CrossMark

Comparison of *Schmallenberg* virus sequences isolated from mammal host and arthropod vector

Julia Kęsik-Maliszewska[1] · Aleksandra Antos[1] · Jerzy Rola[1] · Magdalena Larska[1]

Abstract

Schmallenberg virus (SBV) is the member of Peribunyaviridae family, which comprises pathogens of importance for human and veterinary medicine. The virus is transmitted only between animals and mainly by biting midges of the genus *Culicoides*. This study was performed in order to determine SBV genetic diversity and elucidate the host–vector adaptation. All three viral segments were analysed for sequence variability and phylogenetic relations. The Polish SBV strains obtained from acute infections of cattle, congenital cases in sheep, and from *Culicoides* midges were sequenced using Sanger and next-generation sequencing (NGS) methods. The obtained sequences were genetically similar (99.2–100% identity) to the first-detected strain BH80/11—4 from German cattle. The sampling year and origin of Polish sequences had no effect on molecular diversity of SBV. Considering all analysed Polish as well as European sequences, ovine-derived sequences were the most variable, while the midge ones were more conserved and encompassed unique substitutions located mainly in nonstructural protein S. SBV sequences isolated from *Culicoides* are the first submitted to GenBank and reported.

Introduction

Schmallenberg virus emerged as an unknown *orthobunyavirus*, causing first unspecific signs of fever and diarrhoea in adult cattle in Germany and the Netherlands in 2011, which later unraveled into more serious outbreaks of foetal malformations, stillbirths and abortions. The viruses of *Orthobunyavirus* genus, Peribunyaviridae family, are transmitted by vectors such as mosquitos, biting midges and ticks. Some, such as *Oropouche* virus (OROV), cause human febrile illness in South America, while other, such as *Akabane* virus (AKAV), Aino virus (AINV) and Peaton virus (PEAV), infect animals causing abortions, stillbirths and malformations in domestic ruminants in Australia, Asia and Africa.

SBV-specific antibodies were detected in most domestic, wild and exotic ruminants, camelids, wild boars, and even in dogs. The virus is transmitted by *Culicoides* Latreille (Diptera: Ceratopogonidae), which are also known vectors of other emerging viruses such as Bluetongue (BTV) or African horse sickness (AHSV). Due to the lack of mandatory notifications of SBV-suspected cases, the general impact of SBV on European animal trade remains uncertain; however, during the outbreaks seroprevalence reached up to 100%. At the moment, Poland as most European countries is going through a secondary wave of SBV epidemics [1–3].

SBV genome consists of the three negative single-stranded RNA segments. The large (L) segment encodes RNA-dependent RNA polymerase (RdRp); medium (M) segment encodes a polyprotein precursor, which is post-translationally cleaved into nonstructural protein (NSm) and two glycoproteins (Gn and Gc); and small (S) encodes nucleocapsid (N) protein and a nonstructural protein (NSs) in an overlapping reading frame. Gc and Gn proteins are present on the surface of the virion inserted in the lipid envelope, which makes them the target of host antibodies [4–6]. They play a role in viral attachment and cell fusion [7]. In a few studies, M segment was described as the most divergent part of SBV genome with the hypervariable region (HVR)

Edited by Keizo Tomonaga.

✉ Magdalena Larska
m.larska@piwet.pulawy.pl

1   Department of Virology, National Veterinary Research Institute, Al. Partyzantów 57, 24-100 Puławy, Poland

encoding N-terminus of Gc protein [8–10]. The N protein is responsible for creating viral envelope, and is considered as the best antigen for serological and molecular SBV testing. The NSs is a major virulence factor for orthobunyaviruses [10], but is nonessential to virus viability in mammalian and insect cells [11].

Hitherto, the phylogenetic analyses within SBV were rather limited to S segment studies and a small amount of data originating only from the ruminant host [12–15]. Previous papers presented no geographic- or limited time-dependent variability of SBV genome [1]. In order to extend the knowledge on the diversity of the novel virus as well as to elucidate the host-vector adaptation of SBV, comparative genetic analysis of new Polish plus beforehand reported European SBV sequences was performed in relation to different genome fragments and different host species including, for the first time, *Culicoides* vector.

Materials and methods

Sample origin and RNA purification

The SBV-positive samples initially included in the study were collected between 2012 and 2013 from central nervous system tissue of 27 aborted ovine foetuses; 20 serum samples from cattle herds during acute infection from Silesia and West Pomeranian provinces [16, 17]; a serum sample from elk from Podlasie province [18]; and three positive samples of bull semen collected at Warmia-Masuria province [19]. The *Culicoides* were collected during BTV virus monitoring program using UV light traps placed at the cattle farms throughout the country between 2012 and 2015. Collected insects were entomologically examined, segregated and pooled (approx. 23 individual females) according to their species and parity status (depending on feeding and oviposition). Since *C. obsoletus* and *C. scoticus* females cannot be distinguished morphologically, those pools were marked as *C. obsoletus/scoticus* complex. Sixty-seven SBV-positive pools from the most abundant *Culicoides obsoletus/scoticus* complex or *Culicoides punctatus* species were found. More pools were parous females (with pigmented abdomen, which blood fed and oviposited at least once), than gravid (containing a mature egg batch in abdomen) and freshly blood fed. See Supplementary Table 1 for detailed list of samples successfully sequenced.

Ten percent mammal tissue homogenates were prepared in RLT lysis buffer (Qiagen, Germany) using Lysing Matrix D Tubes (MP Biomedicals, France) and FastPrep-24 instrument (MP Biomedicals, USA) for 90 s at 6.5 m/s speed. The insects were drained from 70% ethanol in which they were previously stored, and homogenized in 500 µl RLT using the same system. Subsequently, total RNA was extracted using RNeasy Mini Kit (Qiagen, Germany) in the QIAcube automatic purification station (Qiagen, Germany). Total RNA from serum was extracted using TRI Reagent (Sigma-Aldrich, USA) and eluted in a final volume of 20 µl of nuclease-free water. Semen was extracted using Viral RNA Mini Kit (Qiagen, Germany) after Qiazol Lysis Reagent (Qiagen, Germany).

Conventional reverse transcription PCR (RT-PCR), molecular cloning and sequencing

RT-PCR was conducted using Transcriptor One-Step RT-PCR Kit (Roche, Switzerland). Complementary DNA (cDNA) of three SBV segments was amplified using 13 primer pairs kindly provided by Dr. Wernike from Friedrich Loeffler Institut (FLI) Insel Riems, Germany [9] and one pair of self-designed primers (forward: TGACATTCCATG AGTCTATG and reverse: GTCGGATTGTCTCCTGCA AAC flanking from 668 to 1874 nt in segment L). Amplification products were analysed in agarose gels, gel purified using QIAquick Gel Extraction Kit (Qiagen, Germany) and subjected to DNA Sanger sequencing by external company using BigDye Terminator kit v3.1 Kit and Cycle Sequencing 3730XL Genetic Analyzer (Applied Biosystems, USA). The cDNA amplicons were purified by the QIAquick PCR Purification kit (Qiagen, Germany), followed by sequencing in the ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, USA), and the results were analysed by GeneScan Analysis Software (Applied Biosystems, USA). For samples that contained insufficient amount of RNA to obtain nucleotide sequences, molecular cloning of gel purified DNA fragments prior to sequencing was performed using Qiagen PCR Cloning Plus Kit with pDrive cloning vector (Qiagen, Germany) and *Escherichia coli* cells (Qiagen, Germany). Purification of plasmid DNA was performed with Pure Link Plasmid Miniprep Kit (Thermo Fisher Scientific, USA). The presence of insert was confirmed by PCR using Jump Start AccuTaq LA DNA polymerase (Sigma-Aldrich, USA).

