sciendo

J Vet Res 66, 2022
DOI: 10.2478/jvetres-2022-0064

Phylogenetic analysis of small ruminant lentiviruses originating from naturally infected sheep and goats from Poland based on the long terminal repeat sequences

Monika Olech[1, 2✉], Jacek Kuźmak[2], Anna Kycko[1], Andrzej Junkuszew[3]

[1]Department of Pathology, [2]Department of Biochemistry,
National Veterinary Research Institute, 24-100 Puławy, Poland
[3]Institute of Animal Breeding and Biodiversity Conservation,
University of Life Sciences in Lublin, 20-950 Lublin, Poland
monika.olech@piwet.pulawy.pl

Received: June 07, 2022 *Accepted: November 03, 2022*

Abstract

Introduction: Previous *gag* and *env* sequence studies placed Polish small ruminant lentiviruses (SRLVs) isolated from sheep and goats in subtypes B1, B2, A1, A5, A12, A13, A16–A18, A23, A24 and A27. This study extended the genetic/phylogenetic analysis of previously identified Polish SRLV strains by contributing long terminal repeat (LTR) sequences. **Material and Methods:** A total of 112 samples were analysed. Phylogenetic analyses were carried out on the LTR fragment using the neighbour-joining, maximum likelihood, and unweighted pair group method with arithmetic mean methods. **Results:** Polish caprine and ovine LTR sequences clustered within group A and grouped in at least 10 clusters (subtypes A1, A5, A12, A13, A16–A18, A23, A24 and A27). Most of the Polish strains (78%) belonged to the same subtype by the indication of the *gag*, *env* and LTR genomic regions. Discrepancies in affiliation depending on the particular sequence were observed in 24 (21%) strains, most of which came from mixed-species flocks where more than one SRLV genotype circulated. Sequences of the LTR reflected subtype-specific patterns. Several subtype-specific markers were identified, *e.g.* a unique substitution of T to A in the fifth position of the TATA box in A17, A27, A20 and B3. **Conclusion:** This study provides valuable insights into the genetic diversity of SRLV field strains in Poland, their phylogenetic relationships and their position in the recently established SRLV classification. Our results confirmed the existence of the ten subtypes listed and the readier emergence of new SRLV variants in mixed-species flocks.

Keywords: SRLV, maedi-visna virus, caprine arthritis encephalitis virus, phylogeny, LTR.

Introduction

Maedi-visna virus (MVV) and caprine arthritis encephalitis virus (CAEV) belong to the *Lentivirus* genus of the *Retroviridae* family. First isolated respectively from sheep and goats, MVV and CAEV were considered to be strictly host specific. Today, these viruses are no longer considered species-specific pathogens, since they cross the species barrier between goats and sheep very efficiently and infect both. Therefore, MVV and CAEV are considered as a single group referred to as small ruminant lentiviruses (SRLVs).

Small ruminant lentiviruses are transmitted vertically *via* infected milk and colostrum uptake and horizontally through respiratory secretion when infected animals are in close contact with uninfected ones (10). These lentiviruses induce a chronic multisystemic disease with inflammatory lesions in the mammary gland, lungs, joints and brain. Symptoms are observed only in one third of the infected animals and are typically pneumonia, arthritis (enlargement of the carpal joints), mastitis or encephalitis (15). The retroviral SRLVs are characterised by a single-stranded RNA genome, which, once retrotranscribed by the viral reverse transcriptase (RT) enzyme, integrates into the

host genome in the provirus form. The provirus can then undergo a productive replicative cycle or remain hidden or dormant inside the cell in a state called latency. The genome of SRLVs is comprised of three structural genes encoding the group-specific antigens (*gag*), the polymerase (*pol*) and envelope (*env*) genes, and accessory genes including *vpr-like* (formerly *tat*), *rev* and *vif*, which have regulatory functions. The *pol* gene encodes the reverse transcriptase (RT), protease (PR) and integrase (IN) enzymes involved in replication and DNA integration. The *gag* gene encodes internal structural proteins, which are the matrix (MA), capsid (CA) and nucleocapsid (NC) proteins. The *env* gene encodes surface and transmembrane glycoproteins. The proviral DNA of SRLVs is flanked by non-coding sequences known as long terminal repeats (LTRs). The LTR is subdivided into three regions, U3, R and U5, and contains principal regulatory sequences such as promoter and enhancer elements, including host transcription factor binding sites required for the replication of the virus and viral gene expression (5).

Based on the *gag* and *pol* nucleotide sequences, the current SRLV phylogeny consists of five groups (A–E) which are further divided into multiple subtypes. Group A SRLVs comprise the prototypic strain K1514 (A1 subtype), which was the first lentivirus strain characterised. This group is the most heterogeneous and holds at least twenty-seven subtypes (A1–A27), while group B is subdivided into five subtypes (B1–B5) containing the prototypic Cork CAEV strain belonging to subtype B1. Groups A and B are widespread, while groups C, D and E are restricted to certain geographical areas. Group C was detected only in Norway, group D was detected in Switzerland and Spain, and group E, which comprises two subtypes (E1 and E2), was isolated only in Italy (5, 10). The *env* gene and LTRs show a high level of variability, while the *pol* and *gag* genes are relatively conserved, and it is these which are mainly used for phylogenetic analyses of SRLVs. However, the phylogenetic analysis of variable regions is more informative than that of conservative regions (24). Moreover, the highest level of phylogenetic information for SRLVs seems to be accumulated in LTR sequences rather than in the other genomic regions of SRLVs (29). In Poland, SRLV infections are widespread, with herd-level seroprevalence of 33.3% and 71.9% in sheep and goats, respectively (8, 9). Previous studies on the genetic diversity of SRLVs from Poland provided evidence for the presence of two groups (A and B) and 12 subtypes within SRLVs. Polish SRLVs isolated from sheep and goats belonged to subtypes B1, B2, A1, A5, A12, A13, A16–A18, A23, A24 and A27 (17). The phylogenetic analysis of Polish strains has so far been performed based on *gag* gene or *gag* and *env* gene sequences; however, phylogenetic analysis of the LTR has never been performed. Thus, the aim of this study was to extend genetic/phylogenetic analysis of the previously identified SRLV Polish strains by adding LTR sequences. This study provides additional insight into the genetic diversity of SRLV field strains in Poland, their phylogenetic relationships, and their position in the recently established SRLV classification.

Material and Methods

A total of 112 samples analysed in this study were collected from six different voivodeships of Poland over the 13-year period of 2008–2021. Previous studies made 61 LTR sequences available (16, 17, 18) while this study obtained 51 new LTR sequences. The samples originated from 51 sheep and 61 goats from 19 flocks of a single animal species (sheep or goats) or mixed species (goats and sheep). Anticoagulated blood was used as a source of peripheral blood leukocytes (PBL) which were isolated by the standard protocol (16). The phylogenetic affiliation of all samples was previously determined on the basis of the *gag* and/or *env* sequences. The sequences were assigned to SRLV subtypes A5, A12, A13, A16, A17, A18, A23, A24, A27, B1 and B2 (Table 1).

The genomic DNA was extracted from the original biological samples (leukocyte pellets) using a NucleoSpin Blood Quick Pure Kit (Macherey-Nagel, Düren, Germany), according to the manufacturer's recommendations. The quality and quantity of DNA were evaluated in a nanophotometer (Implen, Munich, Germany). All methods were performed in accordance with the relevant guidelines and regulations. Specifically, blood collection was approved (no. 37/2016) by the Local Ethics Committee on Animal Testing at the University of Life Sciences in Lublin, Poland. The LTR U3-R region was amplified using a nested PCR protocol already described (25). The primer pair of LTREFW (5′-ACTGTCAGGRCAGAGAACARATGCC-3′) and LTRERV (5′-CTCTCTTACCTTACTTCAGG-3′) was used in the first round of PCR, and the LTRIFW (5′-AAGTCATGTAKCAGCTGATGCTT-3′) and LTRIRV (5′-TTGCACGGAATTAGTAACG-3′) pair was used in the second round. The PCR products were analysed by electrophoresis on a 1.5% agarose gel. Then, the PCR products were purified using NucleoSpin Gel and PCR Clean-up (Marcherey-Nagel) and directly sequenced on a 3730xl DNA Analyzer using the BigDye Terminator v3.1 Cycle Sequencing Kit (both products of Applied Biosystems, Foster City, CA, USA). The obtained SRLV sequences were edited and analysed using Geneious Pro 5.3 software (Biomatters, Auckland, New Zealand). All new sequences reported in this study were submitted to GenBank under accession numbers ON637539–ON637623. The consensus sequences were aligned to each other and with previously published SRLV sequences of genotypes A–E. Multiple sequence alignment was performed using multiple sequence comparison by log-expectation (MUSCLE).

