHR
MHR
Epitope 3
p25<-
->p14
insertions

Supplementary Figure S2 related to Figure 2. Amino acid sequence alignments of the SRLV *gag-pol* fragment (not shown at the main text). The reference sequence is the SRLV sequence of strain BKH1 (accession number: MK098477) from the Iranian province of Chaharmahal-Va-Bakhtiari. Immunodominant epitopes 2 and 3, the major homology region (MHR), the double glycine motif (GG) and insertions (at positions 172 or 173) are delineated with boxes. The major core protein (p25) and nucleic acid-binding protein (p14) are separated with left and right arrows (p25 ≤ position160; p14 ≥ position161). Dashes and dots indicate deletions and identical residues, respectively.



Supplementary Figure S3. Insertions (black box) on *gag-pol* fragment (amino acid at positions 172 or 173; numbering according to prototype strain BKH1, accession number: MK098477) were shown in different lentiviruses: bovine immunodeficiency virus (BIV), human immunodeficiency virus type 1 (HIV1), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV), equine infectious anemia virus (EIAV), human immunodeficiency virus type 2 (HIV2) and small ruminant lentiviruses (SRLVs).

4.1 Research topics

The overall purpose of the present dissertation was to enhance our knowledge regarding host-pathogen genetics and the epidemiology of lentivirus infection in German and Iranian sheep. Previous studies have shown the existence of SRLV infection among German and Iranian sheep (Straub, 2004; Azizi et al., 2012; Sasani et al., 2013; Norouzi et al., 2015). While the effects of SRLV infection among sheep and goats are economically vast (Peterhans et al., 2004), no efforts have been made to investigate the genetic susceptibility of German or Iranian sheep to SRLV infection. Additionally, no genetic knowledge regarding German and Iranian SRLVs is available so far. Two main issues (i and ii) were addressed in the present dissertation: i) Evaluation of the candidate gene(s) of the host associated with SRLV infection and ii) evaluation of the genetic characteristics of the pathogen. The issue i was evaluated using German (studies 1 and 2) and Iranian (study 2) sheep, while issue ii was evaluated using German and Iranian SRLVs (study 3). Moreover, general knowledge regarding the epidemiology of SRLV infection in German and Iranian sheep was provided in studies 1 and 2. However, due to the specific goals of each study, this part will not be discussed intensively. In the current dissertation, additional tables (Tables 2 and 3) were provided as the results have been published in two consecutive studies (1 and 2) and we aimed to provide a comprehensive picture of these studies.

4.2 Issue i

4.2.1 Organisation of sampling

The samples collected for studies 1–3 belonged to 11 German sheep breeds/mixed (772 samples) and 8 Iranian sheep breeds (365 samples). Additionally, 302 samples belonging to pure German sheep breeds were available from previous studies, without ELISA data. Sheep belonging to 7 of the 11 German sheep breeds/mixed and 4 of the 8 Iranian sheep breeds

were found to be positive for SRLV infection. Number, organisation of sampling, and genetic background of positive flocks are given in Table 2.

Table 2. Overview of sampling in studies 1 and 2.

no	country	breed	no. flocks	no pos flocks	Study 1	Study 2	Total
1	Germany	Texel (pure/crossbred)	16+1	16	353	39	392
2	Germany	Merinoland (pure/crossbred)	1+1	2	125	49	174
3	Germany	German Blackheaded Mutton	1	1	39		39
4	Germany	Suffolk	1	0	32		32
5	Germany	Rhoen sheep	1	0	25		25
6	Germany	East Friesian Milk	3	1	25		25
7	Germany	Rough-coated Pomeranian Landrace	1	0	18		18
8	Germany	Lacaune	1	1	11		11
9	Germany	Brown Hair	1	1		25	25
10	Germany	Zackel	1	1		14	14
11	Germany	Mix of several breeds	1	0	17		17
Total	Germany		30	23	645	127	772
1	Iran	Makouee	9	4		74	74
2	Iran	Qezel	2	1		26	26
3	Iran	Moghani	1	0		21	21
4	Iran	Bakhtiari	3	3		50	50
5	Iran	Kaboudeh	4	0		51	51
6	Iran	Kermani	3	2		46	46
7	Iran	Balouchi	4	0		45	45
8	Iran	Karakul	4	0		52	52
Total	Iran		30	10		365	365