Next-generation sequencing (NGS)

NGS for two samples was performed. Isolate PL_13_130-6_ovis was isolated from brain tissue of stillbirth lamb. Virus was propagated in BHK-21 (baby hamster kidney) cell culture. The viral RNA was extracted after the third passage, while the RNA of the second SBV strain PL_12_1018-08_bov was extracted directly from the serum of infected cow. Libraries for NGS were prepared using NEBNext Ultra RNA Library Prep Kit for Illumina and then checked in BioAnalyzer 2100 using Agilent High Sensitivity DNA Kit. Illumina sequencing was performed with MiSeq Reagent Kit v2 mode PE250. Bioinformatic analysis was carried out using trimmomatic (filtering reads, trimming adapters) and CLC

GenomicWorkbench (mapping reads to reference sequences and De novo assembly of NGS data).

Phylogenetic analysis

Alignments of the nucleotide consensus sequences and deduced amino acid (aa) sequences and comparative sequence analysis were carried out with MEGA version 5.2, BioEdit v. 7.2.5 and Clustal Omega. Primers were designed using Primer3 v.0.4.0 and OligoAnalyzer 3.1. Obtained sequences of Polish isolates were submitted to GenBank (Supplementary Table 1). The ratio between nonsynonymous and synonymous substitutions (dN/dS) was used to estimate the patterns of selection [20]. The higher the ratio, the higher positive selection onto the protein. Conversely, the ratio below 1 suggests the protein evolves slowly under negative (purifying) selection, meaning more conservative. Neutral evolution occurs when dN/dS ~ 1.

Results

Phylogenetic analysis of Polish SBV sequences

Seven fragments of SBV detected in lamb brain tissue samples (C_t cut off 19.2–31.3) were obtained: three segment S, two segment M and one segment L as well as one fragment containing HVR corresponding to 410–687 aa of M polyprotein (Supplementary Table 1). From cattle serum (C_t cut off 22.5–27.9), six sequences were assembled: one full S and M segment, additionally two HVR sequences and two complete segment L. Unfortunately, the attempts to sequence the virus from elk serum and bull semen failed.

Twelve of 44 SBV-positive pools from 2012 (C_t 17.6–29.0), four of 13 from 2014 (C_t 22–39.3) and one (C_t 29.3) of 10 positive in 2015 provided seven segment S sequences, one full segment M sequence as well as five sequences within HVR, 7 Gn (1–188 aa) and five Gc (1141–1394 aa) encoding fragments. Only three short SBV sequences of L segment (1137–1390 aa) were obtained from the midge. The geographical distribution of the animals which provided the sequences is presented in Fig. 1.

Among Polish sequences, only 0–2 nucleotide (nt) changes per sequence were observed within segment S sequence, which would result in 0–1 amino acid (aa) substitutions in N protein and 0–2 amino acid substitutions in NSs protein. The S segment nucleotide sequences were highly homologous among Polish isolates (99.3–100% identity) and to German reference strain BH80/11—4 (99.6%–100%), first SBV strain ever identified [19]. Phylogenetic tree of S segments reflected virus genetic stability (Fig. 2a). Most of Polish sequences were distributed throughout the tree, regardless of the origin or year of detection, showing

seemingly random relationship to Belgian and German sequences. However, three strains originating from *Culicoides obsoletus/scoticus* complex midges (PL_12_28_culic, PL_14_962_culic, PL_15_711_culic) formed separate branch. In general, sequences derived from particular host, country or sampling year did not form any clear clades or subgroups either.

In M segment, sequence similarity between Polish isolates and reference strain BH 80/11—4 presented 99.2–99.7% identity and among Polish isolates was 98.6–99.9%. Comparison of 12 Polish and reference HVR sequences showed 57 nt substitutions; 35 of them were nonsynonymous (dN/dS ratio 1.6). Due to the low diversity of segment M, phylogenetic trees constructed on the complete segment or HVR fragment sequences showed again low resolution (Fig. 2b, c). On HVR fragment and full-length M segment trees comprising the same available sequences, four out of six sequences derived from *Culicoides* clustered together with sequence from the first SBV outbreak in cattle herd in Poland creating separate branch from the rest of the sequences. Other Polish sequences were spread across the tree showing highest affinity to German, Swiss and Liechtenstein sequences. The same as for S segment, the origin, year of detection of the Polish strains and animal species had no effect on the pattern of molecular diversity of SBV in our country.

The analysis of aligned sequences has shown that segment L is genetically stable (99.9–100% identity with reference strain BH80/11—4). Among Polish sequences, 24 nt substitutions (eight per sequence) of which four (1.3 per sequence) were nonsynonymous (Asp121Asn, Met1093Thr, Glu1159Gln, Glu1823Lys) were found. On the phylogenetic tree based on L segment, two bovine isolates from Silesia province formed separate branch from the rest of the sequences.

Comparative analysis of SBV sequences isolated from mammal host and arthropod vector

The limited number and spatio-temporal divergence of Polish sequences obtained in the presented study was insufficient to investigate SBV variability in different hosts. Representative reference nucleotide sequences from different countries, studies and species available in GenBank were selected to supplement the analysis (Supplementary Table 2). Only sequences encoding complete ORF were chosen for detailed analysis. Sequences that were identical, originating from the same study, country and species were excluded. If sequences were derived from tissues as well as passaged in cell lines—only ones from original material were chosen. Finally, 62 sequences (11 Polish, 51 European) of segment S, 45 (4 Polish, 41 European) of M segment and 17 (3 Polish, 14 European) coding region of L segment were included in the analysis. To study N-terminal sequence of Gc
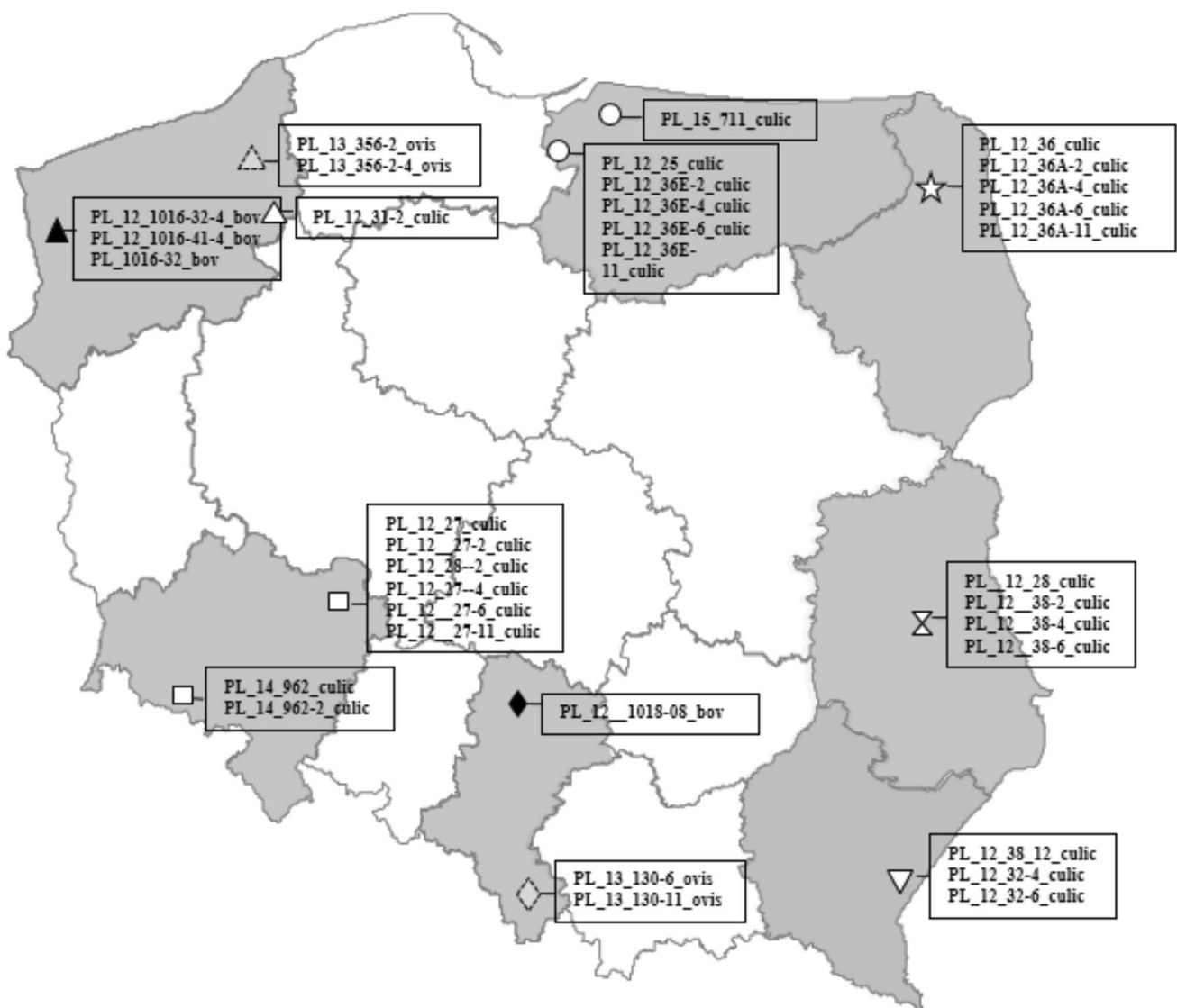
Fig. 1 Map of geographical location of Polish isolates from the present study. To follow the association between localization and molecular diversity of isolates, they were marked with the same symbols on the phylogenetic trees