Table 1. Information on the samples characterised in the present study

Sample	Flock	Type of flock	Voivodeship	Host	Strain	Subtype *gag*	Subtype *env*	Proposed subtype LTR	GenBank accession number
1	1	Single-species	Małopolskie	sheep	0016	A13	A13	unassigned	ON637582
2	2	Single-species	Małopolskie	sheep	209	A1	A1	N/A	N/A
3	3	Single-species	Podkarpackie	goat	90472	A1	A1	A1	ON637581
4				goat	90960	A1	A1	N/A	N/A
5				goat	90281	A1	A1	N/A	N/A
6	4	Single-species	Lubelskie	sheep	11	B2	B2	N/A	N/A
7				sheep	2437	B2	B2	N/A	N/A
8				sheep	4106	B2	B2	N/A	N/A
9				sheep	4084	B2	B2	N/A	N/A
10	5	Mixed-species	Lubelskie	sheep	Tryk2	A12	A12	A12	ON637584
11				sheep	5	A12	N/A	N/A	N/A
12				sheep	Tryk6	A12	A12	A12	ON637583
13				sheep	10	A12	N/A	N/A	N/A
14				sheep	13	B2	A12/B2	N/A	N/A
15				sheep	15	A12	A12	N/A	N/A
16	6	Single-species	Małopolskie	goat	2461	B1	B1	N/A	N/A
17				goat	2462	B1	B1	N/A	N/A
18				goat	2466	B1	B1	N/A	N/A
19	7	Single-species	Podkarpackie	goat	2991	B1/A1	B1	N/A	N/A
20				goat	2993	A16	A16	A16	ON637585
21	8	Single-species	Podlaskie	goat	1202	A12/A	B2/A12	A16	ON637586
22				goat	1203	A	B1/A1/A12	N/A	N/A
23	9	Single-species	Mazowieckie	goat	3085	A17	A17	A17	ON637551
24				goat	1561	A17	A17	A17	ON637545
25				goat	8370	A17	A17	A17	ON637608
26				goat	5616	A17	A17	A17	ON637546
27				goat	0042	A17	A17	A17	ON637541
28				goat	8344	A17	A17	A17	ON637553
29				goat	5675	A17	A17	A17	ON637554
30				goat	1485	A17	A17	A17	ON637547
31				goat	1580	A17	A17	A17	ON637548
32				goat	5686	A17	A17	A17	ON637544
33				goat	8172	A17	A17	A17	ON637549
34				goat	5621	A17	A17	A17	ON637543
35				goat	6909	A17	A17	A17	ON637542
36				goat	9431	A17	A17	A17	ON637540
37				goat	5654	A17	A17	A17	ON637550
38	10	Mixed-species	Małopolskie	sheep	6922	A13	N/A	A13	ON637573
39				sheep	8063	A13	N/A	A13	ON637576
40				sheep	9179	A13	N/A	A13	ON637575
41	11	Mixed-species	Małopolskie	sheep	7041	A18	N/A	A18	ON637589
42				sheep	0090	A18	N/A	A18	ON637588
43				sheep	7010	A18	N/A	N/A	N/A
44				sheep	7020	A18	N/A	A18	ON637587
45	12	Mixed-species	Małopolskie	sheep	6981	A13	N/A	A13	ON637580
46				sheep	9155	A13	N/A	A13	ON637578
47				sheep	1406	A13	N/A	A13	ON637579
48				sheep	4742	A18	N/A	N/A	N/A
49	13	Single-species	Podkarpackie	sheep	1304	A13	N/A	A13	ON637574
50				sheep	1911	A13	N/A	A13	ON637577
51				goat	5826	A5	A5	A5	ON637566
52				goat	5819	A5	A5	A5	ON637568
53				goat	4742	A5	A5	A5	ON637569
54				goat	7592	A5	A5	A5	ON637567
55				goat	5962	A5	A5	A5	ON637565
56				goat	5994	A5	A5	A5	ON637570
57				goat	6038	A5	A5	A5	ON637572
58				goat	3038	A5	A5	A5	ON637571
59				goat	5870	A5	A5	A5	ON637564
60	14	Mixed-species	Wielkopolskie	goat	7134	A12	A12	A12	ON637562
61				goat	7102	A12	A12	A12	ON637561
62				goat	6808	A12	A12	A12	ON637560
63				goat	7096	A12	A12	A12	ON637559
64				goat	8891	A12	A12	A12	ON637557
65				goat	7219	A12	A12	A12	ON637558
66				sheep	0334	B2	N/A	A5	ON637563

67				goat	3540	B2	B2	A5	ON637599
68				goat	0580	B2	B2	A12	ON637552
69				goat	0788	B2	N/A	A5	ON637605
70				goat	9509	A12	A12	A12	ON637606
71				goat	9510	A12/B2	A12	A12	ON637602
72				goat	3533	A12	A12	A12	ON637609
73				goat	3535	A12/B2	A12	A5	ON637607
74				goat	0599	B2	A	A5/A12	ON637603/ON637604
75				goat	8699	A12	A12	A12	ON637601
76				sheep	1	A12	A12	A24	ON637590
77	15	Mixed-species	Lubelskie	sheep	3	A12	A12	unassigned	ON637594
78				sheep	14	A12/B2	B2	A24	ON637600
79				sheep	20	B2	B2	A5	ON637598
80				sheep	21	B2	B2	N/A	N/A
81				sheep	29	B2	N/A	N/A	N/A
82				sheep	4	A12	N/A	A24	ON637591
83				sheep	6	A12	N/A	A24	ON637597
84				sheep	13	A12	A12	unassigned	ON637595
85				sheep	16	A12	A12	unassigned	ON637596
86				sheep	33	A12	A12	unassigned	ON637593
87				sheep	12	A12	A12	N/A	N/A
88				sheep	40	A12	A12	unassigned	ON637592
89				sheep	3225	A24	N/A	A24	ON637623
90	16	Mixed-species	Podkarpackie	sheep	3188	A24	A24	A24	ON637622
91				sheep	3249	A24	A24	A24	ON637621
92				sheep	3201	A24	A24	N/A	N/A
93				sheep	3275	B2	A23	A23	ON637610
94				sheep	3691	B2	B	unassigned	ON637612
95				sheep	4018	A23	B	unassigned	ON637613
96				sheep	2590	A23/B2	A23	A23	ON637614
97				sheep	4315	A23/B2	A23	A23	ON637615
98	17	Mixed-species	Podkarpackie	sheep	1622	A23	A23	A23	ON637616
99				goat	8046	A5	N/A	A5	ON637617
100				goat	8039	A5	N/A	A5	ON637618
101				goat	8008	A5	A5	A5	ON637619
102				goat	9692	A5	A5	A5	ON637620
103				goat	1318	A5	N/A	A5	ON637611
104				sheep	9855	B2	B	A18	ON637539
105	18	Mixed-species	Podkarpackie	goat	4464	A	N/A	A18	ON637556
106				sheep	5023	A24	A	A5	ON637555
107				goat	goat2	A27	N/A	A27	OM517095
108				goat	goat3	A27	N/A	A27	OM517100
109	19	Single-species	Warmińsko-Mazurskie	goat	goat4	A27	N/A	A27	OM517105
110				goat	goat5	A27	N/A	A27	OM517114
111				goat	goat6	A27	N/A	A27	OM517117
112				goat	goat7	A27	N/A	A27	OM517122

LTR – long terminal repeat; N/A– not available

Model testing was performed to select the best-fit evolutionary model based on the Bayesian information criterion (BIC) and Akaike information criterion (AIC). As best accorded with the results, the Kimura 2-parameter model with gamma distribution (+G) and five rate categories was used to construct phylogenetic trees using the neighbour-joining method (NJ), the maximum-likelihood method (ML), and the unweighted pair group method with arithmetic mean (UPGMA). These methods contrast a statistical approach (in ML) with distance-based methods (in UPGMA and NJ). The statistical confidence of the topologies was evaluated by nonparametric bootstrap analysis with 1,000 iterations. Multiple alignment, model testing, tree building and pairwise genetic distances were calculated with MEGA 7 software (11). The RDP4 application was used to perform a recombination analysis (12). The software used seven primary methods: RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan and 3Seq. Putative recombinant events were considered significant when $P \leq 0.01$ was observed for the same event using five or more methods.

Results

The LTR fragment from 15 sheep and 7 goat samples could not be amplified. The phylogenetic analyses of the LTR region of 54 goat and 36 sheep sequences from Polish SRLV strains and 39 reference strains produced three phylogenetic trees with topologies that were almost identical when using the three different methods ML, NJ and UPGMA. The unrooted phylogenetic trees of LTR sequences are shown in Figs 1–3.

Three main clusters were observed in the phylogenetic trees with clear separation between group A, group B and group E according to the reference sequences. Group A represented MVV-like isolates, while group B represented CAEV-like isolates. Group E comprised two subtypes (E1 and E2) isolated in Italy.

