

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Virology 319 (2004) 12–26

VIROLOGY

[www.elsevier.com/locate/yviro](http://www.elsevier.com/locate/yviro)

# 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

Cyril Shah,<sup>a</sup> Jürg Böni,<sup>a</sup> Jon B. Huder,<sup>a</sup> Hans-Rudolf Vogt,<sup>b</sup> Janine Mühlherr,<sup>c</sup>  
Reto Zanoni,<sup>b</sup> Raymond Miserez,<sup>d</sup> Hans Lutz,<sup>e</sup> and Jörg Schüpbach<sup>a,\*</sup>

<sup>a</sup>Swiss National Center for Retroviruses, University of Zurich, Zurich, Switzerland

<sup>b</sup>Institute for Veterinary Virology, University of Berne, Bern, Switzerland

<sup>c</sup>Extension and Health Service for Small Ruminants, Herzogenbuchsee, Switzerland

<sup>d</sup>Swiss Federal Veterinary Office, Bern, Switzerland

<sup>e</sup>Veterinary Laboratory, Department of Veterinary Medicine, University of Zurich, Zurich, Switzerland

Received 21 August 2003; returned to author for revision 25 September 2003; accepted 29 September 2003

## Abstract

We performed a phylogenetic analysis of caprine and ovine lentiviruses using long sequences in *gag* and *pol* of 104 new Swiss isolates and six available corresponding database sequences. Forty-five isolates, forming five sequence clusters, were unclassifiable by the present classification. Pairwise DNA distance analysis indicated different categories of relatedness, requiring a new classification system. We propose four principal sequence groups, A–D, which differ by 25–37%. Groups A and B are further divided into subtypes which differ by 15–27%. Group D and four of the seven group A subtypes, A3, A4, A5 and A7, are formed by new Swiss isolates. Molecular epidemiology revealed that Swiss B1 strains differed no more from French, Brazilian or US strains than from each other, suggesting virus propagation through international livestock trade. Furthermore, infection of goats by subtypes A3 or A4 was significantly associated with documented contact with sheep, which also harbor these subtypes, thus indicating regularly occurring sheep-to-goat transmission.

© 2003 Elsevier Inc. All rights reserved.

**Keywords:** Caprine arthritis-encephalitis virus (CAEV); Maedi-visna virus (MVV); Small ruminant lentiviruses (SRLV); Goat; Sheep; Molecular epidemiology; Sequence analysis; Phylogenetic analysis; Switzerland

## Introduction

Caprine arthritis-encephalitis virus (CAEV) and maedi-visna virus (MVV) are related members of the group of small ruminant lentiviruses (SRLV) which infect goats and sheep (Narayan et al., 1980; Pasick, 1998; Pepin et al., 1998; Sundquist, 1981). The major route of transmission of these viruses is through the ingestion of virus-infected milk (De Boer et al., 1979; East et al., 1987; Ellis et al., 1983; Greenwood et al., 1995; McGuire et al., 1990; Rowe and East, 1997); less efficient routes of transmission are associated with prolonged close contact with infected animals, especially with the maedi form of MVV infection, in which

respiratory exudates may lead to virus transmission (Narayan and Cork, 1985). In sheep, the symptoms of MVV infection frequently include dyspnea, weight loss, mastitis and arthritis (Pepin et al., 1998). CAEV infection of goats in newborn kids may lead to encephalitis and in adult animals to chronic arthritis and interstitial mastitis. Both manifestations may contribute to impaired milk production (Kennedy-Stoskopf et al., 1985; Krieg and Peterhans, 1990). The widespread distribution of these viruses in certain regions and the resulting economic losses have led to segregation-based virus eradication programs for both goats and sheep in several countries (Cutlip and Lehmkuhl, 1986; Greenwood et al., 1995; Houwers, 1980; Houwers et al., 1987; Peretz et al., 1994; Rowe and East, 1997; Rowe et al., 1992; Scheer-Czechowski et al., 2000).

In Switzerland, a CAEV eradication program was started on a voluntary basis in the early 1980s, and since 1998, is mandatory for all cantons of the country. Within less than 15

\* Corresponding author. Swiss National Center for Retroviruses, University of Zurich, Gloriastrasse 30, CH-8028 Zurich, Switzerland. Fax: +41-1-634-4965.

E-mail address: [jschupb@immv.unizh.ch](mailto:jschupb@immv.unizh.ch) (J. Schüpbach).

years, the program has reduced the prevalence of CAEV, which was 83% in 1989, to about 1%. Detailed information on the program and its results is available online ([http://www.bvet.admin.ch/0\\_navigation-e/0\\_index-intern.html](http://www.bvet.admin.ch/0_navigation-e/0_index-intern.html)). The program provides annual serological testing for all goats. Seropositive animals are to be culled, and a herd in which a seropositive animal was detected is put under strict quarantine which is maintained until at least three consecutive annual serological assays have yielded negative results. When these criteria are met, farms are declared CAEV-free and may again engage in purchase or sales relations.

Despite the undisputed successes of the eradication program, goat owners, veterinarians and authorities alike were puzzled time and again by the reemergence of new cases of seropositivity in farms that had been CAEV-free for many years and had always maintained good adherence to the program's guidelines. Because delayed seroconversion has been reported for both goats (Rimstad et al., 1993; Wagter et al., 1998) and sheep (Johnson et al., 1992; Krassnig and Schuller, 1998) and because animals which are seropositive only intermittently or lack detectable antibodies entirely have also been described (Hanson et al., 1996; Johnson et al., 1992), low-grade infection leading to delayed diagnosis was one hypothesis to explain the reappearance of seroconversions in apparently CAEV-free herds. An alternative hypothesis involved new infection of truly CAEV-free herds from unidentified sources, notably sheep, which are not included in the program. Owners of newly infected goats had occasionally reported contact of their animals with sheep.

For elucidation of the two hypotheses molecular amplification and characterization of viral sequences were necessary. PCR for proviral DNA in blood cells (Barlough et al., 1994; Celer et al., 2000; Ding, 1995; Haase et al., 1990; Johnson et al., 1992; Reddy et al., 1993; Wagter et al., 1998) or RT-PCR for viral RNA in the milk (Leroux et al., 1997b) have been used, but the diagnostic sensitivity of these methods appears to be hampered by the high degree of sequence diversity and low copy numbers (Leroux et al., 1995, 1997a; Zanoni, 1998; Zanoni et al., 1992, 1996; Zhang et al., 2000). To overcome these obstacles it is necessary to develop PCR methods able to detect all the different virus variants even if present at low concentration. However, although the high degree of genomic variation is well acknowledged, little sequence information on virus variants circulating in different geographical regions is available. Only four full-length sequences, three from MVV (Braun et al., 1987; Querat et al., 1990; Sonigo et al., 1985) and one from CAEV (Saltarelli et al., 1990), were published when this work was started in 1998. Phylogenetic work, which to date has identified six different clusters of SRLV numbered I–VI, was based on relatively short sequences within *gag*, *pol* or *env* (Rolland et al., 2002; Zanoni, 1998). These short sequences are not necessarily optimal for diagnostic purposes.

The goal of this study was to therefore establish a sequence database of large fragments in the more conserved sequences of the viral genome which should be representa-

tive of the SRLV strains now circulating in goats in Switzerland and also give some first crude information on the viruses present in sheep. This new sequence information should help identify the cause of the reappearance of seroconversions among CAEV-free herds. It can also serve for phylogenetic analysis and may finally permit identification of broadly conserved sequences suitable for sensitive molecular diagnosis of these infections.

## Results

### *Development of an amplification system for long SRLV sequences*

The primary goal of this work was to amplify proviral sequences of sufficient length to be representative of the proviral sequences circulating in the country. Based on the four full-length sequences published at the start of this work in 1998, namely the three MVV isolates EV1 from Scotland (Sonigo et al., 1985), K1514 from Iceland (Braun et al., 1987) and SA-OMVV from South Africa (Querat et al., 1990), the single full-length CAEV sequence CAEV<sub>Cork</sub> from the US (Saltarelli et al., 1990) and a 2.5-kb viral sequence recently isolated from the milk of a seropositive goat (J. Huder et al., unpublished), conserved regions within *gag* and *pol* suitable for positioning PCR amplification primers were identified. Candidate primer pairs were then assessed using the pCAEV (9 kb) plasmid representing positions 21–8661 of CAEV<sub>CO</sub> (37) and six standardized field isolates from which a fragment of 1.5 kb had previously been amplified and sequenced and in which the DNA copy numbers had been determined by endpoint titration. These standardized specimens were tested at a dilution exhibiting 25–80 copies/ $\mu$ l. When results were not satisfactory, new primers were designed and evaluated in the same way. Altogether, a total of 41 primers were evaluated in this manner. Eleven primers were finally chosen which were used in two nested-PCR procedures which permitted, in two rounds of amplification, isolation of a 1756-bp fragment in *gag-pol* (hence called 1.8 kb *gag-pol*) or of a 1175-bp fragment in *pol* (hence called 1.2 kb *pol*) (Fig. 1). To minimize amplification failures due to sequence variability, in the first PCR round the 1.8 kb *gag-pol* region was amplified with three primers and the 1.2 kb *pol* region with four primers each in a multiplex reaction, the primers being selected in a manner that all combinations between the upper and lower primers could form different primer pairs. The estimated detection limit of this nested PCR was 1–4 proviral copies/reaction, as determined by assessing the above-mentioned pCAEV (9 kb) plasmid (data not shown).

