Alongshan virus in Russia

Ivan S. Kholodilov[1,†], Alexander G. Litov[1,†], Alexander S. Klimentov[1,2], Oxana A. Belova[1,3], Alexandra E. Polienko[1], Nikolai A. Nikitin[4], Alexey M. Shchetinin[2], Anna Y. Ivannikova[1], Lesley Bell-Sakyi[5], Alexander S. Yakovlev[1], Sergey V. Bugmyrin[6], Liubov A. Bespyatova[6], Larissa V. Gmyl[1], Svetlana V. Luchinina[7], Anatoly P. Gmyl[1,#], Vladimir A. Gushchin[2,4], Galina G. Karganova[1,8,]*

[1] "Chumakov Institute of Poliomyelitis and Viral Encephalitides" FSBSI "Chumakov FSC R&D IBP RAS", Moscow, Russia

[2] Federal Research Centre for Epidemiology and Microbiology named after the honorary academician N.F. Gamaleya of the Ministry of Health of the Russian Federation, Gamaleya street 18, Moscow 123098, Russia

[3] Martsinovsky Institute of Medical Parasitology, Tropical and Vector Borne Diseases, Sechenov University, Moscow, Russia

[4] Faculty of Biology, Lomonosov MSU, Lenin Hills, 1/12, Moscow 119234, Russia.

[5] Department of Infection Biology, Institute of Infection and Global Health, University of Liverpool, Liverpool L3 5RF, UK

[6] Institute of Biology of Karelian Research Centre of the Russian Academy of Sciences (IB KarRC RAS), Petrozavodsk, Russia.

[7] Russian Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing, Yelkin str., 73, Chelyabinsk 454091, Russia

[8] Institute for Translational Medicine and Biotechnology, Sechenov University, Bolshaya Pirogovskaya st, 2, Moscow, 119991, Russia

[†] These authors contributed equally to this article

[#] Deceased

* Corresponding author: Dr. Galina G. Karganova, karganova@bk.ru; "Chumakov Institute of Poliomyelitis and Viral Encephalitides" FSBSI "Chumakov FSC R&D IBP RAS", prem. 8, k.17, pos. Institut Poliomyelita, poselenie Moskovskiy, Moscow, 108819, Russia; tel. +74958419327, fax +74958419330

Keywords

Alongshan virus, Jingmen tick virus, open reading frame, Flavivirus, Flavi-like virus, *Ixodes persulcatus*, tick cell line

Abstract

In recent decades, many new flavivirus-like viruses have been discovered predominantly in different invertebrates and, as was recently shown, some of them may cause disease in humans. The Jingmen tick virus (JMTV) group holds a special place among flavivirus-like viruses because, in contrast to the "classic" flaviviruses and other flavivirus-like viruses, they have a segmented ssRNA(+) genome. Two segments of the JMTV genome have homology with regions of the flavivirus genome encoding polymerase and helicase-protease proteins. JMTVs have several open reading frames (ORF) in segments encoding glycoprotein and capsid protein and these ORF are specific only to them. JMTVs greatly differ in virion size.

We isolated three strains of Alongshan virus (ALSV), which is a representative of the JMTV group, from adult *Ixodes persulcatus* ticks collected in two geographically-separated Russian regions into the tick cell line IRE/CTVM19. One of the strains persisted in the IRE/CTVM19 cells without cytopathic effect for three years. Most virions purified from tick cells were spherical with a diameter of approximately 40.5 nm. In addition, we found smaller particles of approximately 13.1 nm in diameter with either electron-translucent or electron-dense cores. We obtained full genome sequences of all four segments of two of the isolated ALSV strains, and partial sequences of one segment from the third strain. Phylogenetic analysis on genome segment 2 of the JMTV group clustered our novel strains with other ALSV strains. We found evidence for the existence of a novel upstream ORF in the glycoprotein-coding segment of Alongshan and Jingmen tick viruses.