Group B consisted of subtype B1, B2 and B3 strains. The LTR sequence of the Norwegian strain 1GA of group C clustered with strains belonging to group B. This strain was the most closely related to the Volterra and Fonni strains representing subtype B3; however, the mean genetic distance between these strains was 30%. Furthermore, the It009 strain, a member of subtype A20, was most closely related to subtype B1 sequences (mean genetic distance 18.4%). The highly divergent group A consisted of many branches. Reference sequences representing subtypes A1, A2/A3, A4, A8 and A19 formed separate clusters. The Polish strains were clearly separated from the strains belonging to groups B and E. Our analysis showed that all Polish sequences were affiliated to group A and grouped in at least 10 clusters. The sequences belonged to the known subtypes A1, A5, A12, A13, A16, A17, A18, A23, A24 and A27. In particular, the LTR sequence of the #3g90472 strain was closely related to the K1514 and LV-1 strains (mean genetic distance 1.4%) and clustered in subtype A1. The LTR sequences of the Polish #13g4742, #13g5994, #13g5826, #13g7592, #13g5819, #13g6038, #13g5962, #13g3038, #17g8008, #17g9692, #15g3540, #15g0788, #14s0334, #13g5870, #17g1318, #17g8039, #15g3535, #15g0599, #17g8046, #15s20 and #18s5023 strains clustered within subtype A5 showing a mean genetic distance of 3.3%. Affiliation of this cluster was supported by high bootstrap values of 89% by ML, 100% by NJ and 97% by UPMGA. The sequences of the Polish #15g8699, #15g9509, #15g0580, #5sTryk6, #5sTryk2, #15g0599(2), #15g9510, #15g3533, #14g7096, #14g8891, #14g7219, #14g6808, #14g7102 and #14g7134 strains belonged to subtype A12; however, the existence of this subtype was not confirmed by high bootstrap values. The intra-subtype similarity of sequences belonging to this subtype was 5.0%. The LTR sequences of the #6922, #9179, #8063, #6981, #9155, #1406, #1304 and #1911 strains originating from sheep from flocks no. 10 and 12 clustered together (mean genetic distance 1.6%) and formed subtype A13, supported by high bootstrap values of 99% by NJ and 99% by UPMGA. Strain #7g2993 was closely related to strain #8g1202 (nucleotide distance 9.6%) and together with this strain formed a separate cluster named A16, supported by high bootstrap values of 62% by ML, 79% by NJ and 89% by UPMGA. The sequences originating from goats from flock 9 (#9431, #8344, #0042, #5675, #6909, #5621, #5616, #1561, #3085, #5654, #5686, #1580, #8172, #1485 and #8370) clustered in subtype A17 (mean nucleotide distance 3.3%). Affiliation of this subtype was supported by high bootstrap values of 89% by ML, 99% by NJ and 99% by UPMGA. Sequences of this region of strains #1, #4, #16, #14, #3225, #3249 and #3188 originating from sheep from flocks 15 and 16 formed a separate cluster, subtype A24, supported by high bootstrap values of 69% by NJ and 80% by UPMGA. The sequences of the #goat2, #goat3, #goat4, #goat5, #goat6, and #goat7 strains clustered in subtype A27, supported by high bootstrap values of 91% by ML, 99% by NJ and 100% by UPMGA. The LTR sequences of strains #17s2590, #17s4315, #17s1622 and #17s3275 clustered in subtype A23, while those of strains #18s9855, #18g4464, #10s7041, #11s7020 and #11s0090 clustered in subtype A18.

The affiliation of the A23 subtype was supported by high bootstrap values of 98% by NJ and 83% by UPGMA while affiliation of the A18 subtype was supported by high bootstrap values of 71% by NJ and ML. The LTR sequences of strains #3, #16, #40, #13 and #33 originating from a sheep from flock 15 formed a unique cluster within group A, which was supported by high bootstrap values of 59% by ML, 77% by NJ and 95% by UPMGA.

Most of the analysed strains belonged to the same subtype by the indication of the *gag*, *env* and LTR genomic regions (Table 1). However, the affiliations of some strains were different. In particular, strain #0016 belonged to subtype A13 on the basis of its *gag* and *env* sequences, while on the basis its LTR this strain formed a separate cluster within group A, different from other subtypes. The LTR sequence of this strain was closely related to that of a strain belonging to subtype A12 (mean genetic distance of 10.5%). Strains #14s0334, #15g3540, #15g0788 and #15s20 had *gag/env* sequences which classified the strains as B2 but had LTR sequences which assigned them to subtype A5. Strain #15g0580, which had previously been classified as B2, affiliated to subtype A12 when heeding its LTR sequence. Furthermore, the *gag/env* sequences of strains #15s1, #15s4 and #15s6 were classified as A12 while the LTR sequences of these strains clustered in subtype A24. Strains #15s3, #15s13, #15s16, #15s33 and #15s40, of which the *gag/env* sequences indicated them to be in A12, revealed LTR sequences placing them in a new cluster within group A. The LTR sequence of strain #17s3275 confirmed its affiliation to subtype A23 predicated on its *env* sequence, although the *gag* sequence of this strain belonged to subtype B2. The *gag* sequence of strain #18s5023 matched subtype A24, the *env* sequence was not assigned to any known subtype, while the LTR sequence of this strain corresponded to subtype A5. Discrepancies in affiliation also occurred in strain #17s3691 (B2/B/A), #17s4018 (A23/B/A), #15g0599 (B2/A/A5/A12), #18s9855(B2/B/A18) and #18g4464 (A/A18). No clear affiliation was also observed in the case of strains isolated from animals co-infected with different subtypes. When the *gag* sequences were analysed, subtypes A12 and B2 were found in goat 3535 and sheep 14 from flock 15, while the LTR sequences isolated from these animals clustered in subtypes A5 and A24. Furthermore, goat 1202 from flock 8 co-infected with strains A12 and B2 (*env/gag* sequences) was apparently infected with a subtype A16 strain when the LTR sequence was consulted.