### *Nested PCR amplification of study specimens*

To obtain a collection of SRLV sequences representing the spectrum of sequence variation in Switzerland, blood speci-

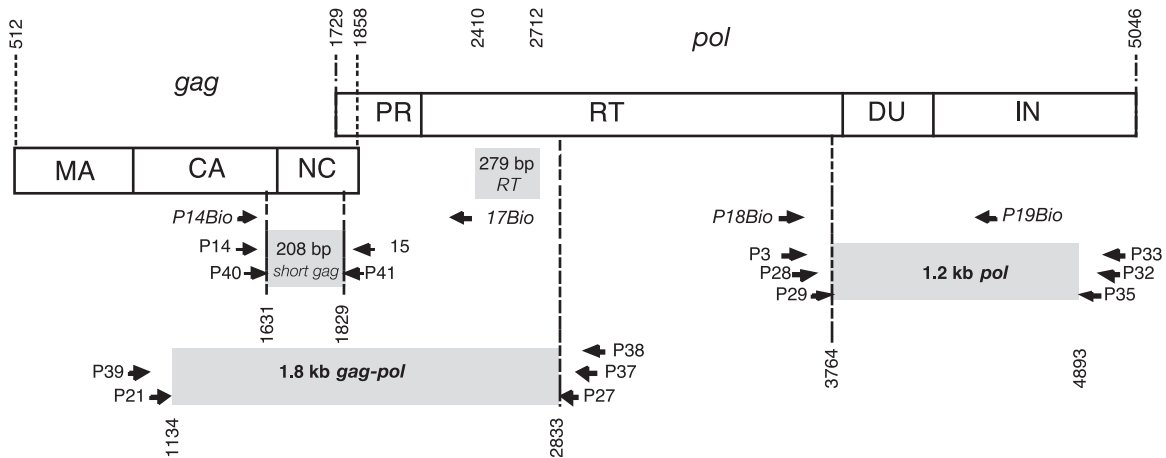


Fig. 1. Overview of primer localizations and phylogenetically analyzed sequences in the *gag* and *pol* regions of goat and sheep lentiviruses. Positions correspond to CAEV<sub>CORK</sub> (Saltarelli et al., 1990). MA, matrix; CA, capsid; NC, nucleocapsid; PR, protease; RT, reverse transcriptase; DU, dUTPase; IN, integrase.

mens were selected as described under Methods. In total, 190 animals from 115 different herds were selected (Table 1). The 169 goats included 121 animals with a Western blot confirmed positive SRLV serology, 21 seroindeterminate animals and 27 negative controls. Sheep specimens assessed included 19 Western blot confirmed seropositives and 2 seroindeterminates. The samples originated from 21 of the 23 cantons of the country. Among goats, the different breeds were represented as follows: Saanen, 30.2%; Chamois e, 29.6%; Toggenburg, 5.3%; Dwarf, 5.3%; Grisons Striped, 4.7%; Nera Verzasca, 3.5%; various other breeds, 8.3%; unknown, 13.0%. This distribution is similar to that found in the herd book (Anonymous, 1999). The 21 sheep included 16 animals of dairy breeds (Eastfriesian or Lacaune), 3 of meat breeds and two of unknown breed.

All 190 samples were initially screened for the short *gag* sequence. In the initial phase of the study, positive samples were subsequently tested for the 1.8 kb *gag-pol* and the 1.2 kb *pol* fragment. If regular PCR for these long sequences yielded a negative result, the analysis was usually repeated using the sequence capture procedure with a DNA input of 100  $\mu$ g (see Methods). At later stages of the work, some samples were, however, directly subjected to the sequence capture procedure.

A total of 39 amplifications with the 23 seroindeterminate and of 51 amplifications with the 27 seronegative samples

were negative (data not shown). Among the 140 seropositive animals, at least one of the three assessed sequences was detected in 104 animals (74.3%), namely in 91 (75.2%) of the 121 goats and in 13 (68.4%) of the 19 sheep (Table 2). The short *gag* fragment was detected most often, followed closely by the 1.8 kb *gag-pol* fragment. In contrast, in both goats and sheep, detection of the 1.2 kb *pol* fragment was significantly less frequent than that of the short *gag* fragment (chi-square 4.985,  $P < 0.05$ ). In 66 animals including 56 goats and 10 sheep, both the *gag-pol* and the *pol* sequence could be established, while in 11 animals only the 1.8 kb *gag-pol* sequence and in four animals only the 1.2 kb *pol* sequence could be isolated. Successful amplification of the short *gag* fragment was highly predictive of successful amplification of the long sequences. Of 90 specimens in which the short *gag* fragment had been positive and in which both long sequences were assessed, at least one long sequence could be established in 74 cases (82%) compared to 7 of 37 cases (19%) in which the short *gag* fragment had been negative ( $2 \times 2$  table chi-square = 42.8;  $P < 0.0001$ ).

#### Nucleic acid sequences

All *gag-pol* and all *pol* amplicons were subjected to sequencing. The short *gag* fragment was sequenced instead of *gag-pol* if the latter was not available. Comparison of

Table 1  
Overview of herds and animals tested

SRLV status of herd	No. of herds		No. of animals						
	Goats	Sheep	Goats			Sheep			
			Total	Seropositive	Seroindeterminate	Seronegative	Total	Seropositive	Seroindeterminate
SRLV-free	36	6	50	33	14	3	11	9	2
SRLV-positive	65	4	104	76	6	22	8	8	–
SRLV-unresolved	3	1	15	12	1	2	2	2	–
Total	104	11	169	121	21	27	21	19	2

Table 2  
Overview of lentivirus sequences amplified from seropositive goats and sheep

Sequence assessed	Seropositive goats		Seropositive sheep	
	<i>n</i>	%	<i>n</i>	%
<i>284 bp gag</i>				
Tested	121	100	19	100
Detectable	84	69.4	13	68.4
<i>1.8 kb gag-pol</i>				
Tested	101	100	17	100
Detectable	67	66.3	11	64.7
<i>1.2 kb pol</i>				
Tested	111	100	19	100
Detectable	61	55.0	11	57.9
<i>Any of the three</i>				
Tested	121	100	19	100
Detectable	91	75.2	13	68.4

epidemiologically unlinked sequences (including database sequences, but excluding closely related sequences amplified from different animals of a given herd) showed a median pairwise distance of 0.223 (range: 0.036–0.367) among *gag-pol* and of 0.225 (range 0.024–0.360) among *pol*. Distance within a 279-bp sequence from the RT region contained within the 1.8 kb *gag-pol* fragment (see Fig. 1), first used for phylogenetic analysis by Leroux (Leroux et al., 1997a) amounted to a median of 0.226 (range: 0.026–0.376).

#### Phylogenetic analysis of the 1.8 kb *gag-pol* fragment

Unrooted phylogenetic trees of these sequences, constructed using the distance-based F84 neighbor-joining model, are shown in Fig. 2. Analysis of the 1.8 kb *gag-pol* fragments together with all available corresponding database sequences showed that, very schematically, the SRLVs of the present investigation could be divided into two numerically dominant groups of sequences, A and B (Fig. 2, panel A). Group A contains highly divergent isolates related to prototype MVV viruses, while group B is less divergent and contains viruses related to the prototype isolate CAEV<sub>Cork</sub>. A majority of the sequences of this study, all isolated from goats of different herds, clustered on a branch-labeled B1 together with CAEV<sub>Cork</sub> (labeled M33677 Co in the figure). Located on a separate branch outside of the origin of cluster B1, as demonstrated by a bootstrap value of 100, were two closely related isolates of which only one, labeled 5720, is shown that was established from two sheep of the same flock. This branch is labeled as B2.

The highly divergent group A consisted of many branches. Branch A1, separated from the other branches by a bootstrap value of 80, contained the three MVV reference isolates SA-OMVV (labeled S31646 SA), K1514 (labeled M10608 Icl) and EV1 (labeled S51392 EV), which are included in cluster I of the classification proposed earlier (Zanoni, 1998). Located within the same cluster were also

sequences recently established from a sheep, labeled 6247, in Switzerland with the typical symptoms of visna (Braun et al., 2001) and another seropositive sheep of the same herd, which is not shown.

Sequence AY101611 from the North American ovine visna virus strain 85/34 (Mwaengo et al., 1997) (Hotzel and Cheever, unpublished), which in the current classification belongs to cluster II, formed a separate branch designated accordingly as A2. Branch A3, distinguished from A2 by a bootstrap value of 100, contained a cluster of more closely related sequences established from both goats and sheep of the present investigation. Branch A4 contained another dense cluster of sequences, which were also established from both goats and sheep and are separated from all other sequences by a bootstrap value of 100. Branch A5 consisted of a single isolate, 5560, which was established from a goat. None of the sequences in group A revealed any evidence for recombination with isolates on any other branch when investigated by the genotyping tool of the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/retroviruses/subtype/subtype.html>). Hence, all new branches were considered to have resulted from independent evolution (data not shown).

No sequences were found with similarity to a single, highly divergent isolate (AF322109 N) established from a goat in Norway which branched off in-between clusters A and B, which in the current classification belongs to cluster III and is marked here as group C.

Trees constructed using parsimony or likelihood methods also supported the differentiation and the branching order of groups A, B and C, and of the branches B1, B2, A3 and A4 shown in Fig. 2A with bootstrap values of 98 or higher. The common stem of the branches A2 and A3, which had a bootstrap value of 73 in the distance-based tree, reached bootstrap values of 70 and 76 with parsimony and likelihood methods, respectively, and was, therefore, supported at a similar level. With bootstrap values of 63, the sequence cluster A1 was less well supported by these other two tree-building methods.