4.2.2 Evaluating the association of TMEM154 and CCR5 variants with SRLV infection status

At first, we examined two previously identified genetic markers (*TMEM154* and *CCR5*) against lentivirus infection in four sample sets over almost 500 German sheep (study 1). Broadly, the results of study 1 showed that *TMEM154* could be a suitable marker for investigating genetic susceptibility among German sheep. On the contrary, the *CCR5* promoter's deletion did not show a consistent association with serological SRLV infection status. Given the results of study 1, we decided to limit the genotyping of samples to

TMEM154 for study 2.

Not all sheep breeds with SRLV positive flocks were suitable for the association analysis. In study 2, we excluded the association analysis of two German sheep breeds (Texel and Zackel) and two Iranian sheep breeds (Bakhtiari and Qezel). Altogether, in studies 1 and 2, we performed association analyses (chi-squared test) on six German sheep breeds/sets and two Iranian sheep breeds. According to studies 1 and 2, among 6 German sheep breeds/sets, the association of *TMEM154* with serological SRLV status was confirmed in all sample sets except purebred German Blackheaded Mutton sheep (GBM). Among Iranian sheep samples, the association was tested only in two sheep breeds, and it was significant in Kermani sheep (study 2). Status of association analyses of *TMEM154* with SRLV infection status in studies 1 and 2 are shown in Table 3.

Both the German Texel and Merinoland sheep are crucial to Germany's sheep industry, and we collected the majority of samples from these two breeds. We tested the association of *TMEM154* with SRLV infection status in samples belonging to these two sheep breeds from different regions, both in studies 1 and 2. The association of *TMEM154* with SRLV infection status in German Texel (n=343 samples) was strongly significant (study 1). However, in study 2, evaluating the association of *TMEM154* with SRLV infection in German Texel was not possible, as all the samples were SRLV positive. However, according to study 1 (chi-squared test and non-parametric analyses) and study 2 (regression analysis), it could be concluded that in Texel sheep, SRLV susceptibility is associated with *TMEM154* variants. For Merinoland sheep, we could evaluate the association of *TMEM154* and SRLV infection in both study 1 and 2 (see Table 3). In study 1, the association between *TMEM154* variants and SRLV infection status in Merinoland sheep was borderline significant, while in study 2,

it was significant. However, unlike German Texel, in the Merinoland breed, we repeatedly observed sheep with *TMEM154* KK protective genotype (44% study 1; 55% study 2) to be serologically SRLV positive. Therefore, as suggested, in studies 1 and 2, a follow-up GWAS study could help find potential additional variants/markers related to SRLV infection status in Merinoland sheep.

Table 3. Association analyses of *TMEM154* with SRLV infection status in German and Iranian sheep breeds/mixed.

Study 1				Study 2			
Breed subset	**Country**	**(n sheep)**	***P* value**	**Breed subset**	**Country**	**(n sheep)**	***P* value**
Texel (pure/cross bred)	Germany	341	<0.001	Makouee	Iran	35	0.402
German Blackheaded Mutton	Germany	39	0.277	Kermani	Iran	36	0.038
Merinoland (pure/cross bred)	Germany	125	0.067	Merinoland	Germany	49	<0001
EFM–LAC	Germany	22	0.006	Brown Hair	Germany	25	0.0311
Total		527				145	

4.2.3 SRLV susceptibility in German and Iranian sheep based on the frequency of allele E

In studies 1 and 2, the relationship between *TMEM154* variants and SRLV infection status was tested in only eight out of 12 data sets. Specifically, four sets of data were not suitable for the association test. In particular, the chi-squared test for Qezel was prevented because of low infected sheep (5%), and in Bakhtiari, German Texel, and Zackel flocks/breeds due to