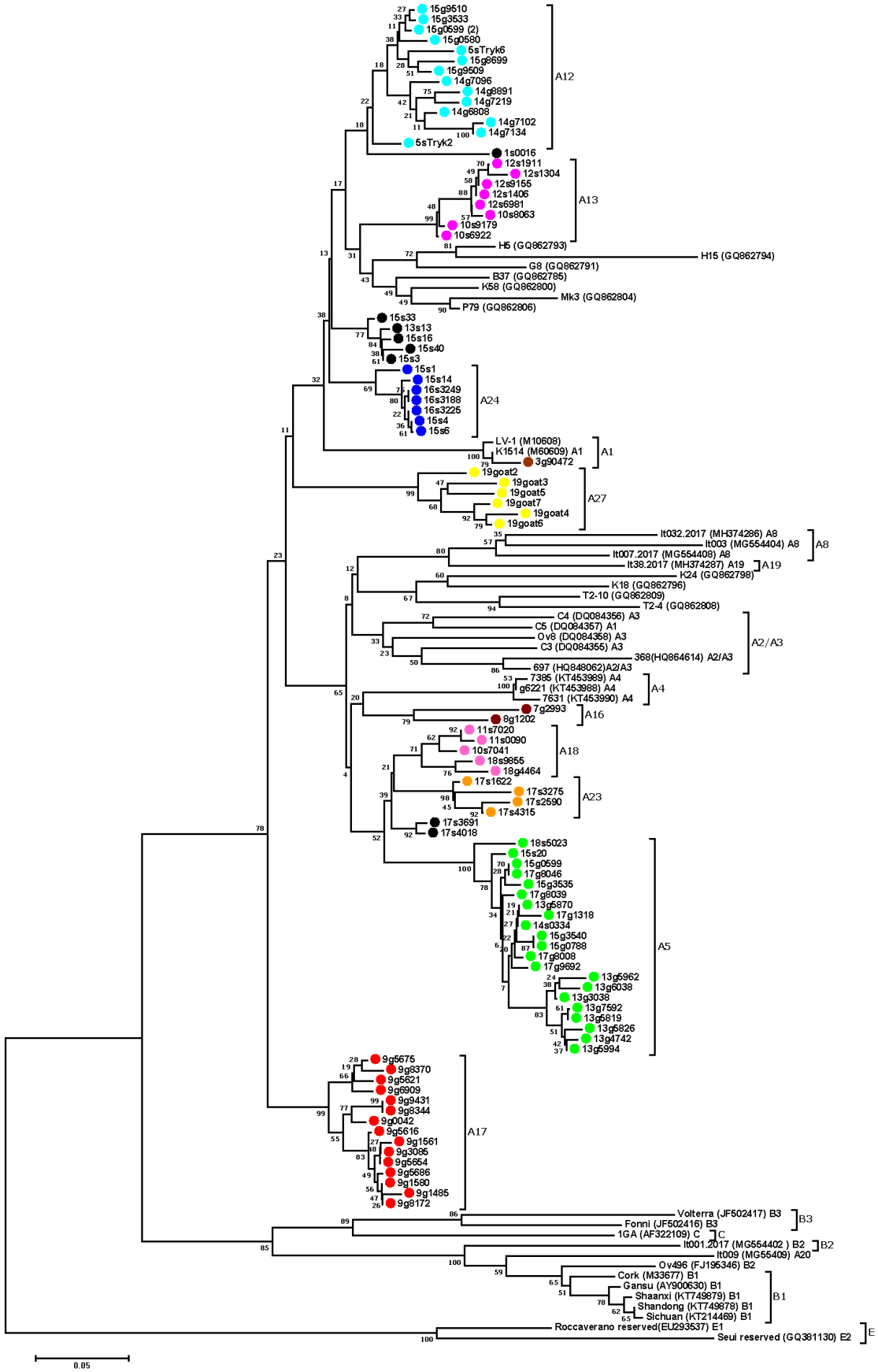


Fig. 1. Neighbour-joining phylogenetic tree based on the alignment of the long terminal repeat fragment. Sequences from this study are labelled by coloured circles (red – subtype A17; green – subtype A5; orange – subtype A23; pink – subtypes A18 and A13; brown – subtypes A16 and A1; yellow – subtype A27; blue – subtype A24; light blue – A12; black – unassigned) and their names are preceded by the flock origin and the animal species (s – sheep; g – goat). Reference small ruminant lentivirus strains are shown by name followed by GenBank accession number and subtype. Numbers at the branches indicate the percentage of bootstrap values obtained from 1,000 replicates

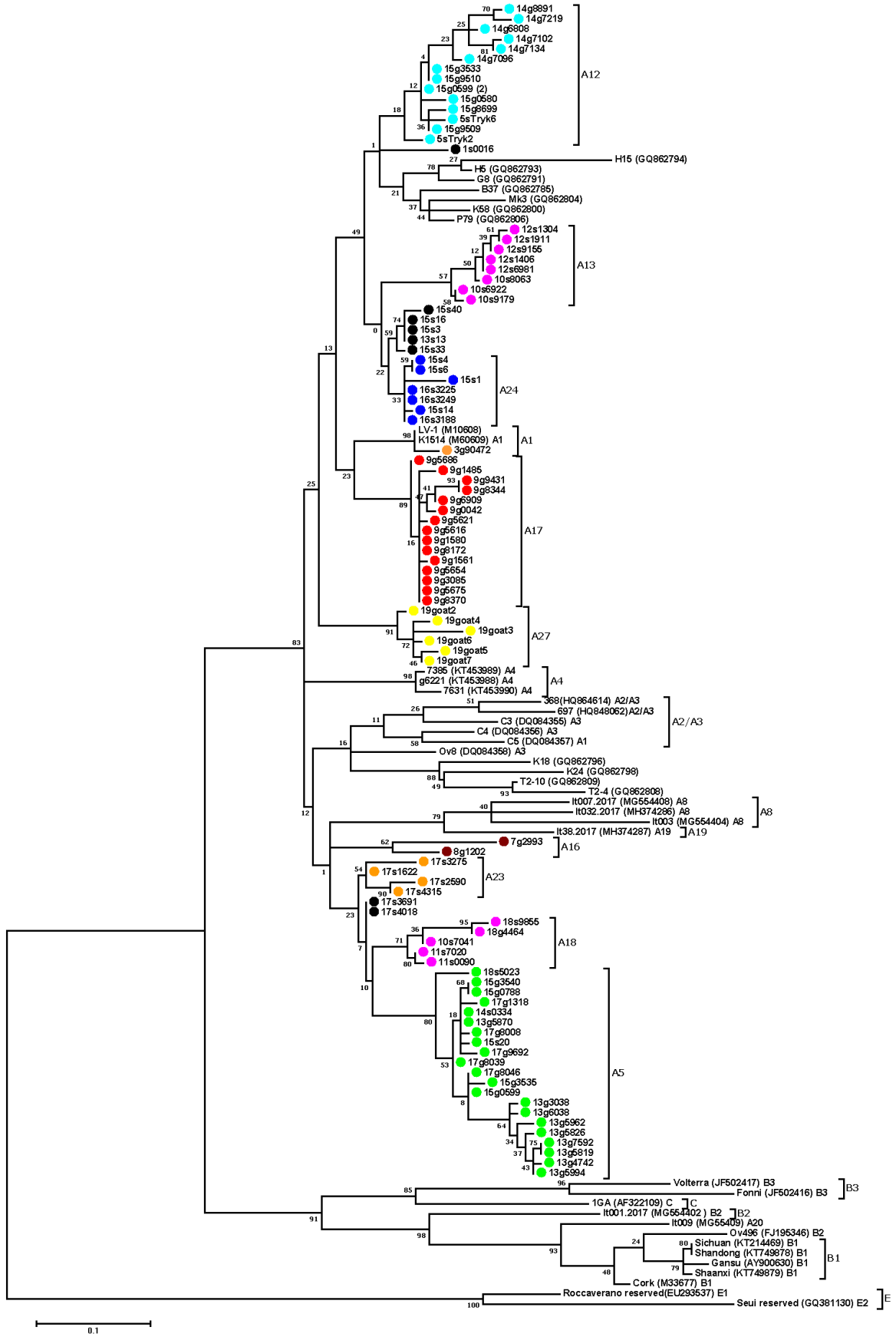


Fig. 2. Maximum-likelihood phylogenetic tree based on the alignment of the long terminal repeat fragment. Sequences from this study are labelled by coloured circles (red – subtype A17; green – subtype A5; orange – subtypes A23 and A1; pink – subtypes A18 and A13; brown – subtype A16; yellow – subtype A27; blue – subtype A24; light blue – A12; black – unassigned) and their names are preceded by the flock origin and the animal species (s – sheep; g – goat). Reference small ruminant lentivirus strains are shown by name followed by GenBank accession number and subtype. Numbers at the branches indicate the percentage of bootstrap values obtained from 1,000 replicates