#### Phylogenetic analysis of the 1.2 kb *pol* fragment

Analysis using the same models and algorithms showed that, in general, all strains for which both the *gag-pol* and the *pol* sequence were available were placed within the same sequence groups or branches (Fig. 2, panel B). The overall tree topology and bootstrap values were, therefore, very similar to those of the tree constructed from the 1.8 kb *gag-pol* sequences. As an exception, the sequence amplified from goat 5692, which in *gag-pol* was affiliated with branch A3, now formed a separate branch, A7, which originated from a similar rooting point as those of the other branches of group A but, with a bootstrap value of 76, was relatively well distinguished from the sequences forming branch A1. Analysis of this sequence by the genotyping tool of the NCBI demonstrated that the diverging position of this strain in the *pol* tree did not result from recombination

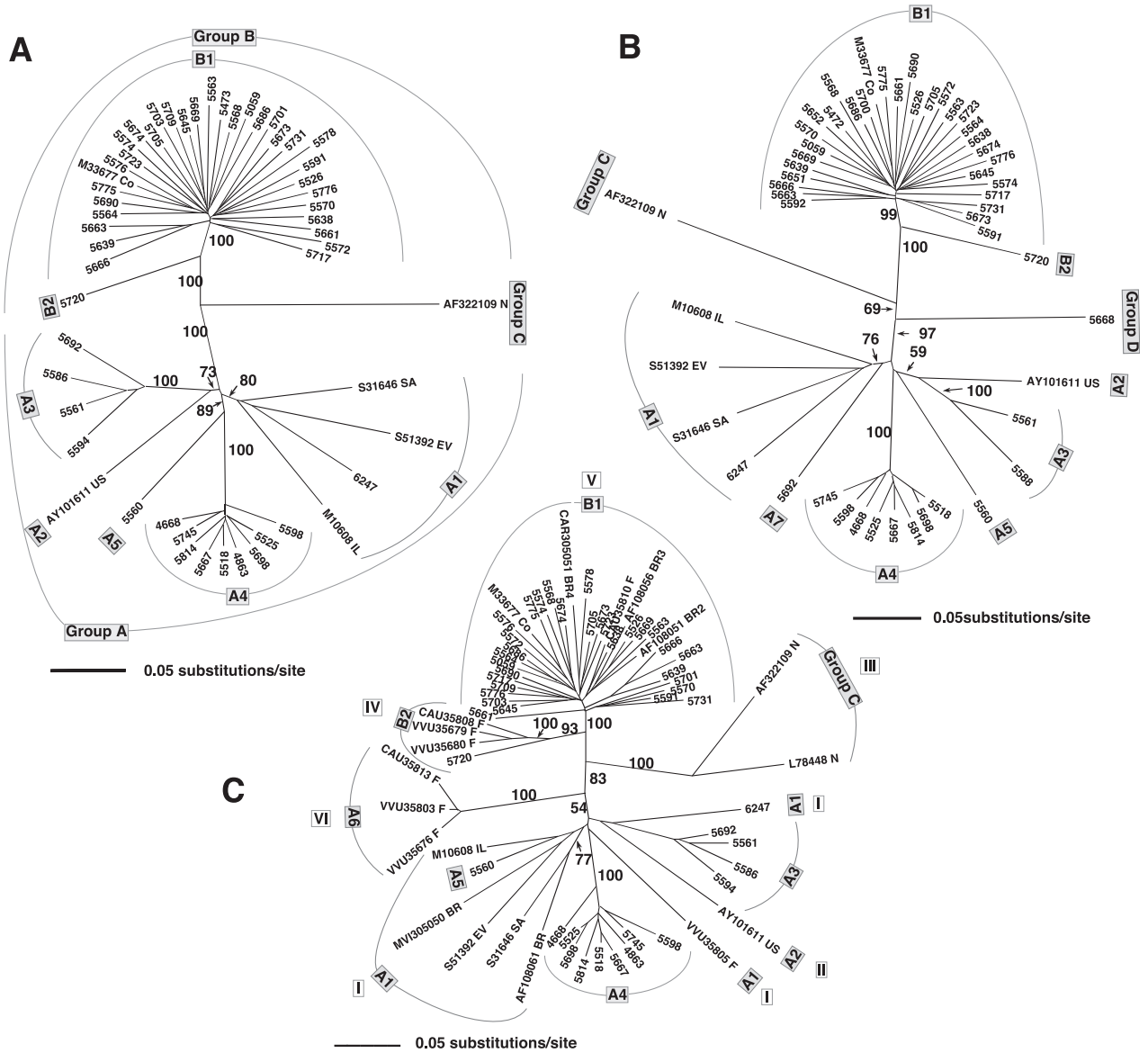


Fig. 2. Phylogenetic relationship of goat and sheep lentiviruses inferred from sequence comparisons in the 1.8 kb *gag-pol* sequence (panel A), the 1.2 kb *pol* sequence (panel B) and the 279 bp RT fragment contained within *gag-pol* (panel C). The unrooted trees were constructed by the neighbor-joining method and are based on the distances calculated with the F84 substitution model, as described under Methods. Bootstrap values are based on 1000 repetitions. Epidemiologically unlinked virus strains of the present study are shown together with available database sequences. Individual sequences are labeled with a four-digit number if they are derived from the present investigation, while database-derived sequences are denoted with their GenBank accession number and a country code (Br, Brazil; F, France; IL, Iceland; N, Norway; SA, South Africa; US, United States). S31646 SA corresponds to SA-OMVV, M10608 IL to K1514, S51392 EV to EV1 and M33677 Co to CAEV<sub>Cork</sub>. Tree branches are labeled according to the proposed new classification based on groups A–D and their subtypes (Arab numbers), where applicable. Roman numbers I–VI in panel C indicate clusters of the classification proposed earlier (Zanoni, 1998).

within this region by viruses from other branches (data not shown).

Between groups A and B, a further branch-labeled D ran off which consisted of a single, highly divergent isolate from a sheep, 5668, for which amplification in *gag-pol* had remained unsuccessful despite repeated attempts using alternative sets of primers. The branchpoint of strain 5668 was placed outside of the common root of cluster A, as supported by a bootstrap value of 97%. Distinction from group

C was, with a bootstrap value of 69, less clear, but bootstrap values of 77 using parsimony and of 78 using likelihood methods were in support of strain 5668 being distinct from group C.

*Phylogenetic analysis of a 279 bp RT fragment inside gag-pol*

To better relate these new sequences to previously published phylogenetic work we also analyzed a 279-bp frag-

ment from the RT region that is contained within the 1.8 kb *gag-pol* amplicon (Leroux et al., 1997a; Ravazzolo et al., 2001; Rolland et al., 2002; Zanoni, 1998). From a total of 137 available sequences, 82 from the present study and 55 from the database, 68 were selected for analysis (Fig. 2, panel C). Although the tree was more complicated, the existence of the two large groups A and B was confirmed. In-between these were placed the Norwegian isolates of group C (cluster III of the current classification) and isolates from southern France currently classified as cluster VI. The low bootstrap value of 54 did not justify a phylogenetic separation of cluster VI from group A; it thus represents another principal branch of group A, which we accordingly label as A6. The position of isolate 5668, which could only be amplified in *pol*, in this tree remains open. Amplification with cluster VI-derived specific primers yielded a negative result suggesting that isolate 5668, classified in *pol* as group D, does not belong to A6.

In group B, branch B1 now also comprised the recently reported SRLV isolates from goats in southern France and Brazil which were classified within cluster V (Leroux et al., 1995; Ravazzolo et al., 2001). Branch B2 was now placed on a branch together with viruses (cluster IV) that had been isolated from sheep and goats in southern France (Leroux et al., 1995).

#### *Pairwise sequence distances*

For better visibility of the genetic distances among the different virus strains, which in the phylogenetic trees are represented by the cumulative length of branches and twigs that connect two given strains, we generated boxplots of pairwise DNA distances among all strains of the present study or available from GenBank, but excluding all pairs of epidemiologically linked, very closely related strains (Fig. 3). Distances could be grouped into different categories. Distances above 25% for both the 1.8 kb *gag-pol* and the 1.2 kb *pol* sequences were found for pairs formed among strains originating from different sequence groups A–D, with pairs involving C as one of the partners exhibiting the highest distances (panels A and B of Fig. 3). Distances of 17–24% in *gag-pol* and of 15–26% for *pol* were found among pairs originating from different branches within sequence groups A or B and also among the different twigs of branch A1. Distances below 15% in both sequences were restricted to pairs located within single branches containing at least two sequences, namely A3, A4 and B1.

Boxplots of distances in the 279 bp *pol* fragment also showed that intergroup distances were highest, intrabranched distances lowest and interbranch distances in between these two. There was, however, more overlap among these different categories, which was due to visibly higher variation of the pairwise distances of a given relationship (intrabranched, interbranch, intergroup) (Fig. 3, panel C). This confirmed that this short sequence was less reliable for phylogenetic analysis than were the two long sequences.

#### *Molecular epidemiology*

The possible association of sequence branches with various epidemiologic variables was first analyzed by contingency testing. An inhomogeneous distribution among the different branches was found concerning species, geographic localization, CAEV status of the herd and goat breed, but not age (Table 3).

#### *Species distribution*

Among the 68 different goat herds from which the virus was amplified in at least one animal, branch B1 (cluster V) was present in 44, A4 in 14 and A3 in 10 herds. Three of the six sheep herds from which viruses were amplified were infected with A4, two with A3 and one with B2. None of the goat or sheep herds which included more than one infected animal was infected by viruses from more than one branch (not shown). When both goats and sheep of a given farm were infected, which was the case in three instances, the branch was the same in both species; in two instances it was A4 and in one it was A3.

#### *Geographic distribution*

Branches B1 and A3 were present in herds from both north and south of the Alps. B1 was isolated from 44 goat herds that were in 11 of the 16 cantons from which at least one virus sequence originated. In contrast, A4 was only found in herds north of the Alps; here it was present in 17 herds in 9 different cantons. Sequences of branch A3 were dominating south of the Alps. A3 was the only branch found in the neighboring southern valleys of the Engiadina and Val Bregaglia and was isolated from all of the six herds that were in the east, middle and west of this region.

#### *CAEV status of goat herds and contact with sheep*

When the sequences isolated from goat herds were analyzed whether the herds had been registered as CAEV-free before the present cases, the CAEV-free herds exhibited a much higher frequency of branch A4. Goats with documented contact with sheep were more frequently infected with a virus belonging to A3 or A4, that is, viruses also found in sheep, while those without written documentation of such contact were more frequently infected with B1, which was found only among goats.