the lack of KK genotypes (Molaee et al., 2019). We noted that in Bakhtiari, German Texel, and Zackel breeds, the high frequency of risk allele was associated with high prevalence, and correspondingly, in Qezel, the low frequency of risk allele was associated with low prevalence. Unlike the chi-squared test, regression analysis showed the association between E allele frequency and SRLV infection status in all data sets, without any prerequisites. Accordingly, an exciting conclusion was obtained to predict SRLV susceptibility based on the frequency of *TMEM154* E allele. The regression analysis showed that in a flock or breed, SRLV susceptibility could be predicted based on allele E frequency with a variation of approximately +/– 20%. In this respect, an SRLV prevalence of about 30% is expected in infected flocks with an E frequency of 20%, while an SRLV prevalence of about 60% is expected in sheep flocks with an E frequency of 70% (Molaee et al., 2019). With regard to the obtained results of association analyses (*TMEM154* with SRLV infection) and regression analysis, in study 2, the degree of susceptibility to SRLV infection in 13 German and eight Iranian sheep breeds was shown based on the frequency of the *TMEM154* E allele. Before study 2, in a US study (Heaton et al., 2013), a genetic susceptibility comparison was made for 74 global sheep populations based on the frequency of allele c *TMEM154*. In study 2, with the example of the Iranian Qezel breed, we were able to show that data on the frequency of the c allele, which is available for a high number of international sheep breeds, can result in a wrong assessment of the SRLV susceptibility of a breed. Thus, comparing SRLV susceptibility based on allele E compared to allele c provide more valuable information on genetic susceptibility to SRLV infection.

4.2.4 The epidemiology of SRLV infection in Germany and Iran

One of the aims of studies 1 and 2 was to provide epidemiological data on the SRLV infection status in Germany and Iran. The epidemiological data were not intensively discussed in studies 1 and 2 because of other specific objectives. Sampling from German sheep flocks was mostly performed using a step-by-step approach. Sheep flocks already known to be infected or suspected to be infected with MV were selected for sampling. Only a few samples were initially collected from these flocks. After the diagnosis was made using these samples, more samples of a flock were collected if the results were positive. Thus, as the flocks were not randomly selected, an accurate epidemiological assessment could not be performed to indicate the prevalence of SRLVs in German sheep. However, it is still valuable to see the percentage of SRLV prevalence in German positive flocks. Additionally, due to the lack of such data, studies 1 and 2 have provided basic knowledge about the distribution of SRLV infection in German states/regions. For example, the majority of positive flocks were detected from Schleswig-Holstein, a state in the north of Germany.

A step-by-step approach was not possible for collecting Iranian sheep samples as we could not diagnose the presence of SRLV infection in these samples in Iran. Blood samples were collected in a single step, mostly from Iran's speculated provinces, without knowing the SRLV infection status of the flocks. The observed SRLV infection status of Iranian sheep flocks in different provinces was consistent with the findings of earlier studies (Azizi et al., 2012; Sasani et al., 2013; Norouzi et al., 2015). Among Iranian sheep flocks, one-third of flocks (10 out of 30 flocks) were SRLV positive and had at least one sample that was SRLV positive. The range of infection within samples of individual flocks varied from 0–89% (Molaee et al., 2019). Interestingly, for the first time, a PCR-based test performed in study 2 revealed SRLV-positive sheep flocks in the Iranian province of Western Azarbaijan.

4.3 Issue ii

For the characterisation of German and Iranian SRLVs, 54 DNA samples were selected from SRLV positive sheep flocks. Positive samples were detected by diagnostic tests performed in studies 1 and 2. The initial goal of issue ii was to provide the phylogenetic classification of German and Iranian SRLVs for the first time. By investigating German and Iranian SRLVs, we discovered two new SRLVs, namely A21 in German and A22 in Iranian sheep. In German flocks also SRLV subtypes A4, A5, A11 and A16 were found. Based on phylogenetic analysis and genetic distance analyses, a noticeable genetic difference was observed between the Iranian SRLVs and other known SRLVs. At the phylogenetic level, a clear difference was noted in genotype A. The phylogeny always showed two distinct clusters of genotype A, even with different models or methods of analyses. In addition, the mean genetic distance between Iranian SRLVs and other SRLVs of genotype A was greater than that observed between other SRLVs of genotype A. Based on these initial observations, a search for clues related to the evolution of these SRLVs was necessary. We hypothesised the following three scenarios regarding Iranian SRLVs: (1) They are recombinant; (2) they belong to genotype A and are not distinctly different from other SRLVs of this genotype and (3) they belong to genotype A; however, they are evolutionarily different from other SRLVs of genotype A.