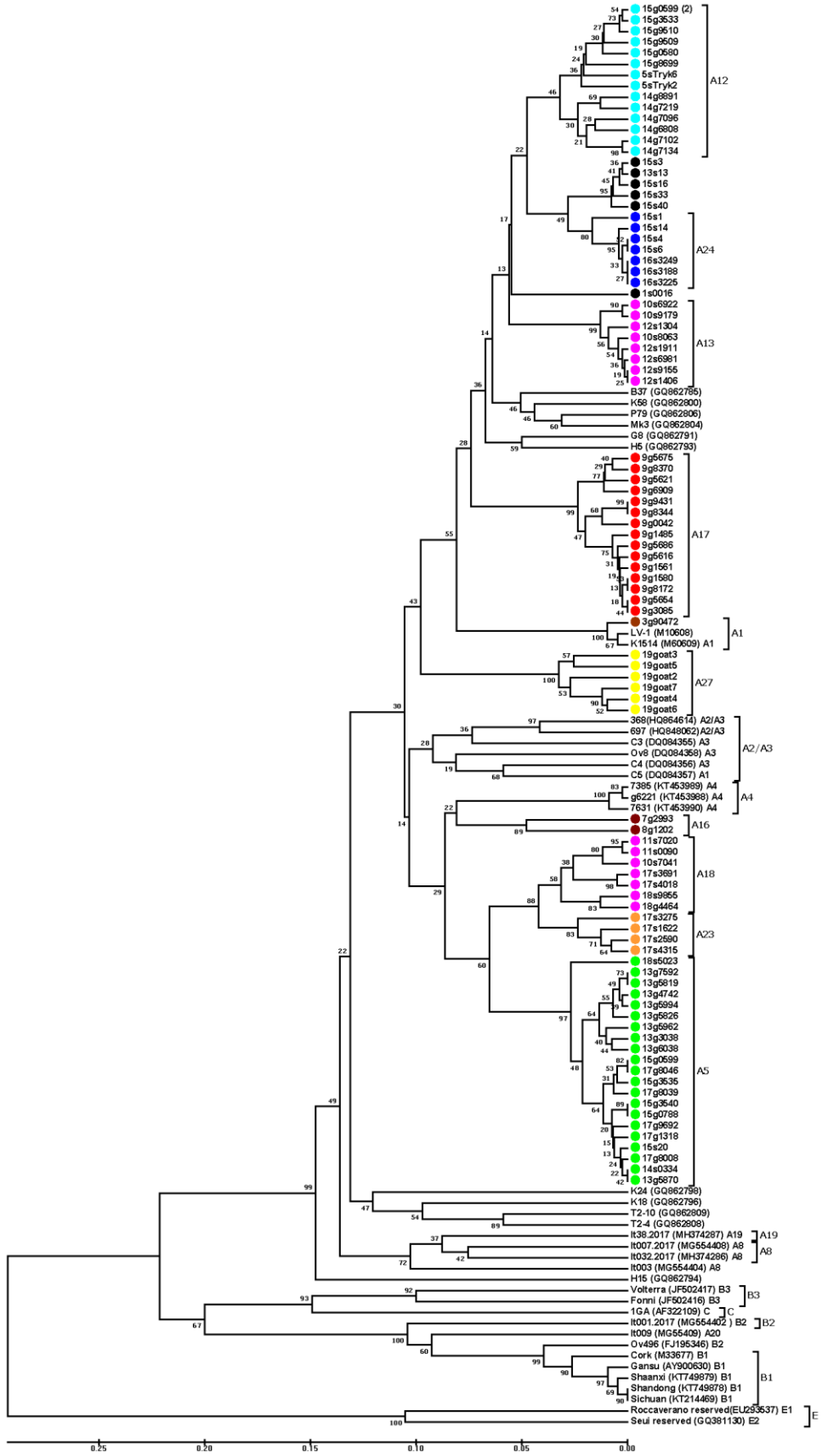


Fig. 3. Unweighted pair group with arithmetic means phylogenetic tree based on the alignment of the long terminal repeat fragment. Sequences from this study are labelled by coloured circles (red – subtype A17; green – subtype A5; orange – subtype A23; pink – subtypes A18 and A13; brown – subtypes A16 and A1; yellow – subtype A27; blue – subtype A24; light blue – A12; black – unassigned) and their names are preceded by the flock origin and the animal species (s – sheep; g – goat). Reference small ruminant lentivirus strains are shown by name followed by GenBank accession number and subtype. Numbers at the branches indicate the percentage of bootstrap values obtained from 1,000 replicates

The WebLogo corresponding to generated alignments of LTR sequences representing SRLV subtypes A1, A2/A3, A4, A5, A8, A12, A13, A16, A17, A18, A19, A23, A24, A27, B1, B2, B3, C, E1 and E2 and the inferred U3/R and R/U5 boundaries are shown in Fig. 4. The LTR region analysed in this study contained two AP-1 sites, one AML(vis), one AP-4 motif and one TATA box in the U3 region and the AATAAA motif in the R region (Figs 4 and 5). This representation of the alignment revealed that the U5 region is generally more conserved than the R and U3 regions. The AP-4 site, the TATA box and the polyadenylation site (AATAAA) were the most conserved elements between all strains belonging to group A, B, C and E SRLVs. Our results revealed that strains of subtypes A17, A20 A27 and B3 had a unique T to A substitution in the fifth position of the TATA box. The CAEV-like sequences did not have an AML sequence close to the TATA box present in any MVV-like sequences. The AP-1 sites revealed rather group-specific conservation. Our results also showed that the LTR sequence reflected subtype-specific patterns of sequence diversity (Fig. 5). The GAGAAGCTTTG and TAAGAGCTTTG insertions next to the TATA box were only found in sequences belonging to subtypes B1 and B2, respectively, and CTTGCTACT and TCAGACGCT insertions were found exclusively in sequences belonging to groups C and E, respectively. Sequences from subtype A1 had a unique TCGAAG GAAAGA insertion in the R region, while the sequences from subtypes A17, B3 and C had respective unique CCGAAGGAAAG, TCGAAGGAAAGA and TCGAAGGAAAGAG insertions. The sequence from subtype A4 had a unique GCTTTGCC insertion, while the sequences from subtype A8 had a unique CTGGTCGC insertion close to the polyA site. All subtype A20, B1, and B2 sequences had a unique GTACCGAGACCT insertion located downstream of the polyA site. Two insertions located upstream of polyA, GATTGCC and TGCCGAGTG, were identified only in sequences belonging to subtypes B1 and B2.

Some indications of recombination events were observed; however, these possible recombination events were confirmed only by two out of seven methods used in this study. On the basis of LTR alignment, two putative recombination events were detected. The MaxChi and SiScan methods detected a recombination event in almost all subtype A12 strains. In this recombination event, the beginning and end breakpoints were located at 88 and 264 nucleotides in alignments and the major and minor parents were the Turkish #Mk3 strain and an unknown strain, respectively (Fig. 6a). The MaxChi and 3Seq methods detected a recombination event in all subtype A5 strains between positions 258 and 114 in alignment with #It007 (subtype A8) as the major parent and the unknown strain as the minor parent (Fig. 6b).

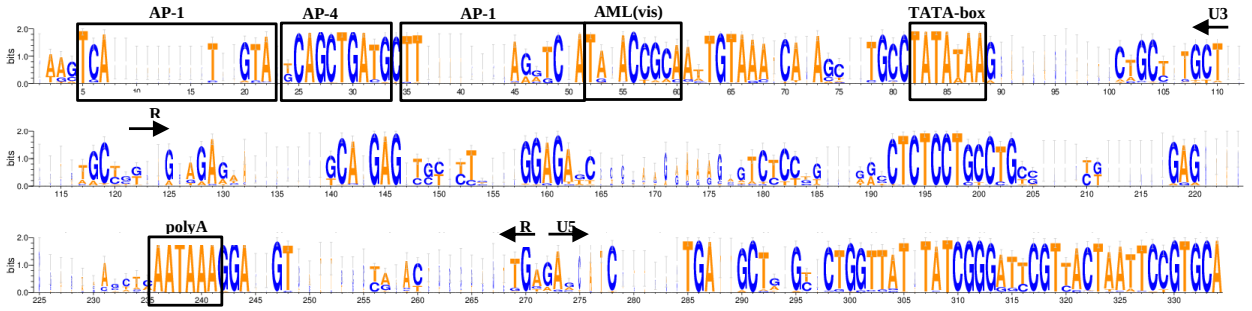


Fig. 4. WebLogo for alignment of long terminal repeat sequences of small ruminant lentivirus strains belonging to groups A, B, C and D. The heights of the letters are a measure of how well conserved the residues are. The boundaries between U3, R and U5 are indicated by arrows. Transcription factor AML (vis), AP-1, AP-4 TATA box and polyadenylation signal (poly A) binding sites are shown in boxes

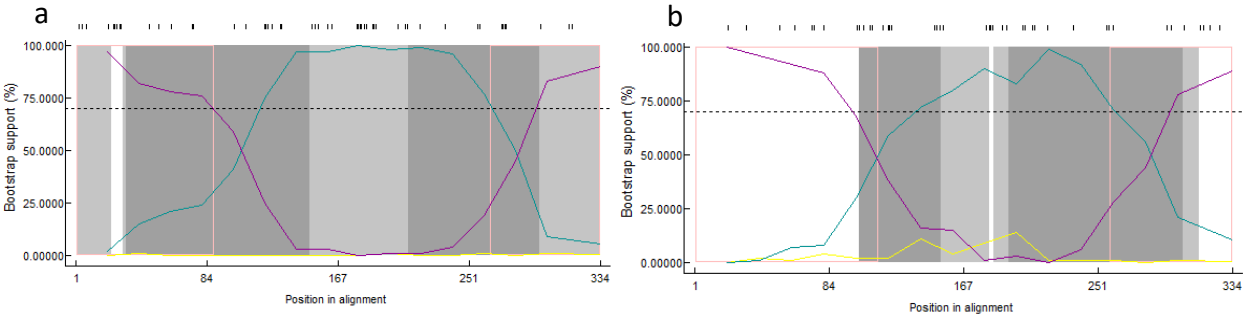
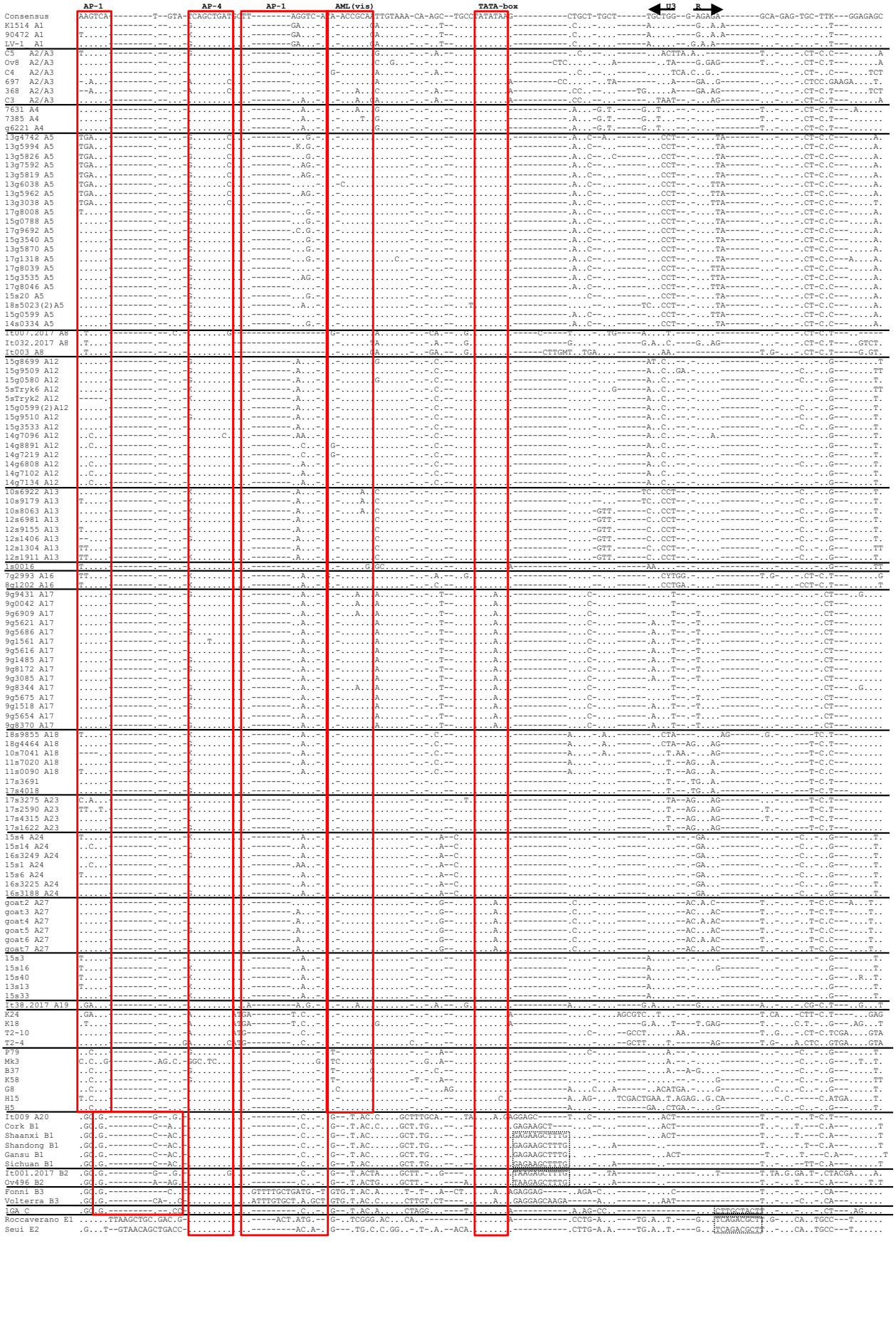