#### *Goat breeds and age*

Each of branches B1, A4 and A3 was present in at least four different registered goat breeds. All three were present at similar proportions in the two major breeds, Chamois with 24 animals and Saanen with 17 animals (data not shown). Nevertheless, there was a weakly significant association of branch and goat breed. In contrast, the distribution of branches in age groups below or, respectively, above 2 years was comparable (data not shown).

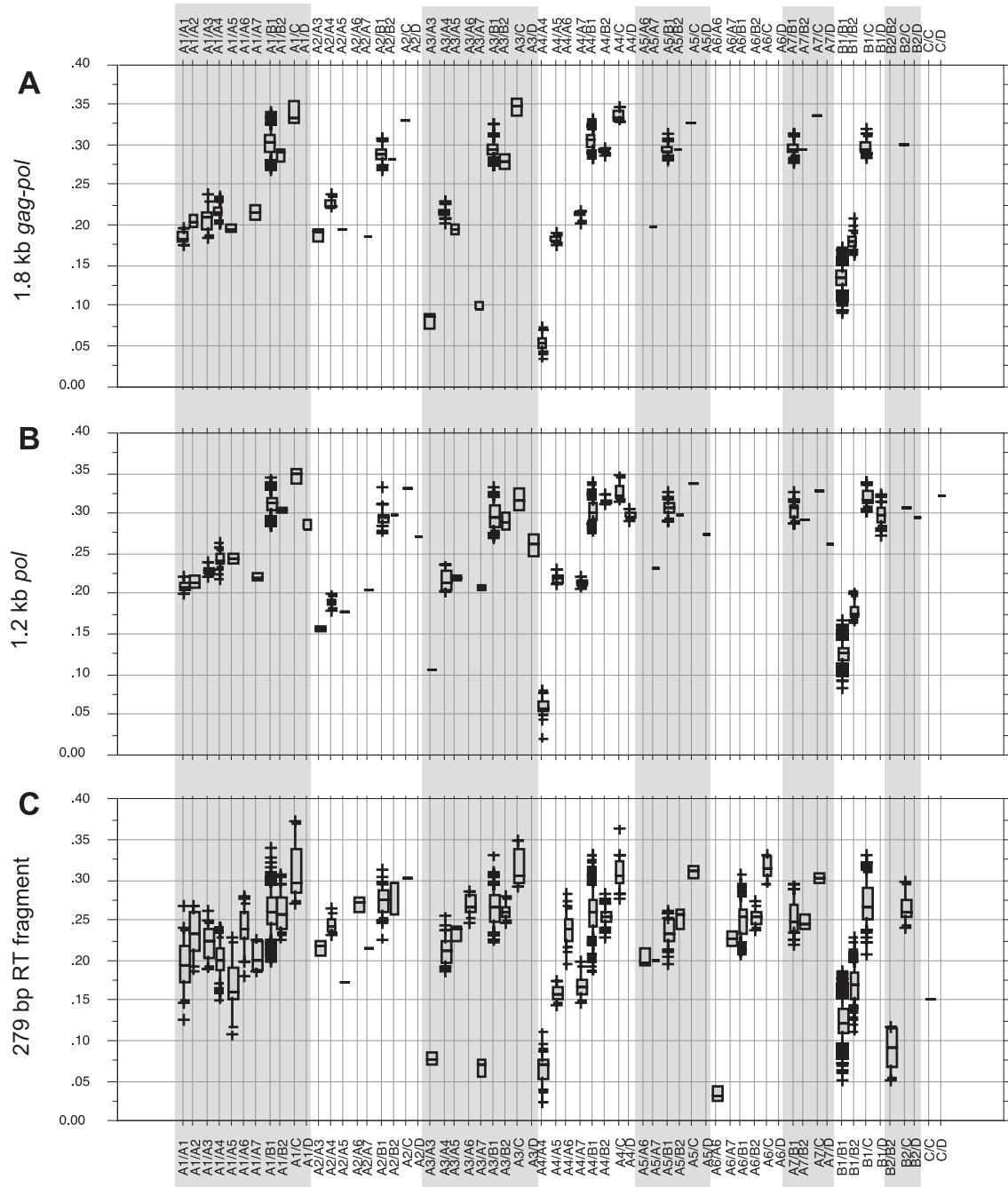


Fig. 3. Pairwise DNA distances as determined by the F84 substitution model of the *gag-pol* fragment (panel A), the 1.2 kb *pol* fragment (panel B) and the 279 bp RT fragment contained in the 1.8 kb *gag-pol* sequence (panel C) (see Methods). Boxplots show the median (line in the center of a box), the 25th and 75th percentiles (lower and, respectively, upper limit of a box), the 10th and 90th percentiles (closest horizontal lines outside of a box) and individually plotted outliers.

#### Logistic regression analysis

To sort out the impact of these variables on branch distribution in goats we employed logistic regression analysis (Table 4). In univariate models, CAEV-free status of the respective goat herd and evidence of contact with sheep were significantly correlated with infection by a virus of a branch other than B1, namely A3 or A4, with odds ratios of

8–9. Geographic localization of the herd, age and breed did not reach significance (Table 4A).

In a multivariate logistic regression model which combined the two variables that were significant by univariate analysis, that is, CAEV status of the herd and contact with sheep, both variables were significant and exhibited similar odds ratios as seen in the univariate analysis. The multivar-

Table 3  
Epidemiologic parameters with a nonhomogenous distribution among different sequence branches

Denominator	Branch B1		Branch A4		Branch A3		Branch B2		All		Contingency testing	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	$\chi^2$	<i>P</i>
Species									74	100	18.472	0.0004
Goat	44	64.7	14	20.6	10	14.7	–	–	68	100		
Sheep	–	–	3	50	2	33	1	17	6	100		
Geography									74	100	19.914	0.0002
Herds north of the Alps	39	62.9	17	27.4	5	8.1	1	1.6	62	100		
Herds south of the Alps	5	41.7	–	–	7	58.3	–	–	12	100		
CAEV status of herd <sup>a</sup>									68	100	19.762	0.0001
Certified CAEV-free	5	27.8	10	55.6	3	16.7	–	–	18	100		
Not certified	39	78.0	4	8.0	7	14.0	–	–	50	100		
Contact with sheep <sup>a</sup>									68	100	18.017	0.0001
Documented	4	26.7	4	26.7	7	46.7	–	–	15	100		
Not documented	40	75.5	10	18.9	3	5.7	–	–	53	100		
Breed <sup>a,b</sup>									53	100	36.475	0.0493

<sup>a</sup> Goats only.

<sup>b</sup> This variable includes 13 registered breeds and a listing up of sequence group frequencies is thus not feasible.

iate model correctly identified 84% of branch B1 and 79% of branches other than B1, resulting in 82.4% correctly identified cases overall (Table 4B). Addition of other variables that had been nonsignificant in the univariate analysis did neither alter the significance of either variable, nor did it improve the whole model (not shown). Thus, both a CAEV-free status of the herd and documented contact of the herd with sheep were strong independent correlates of infection of a goat with a lentivirus with a sequence other than of branch B1. These branches “other than B1” in goats were A3 or A4, which were also the dominant branches among infected sheep herds (Table 3). These data thus suggest an epidemiologic link between reemergence of infection in CAEV-free herds and contact with sheep.

## Discussion

The present survey of lentivirus isolates circulating among goats and sheep in Switzerland underscores the high degree of variability of this group of viruses for which six different lineages or clusters numbered I–VI have hitherto been described (Leroux et al., 1997a; Ravazzolo et al., 2001; Rolland et al., 2002; Zanoni, 1998). The present study increases the sequence information on these viruses by adding 104 new field isolates, thus providing 78 new sequences from a 1.8 kb *gag-pol* fragment and 72 new sequences from a 1.2 kb *pol* fragment (Fig. 1) to the six sequences covering these regions which are available from the GenBank database as of January 2003 (Braun et al.,

Table 4  
Logistic regression analysis of epidemiologic variables for association with infection by a sequence branch also found in sheep

A. Univariate models										
Variable(s)	Odds ratio	95% C.I.	<i>P</i>	Logistic likelihood ratio		Percent of cases correctly identified				
				Chi-square	<i>P</i>	Branch B1	Branches A3 or A4	All		
South of the Alps	2.600	0.700–9.653	0.1534	2.043	0.1529	88.6	25	66.2		
Age >2 years	0.409	0.122–1.367	0.1464	2.133	0.1441	100	0	63.3		
CAEV-free status of herd	9.218	2.696–31.523	0.0004	14.377	0.0002	88.6	54.2	76.5		
Contact with sheep	8.462	2.296–31.185	0.0013	11.848	0.0006	90.9	45.8	75.0		
Breed <sup>a</sup>	n.a.	n.a.	n.a.	17.752	0.1234	94.3	38.9	75.5		
B. Multivariate model										
Variable(s)	Odds ratio	95% C.I.	<i>P</i>	Logistic likelihood ratio		Logistic whole model fit		Percent of cases correctly identified		
				Chi-square	<i>P</i>	Chi-square	<i>P</i>	Branch B1	Branches A3 or A4	All
CAEV-free status of herd and contact with sheep						22.633	<0.0001	84.1	79.2	82.4
CAEV-free status of herd	8.253	2.200–30.965	0.0018	10.785	0.0010					
Contact with sheep	7.425	1.799–30.981	0.0059	8.296	0.0040					

<sup>a</sup> This variable includes 13 registered breeds none of which per se reaches significance.



1987; Querat et al., 1990; Rolland et al., 2002; Saltarelli et al., 1990; Sonigo et al., 1985). With this new wealth of sequence information, it has become possible to perform a more refined analysis of the phylogenetic relationship of the SRLVs and to perform molecular epidemiologic studies.