To find the most probable scenario, different items were considered using a step-by-step approach. First, the database was monitored to find the sequence data most similar to those of Iranian SRLVs. Two SRLV sequence data from Jordan and one from Lebanon were found to be very similar to the Iranian SRLV sequence data. Recombination analysis was then performed and revealed no evidence of recombination for SRLVs belonging to Iran, Jordan and Lebanon. Therefore, scenario 1 was not considered valid.

Assessment of the amino acid alignment revealed that the insertion of amino acid 172 or 173 is typical for this group of SRLVs (e.g. subtype A22) and other lentiviruses but not for all other SRLVs. Hence, scenario 2 was also not considered valid.

As only scenario 3 remained, we tried to define it more precisely. We aimed to determine whether the existence of SRLVs is linked to the year 1933 (the time of importation of Karakul sheep from Germany to Iceland) or is associated with a long duration (10000 years ago), as proposed by Muz et al. (2013).

Our results showed that the evolution and presence of SRLVs date back to ancient times, as suggested by Muz et al. (2013). First, we found no evidence regarding the presence of the Icelandic subtype A1 among 48 SRLV *gag* sequences of German sheep. Therefore, our results did not support the hypothesis about the maedi-visna disease being imported from Germany to Iceland in 1933 with the Karakul flock. Second, phylogenetic analysis revealed two distinct clades of genotype A, consisting of the European SRLVs and another subtype (A22). These distinct clades likely indicate the presence of two different ancestors for genotype A. One cluster of SRLVs belong to Europe, the ancestor of which was proposed to be Turkish SRLVs. The other cluster of SRLVs belongs to Iran, Jordan and Lebanon. Iranian samples came from three provinces close to the Zagros Mountains, known as primary goat and sheep domestication centres. The Jordanian isolates' hosts were sheep of the Awassi breed, a fat-tailed sheep with a history related to long times (about 5,000 years ago). The mere presence of a unique variant of SRLVs (insertion at positions 172 or 173) among sheep from the countries of the ancient Fertile Crescent suggests that subtype A22 (including Iranian, Jordanian and Lebanon strains) may represent an ancient precursor of modern SRLVs.

Accordingly, the most interesting conclusion of study 3 was that the phylogenetic classification of some SRLV sequences fits the domestication pathways of sheep in Europe and the Middle East.

Our results show that the SRLV subtype A5 has likely been carried from Turkey to Germany through the Danube river domestication pathway. Furthermore, the SRLV subtype A22 has likely travelled from Iran to Jordan and Lebanon through the southern Mediterranean domestication pathway. Distribution of the SRLV subtype A11, from Turkey to Italy, is likely to be related to the northern Mediterranean pathway. We have also given examples of other SRLVs that could be associated with sheep domestication pathways (subtypes A3, A9, B3, and genotype E) as well as those not related to domestication pathways (A16, A18 and A21).

4.4 Limitations

For studies 1 and 2, all collected samples were genotyped for *TMEM154* E35K by an allele-specific PCR method (Molaee et al., 2018, 2019). Like other genotyping methods, this may not be 100% accurate, for haplotype analyses derived from direct sequencing results. In a Turkish study, it was shown that rare E35-containing haplotypes could present with lower susceptibility, recessively, to SRLV (i.e., haplotype "4" together with haplotype "16" according to the study of Yaman et al., (2019). Therefore, the findings of studies 1 and 2 could be more precise by using direct sequencing (*TMEM154* haplotype data), rather than by genotyping with allele-specific PCR method. For study 2, different diagnostic testing was used to evaluate individual infection status. The maedi-visna status of the Iranian (set 1) and German (set 2) sheep was determined by PCR and ELISA testing, respectively. Due to animal health requirements, transferring serum samples from Iran to Germany was not

possible. For the Iranian sheep sample data, a semi-nested PCR test that detects the virus was performed versus the indirect ELISA technique that monitors immune response used for the German sheep samples. Considering the known pitfalls of diagnosing the SRLV status of small ruminants, study 2 may have its limitations. Herrmann-Hoesing reviewed that PCR tests compared with indirect ELISA tests have 92–95% positive concordance, and 87–88% negative concordance (Herrmann-Hoesing, 2010). In another review (Ramírez et al. 2013), it

was shown that PCR and serological assays (ELISA and AGID) have 70–95% positive concordance, and 87–100% negative concordance (Ramírez et al., 2013). Based on these evaluations (Herrmann-Hoesing, 2010; Ramírez et al., 2013), we do not believe the SRLV diagnosis test used would have substantially affected the results.