known as the most variable fragment in whole SBV genome (HVR), 52 fragments (12 Polish, 40 from GenBank) were selected. Previous studies had defined the HVR located in Gc protein between 457 and 847 aa [8] or between 493 and 629 aa [9]. In this study, the HVR was slightly modified into a fragment between 410 and 687 aa to present substitutions in the, so far unique, sequences obtained from *Culicoides*. The genome of the first SBV strain BH80/11—4 detected in German cattle was used as a reference in all the analyses [21].

Comparing nucleotide changes in all European and Polish sequences obtained from different species, we can conclude that N protein is the most conserved among cattle (Fig. 3a). In *Culicoides*-derived sequences, however limited in number, less substitutions per sequence were observed. In

sheep, sequence substitutions were the most frequent. Interestingly, nucleotide substitutions observed in SBV isolates from *Culicoides* did not occur in mammalian strains. G354A substitution (Ser111Asn) noted in one bovine sequence was observed also in approximately 38% sheep isolates, which may determine this amino acid the most variable among N protein. NSs protein among cattle strains proved to be the most conserved region, where no nt or aa changes occurred (Fig. 3b). In arthropod sequences, all nt changes were nonsynonymous and were unique to the sequences originating from mammals: Asn25Asp, His84Arg, Asp90Gly. However, His in 84 position was also substituted to Tyr in sheep-derived sequence from Belgium (GenBank Accession Number—KT7951461). The most mutations occurred in sheep-derived sequences, including seven substitutions
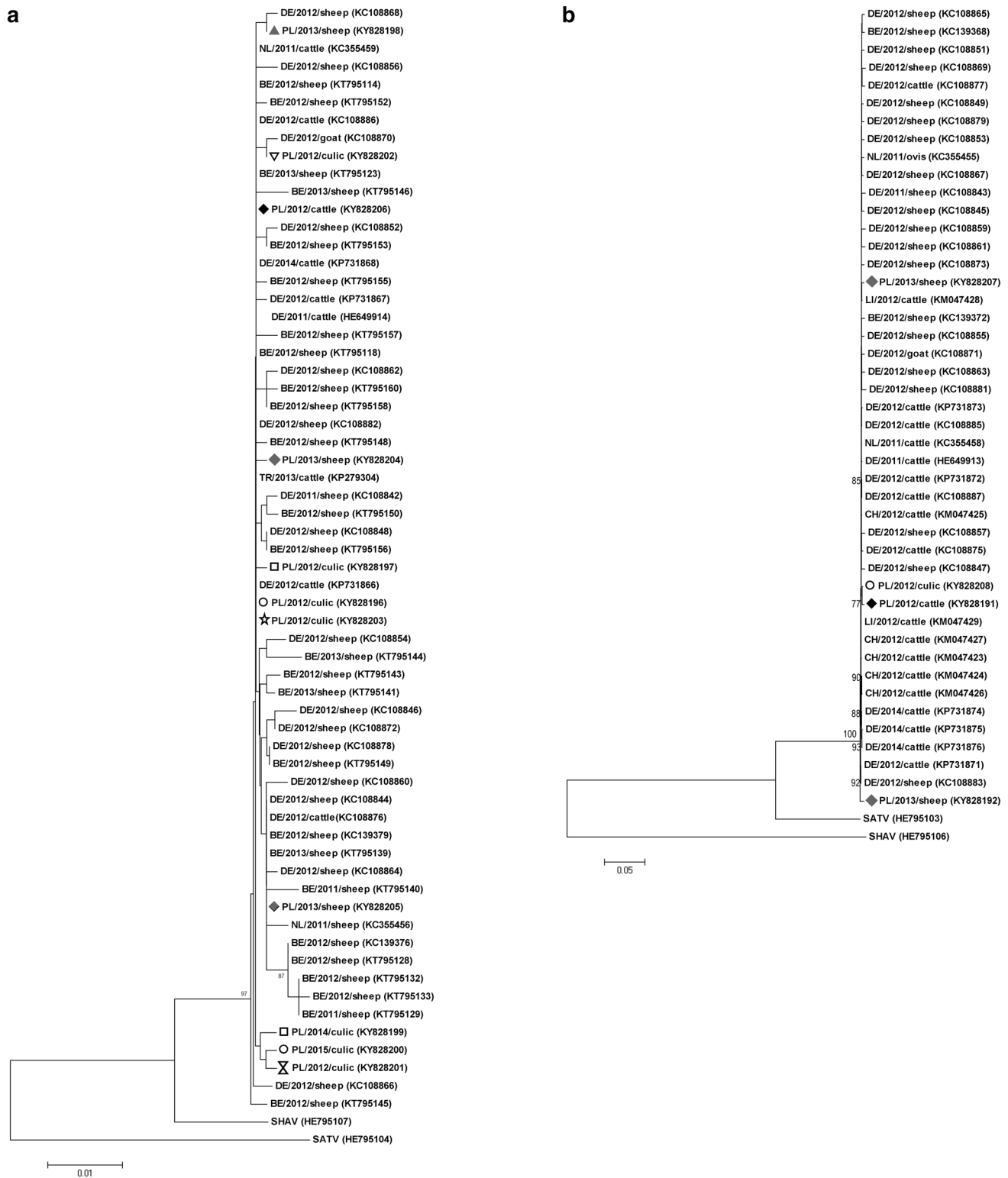
a