Fig. 6. The BootScan analysis of recombination in the alignments. The analysis was performed with the pairwise distance model with a window size of 200, step size of 20 and 1,000 bootstrap replicates by the RPD4 program. a – recombination even in almost all subtype A12 strains; b – recombination event in all subtype A5 strains


AP-1
AP-4
AP-1
AML(vis)
TATA-box
U3
R
Consensus
K1514 A1
90472 A1
LV-1 A1
C5 A2/A3
Ov8 A2/A3
C4 A2/A3
697 A2/A3
368 A2/A3
C3 A2/A3
7631 A4
7385 A4
q6221 A4
13g4742 A5
13g5994 A5
13g5826 A5
13g7592 A5
13g5819 A5
13g6038 A5
13g5962 A5
13g3038 A5
17g8008 A5
15g0788 A5
17g9692 A5
15g3540 A5
13g5870 A5
17g1318 A5
17g8039 A5
15g3535 A5
17g8046 A5
15s20 A5
18s5023(2)A5
15g0599 A5
14s0334 A5
It007.2017 A8
It032.2017 A8
It003 A8
15g8699 A12
15g9509 A12
15g0580 A12
5sTryk6 A12
5sTryk2 A12
15g0599(2)A12
15g9510 A12
15g3533 A12
14g7096 A12
14g8891 A12
14g7219 A12
14g6808 A12
14g7102 A12
14g7134 A12
10s6922 A13
10s9179 A13
10s8063 A13
12s6981 A13
12s9155 A13
12s1406 A13
12s1304 A13
12s1911 A13
1s0016
7g2993 A16
8g1202 A16
9g9431 A17
9g0042 A17
9g6909 A17
9g5621 A17
9g5686 A17
9g1561 A17
9g5616 A17
9g1485 A17
9g8172 A17
9g3085 A17
9g8344 A17
9g5675 A17
9g1518 A17
9g5654 A17
9g8370 A17
18s9855 A18
18g4464 A18
10s7041 A18
11s7020 A18
11s0090 A18
17s3691
17s4018
17s3275 A23
17s2590 A23
17s4315 A23
17s1622 A23
15s4 A24
15s14 A24
16s3249 A24
15s1 A24
15s6 A24
16s3225 A24
16s3188 A24
goat2 A27
goat3 A27
goat4 A27
goat5 A27
goat6 A27
goat7 A27
15s3
15s16
15s40
13s13
15s33
It38.2017 A19
K24
K18
T2-10
T2-4
P79
Mk3
B37
K58
G8
H15
H5
It009 A20
Cork B1
Shaanxi B1
Shandong B1
Gansu B1
Sichuan B1
It001.2017 B2
Oy496 B2
Fonni B3
Volterra B3
1GA C
Roccaverano E1
Seui E2

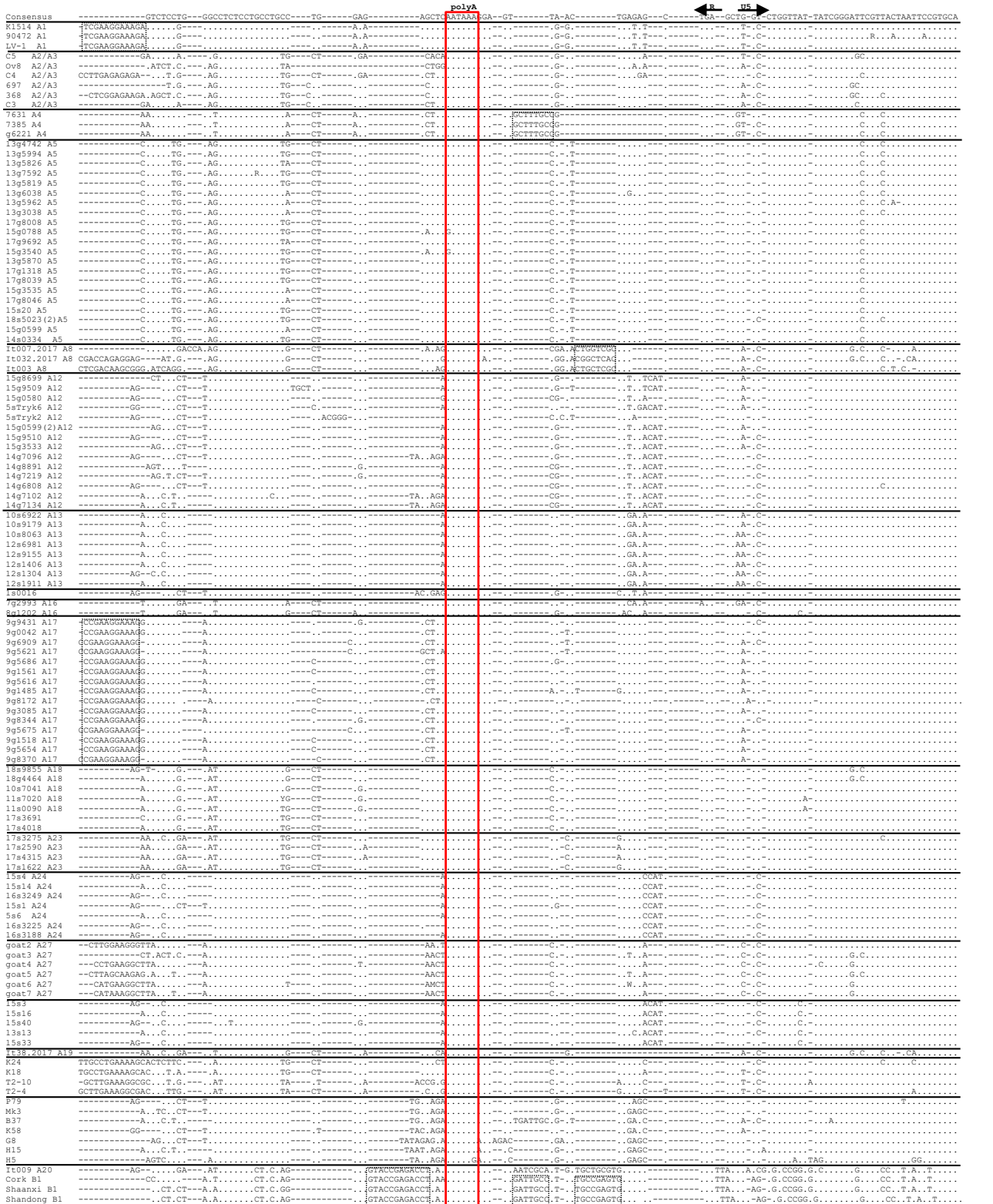


Fig. 5. Alignment of U3-R sequences of the long terminal repeat region from Polish small ruminant lentivirus (SRLV) strains. Sequences are aligned against the prototype K1514 and Cork strains representative of SRLV groups A and B, respectively. Dots indicate identity with Cork, and dashes represent gaps. Boundaries between U3, R and U5 are indicated by straight arrows. AP-1, AP-4 and AML (vis) motifs, the TATA box and polyadenylation signal (poly A) are marked by boxes