The principal limitation of study 3 is the restricted number of sequences obtained. For the present studies (1–3), over 1000 samples were collected from 60 flocks (30 from Germany and 30 from Iran). Although the sampling was relatively comprehensive, the distribution of SRLV positive samples among flocks was imbalanced. Depending on the genetic background (breed) of the sheep and on the geographical region, we found more positive samples in some flocks and fewer or no positive samples in others. We were able to sequence 54 samples and, eventually, based on this sequence data, the conclusions of study 3 were obtained. We think that performing follow-up studies on the characterisation of SRLVs in other German and Iranian sheep flocks is necessary so that a clearer picture of SRLV variability in both countries can be obtained.

4.5 Suggestions for future research

Firstly; where the *TMEM154* does not provide significant results with SRLV infection, other candidate genes in sheep associated with SRLV infection status including *ZNF389* (White et al., 2014), *MHC* class ii (Herrmann-Hoesing et al., 2008; Larruskain et al., 2010), and *APOBEC3* (Larruskain and Jugo, 2013) should be investigated.

Secondly, in study 3, it was shown that genotype A of SRLVs potentially has two ancestors, the Iranian clade and the Turkish clade. In a follow-up study, the *env* fragment of German and Iranian SRLVs could be investigated, to evaluate the situation of the phylogeny of SRLVs and to confirm the existence of two clades for genotype A.

Thirdly, the results of study 3 showed that, due to the Danube domestication pathway (Zeder, 2017), further investigation of domestication in German sheep samples is warranted. In a follow-up study, the Danube domestication pathway could be investigated comprehensively using endogenous lentiviruses (Chessa et al., 2009), with German and Turkish sheep samples. Sheep samples from other domestication sites, as well as sheep samples from those countries lying between Germany and Turkey, are recommended.

Finally, establishment of a real-time test for quantification and diagnosis of German SRLVs is strongly recommended. While quantification of SRLVs provides quantified phenotypic data for the analyses of candidate genes, high sensitivity and repeatability of a real-time PCR test compared to the conventional PCR or ELISA test (Herrmann-Hoesing et al., 2007; Herrmann-Hoesing, 2010) make a precise diagnosis more likely through a real-time PCR test. Accordingly, the establishment of a reliable and practical protocol for the diagnosis of German SRLVs, using one of the diagnostic tests (ELISA or nested-PCR or real-time PCR), or combination of two or three tests, could be evaluated and the results reported for practical use.

De Andres, D., D. Klein, N.J. Watt, E. Berriatua, S. Torsteinsdottir, B.A. Blacklaws, and G.D. Harkiss. 2005. Diagnostic tests for small ruminant lentiviruses. Vet. Microbiol. 107:49–62.

Azizi, S., E. Tajbakhsh, F. Fathi, A. Oryan, H. Momtaz, and M. Goodarzi. 2012. Maedi in slaughtered sheep: A pathology and polymerase chain reaction study in southwestern Iran. Trop. Anim. Health Prod. 44:113–118.

Benavides, J., N. Gomez, D. Gelmetti, M.C. Ferreras, C. Garcia-Pariente, M. Fuertes, J.F. Garcia-Marin, and V. Perez. 2006. Diagnosis of the nervous form of Maedi-Visna infection with a high frequency in sheep in Castilla y Leon, Spain. Vet. Rec. 158:230–235.

Bertolotti, L., M. Mazzei, G. Puggioni, M.L. Carrozza, S. Dei Giudici, D. Muz, M. Juganaru, C. Patta, F. Tolari, and S. Rosati. 2011. Characterization of new small ruminant lentivirus subtype B3 suggests animal trade within the Mediterranean Basin. J. Gen. Virol. 92:1923–1929.

Blacklaws, B. 2012. Small ruminant lentiviruses: immunopathogenesis of visna-maedi and caprine arthritis and encephalitis virus. Comp. Immunol. Microbiol. Infect. Dis. 35:259–269.