Introduction

In recent decades, many new flavivirus-like viruses have been discovered in different invertebrates. All of them share homology in the nonstructural proteins with the well-studied Flavivirus RNA-dependent RNA polymerase NS5 and RNA helicase-protease NS3, but they differ greatly from each other regarding other proteins, as well as in overall genome size and structure (1).

The Jingmen tick virus (JMTV) group holds a special place among flavivirus-like viruses because in contrast to the "classic" flaviviruses and other flavivirus-like viruses, they have a segmented ssRNA(+) genome (1–5), which is more commonly observed in viruses of fungi and plants (6). The only other known viruses with a segmented ssRNA(+) genome are

representatives of the family Nodaviridae and infect insects (Alphanodavirus) and fish (Betanodavirus) (7).

Two segments of the JMTV group genome are homologous to the regions of the Flavivirus genome encoding NS3 and NS5 proteins. The remaining segments are specific only to JMTVs. Additionally, in contrast to the "classic" flaviviruses that have one long open reading frame (ORF) and virion size of 40-60 nm, these viruses have several ORFs within single segments and virion sizes varying from 30-35 nm to 70-100 nm (2, 4, 5). It has also been suggested that genomic RNAs of a multicomponent virus can be packaged in separate virions (2), although this has not yet been demonstrated.

The geographic distribution of the JMTV group is very wide, encompassing Asia (1, 4, 8, 9), Europe (9–12), Central and South America (2, 3, 12–14) and Africa (2, 12). Virus RNAs of this group were detected not only in different species of insects, e.g. mosquitoes (5) and *Drosophila melanogaster* (15), but also in the nematode *Toxocara canis* (16, 17) and ticks (1, 4, 9) including *Rhipicephalus microplus* (12, 13), *Ixodes ricinus* (11, 12) and *Ixodes persulcatus* (5, 8). JMTV were isolated from *Rhipicephalus microplus* ticks with restricted *in vitro* multiplication in tick and mammalian cells (3, 4), and from *Amblyomma javanense* ticks with continuous cultivation in a tick cell line (8). Some representatives of the JMTV group were detected in patients with other infections transmitted by arthropods (10).

Alongshan virus (ALSV), a representative of the JMTV group, was detected in *I. persulcatus* ticks, *Culex tritaeniorhynchus* and *Anopheles yatsushiroensis* mosquitoes in China (5) and later in *I. ricinus* ticks in Finland and France (11, 12). ALSV was suspected to infect mammals and cause illness in humans (5, 18).

Our work presents data on the isolation of ALSV in the Russian Federation and new data on the genome organization of this and other known representatives of the JMTV group.

Materials and Methods

Collection and Processing of Ticks

In 2014, from the beginning of May to the middle of June, unfed adult ticks were collected by flagging from vegetation in the Republic of Karelia and Chelyabinsk region (Supplementary file 1). Ticks were identified using taxonomic keys (19, 20). Adult ticks were homogenised in pools of 3-5 individuals according to species composition, location, and route of collection using the laboratory homogeniser TissueLyser II (QIAGEN, Germany) in Medium 199 supplemented with Earle's salts and antibiotics (FSBSI Chumakov FSC R&D IBP RAS, Russia). All homogenates were tested for the presence of flavivirus RNA using a heminested RT-

PCR and primers targeting a conserved region of the flavivirus NS5 gene, as described earlier (21).

Infection of Tick Cell Line

We used a cell line derived from embryos of the tick *I. ricinus* − IRE/CTVM19 (22) provided by the Tick Cell Biobank. The tick cell line was maintained at 28 °C in L-15 (Leibovitz) medium supplemented with 10% tryptose phosphate broth, 20% fetal bovine serum, 2 mM L-glutamine and antibiotics as described earlier (23). Prior to infection, IRE/CTVM19 cells were seeded in flat-sided culture tubes (Nunc, ThermoFisher Scientific) in 2.2 ml of complete medium and incubated at 28 °C. A week later, cells were infected by adding 200 µl of virus-containing material (unfiltered tick homogenate or culture supernate) and incubated at 28 °C. Medium was changed at fortnightly intervals by removal and replacement of 1.5 ml; the spent medium was used to harvest the virus as described below. Occasional subcultures were carried out at intervals between 1 month and >1 year by adding 2.2ml of fresh medium, resuspending the tick cells and transferring half of the cell suspension to a new culture tube.