Discussion

The first attempt to establish the phylogeny of SRLVs was published in 1998 by Zanoni (29). At that time, SRLVs were phylogenetically classified into at least six clades, with no clear separation by host species or geographical origin (24, 29). The phylogenetic organisation of SRLVs proposed by Shah *et al.* (27) based on sequences of 1.2 kilobase (kb) *pol* and 1.8 kb *gag-pol* fragments sorted strains into four main groups, termed A–D. However, detection of these fragments with classifiable sequences often could not be achieved, most probably because of the high genetic heterogeneity of SRLV strains. Therefore, many other strains have been added to the phylogenetic tree using different fragments and regions. This led to the new classification that currently affiliates SRLVs into five groups (A–E) and at least 34 subtypes (A1–A27, B1–B5 and E1–E2). Previous studies using *gag* and *env* gene sequences revealed that Polish SRLVs isolated from sheep and goats belonged to subtypes B1, B2, A1, A5, A12, A13, A16–A18, A23, A24 and A27 (17). The LTR sequence analysis in this research elaborates the phylogeny of known SRLVs from this country and augments the genetic information for subtyping.

The phylogenetic trees constructed in this study showed the distribution of sheep- and goat-derived sequences in three out of five clades in the SRLV phylogeny (groups A, B and E). Group A was the most heterogeneous and consisted of many subtypes. Group B consisted of subtype B1, B2 and B3 strains, while group E comprised two subtypes isolated in Italy (E1 and E2). The phylogram of the LTR did not include group D because LTR sequences of strains belonging to this group were not available. Our results revealed that the LTR sequence of the Norwegian 1GA strain classified into group C by Shah *et al.* (27) clustered with strains in group B. This strain was the most closely related to the subtype B3 Volterra and Fonni strains. In this aspect, our results are in line with those of Mendiola *et al.* (13), who showed that the LTR sequences of the strains from genotypes C and B3 had a common root. Furthermore, in the phylogeny based on the *gag* or *gag-pol* sequence, the Norwegian 1GA strain also belonged to group B (14). Analysing the phylogenetic trees of SRLVs, it can be observed that the affiliation of some strains is inconstant and depends on the analysed fragment. For example, on the basis of the *pol* fragment, the Belgian VB.5.6 and LB.2.3 strains belong to the B5 subtype, while on the basis of the *gag-pol* region these strains belong to subtype B1 (14). In the *gag/pol* phylogenetic tree, the It009.2017 strain affililates to subtype A20, while in the *pol* tree it finds a place in subtype A1. Moving to the LTR sequence of this strain, it was most closely related to subtype B1 sequences, which may suggest that it is in fact a recombinant strain. A similar situation was observed in the case of B4 strains, which transpired to be A/B recombinant strains (26). In addition, sequences purported to belong to genotype D turned out to be genotype A strains, exhibiting divergence in the *pol* gene (21). Our results revealed that the Polish caprine and ovine LTR sequences clustered within group A and grouped into at least 10 clusters. Most of the Polish strains (78%) analysed in this study belonged to the same subtype when the *gag*, *env* and LTR genomic regions were scrutinised. Reina *et al.* (23) and Zanoni (29) also plotted phylogenetic trees on the basis of different regions (*gag*, *pol*, *env* and LTR) which had topologies compatible with one another. Phylogenetic analysis of the LTR fragment confirmed the high heterogeneity of Polish strains which clustered in subtypes A1, A5, A12, A13, A16–A18, A23, A24 and A27. The identical affiliations of most Polish strains emerging from analysis of three different regions confirmed the existence of these subtypes. The topologies of LTR trees were almost identical even when derived from three different methods (NJ, ML and UPGMA), which also increases the reliability of the obtained results. The agreement of the *gag*, *env*, and LTR phylogenies suggests that these genomic fragments have co-evolved. The high SRLV heterogeneity in Poland is probably the result of the absence of any SRLV control programmes, and the lack of the data such a programme might have provided may cause molecular and serological misdiagnosis of infections with some divergent strains possessing a high number of mutations.

Discrepancies in affiliation were nevertheless observed in 24 (21%) Polish strains, most of which (22 strains) came from mixed flocks where more than one SRLV genotype circulated. This confirms that mixed flocks are an excellent environment for the emergence of new SRLV variants. The different affiliation of LTR versus *gag/env* sequences of some Polish strains may suggest that these strains resulted from cross-species transmission and later evolved to adapt to a new host. Erhouma *et al.* (2) showed that after transmission from goats to wild ibexes *via* natural contact, important genetic differences were present mainly in the LTR region of proviral SRLV sequences. Circulation of different subtypes in one flock provides an opportunity for co-infection and recombination. On the basis of the LTR sequence, co-infection with strains belonging to subtypes A5 and A12 was evidenced in goat 0599 from flock 15. Double infections are a prerequisite for recombination between heterologous or homologous SRLV strains and an important source of genetic variability. In this article, some evidence of recombination was seen in strains from subtypes A5 and A12; however, possible recombination events were confirmed only by two out of seven methods.

Our results revealed that the LTR sequence of SRLV genotypes A, B, C and E have several homologous regions, such as those encompassing the TATA box, the polyA-signal, the AP-4 site, and some stretches of the U5 region. The observed conservation of these sites confirmed previous findings suggesting

that maintenance of these sites is essential to the transcriptional activity of the virus (3). Our results also showed that variations in the LTR sequence were subtype specific. Considerable sequence variation was also observed between different HIV-1 subtypes, and several subtype-specific sequence motifs were identified in the HIV-1 LTR region (22). Furthermore, it was revealed that the subtype-specific LTRs dictated a different replication rate and expression of HIV-1, which resulted in biological differences between subtypes (7). It is known that minor changes or rearrangements within the transcription factor binding sites in the LTR of retroviruses can have a significant impact on cell tropism and pathogenicity and affect viral fitness (7, 28). We suppose that the sequence variation between SRLV subtypes may also correspond to the different transcriptional activity of these strains, which may affect their phenotypes. The relationship between SRLV subtype, biological properties and pathogenicity is still unclear; however, in some cases a certain relationship can be observed. The presence of non-pathogenic SRLVs was reported in goats infected with viruses belonging to the A4 subtype, and it was evidenced that particular LTR mutations may explain their attenuated phenotype (1). The highly divergent SRLV E1 genotype was characterised in Italy as a low pathogenic caprine lentivirus, since it did not cause clinical symptoms in goats (6). Sequences of SRLVs isolated from Spanish sheep in an arthritis outbreak were assigned to subtype B2 (20), while proviral sequences isolated from Spanish sheep with neurological signs were assigned to genotype A2/A3 (4). Our results also revealed that strains representing subtypes A17, A20, A27 and B3 had a unique T to A substitution in the fifth position of the TATA box. The significance of this mutation is still unknown. An identical mutation (TATAAAA) was observed in subtypes E and J of HIV-1 strains (28). On the one hand, *in vitro* and *in vivo* transcription analyses indicated that mutation in the TATA box dramatically decreases HIV gene expression (28). On the other hand, Jeeninga *et al.* (7) revealed that mutation in the TATA box did not affect the activity of the subtype E LTR of HIV-1, which may indicate that the TATAAAA motif functions as an alternative TATA box. Additionally, it was also revealed that the sequences flanking the TATA box could contain multiple recognition sequences that may impart function to the TATA element (19).

In summary, the results of this study extend the current knowledge on the genetic diversity of SRLV field strains in Poland. Our results confirmed the existence of the A1, A5, A12, A13, A16–A18, A23, A24 and A27 subtypes and revealed that the phylogeny of a single genomic fragment does not necessarily reflect the phylogeny of the whole genome sequence because it reflects only a portion of it. Furthermore, we confirmed that mixed flocks favoured the emergence of new SRLV variants. Genetic analysis of LTR sequences revealed variability across SRLV subtypes. Several subtype-specific markers were identified; however, further studies should be conducted to assess whether these variations determine different transcription, viral fitness, and pathogenicity levels among distinct SRLV subtypes.