Blacklaws, B., E. Berriatua, S. Torsteinsdottir, N.J. Watt, D. De Andres, D. Klein, and G.D. Harkiss. 2004. Transmission of small ruminant lentiviruses. Vet. Microbiol. 101:199–208.

Chessa, B., F. Pereira, F. Arnaud, A. Amorim, F. Goyache, I. Mainland, R.R. Kao, J.M. Pemberton, D. Beraldi, and M.J. Stear. 2009. Revealing the history of sheep domestication using retrovirus integrations. Science. 324:532–536.

Colitti, B., E. Coradduzza, G. Puggioni, M.T. Capucchio, R. Reina, L. Bertolotti, and S. Rosati. 2019. A new approach for Small Ruminant Lentivirus full genome

characterization revealed the circulation of divergent strains. PLoS One. 14:e0212585.

Extramiana, A.B., L. González, N. Cortabarrıa, M. Garcıa, and R.A. Juste. 2002. Evaluation of a PCR technique for the detection of Maedi-Visna proviral DNA in blood, milk and tissue samples of naturally infected sheep. Small Rumin. Res. 44:109–118.

Gifford, R.J. 2012. Viral evolution in deep time: lentiviruses and mammals. Trends Genet. 28:89–100.

Gomez-Lucia, E., N. Barquero, and A. Domenech. 2018. Maedi-Visna virus: current perspectives. Vet. Med. Res. Reports 9:11.

Hamza, L.A., and C. Özkan. 2017. Serological investigation of Maedi-Visna in sheep with chronic respiratory disease in Erbil, Iraq. Atatürk Üniv Vet Bil Derg 12:227–234.

Heaton, M.P., M.L. Clawson, C.G. Chitko-Mckown, K.A. Leymaster, T.P.L. Smith, G.P. Harhay, S.N. White, L.M. Herrmann-Hoesing, M.R. Mousel, G.S. Lewis, T.S. Kalbfleisch, J.E. Keen, and W.W. Laegreid. 2012. Reduced Lentivirus Susceptibility in Sheep with TMEM154 Mutations. PLoS Genet. 8:e1002467.

Heaton, M.P., T.S. Kalbfleisch, D.T. Petrik, B. Simpson, J.W. Kijas, M.L. Clawson, C.G. Chitko-McKown, G.P. Harhay, and K.A. Leymaster. 2013. Genetic Testing for TMEM154 Mutations Associated with Lentivirus Susceptibility in Sheep. PLoS One 8:e55490.

Herrmann-Hoesing, L.M. 2010. Diagnostic assays used to control small ruminant lentiviruses. J. Vet. diagnostic Investig. 22:843–855.

Herrmann-Hoesing, L.M., S.N. White, G.S. Lewis, M.R. Mousel, and D.P. Knowles. 2007. Development and Validation of an Ovine Progressive Pneumonia Virus Quantitative PCR. Clin. Vaccine Immunol. 14:1274–1278.

Herrmann-Hoesing, L.M., S.N. White, M.R. Mousel, G.S. Lewis, and D.P. Knowles. 2008.

Ovine progressive pneumonia provirus levels associate with breed and Ovar-DRB1. Immunogenetics 60:749–758.

Kaewthamasorn, M., M. Takeda, T. Saiwichai, J.N. Gitaka, S. Tiawsirisup, Y. Imasato, E. Mossaad, A. Sarani, W. Kaewlamun, and M. Channumsin. 2018. Genetic homogeneity of goat malaria parasites in Asia and Africa suggests their expansion with domestic goat host. Sci. Rep. 8:5827.

Kuhar, U., D. Barlič-Maganja, and J. Grom. 2013. Phylogenetic analysis of small ruminant lentiviruses detected in Slovenia. Vet. Microbiol. 162:201–206.

Larruskain, A., and B. Jugo. 2013. Retroviral Infections in Sheep and Goats: Small Ruminant Lentiviruses and Host Interaction. Viruses 5:2043–2061.

Larruskain, A., E. Minguijón, K. García-Etxebarria, B. Moreno, I. Arostegui, R.A. Juste, and B.M. Jugo. 2010. MHC class II DRB1 gene polymorphism in the pathogenesis of Maedi–Visna and pulmonary adenocarcinoma viral diseases in sheep. Immunogenetics 62:75–83.