For transmission electron microscopy and high-throughput sequencing, cell culture supernate (spent medium) was centrifuged at 25,000 rpm for 6 h at 4 °C using an SW-28 rotor in an Optima L-90K Ultracentrifuge (Beckman Coulter, USA).

Transmission Electron Microscopy

The pellet obtained following ultracentrifugation of culture supernate was dissolved overnight in borate saline buffer (0.05M H_3BO_4, 0.12M NaCl, 0.24M NaOH), pH 9.0 (BSB), at 4 °C. A portion of the dissolved pellet was applied to a sucrose density gradient (15-60%) and centrifuged at 30,000 rpm for 4 h at 4°C using an SW-40 rotor in the Optima L-90K Ultracentrifuge (Beckman Coulter, USA). To identify resultant density gradient fractions containing virus particles, aliquots of each fraction were subjected to RT-PCR and sequencing of amplified products as described below.

The specimens after ultracentrifugation were adsorbed onto Formvar film attached to 200-mesh nickel EM grids and contrasted with 2% uranyl acetate (24). The grids were exposed to UV light for one hour to inactivate the virus, air-dried and examined in a JEOL JEM-1400 transmission electron microscope (JEOL, Tokyo, Japan) operated at 80 kV. Images were acquired with an Olympus Quemesa digital camera using iTEM software (Olympus Soft Imaging Solutions GmbH, Munster, Germany).

Reverse-transcriptase PCR (RT-PCR) and Sequencing of Amplified Products

Viral RNA from tick suspensions and culture supernate of infected cells was isolated with TRI Reagent LS (Sigma-Aldrich, USA) according to the manufacturer's protocol. Reverse transcription was performed with random hexamer primers (R6) and M-MLV reverse transcriptase (Promega, Madison, WI) according to the manufacturer's protocol. Viral genomic cDNA was amplified by PCR using specific primers MiassF and MiassR (Supplementary file 2). Sequencing was carried out in both directions directly from PCR-amplified DNA on the ABI PRISM 3730 (Applied Biosystems) sequencer using ABI PRISM® BigDye™ Terminator v. 3.1. Genomic sequences were assembled using SeqMan software (DNAstar, USA).

High-Throughput Sequencing

Total RNA from an ultracentrifugation pellet was fragmented and reverse-transcribed into cDNA with R6 primers using RevertAid reverse transcriptase (ThermoFisher Scientific) followed by second strand synthesis with the NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module (New England Biolabs Inc. USA). Resultant double-stranded DNA was purified using Ampure XP (Beckman Coulter) and subjected as an input for library preparation process using the NEBNext® Fast DNA Library Prep Set for Ion Torrent™ (New England Biolabs Inc. USA) following the manufacturer's instructions. The resultant DNA library was quantified with the Ion Library TaqMan™ Quantitation Kit (ThermoFisher Scientific) followed by templating on the Ion Chef instrument (ThermoFisher Scientific) and sequencing on the Ion S5XL instrument with the viral library constituting a part of the Ion 530 Chip. Raw reads were quality-controlled using FaQCs v2.09 (25) and assembled into contigs using SPAdes v3.13.0 (26) with iontorrent flag. The resultant contigs were screened for viral sequences using Blastn algorithm in BLAST v2.9.0+ with nt database and contigs corresponding to virus were extracted for further investigation.

Phylogenetic Analysis

RNA sequences of all published strains of ALSV, representatives of the JMTV group and the three strains described in this article were used in the phylogenetic analysis. The nucleotide sequences of the genome coding region of segment 2 were aligned using Clustal-X 2.0.11. Phylogenetic analysis was conducted using the neighbour-joining method using MEGA 6.0 with 1000 bootstrap replications.