#### *Efficacy of sequence amplification*

Although both the 1.8 kb *gag-pol* and the 1.2 kb *pol* fragment were amplified by nested PCR using carefully evaluated primer combinations which had an estimated detection limit of 1–4 provirus copies/reaction (data not shown) and despite use of a sequence capture method, which further lowered the detection limit to 1–4 copies/100  $\mu$ g of DNA (data not shown), the overall success rate for amplification of at least one sequence was only 75% (Table 2). Of note, use of the sequence capture method for sequences not amplifiable by the regular nested procedure was rarely more successful, suggesting that low copy number was not the major factor for an amplification failure. If amplification of the short *gag* fragment was successful, proving the presence of provirus at detectable concentrations, successful amplification of at least one of the long fragments was with 82% relatively high, but still far from satisfactory. The large size of the amplicons probably did not significantly contribute to amplification failure, because the success rate for amplification of the 1.8 kb *gag-pol* fragment was not significantly lower than that of the short *gag* sequence and even higher than that of the 1.2 kb *pol* fragment. This leads us to conclude that sequence variation was probably the major cause of amplification failure.

The absence of amplifiable sequences in 23 seroindeterminate and 27 seronegative animals in a total of 90 reactions demonstrates the specificity of our PCR procedures. Although overall diagnostic sensitivity of PCR was only 75%, the entirely negative results among seroindeterminate animals can be interpreted as providing no support for a low-grade viral infection in these animals; at least a few should have been positive if that had been the cause of their indeterminate serologic results.

#### *Phylogenetic organization of SRLVs*

For human immunodeficiency virus type 1 (HIV-1), which is the best-studied lentivirus, subtypes A to H, J and K together form a “main” group M of phylogenetically related sequences which are distinguished from the “outlier” sequences of group O and a more recently identified group N (non-M/non-O). A recent reevaluation of the phylogenetic relationship of these viruses has resulted in a consensus nomenclature proposal which recommends to define subtypes as roughly equidistant from all previously characterized subtypes in all regions of the genome with a distinct pre-subtype branch similar to those of other subtypes (Robertson et al., 1999). The phylogenetic trees of Fig. 2 and the supplemental pairwise distance information of Fig. 3 indicate

that the various SRLV strains cannot be classified on the level of clades or subtypes alone, as is the case with the system of clusters I–VI proposed 5 years ago (Zanoni, 1998), because the different clades are not equidistant. Cluster V, for example, is much more closely related to cluster IV than it is to clusters I, II, III or VI, and clusters I, II and VI are more closely related to each other than they are to cluster III. It is thus necessary to superimpose sequence groups of a higher hierarchy by which the different branches can be classified into a phylogenetic relationship which satisfies the criterion of sequence equidistance among subtypes within a given group.

#### *Proposed new sequence groups A–D*

Due to historical reasons (Sigurdsson et al., 1957), we propose the designation group A for the large heterogeneous group which clusters around the MVV prototype isolates EV11514 and SA-OMVV. We propose the designation group B for the subsequently discovered more CAEV<sub>Cork</sub>-related viruses (Cork et al., 1974), and group C for the Norwegian virus isolates represented by the full-length sequence AF322109. These three groups are roughly equidistant to each other, as shown by pairwise distances of the respective virus strains that amount to around 30% for both the 1.8 kb *gag-pol* and the 1.2 kb *pol* sequence (Fig. 3). The high divergence of sheep isolate 5668, which is as distant from the other groups as they are among themselves, suggests that it represents a further principal group, D. Analysis of this sequence with a subtyping tool yields no evidence that it is the result of recombination of strains in the other listed groups. However, with just a single isolate and sequence information only available for the 1.2 kb *pol* fragment, the criteria recommended in the HIV field for definition of a new phylogenetic group or subtype (at least two isolates should be sequenced in their entirety, they should resemble each other but no other existing subtype throughout the genome and they should have been found in at least two epidemiologically unlinked individuals (Robertson et al., 1999)) are not yet met. Given the fact that the 1.8 kb *gag-pol* fragment of this isolate could not be amplified despite repeated attempts involving different sets of primers, it is likely, however, that the unidentified parts of this viral sequence will also be highly divergent, thus justifying classification into a separate group D. With respect to the fact that no viral sequence could be amplified from 25% of the seropositive samples, there is certainly room for the presence of still other, divergent viruses in the investigated population. It will also be important to get a more detailed knowledge of the virus strains that circulate in the Swiss sheep population, which is estimated to number 7-fold higher than the approximately 60,000 goats.

#### *Proposed subtypes*

Our data indicate that groups A and B should be further divided into different subtypes or clades. For HIV, the various subtypes of the M group exhibit a mean pairwise

DNA distance of 13% from each other in the *gag* gene and a mean of 11% in *pol* (Robertson et al., 1999). Based on these figures as a reference, the separation of group A into at least seven different subtypes, A1–A7, and of group B into two subtypes, B1 and B2, as indicated in Fig. 2, is justified because their inter-subtype genetic distances are clearly above 15% for both the 1.8 kb *gag-pol* and the 1.2 kb *pol* fragment (Fig. 3). Given the fact that pairwise distances among the isolates located within branch A1 (cluster I) are all above 15% and higher than the inter-subtype distances B1/B2 (clusters V–IV), the question arises whether clade A1 should actually be further divided into several different phylogenetic subtypes. This can only be resolved after further characterization of long sequences originating from the geographical regions from which the respective prototype viruses were isolated.

Pairwise distances and phylogenetic analysis both suggest that the viruses of the new Swiss subtype A3 and the North American ovine lentiviruses of A2 (cluster II) (Karr et al., 1996; Mwaengo et al., 1997) may be phylogenetically linked. While the bootstrap value of 100 clearly distinguishes the Swiss from the American isolates, the bootstrap values of 73 for *gag-pol* and of 59 for *pol*, which distinguish A2 and A3 from the other viruses of group A, represent insufficient proof of a common phylogenetic origin of A2 and A3. Our analysis certainly confirms the conclusion of Mwaengo et al. who found the North American isolates more closely related to MVV than to CAEV (Karr et al., 1996; Mwaengo et al., 1997).

The new Swiss subtypes A3 and A4 are documented by several epidemiologically unlinked isolates that were made from both goats and sheep. In contrast, the proposed new Swiss subtype A5 is currently represented by a single isolate, 5560. Isolates of French cluster VI, named accordingly subtype A6, were not found in the present investigation, and sequences of this subtype corresponding to the 1.8 kb *gag-pol* and the 1.2 kb *pol* fragments are not available from the database. Although the DNA distance in the 279 bp RT fragment of *gag-pol* to other clades of group A is

somewhat higher than among these other clades themselves, the low bootstrap value of 54 does not justify classification of these French isolates as a separate sequence group.

With regard to the proposed new Swiss subtype A7, the amplification of two very different sequences in *gag-pol* and, respectively, *pol* from goat 5692 suggests either a virus with a recombination in-between these two regions, or the simultaneous presence in this animal of two viruses that belong to two different clades, A3 and A7, from which only one region each, *gag-pol* or *pol*, could be amplified. Mix-up with a sample that remained negative for *gag-pol* amplification can also not be excluded, however. Whatever the true origin, there is no doubt that this *pol* sequence exists and that it is sufficiently distant from the other sequences of group A to form a separate subtype. Analysis with the subtyping tool yielded no evidence for the possibility that the aberrant classification of this *pol* sequence was the result of recombination within the *pol* of other viruses found in this study.

In subtype B2, the situation is also not fully clear because the affiliation of the single sequence of the present study, 5720, to the isolates of cluster IV is only based on an analysis of the 279 bp RT fragment (Fig. 2, panel C), which we found less suitable for phylogenetic classification than the much longer *gag-pol* and *pol* sequences (Fig. 3). The fact that bootstrap analysis of the long sequences clearly separated strain 5720 from subtype B1 confirms, however, that this strain belongs to a distinct subtype, and the high bootstrap value of 95 for the 279 bp RT fragment analysis supports placement of strain 5720 and of French cluster IV within a common subtype, B2.

Table 5 gives an overview of the proposed new classification and its relationship to that proposed earlier (Zanoni, 1998). The new classification proposed adheres to the current cluster designations as far as this was possible and reasonable. Thus, clusters I, II and VI have become subtypes A1, A2 and A6, and cluster III has become group C. The classification is open and provides room for additional groups and subtypes to be identified in the future.

Table 5  
Proposed new classification of small ruminant lentiviruses

Proposed new group designation	A					B		C	D		
Proposed new subtype or clade designation	A1	A2	A3	A4	A5	A6	A7	B1	B2	–	–
Lineage in current nomenclature (Zanoni, 1998)	Cluster I	Cluster II	–	–	–	Cluster VI	–	Cluster V	Cluster IV	Cluster III	–
Characteristics and comments	MVV-like; may need division into more than one subtype	Sheep from North America (Karr et al., 1996)	Isolated from goats and sheep of present study	Isolated from goats and sheep of present study	Single isolate (5560) from a goat of present study	Isolated from goats and sheep in southern France	Single isolate (5692) from a goat of present study	CAEV <sub>Cork</sub> -like, isolated from goats worldwide	Isolated from goats and sheep in southern France and two sheep of present study	Isolated from goats in Norway	Single isolate (5668) from a goat of present study

### *Molecular epidemiology*

A remarkable feature of the subtype B1 tree of Fig. 2 is the homogenous length of its twigs and their common rooting point. Provided that the sequence evolution rate of the various strains of subtype B1 in goats is similar, one would expect a similar timepoint of infection for all the herds represented by these sequences and infection by closely related source viruses. After the original seeding event, the timepoint of which is unknown, these closely related virus strains would then have evolved independently in the different herds. As the major way of transmission of CAEV under natural conditions is from mother to offspring through milk, viruses would be passed from generation to generation without much further genetic exchange with other strains.

It is striking that isolates of the present study, which were collected in a small country and during a narrow time window, generally differ no less from each other than they differ from strains prevalent in other world regions. Strains from southern France (U35809) or Brazil (AF108055 and AF108059) fit perfectly into the cluster formed by the Swiss isolates (Fig. 2C). Most strikingly, isolates 5576 and 5775 that were established from a Chamoisé and a dwarf goat, respectively, were in all analyzed sequences and with all models more closely related to CAEV<sub>Cork</sub>, which was isolated three decades ago in the US from a Saanen goat of Swiss origin (Cork et al., 1974), than to any other of the Swiss strains (Figs. 2A–C), thus suggesting a closer phylogenetic relationship of the CAEV<sub>Cork</sub> prototype with these two isolates, which originate from two villages of northern Switzerland located only 10 km apart, than with other Swiss isolates. Such similarity suggests that international livestock trade may have contributed significantly to the worldwide distribution of SRLV.