DE/2012/sheep (KC108868)
▲ PL/2013/sheep (KY828198)
NL/2011/cattle (KC355459)
DE/2012/sheep (KC108856)
BE/2012/sheep (KT795114)
BE/2012/sheep (KT795152)
DE/2012/cattle (KC108886)
DE/2012/goat (KC108870)
▽ PL/2012/culic (KY828202)
BE/2013/sheep (KT795123)
BE/2013/sheep (KT795146)
◆ PL/2012/cattle (KY828206)
DE/2012/sheep (KC108852)
BE/2012/sheep (KT795153)
DE/2014/cattle (KP731868)
BE/2012/sheep (KT795155)
DE/2012/cattle (KP731867)
DE/2011/cattle (HE649914)
BE/2012/sheep (KT795157)
BE/2012/sheep (KT795118)
DE/2012/sheep (KC108862)
BE/2012/sheep (KT795160)
BE/2012/sheep (KT795158)
DE/2012/sheep (KC108882)
BE/2012/sheep (KT795148)
◆ PL/2013/sheep (KY828204)
TR/2013/cattle (KP279304)
DE/2011/sheep (KC108842)
BE/2012/sheep (KT795150)
DE/2012/sheep (KC108848)
BE/2012/sheep (KT795156)
□ PL/2012/culic (KY828197)
DE/2012/cattle (KP731866)
○ PL/2012/culic (KY828196)
☆ PL/2012/culic (KY828203)
DE/2012/sheep (KC108854)
BE/2013/sheep (KT795144)
BE/2012/sheep (KT795143)
BE/2013/sheep (KT795141)
DE/2012/sheep (KC108846)
DE/2012/sheep (KC108872)
DE/2012/sheep (KC108878)
BE/2012/sheep (KT795149)
DE/2012/sheep (KC108860)
DE/2012/sheep (KC108844)
DE/2012/cattle(KC108876)
BE/2012/sheep (KC139379)
BE/2013/sheep (KT795139)
DE/2012/sheep (KC108864)
BE/2011/sheep (KT795140)
◆ PL/2013/sheep (KY828205)
NL/2011/sheep (KC355456)
BE/2012/sheep (KC139376)
BE/2012/sheep (KT795128)
BE/2012/sheep (KT795132)
BE/2012/sheep (KT795133)
BE/2011/sheep (KT795129)
87
97
□ PL/2014/culic (KY828199)
○ PL/2015/culic (KY828200)
✕ PL/2012/culic (KY828201)
DE/2012/sheep (KC108866)
BE/2012/sheep (KT795145)
SHAV (HE795107)
SATV (HE795104)

0.01

b

DE/2012/sheep (KC108865)
BE/2012/sheep (KC139368)
DE/2012/sheep (KC108851)
DE/2012/sheep (KC108869)
DE/2012/cattle (KC108877)
DE/2012/sheep (KC108849)
DE/2012/sheep (KC108879)
DE/2012/sheep (KC108853)
NL/2011/ovis (KC355455)
DE/2012/sheep (KC108867)
DE/2011/sheep (KC108843)
DE/2012/sheep (KC108845)
DE/2012/sheep (KC108859)
DE/2012/sheep (KC108861)
DE/2012/sheep (KC108873)
◆ PL/2013/sheep (KY828207)
LI/2012/cattle (KM047428)
BE/2012/sheep (KC139372)
DE/2012/sheep (KC108855)
DE/2012/goat (KC108871)
DE/2012/sheep (KC108863)
DE/2012/sheep (KC108881)
DE/2012/cattle (KP731873)
DE/2012/cattle (KC108885)
NL/2011/cattle (KC355458)
DE/2011/cattle (HE649913)
85 DE/2012/cattle (KP731872)
DE/2012/cattle (KC108887)
CH/2012/cattle (KM047425)
DE/2012/sheep (KC108857)
DE/2012/cattle (KC108875)
DE/2012/sheep (KC108847)
○ PL/2012/culic (KY828208)
77 ◆ PL/2012/cattle (KY828191)
LI/2012/cattle (KM047429)
CH/2012/cattle (KM047427)
CH/2012/cattle (KM047423)
90 CH/2012/cattle (KM047424)
CH/2012/cattle (KM047426)
88 DE/2014/cattle (KP731874)
100 DE/2014/cattle (KP731875)
93 DE/2014/cattle (KP731876)
DE/2012/cattle (KP731871)
92 DE/2012/sheep (KC108883)
◆ PL/2013/sheep (KY828192)
SATV (HE795103)
SHAV (HE795106)

0.05

Fig. 2 Phylogenetic trees of complete segment S (**a**), segment M (**b**), hypervariable region (HVR) corresponding to 410–687 aa protein in M segment (**c**), and complete segment L (**d**) nucleotide sequences constructed using maximum likelihood method, Kimura 2-parameter model with 1000 bootstrap replication. Polish strains marked with different shapes corresponding to geographic origin (Fig. 1). SBV reference sequences representing different countries of origin from different animal species were obtained from GenBank (Supplementary Table 2) and indicated by accession numbers. Numbers at nodes represent percentage of 1000 bootstrap replicates (values < 70 were not shown). Scale bars indicate nt substitution per site. All the strains are marked with their ISO country code, year of detection and animal species. GenBank accession numbers in parenthesis
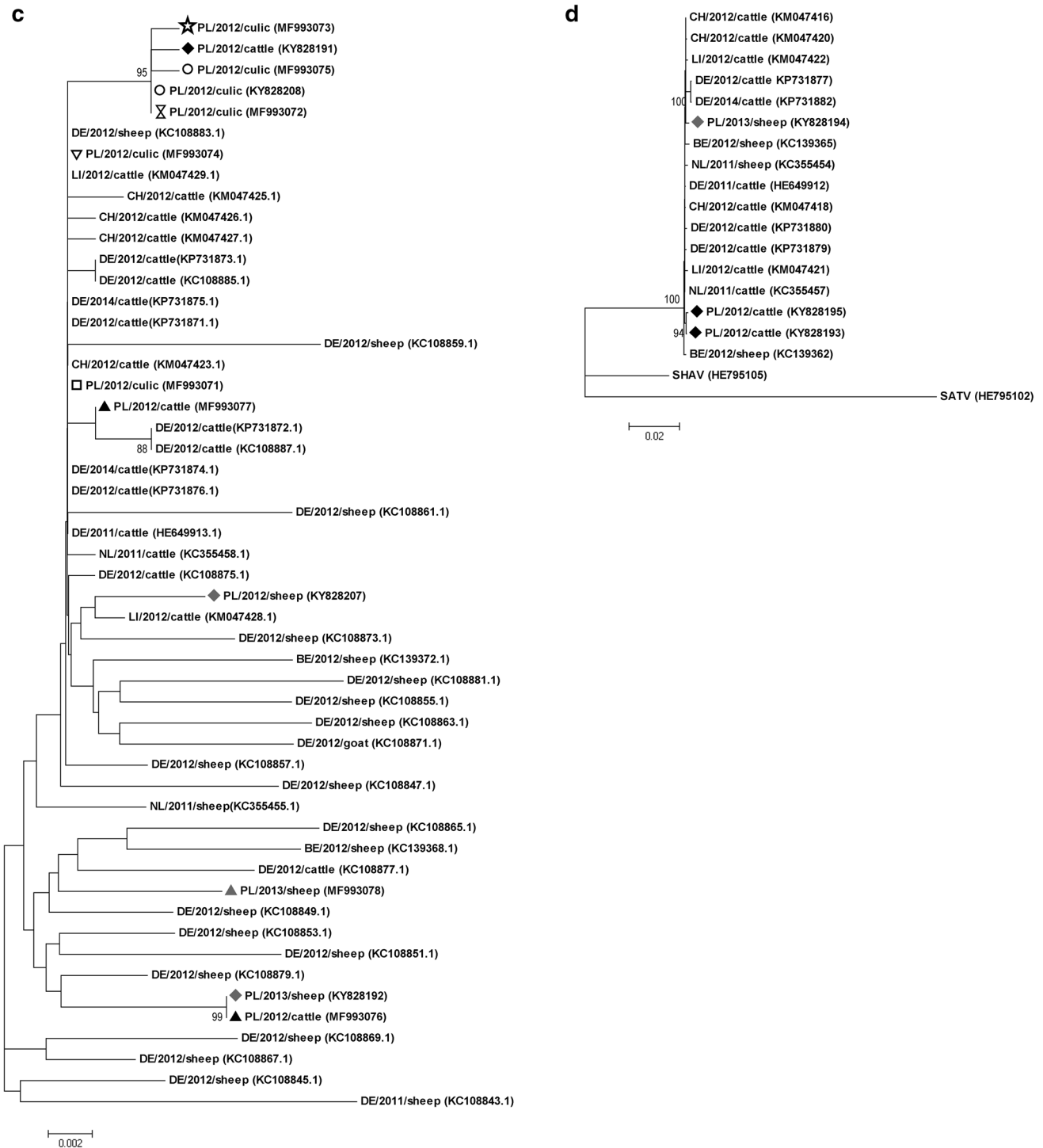
c

d

Fig. 2 (continued)

leading to truncated NSs sequence. Variability in all analysed sequences of N protein was 99.9–100%, while of NSs 98.9–100%. The analysis of the nonsynonymous-to-synonymous substitution ratio (dN/dS) revealed that nucleoprotein is much more conserved than NSm S, respectively, 0.8 (43 nonsynonymous/55 synonymous) and 6.1 (49/8).