Bioinformatics tools used for discovery of overlapping elements

For the analysis we used sequences of the full segment 2 of ALSV, JMTV and Yanggou virus extracted from GenBank (08.11.19), as well as sequences of the ALSV strains obtained in

the present study (Supplementary file 3). Sequences were divided into three groups (named "JMTV", "ALSV" and Yanggou virus) based on their positions in the phylogenetic tree (Supplementary file 4). Sequences obtained in the current study were included in the ALSV alignment.

Analyses of Yanggou virus, ALSV and JMTV were performed separately, because of a low level of similarity between them. Nucleotide sequences of the ORF VP1a (in the case of ALSV and Yanggou virus) and VP1 (in the case of JMTV) were extracted from the full segments. The resultant sequences were codon-aligned using the online tool «Codon Alignment v2.1.0» (available at https://hcv.lanl.gov/content/sequence/CodonAlign/codonalign.html) with default parameters. The resultant codon alignments were used as an input for the Synplot2 (27) online tool (available at http://guinevere.otago.ac.nz/cgi-bin/aef/synplot_c.pl) v. 2014-12-05/03:10:08 with default parameters (no reference sequence, n=12 (sliding window of 25 codons)).

Genome annotation and visualization

Putative N-glycosylation sites were predicted using the NetNGlyc 1.0 Server (available online at http://www.cbs.dtu.dk/services/NetNGlyc/) with a standard set of parameters (28). Transmembrane region prediction was done using the TMHMM Server v. 2.0 (available online at http://www.cbs.dtu.dk/services/TMHMM/) (29). Aligned ALSV sequences were examined with the pAliKiss web server (build in June 25 2015, available https://bibiserv.cebitec.uni-bielefeld.de/palikiss) with default parameters (30) to find possible conserved RNA structures downstream of putative frameshift sites. ORFs, transmembrane regions and N-glycosylation sites were mapped on the genome using SnapGene Viewer v. 4.3.10. Further image editing was performed in the GNU Image manipulation program v. 2.10 software.

Results

Detection and Isolation of Alongshan virus strains

We collected unfed adult *I. persulcatus* ticks in Chelyabinsk region and in the Republic of Karelia in May–June 2014 (Supplementary file 1). The ticks were pooled, homogenised and the homogenates screened for flavivirus NS5 RNA. The PCR products from three flavivirus-positive tick pools were Sanger-sequenced. The resultant sequences were analysed with BLAST. All three isolates, named Miass519, Miass527 and Galozero-14-T20426, were found to be 90.48 – 91.03% similar to Alongshan virus.

To isolate the virus from tick homogenate, we inoculated cultures of the IRE/CTVM19 tick cell line. An IRE/CTVM19 culture inoculated with the Miass527 sample was kept for a year with regular medium changes but without passaging or subculture; individual tick cell cultures can be maintained for long periods without subculture (22). Since the virus did not cause any cytopathic effect in these cells, we changed the medium once every 14 days. Once every 1-2 months we checked the culture medium by RT-PCR for presence of the virus, using specific primers MiassF and MiassR (Supplementary file 2). After a year of maintenance, 200µl of supernate from the Miass527-infected IRE/CTVM19 culture was passaged into a new flat-sided culture tube with uninfected IRE/CTVM19 cells. This new Miass527-infected culture was subcultured by splitting the cells at a ratio of 1:1 twice in the first 6 months of infection and after that was kept without passaging or subculture. In this condition Miass527-infected cells have been monitored for 3 years (observation is still continuing). IRE/CTVM19 cells inoculated with samples Miass519 and Galozero-14-T20426 were each subcultured 1:1 twice in the first 2 months after infection and, at the time of writing, have been monitored for 3.5 months .

RT-PCR analysis revealed that the strain Miass527 successfully persisted in the tick cell line for a year, followed by one passage and further three years persistence at the time of writing. The other two strains, Miass519 and Galozero-14-T20426, had persisted for at least 3.5 months at the time of writing.