### *Evidence for regular interspecies transmission*

As mentioned in the Introduction, Switzerland has an ongoing, very successful CAEV eradication program. The reemergence of infections in previously CAEV-negative goat herds is not only puzzling, but also places severe restrictions upon the herd owners which result in considerable economic damage. The fact that, as shown in Table 4, infection of a goat by subtypes A3 or A4, which were also found to be the dominant clades in the few sheep of this study, was significantly and independently associated with (i) a CAEV-free status of the goat herd and (ii) a documented contact of the herd with sheep provides strong statistical evidence for the hypothesis that reemergence of SRLV infection in certified CAEV-free goat herds is due to viruses that were transmitted from infected sheep. Although such events of natural sheep-to-goat interspecies transmission of SRLV may be infrequent, they must occur regularly because truly exceptional events would not be amenable to statistics. Although MVV can be experimentally transmitted to goats and, vice versa, CAEV to sheep (Banks et al., 1983), interspecies transmis-

sion under natural conditions was not observed when transmission of CAEV from goats to sheep was studied (Banks et al., 1983; Oliver et al., 1985; Rowe and East, 1997; Smith et al., 1985). No publications on naturally occurring SRLV transmissions from sheep to goats were found in the scientific literature. Although viruses from southern France now classified as subtypes A6 and B2 are found in both goats and sheep (Table 5), thus presenting unequivocal evidence for interspecies transmission (Leroux et al., 1997a), it remains open whether this represented repeated, ongoing transmissions or was due to single events that occurred in the past. Moreover, the direction of transmission remained unclear. Here we present evidence for regularly occurring sheep-to-goat transmission of SRLV.

Under the current regulations of the Swiss CAEV eradication program, viruses such as subtype B1 (cluster V) which were found only in goats in this and other studies (Ravazzolo et al., 2001; Valas et al., 1997; Zanoni, 1998) will be successfully eliminated because reinfection of a CAEV-free herd from a CAEV-contaminated goat herd will be prevented. Strains that infect both sheep and goats, such as subtypes A3, A4, A6 (cluster VI), B2 (cluster IV) and possibly others, will, however, continue to be transmitted from infected sheep to goats because the current federal regulations do not consider this possibility and because, depending on the region, 30–60% of Swiss goat owners keep goats and sheep in mixed herds. Complete eradication of lentivirus infection in the goat population will probably not be possible without strict and permanent separation of goats and sheep or inclusion of the sheep population into the eradication program.

## **Methods**

### *Animals and blood samples*

Animals for the investigation were selected by the federally supported Extension and Health Service for Small Ruminants (BGK), which is involved in the CAEV eradication program in Switzerland. The selection of SRLV-positive animals for the present study took place in connection with the country-wide annual CAEV screening of 1999 which was mandatory for all goats. Basically, it included three types of herds, namely (i) herds which had already been seronegative in at least three consecutive annual testings (CAEV-free herds), (ii) herds with a history of CAEV infection that were newly entered into the program (CAEV-positive herds) and (iii) herds that had maintained a status of seronegativity in less than three annual testings (herds of unresolved CAEV status).

From CAEV-free herds, all goats newly found seropositive (positive ELISA result confirmed by a positive Western blot result) or seroindeterminate (positive ELISA result followed by an indeterminate Western blot result) were selected for the study. Seropositive or seroindeterminate sheep kept in such farms and enrolled in the voluntary

MVV control program for milking sheep were also selected. From CAEV-positive herds or herds of unresolved CAEV status, seropositive goats were selected in a manner adequately representing the various breeds, geographic regions and age groups. Control animals were chosen among seronegative goats from herds with negative, positive or unresolved CAEV status. Additional seropositive or seroindeterminate sheep were selected from long-time MVV-negative herds enrolled in the voluntary MVV control program.

From all selected animals, 100 ml blood samples with EDTA as anti-coagulant was drawn by venipuncture. Peripheral blood mononuclear cells (PBMC) were isolated by Ficoll density gradient centrifugation and aliquots of  $5 \times 10^6$  or  $50 \times 10^6$  cells were frozen at  $-80^\circ\text{C}$  as nonviable cell pellets.

#### *DNA extraction*

DNA from cell pellets containing  $5 \times 10^6$  cells was extracted using commercial silica-gel spin-columns selective for genomic DNA (QIAamp DNA Blood Mini Kit, QIAGEN) according to the manufacturer's instructions. Bound DNA was eluted in 200  $\mu\text{l}$  of buffer AE (QIAGEN) by a final centrifugation step at  $6000 \times g$  for 1 min. Purified DNA stocks were stored at  $-70^\circ\text{C}$  until used.

DNA from cell pellets of  $\geq 50 \times 10^6$  cells was extracted using a different protocol. Briefly, the cells were lysed in 600  $\mu\text{l}$  proteinase K buffer (100 mM Tris pH 7.5; 12.5 mM EDTA- $\text{Na}_2$ ; 150 mM NaCl; 0.5% SDS) and incubated at  $65^\circ\text{C}$  for 30 min on a thermomixer. After the addition of 600  $\mu\text{l}$  proteinase K buffer with 50  $\mu\text{g}$  of proteinase K, proteins were digested at  $55^\circ\text{C}$  for 2 h and insoluble material removed by centrifugation at  $14,000 \times g$  for 2 min. The supernatant was then transferred (2 aliquots of 600  $\mu\text{l}$ ) to two new tubes each containing 600  $\mu\text{l}$  100% isopropanol and mixed. After standing for 2 min at room temperature, the precipitates were centrifuged at  $14,000 \times g$  for 15 min. Each DNA pellet was washed once with 500  $\mu\text{l}$  80% ethanol and then dissolved in 210  $\mu\text{l}$   $\text{H}_2\text{O}$  at  $55^\circ\text{C}$  for 3 h on a thermomixer. Finally, the undissolved material was pelleted at  $14,000 \times g$  for 2 min and the two supernatants were pooled. The purified DNA stocks were stored at  $-70^\circ\text{C}$  until used.

#### *Capture oligonucleotides, primers and probes*

All oligonucleotides used in this work as well as their function and location are shown in Table 6. The location of the different amplicons is indicated in Fig. 1. All oligonucleotides were selected by the Oligo 6.1 Primer Analysis software (Molecular Biology Insights) and synthesized by Microsynth GmbH, Balgach, Switzerland.

#### *Short gag PCR*

A first PCR procedure for SRLV sequences was developed based on sequences in *gag* that were conserved among

the three available full-length sequences from MVV (Braun et al., 1987; Querat et al., 1990; Sonigo et al., 1985), the single full-length sequence available from CAEV (Saltarelli et al., 1990) and a 2.5 kb *gag-pol* sequence from a Swiss isolate related to MVV previously established from a seroconverting goat (J. B. Huder et al., unpublished). Forward and reverse primers were in a conserved portion of the *gag* gene dubbed "short *gag*" (Table 6, Fig. 1).

The PCR reaction mix consisted of  $1 \times$  PCR Buffer (Qiagen), 200  $\mu\text{M}$  each of dATP, dCTP, dGTP and TTP (Pharmacia), 300 nM of each primer, 0.025 U/ $\mu\text{l}$  HotStarTaq DNA polymerase (Qiagen). Reaction mix was aliquoted in 48  $\mu\text{l}$  amounts into each 0.5 ml thin-wall tube (Axygen) followed by addition of 2  $\mu\text{l}$  ( $\approx 1 \mu\text{g}$ ) of sample.

Activation of HotStarTaq DNA polymerase was then undertaken at  $95^\circ\text{C}$  for 15 min, followed by 45 cycles of DNA amplification consisting of denaturation at  $94^\circ\text{C}$  for 30 s, annealing at  $57^\circ\text{C}$  for 1 min and extension at  $72^\circ\text{C}$  for 1 min. Detection of specific amplicons was achieved by agarose gel electrophoresis and ethidium bromide staining. Whenever a specific band was not clearly visible or too weak, a second PCR round with 2  $\mu\text{l}$  of the first round PCR reaction and the nested primers, P40 and P41, was performed (Fig. 1).

#### *Nested PCR for the 1.8 kb gag-pol and the 1.2 kb pol fragment*

Reaction conditions for the first and second round of PCR were identical to those described for the short *gag* PCR, with the exception that the specific PCR primers indicated in Table 6 were used and that the total reaction volume was increased to 100  $\mu\text{l}$ . For the first round of PCR activation of HotStarTaq DNA polymerase was undertaken at  $95^\circ\text{C}$  for 15 min, followed by 35 cycles of denaturation at  $94^\circ\text{C}$  for 30 s, annealing at  $55^\circ\text{C}$  for 1 min and extension at  $72^\circ\text{C}$  for 2 min. For the second-round of nested PCR, 2  $\mu\text{l}$  of the first round PCR was transferred to the new reaction mix. After activation of HotStarTaq DNA polymerase at  $95^\circ\text{C}$  for 15 min, 45 cycles of denaturation at  $94^\circ\text{C}$  for 30 s, annealing at  $57^\circ\text{C}$  for 1 min and extension at  $72^\circ\text{C}$  for 2 min were carried out.