Leymaster, K.A., C.G. Chitko-McKown, M.L. Clawson, G.P. Harhay, and M.P. Heaton. 2013. Effects of TMEM154 haplotypes 1 and 3 on susceptibility to ovine progressive pneumonia virus following natural exposure in sheep1,2,3. J. Anim. Sci. 91:5114–5121.

Mikula Jr, I., and I. Mikula Sr. 2011. Characterization of ovine Toll-like receptor 9 protein coding region, comparative analysis, detection of mutations and maedi visna infection. Dev. Comp. Immunol. 35:182–192.

Minardi da Cruz, J., D. Singh, A. Lamara, and Y. Chebloune. 2013. Small ruminant lentiviruses (SRLVs) break the species barrier to acquire new host range. Viruses 5:1867–1884.

Molaee, V., M. Bazzucchi, G. De Mia, V. Otarod, D. Abdollahi, S. Rosati, and G. Lühken. 2020.

phylogenetic analysis of small ruminant lentiviruses in Germany and iran suggests their expansion with domestic sheep. Sci. Rep. 10:1–12.

Molaee, V., M. Eltanany, and G. Lühken. 2018. First survey on association of TMEM154 and CCR5 variants with serological maedi-visna status of sheep in German flocks. Vet. Res.

Molaee, V., V. Otarod, D. Abdollahi, and G. Lühken. 2019. Lentivirus susceptibility in Iranian and German sheep assessed by determination of TMEM154 E35K. Animals. 9:685

Mühlemann, B., A. Margaryan, P. de Barros Damgaard, M.E. Allentoft, L. Vinner, A.J. Hansen, A. Weber, V.I. Bazaliiskii, M. Molak, and J. Arneborg. 2018. Ancient human parvovirus B19 in Eurasia reveals its long-term association with humans. Proc. Natl. Acad. Sci. 115:7557–7562.

Muz, D., T.Ç. Oğuzoğlu, S. Rosati, R. Reina, L. Bertolotti, and I. Burgu. 2013. First molecular characterization of visna/maedi viruses from naturally infected sheep in Turkey. Arch. Virol. 158:559–570.

Norouzi, B., A.T. Razavizadeh, M. Azizzadeh, A. Mayameei, and V.N.N. Mashhadi. 2015. Serological study of small ruminant lentiviruses in sheep population of Khorasan-e-Razavi province in Iran. Veterinary Research Forum. 6:245-249.

Olech, M., M. Murawski, and J. Kuźmak. 2019. Molecular analysis of small-ruminant lentiviruses in Polish flocks reveals the existence of a novel subtype in sheep. Arch. Virol. 164:1193–1198.

Olech, M., S. Valas, and J. Kuźmak. 2018. Epidemiological survey in single-species flocks from Poland reveals expanded genetic and antigenic diversity of small ruminant lentiviruses. PLoS One 13:e0193892.

Otchere, I.D., M. Coscollá, L. Sánchez-Busó, A. Asante-Poku, D. Brites, C. Loiseau, C.

Meehan, S. Osei-Wusu, A. Forson, and C. Laryea. 2018. Comparative genomics of Mycobacterium africanum Lineage 5 and Lineage 6 from Ghana suggests distinct ecological niches. Sci. Rep. 8:11269.

Pépin, M., C. Vitu, P. Russo, J.-F. Mornex, and E. Peterhans. 1998. Maedi-visna virus infection in sheep: a review. Vet. Res. 29:341–367.

Peterhans, E., T. Greenland, J. Badiola, G. Harkiss, G. Bertoni, B. Amorena, M. Eliaszewicz, R.A. Juste, R. Kraßnig, J.-P. Lafont, P. Lenihan, G. Pétursson, G. Pritchard, J. Thorley, C. Vitu, J.-F. Mornex, and M. Pépin. 2004. Routes of transmission and consequences of small ruminant lentiviruses (SRLVs) infection and eradication schemes. Vet. Res. 35:257–274.

Ramírez, H., R. Reina, B. Amorena, D. de Andrés, and H.A. Martínez. 2013. Small Ruminant Lentiviruses: Genetic Variability, Tropism and Diagnosis. Viruses. 5:1175-1207.