Conflict of Interests Statement: The authors declare that there is no conflict of interests regarding the publication of this article.

Financial Disclosure Statement: This research received no external funding.

Animal Rights Statement: None required.

References

1. Blatti-Cardinaux L., Sanjosé L., Zahno M.L., Zanoni R., Reina R., Bertoni G.: Detailed analysis of the promoter activity of an attenuated lentivirus. J Gen Virol 2016, 97, 1699–1708, doi: 10.1099/jgv.0.000489.
2. Erhouma E., Guiguen F., Chebloune Y., Gauthier D., Lakhal L.M., Greenland T., Mornex J.F., Leroux C., Alogninouwa T.: Small ruminant lentivirus proviral sequences from wild ibexes in contact with domestic goats. J Gen Virol 2008, 89, 1478–1484, doi: 10.1099/vir.0.2008/000364-0.
3. Glaria I., Reina R., Crespo H., de Andres X., Ramirez H., Biescas E., Perez M.M., Badiola J., Lujan L., Amorena B., de Andrés D.: Phylogenetic analysis of SRLV sequences from an arthritic sheep outbreak demonstrates the introduction of CAEV-like viruses among Spanish sheep. Vet Microbiol 2009, 138, 156–162, doi: 10.1016/j.vetmic.2009.03.002.
4. Glaria I., Reina R., Ramírez H., de Andrés X., Crespo H., Jauregui P., Salazar E., Luján L., Pérez M.M., Benavides J., Pérez V., Polledo L., García-Marín J.F., Riezu J.I., Borrás F., Amorena B., de Andrés D.: Visna/Maedi virus genetic characterization and serological diagnosis of infection in sheep from a neurological outbreak. Vet Microbiol 2012, 155, 137–146, doi: 10.1016/j.vetmic.2011.08.027.
5. Gomez-Lucia E., Barquero N., Domenech A.: Maedi-Visna virus: current perspectives. Vet Med (Auckl) 2018, 9, 11–21, doi: 10.2147/VMRR.S136705.
6. Grego E., Reina R., Lanfredini S., Tursi M., Favole A., Profiti M., Lungu M.M., Perona G., Gay L., Stella M.C., DeMeneghi D.: Viral load, tissue distribution and histopathological lesions in goats naturally and experimentally infected with the Small Ruminant Lentivirus Genotype E (subtype E1 Roccaverano strain). Res Vet Sci 2018, 118, 107–114, doi: 10.1016/j.rvsc.2018.01.008.
7. Jeeninga R.E., Hoogenkamp M., Armand-Ugon M., de Baar M., Verhoef K., Berkhout B.: Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type 1 subtypes A through G. J Virol 2000, 74, 3740–3751, doi: 10.1128/jvi.74.8.3740-3751.2000.
8. Junkuszew A., Dudko P., Bojar W., Olech M., Osiński Z., Gruszecki T.M., Kania M.G., Kuźmak J., Czerski G.: Risk factors associated with small ruminant lentivirus infection in eastern Poland sheep flocks. Prev Vet Med 2016, 127, 44–49, doi: 10.1016/j.prevetmed.2016.03.011.
9. Kaba J., Czopowicz M., Ganter M., Nowicki M., Witkowski L., Nowicka D., Szaluś-Jordanow O.: Risk factors associated with seropositivity to small ruminant lentiviruses in goat herds. Res Vet Sci 2013, 94, 225–227, doi: 10.1016/j.rvsc.2012.09.018.
10. Kalogianni A.I., Bossis I., Ekateriniadou L.V., Gelasakis A.I.: Etiology, Epizootiology and Control of Maedi-Visna in Dairy

Sheep: A Review. Animals 2020, 10, 616, doi: 10.3390/ani10040616.
11. Kumar S., Stecher G., Tamura K.: MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 2016, 33, 1870–1874, doi: 10.1093/molbev/msw054.
12. Martin D.P., Murrell B., Golden M., Khoosal A., Muhire B.: RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol 2015, 1, vev003, doi: 10.1093/ve/vev003.
13. Mendiola W.P.S., Tórtora J.L., Martínez H.A., García M.M., Cuevas-Romero S., Cerriteño J.L., Ramírez H.: Genotyping Based on the LTR Region of Small Ruminant Lentiviruses from Naturally Infected Sheep and Goats from Mexico. Biomed Res Int 2019, 4279573, doi: 10.1155/2019/4279573.
14. Michiels R., Adjadj N.R., De Regge N.: Phylogenetic Analysis of Belgian Small Ruminant Lentiviruses Supports Cross Species Virus Transmission and Identifies New Subtype B5 Strains. Pathogens 2020, 9, 183, doi: 10.3390/pathogens9030183.
15. Minguijón E., Reina R., Pérez M., Polledo L., Villoria M., Ramírez H., Leginagoikoa I., Badiola J.J., García-Marín J.F., de Andrés D., Luján L., Amorena B., Juste R.A.: Small ruminant lentivirus infections and diseases. Vet Microbiol 2015, 181, 75–89, doi: 10.1016/j.vetmic.2015.08.007.
16. Olech M., Kuźmak J.: Molecular Characterization of Small Ruminant Lentiviruses of Subtype A5 Detected in Naturally Infected but Clinically Healthy Goats of Carpathian Breed. Pathogens 2020, 9, 992, doi: 10.3390/pathogens9120992.
17. Olech M., Kuźmak J.: Molecular Characterization of Small Ruminant Lentiviruses in Polish Mixed Flocks Supports Evidence of Cross Species Transmission, Dual Infection, a Recombination Event, and Reveals the Existence of New Subtypes within Group A. Viruses 2021, 13, 2529, doi: 10.3390/v13122529.
18. Olech M., Kycko A., Kuźmak J.: Molecular Characterization of Small Ruminant Lentiviruses Isolated from Polish Goats with Arthritis. Viruses 2022, 14, 735, doi: 10.3390/v14040735.
19. Ou S.H., Garcia-Martínez L.F., Paulssen E.J., Gaynor R.B.: Role of flanking E box motifs in human immunodeficiency virus type 1 TATA element function. J Virol 1994, 68, 7188–7199, doi: 10.1128/JVI.68.11.7188-7199.1994.
20. Pérez M., Biescas E., Reina R., Glaria I., Marín B., Marquina A., Salazar E., Álvarez N., de Andrés D., Fantova E., Badiola J.J., Amorena B., Luján L.: Small ruminant lentivirus-induced arthritis: clinicopathologic findings in sheep infected by a highly replicative SRLV B2 genotype. Vet Pathol 2015, 52, 132–139, doi: 10.1177/0300985813519654.
21. 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, doi: 10.3390/v5041175.
22. Ramírez de Arellano E., Soriano V., Alcamil J., Holguín A.: New findings on transcription regulation across different HIV-1 subtypes. AIDS Rev 2006, 8, 9–16.
23. Reina R., Mora M.I., Glaria I., García I., Solano C., Luján L., Badiola J.J., Contreras A., Berriatua E., Juste R., Mamoun R.Z., Rolland M., Amorena B., de André D.: Molecular characterization and phylogenetic study of Maedi Visna and Caprine Arthritis Encephalitis viral sequences in sheep and goats from Spain. Virus Research 2006, 121, 189–198, doi: 10.1016/j.virusres.2006.05.011.
24. Rolland M., Mooney J., Valas S., Perrin G., Mamoun R.Z.: Characterisation of an Irish caprine lentivirus strain–SRLV phylogeny revisited. Virus Res 2002, 85, 29–39, doi: 10.1016/s0168-1702(02)00015-1.
25. Ryan S., Tiley L., McConnell I., Blacklaws B.: Infection of dendritic cells by the Maedi-Visna lentivirus. J Virol 2000, 74, 10096–10103, doi: 10.1128/jvi.74.21.10096-10103.2000.
26. Santry L.A., de Jong J., Gold A.C., Walsh S.R., Menzies P.I., Wootton S.K.: Genetic characterization of small ruminant lentiviruses circulating in naturally infected sheep and goats in Ontario, Canada. Virus Res 2013, 175, 30–44, doi: 10.1016/j.virusres.2013.03.019.
27. Shah C., Böni J., Huder J.B., Vogt H.R., Mühlherr J., Zanoni R., Miserez R., Lutz H., Schüpbach J.: 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 2004, 319, 12–26, doi: 10.1016/j.virol.2003.09.047.
28. van Opijnen T., Kamoschinski J., Jeeninga R.E., Berkhout B.: The human immunodeficiency virus type 1 promoter contains a CATA box instead of a TATA box for optimal transcription and replication. J Virol 2004, 78, 6883–6890, doi: 10.1128/JVI.78.13.6883-6890.2004.
29. Zanoni R.G.: Phylogenetic analysis of small ruminant lentiviruses. J Gen Virol 1998, 79, 1951–1961, doi: 10.1099/0022-1317-79-8-1951.