#### *Sequence capture procedure*

Forward and reverse capture probes were in conserved sections of the *gag* and *pol* region and consisted of the sequences listed in Table 6. Selective substitution of dUMP for TMP was undertaken to facilitate eventual destruction of the capture oligonucleotides (Shah et al., 2003). Capture probes were biotinylated at the 5' end during synthesis. Enrichment of target sequences from 100  $\mu\text{g}$  DNA was achieved by sequence capture with all four capture probes. Solution hybridization, coupling to streptavidin-coated magnetic beads and washing were carried out exactly as described in detail for porcine endogenous retrovirus capture

Table 6  
List of oligonucleotides with orientation and application

Designation	Sequence (5'–3')	Orientation <sup>a</sup>	Location <sup>b</sup>	PCR <sup>c</sup>
<i>Short gag PCR</i>				
P14	AGATGCAAGC ATGTCGAGAT GTAGGATCAG AAGG	F	1569–1602	1
P15	GTTACAACAT AGGAGGCGCG GACGGCACCA	R	1830–1859	1
P40	TCAAATGCA ACTGTTAGCA CAAGC	F	1605–1629	2
P41	GCGGACGGCA CCACACG	R	1826–1842	2
<i>Gag-pol PCR</i>				
P39	GCAATGCAGC ATGGMCTTGT GT	F	1019–1140	1
P37	GATTGCCTCC AWACTAACTC TCCTACTAA	R	2938–2966	1
P38	CCWATTAAGG ATTGCCTCCA WACTA	R	2951–2975	1
P21	GCCATGATGC CTGGAAATAG AG	F	1112–1133	2
P27	TTCTCCGGGT GTAGTTCAAAA TC	R	2846–2867	2
<i>Pol PCR</i>				
P28	CATGAAGAGG GGACAAATCA GCA	F	3535–3557	1
P34	TACATTGGGT GCCTGGACAT AA	F	3734–3755	1
P32	TACCTGDGTT GGTCCYWGCC ACT	R	4909–4932	1
P33	CTTCCCAVAG TACCTGDGTT GGTC	R	4918–4941	1
P29	GGTGCCTGGA CATAAAGGGA TTC	F	3741–3763	2
P35	GCCACTCTCC TGRATGTCTT CT	R	4894–4915	2
<i>Capture probes</i>				
P14Bio	AGAUGCAAGC AUGUCGAGAU GUAGGAUCAG AAGG	F	1569–1602	
P17Bio	GGAUGCGGUA GUCCUACUG UGCUUCUGCU A	R	2414–2444	
P18Bio	CAUUGGGUGC CUGGACAUAA AGGGAUUCU CA	F	3736–3767	
P19Bio	GCUACAAAUG CUGGUCCA AUUCUGACUGC AAUGA	R	4543–4577	
<i>Sequencing Gag-pol</i>				
P3	GTAGAAGAAA AAATGCAAGC ATG	F	1559–1581	
P6	CCAGTAGAGG TATTAGGACG AGAT	F	2092–2115	
P7	GGAGTATTAC ATGTCCAATG TGG	R	2311–2333	
P8	GTAAAATATG CATCCCCTAT GTC	R	2482–2504	
P48CAEV	CCTCATGTYC CRCAATGGC	F	2212–2230	
P48MVV	CAAAATRTRG CACAATGGC	F	2212–2230	
<i>Sequencing Pol</i>				
P46	GCAGGTCCAT TGTCTGAYTG	R	4549–4568	
P47	TGGGTAGAAA CAAATTCAGG	F	4435–4454	

<sup>a</sup> Orientation: forward (F); reverse (R).

<sup>b</sup> Localization number corresponding to the Cork-isolate (37).

<sup>c</sup> The use in the first (1.) or second (2.) round PCR.

(Shah et al., 2003). After washing the sequence-coated beads were resuspended in 100 µl of the first-round PCR reaction mix containing 0.5 U of uracyl-*N*-glycosylase, and DNA was amplified as described above after an initial incubation at 50 °C for 15 min.

#### DNA sequence analysis

Specific PCR products were purified from agarose gel with the QIAquick Gel Extraction Kit (Qiagen). Both strands of the DNA were sequenced with the amplification primers and with the sequencing primers listed in Table 6 using dye-terminator chemistry (Applied Biosystems). Capillary electrophoresis was carried out on a ABI 310 Sequence Detector (Applied Biosystems). Nucleic acid sequences were assembled and edited with AutoAssembler 2.1 (Applied Biosystems) and analyzed with MacVector 7.0 (Accelrys Inc.).

#### Phylogenetic analysis

Nucleic acid sequences were aligned using the ClustalW algorithm of the MacVector software 7.0 (Accelrys Inc.), and the aligned, length-adjusted sequences were imported as a nexus-file into the “Phylogenetic Analysis Using Parsimony” software package (PAUP\* 4.0 Beta 10 for power PC, 2001; Sinauer Associates, Sunderland, Massachusetts). All further calculations and tree constructions were carried out with this software.

Pairwise genetic distances were calculated using the F84 substitution model with default settings with the exception that all sites with ambiguous codes and gaps were ignored. All unrooted phylogenetic trees shown were constructed by the neighbor-joining method and were based on the distances calculated with the F84 substitution model, again ignoring all sites with ambiguous codes and gaps. Bootstrap values

are based on 1000 repetitions. Unrooted trees were also constructed using parsimony with default settings and bootstrap values based on 1000 repetitions. Additional unrooted trees from selected isolates were constructed based on the likelihood algorithm with default settings but with the F84 substitution model and the estimated transition/transversion ratios of 2.53 for the 1.8 kb *gag-pol* and 2.11 for the 1.2 kb *pol* fragment, respectively, which were calculated from the entire set of isolates. Bootstrap values were based on 100 repetitions.

#### Nucleotide sequence accession numbers

All nucleotide sequences were deposited in the GenBank database and are available under accession nos. AY454161 to Y454232 for the 78 1.8 kb *gag-pol* fragments and under accession nos. AY454233 to Y454269 for the 72 1.2 kb *pol* fragments.

#### Statistical analysis

All statistical evaluations were performed by the contingency table and logistic regression functions contained in the StatView Version 5.0 software for Macintosh (SAS Institute Inc., Cary, North Carolina).

#### Acknowledgments

We thank Antonietta Baumgartner and Lucia Bertodatto for their excellent technical assistance and Drs. Giuseppe Bertoni, Lukas Perler, Ernst Peterhans, Marietta Schönmann and Alfred Zaugg for valuable discussions. This work was financially supported by grants from the Swiss Federal Office of Veterinary Health, Berne, and Schweizerischer Vorort, Viehandelskonkordat/Concordat Sur Le Commerce Du Bétail, Aarau, Switzerland.

#### References

- Anonymous, 1999. Control of caprine arthritis-encephalitis in Switzerland (article downloadable from [http://www.bvet.admin.ch/0\\_navigation-e/0\\_index-intern.html](http://www.bvet.admin.ch/0_navigation-e/0_index-intern.html)). Animal and Public Health Reports of the Swiss Federal Veterinary Office. 3, 3–9.
- Banks, K.L., Adams, D.S., McGuire, T.C., Carlson, J., 1983. Experimental infection of sheep by caprine arthritis-encephalitis virus and goats by progressive pneumonia virus. *Am. J. Vet. Res.* 44 (12), 2307–2311.
- Barlough, J., East, N., Rowe, J.D., Van Hoosear, K., DeRock, E., Bigornia, L., Rimstad, E., 1994. Double-nested polymerase chain reaction for detection of caprine arthritis-encephalitis virus proviral DNA in blood, milk, and tissues of infected goats. *J. Virol. Methods* 50 (1–3), 101–113.
- Braun, M.J., Clements, J.E., Gonda, M.A., 1987. The visna virus genome: evidence for a hypervariable site in the *env* gene and sequence homology among lentivirus envelope proteins. *J. Virol.* 61 (12), 4046–4054.
- Braun, U., Camenzind, D., Meli, M., Boni, J., Ossent, P., 2001. Clinical findings and diagnostic procedure in a dairy sheep with visna. *Schweiz. Arch. Tierheilkd.* 143 (11), 550–553.
- Celer Jr., V., Celer, V., Nejedla, E., Bertoni, G., Peterhans, E., Zanoni, R.G., 2000. The detection of proviral DNA by semi-nested polymerase chain reaction and phylogenetic analysis of Czech Maedi-Visna isolates based on *gag* gene sequences. *Zentralbl. Veterinärmed., Reihe B* 47 (3), 203–215.
- Cork, L.C., Hadlow, W.J., Gorham, J.R., Pyper, R.C., Crawford, T.B., 1974. Infectious leukoencephalomyelitis of goats. *J. Infect. Dis.* 129, 134–141.
- Cutlip, R.C., Lehmkuhl, H.D., 1986. Eradication of ovine progressive pneumonia from sheep flocks. *J. Am. Vet. Med. Assoc.* 188 (9), 1026–1027.
- De Boer, G.F., Terpstra, C., Houwers, D.J., Hendriks, J., 1979. Studies in epidemiology of maedi/visna in sheep. *Res. Vet. Sci.* 26 (2), 202–208.
- Ding, E., 1995. Analysis and PCR detection of antigen compositions of ovine progressive pneumonia virus. *Sci. China, Ser. B: Chem.* 38 (5), 573–579.
- East, N.E., Rowe, J.D., Madewell, B.R., Floyd, K., 1987. Serologic prevalence of caprine arthritis-encephalitis virus in California goat dairies. *J. Am. Vet. Med. Assoc.* 190 (2), 182–186.
- Ellis, T., Robinson, W., Wilcox, G., 1983. Effect of colostrum deprivation of goat kids on the natural transmission of caprine retrovirus infection. *Aust. Vet. J.* 60 (11), 326–329.
- Greenwood, P.L., North, R.N., Kirkland, P.D., 1995. Prevalence, spread and control of caprine arthritis-encephalitis virus in dairy goat herds in New South Wales. *Aust. Vet. J.* 72 (9), 341–345.
- Haase, A.T., Retzel, E.F., Staskus, K.A., 1990. Amplification and detection of lentiviral DNA inside cells. *Proc. Natl. Acad. Sci. U.S.A.* 87 (13), 4971–4975.
- Hanson, J., Hydbring, E., Olsson, K., 1996. A long term study of goats naturally infected with caprine arthritis-encephalitis virus. *Acta Vet. Scand.* 37 (1), 31–39.
- Houwers, D.J., 1980. (Maedi and maedi control).” (author’s translation). *Tijdschr. Diergeneeskd.* 105 (16), 661–664 (in Dutch).
- Houwers, D.J., König, C.D., Bakker, J., de Boer, M.J., Pekelder, J.J., Sol, J., Vellema, P., de Vries, G., 1987. Maedi-visna control in sheep: III. Results and evaluation of a voluntary control program in The Netherlands over a period of four years. *Vet. Q.* 9 (1), 29S–36S.
- Johnson, L.K., Meyer, A.L., Zink, M.C., 1992. Detection of ovine lentivirus in seronegative sheep by in situ hybridization, PCR, and cocultivation with susceptible cells. *Clin. Immunol. Immunopathol.* 65 (3), 254–260.
- Karr, B.M., Chebloune, Y., Leung, K., Narayan, O., 1996. Genetic characterization of two phenotypically distinct North American ovine lentiviruses and their possible origin from caprine arthritis-encephalitis virus. *Virology* 225 (1), 1–10.
- Kennedy-Stoskopf, S., Narayan, O., Strandberg, J.D., 1985. The mammary gland as a target organ for infection with caprine arthritis-encephalitis virus. *J. Comp. Pathol.* 95 (4), 609–617.
- Krassnig, R., Schuller, W., 1998. Continuation of the observation and serological investigation of a Maedi-Visna virus infected sheep flock from January 1990 to June 1996. *Dtsch. Tierärztl. Wochenschr.* 105 (2), 50–53.
- Krieg, A., Peterhans, E., 1990. Caprine arthritis-encephalitis in Switzerland: epidemiologic and clinical studies. *Schweiz. Arch. Tierheilkd.* 132 (7), 345–352.
- Leroux, C., Vuillermoz, S., Mornex, J.F., Greenland, T., 1995. Genomic heterogeneity in the *pol* region of ovine lentiviruses obtained from bronchoalveolar cells of infected sheep from France. *J. Gen. Virol.* 76 (Pt. 6), 1533–1537.
- Leroux, C., Chastang, J., Greenland, T., Mornex, J.F., 1997a. Genomic heterogeneity of small ruminant lentiviruses: existence of heterogeneous populations in sheep and of the same lentiviral genotypes in sheep and goats. *Arch. Virol.* 142 (6), 1125–1137.
- Leroux, C., Lerondelle, C., Chastang, J., Mornex, J.F., 1997b. RT-PCR detection of lentiviruses in milk or mammary secretions of sheep or goats from infected flocks. *Vet. Res.* 28 (2), 115–121.
- McGuire, T.C., O’Rourke, K.I., Knowles, D.P., Cheevers, W.P., 1990. Caprine arthritis encephalitis lentivirus transmission and disease (Review) (79 refs). *Curr. Top. Microbiol. Immunol.* 160, 61–75.