For two strains, Miass519 and Miass527, high-throughput sequencing of virus purified from IRE/CTVM19 cell supernate was carried out to obtain full genomes, while for the Galozero-14-T20426 strain partial sequences of segment 2 were obtained via Sanger sequencing using primers JMun1S and Miass_gly_1R (Table 1, Supplementary file 2). For the phylogenetic analysis, we used RNA sequences from 18 GenBank entries of representatives of the JMTV group and the three ALSV strains isolated in the present study (Table 1). The phylogenetic analysis was performed on a 1732 bp fragment of segment 2. Interestingly, one of the JMTV isolates (MN095520 – isolate JMTV/I.ricinus/France) clustered with the ALSV sequences. For consistency, we included this JMTV sequence in all our analyses of ALSV (see below). All three of our strains belonged to ALSV and formed one monophyletic group with the strain H3 isolated in China (Figure 1).

Transmission Electron Microscopy

We used cell culture supernate collected after one passage of strain Miass527 in IRE/CTVM19 cells, at 17, 19, 20 and 32 months after initial isolation for transmission electron microscopic examination. The IRE/CTVM19 tick cell line was recently reported to contain RNA sequences with similarity to rhabdoviruses (31). To prevent confusion between ALSV and this

putative rhabdovirus in our microscopic examination, we first screened the pellet obtained after ultracentrifugation of uninfected IRE/CTVM19 for rhabdovirus RNA using specific primers (Supplementary file 2). We detected an unknown rhabdovirus and named it IRE/CTVM19-associated rhabdovirus. Since this new rhabdovirus was also detected in the Miass527-infected IRE/CTVM19 cell pellet, we subjected the dissolved pellet to sucrose density gradient separation and selected the fraction with the largest amount of ALSV RNA and very low levels of rhabdovirus RNA (Supplementary file 5). The microscopy revealed that most virions of the strain Miass527 were enveloped spherical particles with a diameter of 40.5±3.7 nm (Figure 2, A, B, Supplementary file 6). In the same gradient fraction we found small spherical particles with a diameter of 13.1±2.1 nm, which had electron-translucent or electron-dense cores (Figure 2, C, D, Supplementary file 7). The small electron-translucent particles could be virions with incomplete genome, protein structures or some other unidentified component of IRE/CTVM19 cells. The viral particles with an electron-dense core could have lost genomic RNA (Figure 2, B, D, Supplementary file 7).

Full genome, genome structure and proposed novel elements

We sequenced full genomes of the strains Miass527 and Miass519. Segments 1, 3 and 4 of the newly-isolated Miass strains were more similar to each other than to previously-described ALSV strains. However, segment 2 of the strain Miass527 was more similar to the Chinese strain H3 than to Miass519, while segment 2 of the strain Miass519 was closer to JMTV/I.ricinus/France than to Miass527 (Table 2).

Overall, the full genomes of the Miass strains demonstrated >88% nucleotide identity to the other strains of ALSV known to date, with glycoprotein-coding segment 2 having the highest similarity and NS5-like protein coding segment 1 being the most divergent.

The two Miass strains were similar in their genome organization to each other and to known ALSV. Segments 1 and 3 were monocistronic. Segment 1 contained a flavivirus-NS5-like protein with polymerase and putative methyltransferase domains and a predicted signal peptide on the N-terminus of the protein. Segment 3 contained a flavivirus NS3-like protein with proposed functions of serine protease and RNA helicase (Figure 3). Both proteins had transmembrane regions corresponding to the previously mapped strain H3. Segment 4 contained two non-Flavivirus-related ORFs. There was a slippery sequence (G GTT TTT) and some evidence of an RNA-structured region (Supplementary file 8) and we believe that the VP3 ORF may be expressed via -1 ribosome frameshifting (Figure 3), as was previously shown for other ALSV strains. Regarding the VP1a and VP1b ORFs in segment 2, transmembrane regions and N-glycosylation sites were also in positions in the predicted amino acid sequences similar to the

other ALSV strains (5). Interestingly, we were not able to identify signal peptides in either VP2 or VP1a using TMHMM v.2.0 software.