Reina, R., E. Berriatua, L. Luján, R. Juste, A. Sánchez, D. de Andrés, and B. Amorena. 2009. Prevention strategies against small ruminant lentiviruses: An update. Vet. J. 182:31–37.

Reina, R., L. Bertolotti, S. Dei Giudici, G. Puggioni, N. Ponti, M. Profiti, C. Patta, and S. Rosati. 2010. Small ruminant lentivirus genotype E is widespread in Sarda goat. Vet. Microbiol. 144:24–31.

Reina, R., M.I. Mora, I. Glaria, I. García, C. Solano, L. Lujan, J.J. Badiola, A. Contreras, E. Berriatua, and R. Juste. 2006. Molecular characterization and phylogenetic study of Maedi Visna and Caprine Arthritis Encephalitis viral sequences in sheep and goats from Spain. Virus Res. 121:189–198.

Sarafidou, T., C. Stamatis, G. Kalozoumi, V. Spyrou, G.C. Fthenakis, C. Billinis, and Z. Mamuris. 2013. Toll like receptor 9 (TLR9) polymorphism G520R in sheep is associated with seropositivity for small ruminant lentivirus. PLoS One 8:e63901.

Sasani, F., J. Javanbakht, F. Hemmatzadeh, M.R. Moghadam, and M.A.M. Hassan. 2013. Evaluation of histopathological on maedi disease with serological confirmation in North- East of Iran. Res. J. Infect. Dis. 1:1-5.

Sayari, M., and S. Soltanian. 2001. Serological and pathological study of Maedi-like diseases in mammary glands of sheep of Ahwaz region. Pajouhesh Sazandegi (in Persian) 2–7.

Shah, C., J. Böni, J.B. Huder, H.-R. Vogt, J. Mühlherr, R. Zanoni, R. Miserez, H. Lutz, and J. Schüpbach. 2004. Phylogenetic analysis and reclassification of caprine and ovine lentiviruses based on 104 new isolates: evidence for regular sheep-to-goat transmission and worldwide propagation through livestock trade. Virology 319:12–26.

Sigurdsson, B., H. Grímsson, and P.A. Pálsson. 1952. Maedi, a chronic, progressive infection of sheep's lungs. J. Infect. Dis. 90:233–241.

Straub, O.C. 2004. Maedi–Visna virus infection in sheep. History and present knowledge. Comp. Immunol. Microbiol. Infect. Dis. 27:1–5.

Thormar, H. 2005. Maedi-visna virus and its relationship to human immunodeficiency virus. AIDS Rev. 7:233–245.

White, S., and D. Knowles. 2013. Expanding Possibilities for Intervention against Small Ruminant Lentiviruses through Genetic Marker-Assisted Selective Breeding. Viruses 5:1466–1499.

White, S.N., M.R. Mousel, J.O. Reynolds, L.M. Herrmann-Hoesing, and D.P. Knowles. 2014. Deletion variant near ZNF 389 is associated with control of ovine lentivirus in multiple sheep flocks. Anim. Genet. 45:297–300.

Zeder, M.A. 1999. Animal domestication in the Zagros: a review of past and current research. Paléorient. 25: 11–25.

Zeder, M.A. 2008. Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. Proc. Natl. Acad. Sci. 105:11597–11604.

Zeder, M.A. 2017. Out of the Fertile Crescent: The dispersal of domestic livestock through Europe and Africa. Hum. Dispersal Species Mov. From Prehistory to Present 261-296.

Zeder, M.A., and B. Hesse. 2000. The initial domestication of goats (Capra hircus) in the Zagros Mountains 10,000 years ago. Science. 287:2254–2257.

I declare: this dissertation submitted is a work of my own, written without any illegitimate help by any third party and only with materials indicated in the dissertation. I have indicated in the text where I have used texts from already published sources, either word for word or in substance, and where I have made statements based on oral information given to me. At any time during the investigations carried out by me and described in the dissertation, I followed the principles of good scientific practice as defined in the "Statutes of the Justus Liebig University Giessen for the Safeguarding of Good Scientific Practice".

Vahid Molaee

Gießen, January 2021