- Mwaengo, D.M., Grant, R.F., DeMartini, J.C., Carlson, J.O., 1997. Envelope glycoprotein nucleotide sequence and genetic characterization of North American ovine lentiviruses. *Virology* 238 (1), 135–144.
- Narayan, O., Cork, L.C., 1985. Lentiviral diseases of sheep and goats: chronic pneumonia leukoencephalomyelitis and arthritis. *Rev. Infect. Dis.* 7 (1), 89–98.
- Narayan, O., Clements, J.E., Strandberg, J.D., Cork, L.C., Griffin, D.E., 1980. Biological characterization of the virus causing leukoencephalitis and arthritis in goats. *J. Gen. Virol.* 50 (1), 69–79.
- Oliver, R., Cathcart, A., McNiven, R., Poole, W., Robati, G., 1985. Infection of lambs with caprine arthritis encephalitis virus by feeding milk from infected goats. *Vet. Rec.* 116 (3), 83.
- Pasick, J., 1998. Maedi-visna virus and caprine arthritis-encephalitis virus: distinct species or quasispecies and its implications for laboratory diagnosis. *Can. J. Vet. Res.-Rev. Can. Rech. Vet.* 62 (4), 241–244.
- Pepin, M., Vitu, C., Russo, P., Mornex, J.F., Peterhans, E., 1998. Maedi-visna virus infection in sheep: a review (Review) (182 refs). *Vet. Res.* 29 (3–4), 341–367.
- Peretz, G., Bugnard, F., Calavas, D., 1994. Study of a prevention programme for caprine arthritis-encephalitis. *Vet. Res.* 25 (2–3), 322–326.
- Querat, G., Audoly, G., Sonigo, P., Vigne, R., 1990. Nucleotide sequence analysis of SA-OMVV, a visna-related ovine lentivirus: phylogenetic history of lentiviruses. *Virology* 175 (2), 434–447.
- Ravazzolo, A.P., Reischak, D., Peterhans, E., Zanoni, R., 2001. Phylogenetic analysis of small ruminant lentiviruses from Southern Brazil. *Virus Res.* 79 (1–2), 117–123.
- Reddy, P.G., Sapp, W.J., Heneine, W., 1993. Detection of caprine arthritis-encephalitis virus by polymerase chain reaction. *J. Clin. Microbiol.* 31 (11), 3042–3043.
- Rimstad, E., East, N.E., Torten, M., Higgins, J., DeRock, E., Pedersen, N.C., 1993. Delayed seroconversion following naturally acquired caprine arthritis-encephalitis virus infection in goats. *Am. J. Vet. Res.* 54 (11), 1858–1862.
- Robertson, D.L., Anderson, J.P., Bradac, J.A., Carr, J.K., Foley, B., Funkhouser, R.K., Gao, F., Hahn, B.H., Kuiken, C., Learn, G.H., Leitner, T., McCutchan, F., Osmanov, S., Peeters, M., Pieniazek, D., Kalish, M.L., Salminen, M., Sharp, P.M., Wolinsky, S., Korber, B., 1999. HIV-1 nomenclature proposal: a reference guide to HIV-1 classification. *The Human Retroviruses and AIDS 1999 Compendium. PART VI. Reviews.* Los Alamos National Laboratory, Los Alamos, pp. 492–505.
- Rolland, M., Mooney, J., Valas, S., Perrin, G., Mamoun, R.Z., 2002. Characterisation of an Irish caprine lentivirus strain—SRLV phylogeny revisited. *Virus Res.* 85 (1), 29–39.
- Rowe, J.D., East, N.E., 1997. Risk factors for transmission and methods for control of caprine arthritis-encephalitis virus infection. *Vet. Clin. North Am., Food Anim. Pract.* 13 (1), 35–53.
- Rowe, J.D., East, N.E., Thurmond, M.C., Franti, C.E., Pedersen, N.C., 1992. Cohort study of natural transmission and two methods for control of caprine arthritis-encephalitis virus infection in goats on a California dairy. *Am. J. Vet. Res.* 53 (12), 2386–2395.
- Saltarelli, M., Querat, G., Konings, D.A., Vigne, R., Clements, J.E., 1990. Nucleotide sequence and transcriptional analysis of molecular clones of CAEV which generate infectious virus. *Virology* 179 (1), 347–364.
- Scheer-Czechowski, P., Vogt, H.R., Tontis, A., Peterhans, E., Zanoni, R., 2000. Pilot project for eradicating maedi-visna in Walliser blacknose sheep. *Schweiz. Arch. Tierheilkd.* 142 (4), 155–164.
- Shah, C.A., Böni, J., Bisset, L.R., Seebach, J.D., Schüpbach, J., 2003. Ultra-sensitive and specific detection of porcine endogenous retrovirus (PERV) using a sequence-capture real-time PCR approach. *J. Virol. Methods* 109, 209–216.
- Sigurðsson, B., Pálsson, P.A., Grimsson, H., 1957. Visna, a demyelinating transmissible disease of sheep. *J. Neuropathol. Exp. Neurol.* 16, 389–403.
- Smith, V.W., Dickson, J., Coackley, W., Carman, H., 1985. Response of merino sheep to inoculation with a caprine retrovirus. *Vet. Rec.* 117 (3), 61–63.
- Sonigo, P., Alizon, M., Staskus, K., Klatzmann, D., Cole, S., Danos, O., Retzel, E., Tiollais, P., Haase, A., Wain-Hobson, S., 1985. Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus. *Cell* 42 (1), 369–382.
- Sundquist, B., 1981. Goat visna virus: isolation of a retrovirus related to visna virus of sheep. *Arch. Virol.* 68 (2), 115–127.
- Valas, S., Benoit, C., Guionaud, C., Perrin, G., Mamoun, R.Z., 1997. North American and French caprine arthritis-encephalitis viruses emerge from ovine maedi-visna viruses. *Virology* 237 (2), 307–318.
- Wagter, L.H., Jansen, A., Bleumink-Pluym, N.M., Lenstra, J.A., Houwers, D.J., 1998. PCR detection of lentiviral GAG segment DNA in the white blood cells of sheep and goats. *Vet. Res. Commun.* 22 (5), 355–362.
- Zanoni, R.G., 1998. Phylogenetic analysis of small ruminant lentiviruses. *J. Gen. Virol.* 79 (Pt. 8), 1951–1961.
- Zanoni, R.G., Nauta, I.M., Kuhnert, P., Pauli, U., Pohl, B., Peterhans, E., 1992. Genomic heterogeneity of small ruminant lentiviruses detected by PCR. *Vet. Microbiol.* 33 (1–4), 341–351.
- Zanoni, R.G., Cordano, P., Nauta, I.M., Peterhans, E., 1996. PCR for the detection of lentiviruses from small ruminants. *Schweiz. Arch. Tierheilkd.* 138 (2), 93–98.
- Zhang, Z., Watt, N.J., Hopkins, J., Harkiss, G., Woodall, C.J., 2000. Quantitative analysis of maedi-visna virus DNA load in peripheral blood monocytes and alveolar macrophages. *J. Virol. Methods* 86 (1), 13–20.