Our analysis confirmed that, as described for other orthobunyaviruses, M segment was the most variable segment in SBV genome. Comparing all analysed sequences, identity of Gn was 99.2–99.9%, of NSm − 99.4 to 100% and of Gc − 99.0 to 99.9%. Nucleotide sequence analysis revealed that, as well as in S segment, sequences obtained

from sheep were more variable than those from cattle (respectively 22.3 and 8.5 nt per sequence, resulting in 16.5 and 4 aa substitution per deduced aa sequence). The overall dN/dS ratio reached 2.1 (479/233) for all analysed sequences. However, it was significantly higher for ovine sequences − 2.9 (380/133) than for bovine ones − 0.9 (79/90), which suggested positive selection only in sheep-derived sequences. One complete M segment sequence from *Culicoides* with only four and one from goat sequence with 16 aa substitution did not allow to draw firm conclusions.

Comparative analysis of all deduced amino acid sequences ($n = 52$) showed that positive selection pressure on HVR in Gc protein is more marked in sheep (dN/dS 3.3) than in cattle (dN/dS 1.6), while in *Culicoides* only negative selection occurred (dN/dS 0.1) (see Fig. 3c). This reflected sequence similarity for sheep, cattle and *Culicoides* sequences to the reference sequence, respectively, 97.9–100%, 98.7–100% and 99.9–100%. In one available goat sequence, 9 nt substitution arose, all nonsynonymous (99% identity with reference strain) from which only two were detected also in sheep from the same study (9). Nonsynonymous-to-synonymous substitution ratio for all sequences in HVR was 2.6 and differed slightly from calculated for the complete segment M (2.1). Accumulation of nonsynonymous mutations in sequences from mammalian host occurred between 470 and 640 aa (Fig. 3c). A single aa change (Arg432Ile) was observed in only one PL_12_36_culic strain originating from *C. obsoletus*/*scoticus* complex pool and was unique among other sequences. The most frequent aa variation was Gln567Lys which was observed in half ovine and one-tenth of bovine SBV sequences. Second, most common change was observed in 575 aa, where Phe was replaced by Ser in 10% of sequences from cattle and 29% from sheep and by Pro in one sheep sequence. The indel mutations were only observed in HVR, which included insertion of Ile and Phe (between 574 and 575 aa in sheep sequence, GenBank accession number—KC1088591) and the deletion of 12 aa (from 532 to 543 in sheep sequence KC1088531) [9].

The detailed analysis of M segment in arthropod vector included eight sequences of 12–564 nt (1–188 aa) fragment which corresponded to N-terminal fragment of Gn and six sequences of 3465–4195 nt fragment overlapping Gc (1141–1394 aa). In both fragments, sequences were highly conserved with 100% and 99.6–100% homology, respectively. The substitutions observed in Gc fragment isolated from Culicoides were not described in any mammal sequence: nonsynonymous change G3757C (S1249T) in one strain and synonymous A3920G in two strains.

In L segment, comparison of all analysed 17 sequences revealed 79 nt (6.1 per sequence) and 13 aa (1 per sequence) substitutions in sequences from cattle, as well as 34 nt (8.5 per sequence) and 13 aa (3.25 per sequence) in sequences

from sheep. The dN/dS ratio was only 0.2 and 0.6 for bovine and ovine sequences, respectively, suggesting negative selection in this segment. In all sheep sequences, a difference (Glu/Gln or Glu/Asp) at 1159 position in respect to reference strain was observed. Unfortunately, only three fragments (1137–1390 aa) from segment L were obtained from insect-derived material. However, the analysed fragment seemed to be highly homological among the isolates (99.9–100% identity to reference strain). Two point mutations (T3721C, T4012C) occurred resulting in one aa substitution (Tyr1236His) in one culicoid sequence (PL_12_27—11_culic) from *C. obsoletus*/*scoticus* pool, while no changes were observed in the remaining two.

Discussion

Collective analysis of Polish, as well as available European SBV isolate sequences, confirmed high genetic stability of S and L segments in the mammal host and presumably in arthropod vector. As described previously, segment M containing glycoprotein C hypervariable ectodomain was the most variable fragment of viral genome [9, 10]. Still, comparing to better known *orthobunyavirus* of veterinary importance—Akabane virus, SBV presented 25 times higher mutation rate, as it is understood for the virus in its epidemic stage [22]. Our results have revealed some quantitative and qualitative differences of nt as well as aa mutations mostly in ruminants. Interestingly, the sequences derived from *Culicoides*, which should be emphasized that these were described for the first time, seemed less variable. Characteristically, the ovine SBV sequences in all viral segments were the most divergent, including sequences originating from the same sheep flock, which differed significantly. Furthermore, some researchers reported cases of two or three SBV variants found in single animal, which with respect to overall low variability of SBV resulted more probably from co-infection with different SBV strains [12, 23]. Such situations were not reported in cattle. Interestingly, SBV in sheep is more pathogenic causing important loses in breeding but the seroprevalence is lower than in cattle [24]. It may be speculated that in sheep SBV multiplies more efficiently or it interacts differently with ovine immune defense, and thus antigenic escape, expressed in mutations, results in higher morbidity than in cattle. SBV is transmitted to vertebrae host by an insect vector and direct transition between mammals is unlikely. Therefore, the differences in SBV sequences obtained from insect and vertebrae host may refer to formation of virus quasispecies as reported for other virus vector systems [25]. To unravel this, the comparison of nucleotide sequences originating from host and vector from one disease outbreak would be necessary.
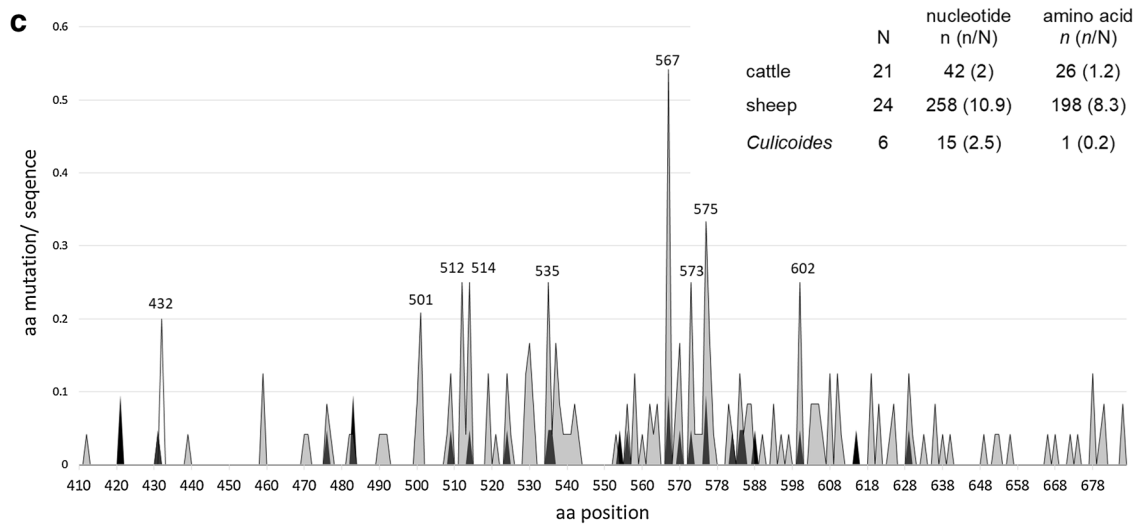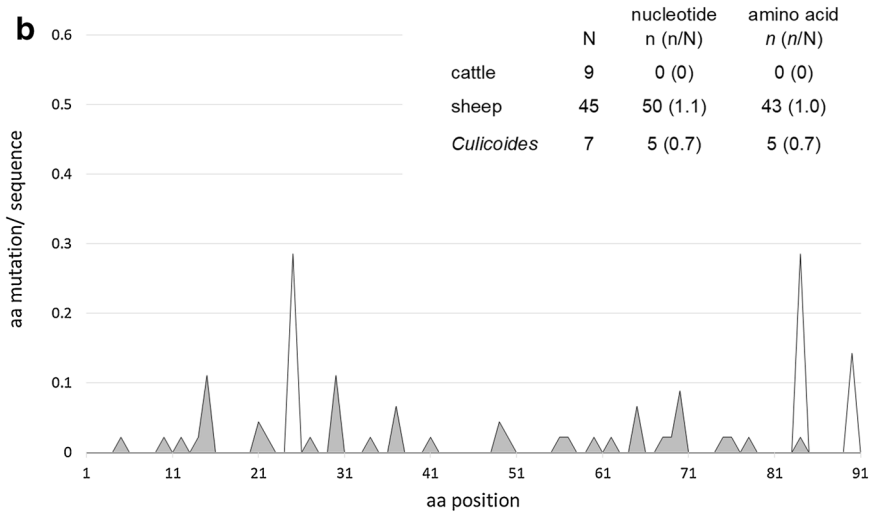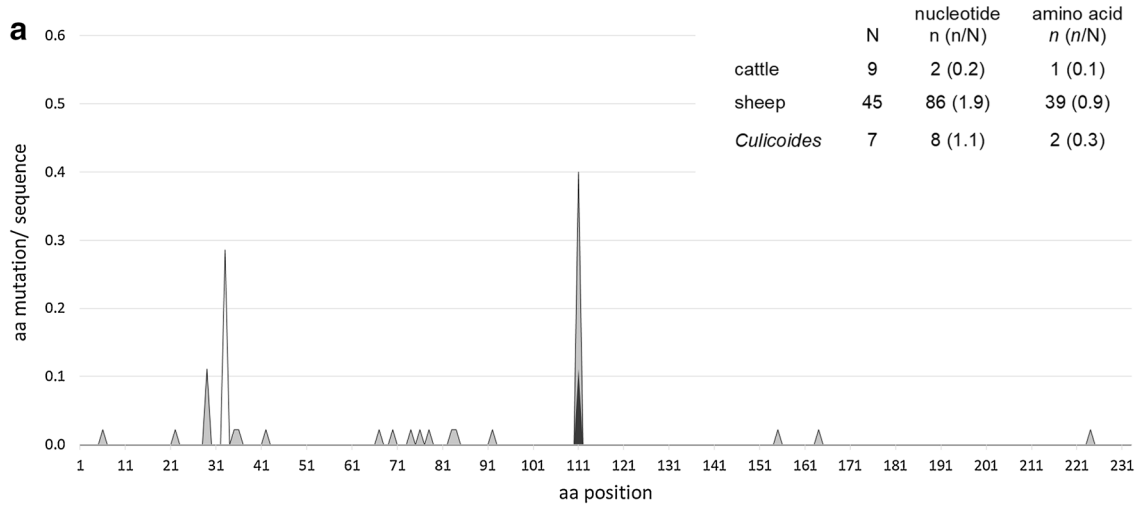
a

	N	nucleotide n (n/N)	amino acid n (n/N)
cattle	9	2 (0.2)	1 (0.1)
sheep	45	86 (1.9)	39 (0.9)
Culicoides	7	8 (1.1)	2 (0.3)

b

	N	nucleotide n (n/N)	amino acid n (n/N)
cattle	9	0 (0)	0 (0)
sheep	45	50 (1.1)	43 (1.0)
Culicoides	7	5 (0.7)	5 (0.7)

c

	N	nucleotide n (n/N)	amino acid n (n/N)
cattle	21	42 (2)	26 (1.2)
sheep	24	258 (10.9)	198 (8.3)
Culicoides	6	15 (2.5)	1 (0.2)

◄**Fig. 3** Distribution of aa substitutions or indel mutations within deduced aa sequences of the most variable of viral proteins: nucleoprotein (**a**), nonstructural protein S–NSs (**b**), and HVR of glycoprotein C–Gc (**c**) with respect to different hosts. In the tables, number of nucleotide (n) or amino acid (*n*) variation was divided per number (N) of sequences from respective species included in the analysis (including Polish and reference sequences). The goat sequence was excluded. The black arrow represents two aa insertions

Based on evolutionary and phenotypic studies on arboviruses, it was concluded that they evolved from arthropod-specific ancestors [26]. Assumed high genetic stability in arthropod vector may be the result of SBV better and more conserved adaptation to counteracting insect host immune defense and virus replication in insects. The argument for this adaptation may be the fact that peribunyaviruses that are teratogenic in vertebrae are harmful to developing arthropod oocyte and embryo at the same time [27]. On the other hand, it seems that the virus does not replicate in the insects as efficiently as in mammal host, meaning lower number of transcriptions equals fewer RNA polymerase mistakes and thereof nucleotide mutations. As discussed earlier [28], bimodal distribution of C_t values in *Culicoides* pools may result from different patterns of virus replication: transmissible, active virus multiplication in parous and gravid females (C_t around 22) and sub-transmissible, probably persistent (C_t around 35) infection in nulliparous or passive carriage of the virus in *C. punctatus*. This finding may be supported by experiments on *C. sonorensis* with BTV where insects with infectious virus had lower C_t values than those with persistent, sub-transmissible infections or passively carrying the viral RNA in engorged blood [29]. It was also reported that viral transcription rate of La Crosse virus (LACV) is strictly connected and co-regulated with mosquito cell transcription due to cap snatching occurring in all bunyaviruses. Insects in diapause dramatically reduce metabolic activity and cell replication. It could modulate virus replication favouring persistent infections and protect vector from virus detrimental effect during dormancy and thus enhancing the overwintering efficiency of LACV [27]. In our study, NSs was the most variable SBV protein in *Culicoides* SBV strains. All detected substitutions were nonsynonymous and independent from the mammalian ones. This could suggest specific role of this protein in infection of arthropod vector and reflect virus adaptation to *Culicoides* immune defense. NSs protein counteracts the immune system of vertebrae host. It shuts down antiviral protein, such as IFN alpha, and regulates translation, apoptosis and RNA polymerase II [11]. In arthropod vector, NSs is also suspected to shut down insect immune response; however, details are still unknown [30]. In *Bunyamwera* virus (BUNV), NSs sequence is required for virus replication in vitro and in vivo [31]. However, some naturally occurring insect-borne orthobunyaviruses with truncated or lacking NSs were reported [32]. This protein may be

crucial for virulence in insects, and hence in plant-infecting *Tospovirus* (*Bunyavirales* order) called Tomato spotted wilt virus, thrips infected with NSs-lack mutant viruses were not able to transmit and amplify the virus efficiently [33]. In LACV and BUNV, NSs does not affect RNA polymerase II or inhibit protein synthesis, as it does in mammalian cells, thus establishing persistent infection instead of apoptosis [27, 31, 34]. Apoptosis of the arbovirus-infected insect cells can have strong antiviral effect, but may also reduce vector vitality necessary for efficient feeding on mammal host and transmitting the virus. Therefore, in competent vector apoptosis is suspected to be inhibited or in low level favouring persistent infection [35]. Moreover, researches presented different effect of NSs on replication in cell lines isolated from distinct insect species [31]. It may be speculated that variation in NSs of *Culicoides*-derived SBV strains may be the result of virus adaptation to new European species of insect vectors.

As discussed by Coupeau et al. [12], NSs protein may not be essential in congenital infections in ruminants since NSs-truncated sequences occurred in lambs. Such shortened NSs was detected in lambs from Germany in 2011 and 2012 [9], Belgium in 2012 [12] and in Great Britain in 2012 and 2017 [1]; in Polish sequences such deletions were not observed. Experiments with SBV NSs deletion mutants indicated no virus replication or clinical disease in cattle [11]. NSs is efficient in shutting down most known major antiviral proteins, including INF [30, 36]. Research on ovine foetuses infected with Cache Valley Virus revealed upregulation of some INF type I-induced proteins involved in bunyavirus immune defense [37]. It proves that foetuses in vulnerable gestation period produce an immune response against bunyavirus infection controlled by NSs protein, and hence efficient replication of NSs-truncated mutants in congenital cases needs elucidation. Probably the answer may be hyperinfection reported in some of these cases [12].

Segment M coding Gc, NSm and Gn is proposed as the most variable within whole SBV genome. The N-terminus of Gc protein encodes ectodomain exposed in the virion envelope and is not essential for virus replication in tissue culture, while the C-terminus contains fusion peptide [7]. The identified HVR is probably a result of viral evasion of host immunity [9] and Gc adaptation to new host to get a better entry into host cells [10]. The M segment (precisely Gc) could also be a virulent factor and is involved in host protein shutoff as described by Varela et al. [10]. Greater variation in HVR in sheep may result from the enhanced teratogenic effect observed in this species. Postulated N-linked glycosylation sites in Gc (493, 686, 1353 aa) remains stable in most analysed sequences [23]. However, mutation in 686 aa (from Asn to Asp) in HL-1 atypical isolate [23] was also found in BH200/12—2 [9] both from lamb's brain.

The HVR of Gc protein in arthropod-derived sequences presented almost no mutations in contrast to mammalian-derived sequences. Possibly lack of immune system pressure and arms race on the Gc ectodomain proves this protein does not play essential role in vector evasion, virulence or adaptation to arthropod host. Probably, as defined for LACV, Gn glycoprotein instead of Gc plays a role in attachment to mosquito cells [38]. Gn glycoprotein is responsible for folding, heterodimerization, and intracellular trafficking of both Gn and Gc proteins [39]. The analysis of N-terminus of Gn, where functional signal peptide for membrane translocation is localized [40], revealed high stability of this glycoprotein fragment in the samples from *Culicoides*.

NSm M takes part in protein synthesis shutoff [41], virus assembly and budding. Mutant lacking NSm can replicate in cell culture but has reduced virulence in mice, in cattle produce viremia and seroconversion [11]. In Rift Valley fever virus (RVFV), deletion of NSm reduced infection, dissemination (possibly by modulation of a mosquito midgut infection barrier) and transmission rates in mosquito proving the essential role of NSm in viral-vector interactions [2, 42]. In Polish Culicoides-derived sequences, similarity was reaching 99.8% with only one synonymous mutation.

RdRp located in the L segment has four conserved regions among peribunyaviruses [43]. Regions 1 and 2 located on amino termini and region 3 with polymerase module (pre-motif A, motif A to E) and region 4 downstream. As in other studies, SBV L segment revealed high stability due to its crucial function in virus replication. As in residual SBV segments, sheep-derived sequences had more mutations than those isolated from cattle.

Analysis of Polish together with European SBV sequences confirmed that the most variable region of the virus genome is N-terminus fragment in Gc within M segment. High sequence identity among strains may simplify the diagnostic testing and the implementation of vaccination scheme. The virus genetic stability found in vector-derived sequences may be the result of prominent virus adaptation and speculated low replication potential in insect organism, while high mutation rate in sheep may coexplain high morbidity in this species. To our knowledge, this is the first report of SBV sequences isolated from the insect vector.

Compliance with ethical standards

References

1. McGowan SL, La Rocca SA, Gierson SS, Dastjerdi A, Choudhury B, Steinbach F (2018) Incursion of *Schmallenberg* virus Into Great Britain in 2011 and emergence of variant sequences in 2016. Vet J 234:77–84. https://doi.org/10.1016/j.tvjl.2018.02.001
2. Collins ÁB, Barrett DJ, Doherty ML, McDonnell M, Mee JF (2017) Significant re-emergence and recirculation of *Schmallenberg* virus in previously exposed dairy herds in Ireland in 2016. Transbound Emerg Dis 64:1359–1363. https://doi.org/10.1111/tbed.12685
3. Wernike K, Hoffmann B, Conraths FJ, Beer M (2015) *Schmallenberg* virus recurrence, Germany, 2014. Emerg Infect Dis 21:1202–1204. https://doi.org/10.3201/eid2107.150180
4. Roman-Sosa G, Brocchi E, Schirrmeier H, Wernike K, Schelp C, Beer M (2016) Analysis of the humoral immune response against the envelope glycoprotein Gc of Schmallenberg virus reveals a domain located at the amino terminus targeted by mAbs with neutralizing activity. J Gen Virol 97:571–580. https://doi.org/10.1099/jgv.0.000377
5. Boshra HY, Charro D, Lorenzo G, Sánchez I, Lazaro B, Brun A, Abrescia NG (2017) DNA vaccination regimes against Schmallenberg virus infection in IFNAR(-/-) mice suggest two targets for immunization. Antivir Res 141:107–115. https://doi.org/10.1016/j.antiviral.2017.02.013
6. Wernike K, Aebischer A, Roman-Sosa G, Beer M (2017) The N-terminal domain of *Schmallenberg* virus envelope protein Gc is highly immunogenic and can provide protection from infection. Sci Rep 7:42500. https://doi.org/10.1038/srep42500
7. Shi X, Goli J, Clark G, Brauburger K, Elliott RM (2009) Functional analysis of the *Bunyamwera orthobunyavirus* Gc glycoprotein. J Gen Virol 90:2483–2492. https://doi.org/10.1099/vir.0.013540-0
8. Coupeau D, Claine F, Wiggers L, Kirschvink N, Muylkens B (2013) In vivo and in vitro identification of a hypervariable region in *Schmallenberg* virus. J Gen Virol 94:1168–1174. https://doi.org/10.1099/vir.0.051821-0
9. Fischer M, Hoffmann B, Goller KV, Höper D, Wernike K, Beer M (2013) A mutation 'hot spot' in the *Schmallenberg* virus M segment. J Gen Virol 94:1161–1167. https://doi.org/10.1099/vir.0.049908-0
10. Varela M, Schnettler E, Caporale M et al (2013) *Schmallenberg* virus pathogenesis, tropism and interaction with the innate

immune system of the host. PLoS Pathog 9(1):e1003133. https://doi.org/10.1371/journal.ppat.1003133

11. Kraatz F, Wernike K, Hechinger S, König P, Granzow H, Reimann I, Beer M (2015) Deletion mutants of *Schmallenberg* virus are avirulent and protect from virus challenge. J Virol 89:1825–1837. https://doi.org/10.1128/JVI.02729-14

12. Coupeau D, Claine F, Wiggers L, Kirschvink N, Muylkens B (2016) S segment variability during the two first years of the spread of *Schmallenberg* virus. Arch Virol 161:1353–1358. https://doi.org/10.1007/s00705-016-2787-x

13. Fehér E, Marton S, Tóth ÁG, Ursu K, Wernike K, Beer M, Dán Á, Bányai K (2017) Sequence analysis of *Schmallenberg* virus genomes detected in Hungary. Acta Microbiol Immunol Hung 64:373–384. https://doi.org/10.1556/030.64.2017.038

14. Izzo F, Cosseddu GM, Polci A, Iapaolo F, Pinoni C, Capobianco Dondona A, Valleriani F, Monaco F (2016) Genetic characterization of Italian field strains of *Schmallenberg* virus based on N and NSs genes. Virus Genes 52:582–585. https://doi.org/10.1007/s11262-016-1335-9

15. Yilmaz H, Hoffmann B, Turan N, Cizmecigil UY, Richt JA, Van der Poel WH (2014) Detection and partial sequencing of *Schmallenberg* virus in cattle and sheep in Turkey. Vector Borne Zoonotic Dis 14:223–225. https://doi.org/10.1089/vbz.2013.1451

16. Larska M, Polak MP, Grochowska M, Lechowski L, Związek JS, Żmudziński JF (2013) First report of *Schmallenberg* virus infection in cattle and midges in Poland. Transbound Emerg Dis 60:97–101

17. Larska M, Tarkowska K, Kuta A, Fidler-Kwiatek E, Ciastek M, Żmudziński JF (2013) Clinical picture of *Schmallenberg* virus infections. Życie Wet 88:488–492

18. Larska M, Krzysiak M, Smreczak M, Polak MP, Żmudziński JF (2013) First detection of *Schmallenberg* virus in elk (Alces alces) indicating infection of wildlife in Białowieża National Park in Poland. Vet J 2198:279–281. https://doi.org/10.1016/j.tvjl.2013.08.013

19. Kęsik-Maliszewska J, Larska M (2016) Detection of *Schmallenberg* virus RNA in bull semen in Poland. Pol J Vet Sci 19:655–657. https://doi.org/10.1515/pjvs-2016-0083

20. Kimura M (1977) Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. Nature 267:275–276

21. Hoffmann B, Scheuch M, Höper D et al (2012) Novel *orthobunyavirus* in cattle, Europe. Emerg Infect Dis 18:469–472. https://doi.org/10.3201/eid1803.111905

22. Padhi A, Ma L (2015) Time-dependent selection pressure on two arthropod-borne RNA viruses in the same serogroup. Infect Genet Evol 32:255–264. https://doi.org/10.1016/j.meegid.2015.03.019

23. Hulst M, Kortekaas J, Hakze-van der Honing R et al (2013) Genetic characterization of an atypical *Schmallenberg* virus isolated from the brain of a malformed lamb. Virus Genes 47:505–514. https://doi.org/10.1007/s11262-013-0975-2

24. Afons A, Abrahantes JC, Conraths F et al (2014) The *Schmallenberg* virus epidemic in Europe-2011-2013. Prev Vet Med 116:391–403. https://doi.org/10.1016/j.prevetmed.2014.02.012

25. Asghar N, Lindblom P, Melik W, Lindqvist R, Haglund M, Forsberg P, Överby AK, Andreassen Å, Lindgren PE, Johansson M (2014) Tick-borne encephalitis virus sequenced directly from questing and blood-feeding ticks reveals quasispecies variance. PLoS ONE 9(7):e103264. https://doi.org/10.1371/journal.pone.0103264. eCollection 2014

26. Marklewitz M, Zirkel F, Kurth A, Drosten C, Junglen S (2015) Evolutionary and phenotypic analysis of live virus isolates suggests arthropod origin of a pathogenic RNA virus family. Proc Natl Acad Sci USA 112:7536–7541. https://doi.org/10.1073/pnas.1502036112

27. Borucki MK, Kempf BJ, Blitvich BJ, Blair CD, Beaty BJ (2002) La Crosse virus: replication in vertebrate and invertebrate hosts. Microbes Infect 4:341–350. https://doi.org/10.1016/S1286-4579(02)01547-2

28. Larska M, Lechowski L, Grochowska M, Żmudziński JF (2013) Detection of the *Schmallenberg* virus in nulliparous Culicoides obsoletus/scoticus complex and C. punctatus-the possibility of transovarial virus transmission in the midge population and of a new vector. Vet Microbiol 166:467–473. https://doi.org/10.1016/j.vetmic.2013.07.015

29. Veronesi E, Antony F, Gubbins S, Golding N, Blackwell A, Mertens PP, Brownlie J, Darpel KE, Mellor PS, Carpenter S (2013) Measurement of the infection and dissemination of bluetongue virus in culicoides biting midges using a semi-quantitative rt-PCR assay and isolation of infectious virus. PLoS ONE 8(8):e70800. https://doi.org/10.1371/journal.pone.0070800

30. Blakqori G, Delhaye S, Habjan M et al (2007) La crosse bunyavirus nonstructural protein NSs serves to suppress the type I interferon system of mammalian hosts. J Virol 81:4991–4999. https://doi.org/10.1128/JVI.01933-06

31. Szemiel AM, Failloux AB, Elliott RM (2012) Role of *Bunyamwera orthobunyavirus* NSs protein in infection of mosquito cells. PLoS Negl Trop Dis 6(9):e1823. https://doi.org/10.1371/journal.pntd.0001823

32. Mohamed M, McLees A, Elliott RM, Anopheles B (2009) Viruses in the Anopheles A and Tete serogroups in the *orthobunyavirus* genus (family Bunyaviridae) do not encode an NSs protein. J Virol 83:7612–7618. https://doi.org/10.1128/JVI.02080-08

33. Margaria P, Bosco L, Vallino M, Ciuffo M, Mautino GC, Tavella L, Turina M (2014) The NSs protein of tomato spotted wilt virus is required for persistent infection and transmission by *Frankliniella occidentalis*. J Virol 88:5788–5802. https://doi.org/10.1128/JVI.00079-14

34. Verbruggen P, Ruf M, Blakqori G, Överby AK, Heidemann M, Eick D, Weber F (2011) Interferon antagonist NSs of La Crosse virus triggers a DNA damage response-like degradation of transcribing RNA polymerase II. J Biol Chem 286:3681–3692. https://doi.org/10.1074/jbc.M110.154799

35. Clem RJ (2016) Arboviruses and apoptosis: the role of cell death in determining vector competence. J Gen Virol 97:1033–1036. https://doi.org/10.1099/jgv.0.000429

36. Blomström AL, Gu Q, Barry G et al (2015) Transcriptome analysis reveals the host response to *Schmallenberg* virus in bovine cells and antagonistic effects of the NSs protein. BMC Genom 16:324. https://doi.org/10.1186/s12864-015-1538-9

37. Rodrigues Hoffmann A, Dorniak P, Filant J, Dunlap KA, Bazer FW, de la Concha-Bermejillo A, Welsh CJ, Varner P, Edwards JF (2013) Ovine fetal immune response to Cache Valley virus infection. J Virol 87:5586–5592. https://doi.org/10.1128/JVI.01821-12

38. Ludwig GV, Israel BA, Christensen BM, Yuill TM, Schultz KT (1991) Role of La Crosse virus glycoproteins in attachment of virus to host cells. Virology 181:564–571

39. Shi X, Brauburger K, Elliott RM (2005) Role of N-linked glycans on *Bunyamwera* virus glycoproteins in intracellular trafficking, protein folding, and virus infectivity. J Virol 79:13725–13734. https://doi.org/10.1128/JVI.79.21.13725-13734.2005

40. Savji N, Palacios G, Travassos da Rosa A, Hutchison S, Celone C, Hui J, Briese T, Calisher CH, Tesh RB, Lipkin WI (2011) Genomic and phylogenetic characterization of Leanyer virus, a novel *orthobunyavirus* isolated in northern Australia. J Gen Virol 92:1676–1687. https://doi.org/10.1099/vir.0.028308-0

41. Varela M, Pinto RM, Caporale M et al (2016) Mutations in the *Schmallenberg* virus Gc glycoprotein facilitate cellular protein synthesis shutoff and restore pathogenicity of NSs deletion

mutants in mice. J Virol 90:5440–5450. https://doi.org/10.1128/JVI.00424-16

42. Crabtree MB, Kent Crockett RJ, Bird BH, Nichol ST, Erickson BR, Biggerstaff BJ, Miller BR (2012) Infection and transmission of Rift valley fever viruses lacking the NSs and/or NSm genes in mosquitoes: potential role for NSm in mosquito infection. PLoS Negl Trop Dis 6(5):e1639

43. Ogawa Y, Kato K, Tohya Y, Akashi H (2007) Sequence determination and functional analysis of the *Akabane* virus (family Bunyaviridae) L RNA segment. Arch Virol 152:971–979