We performed a search of the functional elements in the VP1a region with the Synplot2 program (27). Using codon alignment of four previously-described ALSV strains (MH158416, MN107154, MN107158, MN095520) and the three strains sequenced in this study, we discovered two regions of high conservation in the VP1a ORF (Figure 4A). The first one corresponded to the 3`-terminal region of the ORF. We believe that such high conservation in this region was caused by an overlap with the VP1b ORF. We managed to find a slippery motif (A AAA AAC) followed by a spacer and an RNA structure (Supplementary file 9) in this region of the VP1a ORF that theoretically would allow VP1b to be translated via -1 ribosome frameshift (Figure 3).

Additionally, Synplot2 allowed us to identify an unusually high conservation rate in the first ~110 codons of VP1a ORF in segment 2. We believe that this conservation was caused by an additional ORF that was found in this region. This novel upstream ORF (nuORF) is 399 nt long, is located upstream of the VP1a ORF and has a 365 nt overlap with it (Figure 4A).

Application of the tblastn algorithm using the protein sequence encoded by the Miass527 nuORF as an entry revealed homologs within the JMTV group (Supplementary file 3). Upon further inspection, all of the resultant sequences were shown to have a nuORF. Thus we decided to test known JMTV sequences for existence of the additional ORF upstream of the VP1 region with Synplot2.

We searched for a full-length segment 2 of JMTV using blast algorithm and strain SY84 JMTV full segment 2 sequence as an entry. Since relationships within the JMTV group are not yet decided and some representatives of this group have different names, we could not use a GenBank search. Sequences MN095532 and KX377514 were found to have unknown nucleotides in the middle of the segment and were not used in the analysis. For these sequences, nuORF was confirmed to be intact. We then constructed a phylogenetic tree (Supplementary file 4) that allowed us to identify three groups (Jingmen, Yanggou and Alongshan viruses) that should be analysed separately due to phylogenetic distance. All JMTV sequences with full segment 2 (21 entries) known to date (November 8, 2019) have a nuORF. We performed analysis of the JMTV segment 2 sequences (Figure 4B) and Synplot2 showed high conservation in the first ~100-codon region. This demonstrates that the region associated with nuORF is conserved more highly than would be expected under neutral evolution in both ALSV and JMTV sequences.

We also tested the Yanngou virus sequences and, although all of the full genome sequences had a nuORF, Synplot2 program did not find a region of higher-than-expected

conservation in the first ~100 codons of Yanngou virus VP1a (Supplementary file 10). However there are only three known full-length segment 2 sequences of Yanggou virus and the number of sequences is important to acquire statistics needed for a prediction.

Overall, the ALSV nuORF encoded 132 amino acids. According to the prediction, this protein may have an N-terminal signal sequence and three transmembrane regions with a C-terminal tail, thus making the nuORF product a small membrane protein (Figure 5A).

The predicted protein encoded by the nuORF of JMTVs was shorter (~106 aa) and had ~50% identity to the ALSV nuORF product. It had an overall structure similar to that of ALSV, with a proposed N-terminal signal sequence and three predicted transmembrane regions, but, in contrast to the ALSV nuORF, it had no C-terminal tail (Figure 5B).

Discussion

Due to the development of molecular virology, high-throughput sequencing and viral metagenomics, many new flavivirus-like viruses have been discovered in recent decades (1, 4, 14). A large number of flavivirus-like viruses have been detected in or isolated from arthropods (1, 32, 33), but their ability to cause infection and disease in vertebrates and their medical significance are not known.

The JMTV group is a recently-discovered group of segmented ssRNA(+) viruses that infect different species of invertebrates, including insects, ticks and nematodes (1, 4, 8, 11, 15–17). Moreover, some representatives of the new flavivirus-like viruses were detected in patients who had other severe infectious diseases (10), and these viruses might themselves cause illness in humans (5, 8). There is still little information about the replication and overall coding strategy of this novel virus group.

ALSV is a representative of the JMTV group. It was first isolated in China from a patient suffering fever of unknown aetiology (5). Subsequently, the virus was detected in *I. persulcatus* ticks and *Culex tritaeniorhynchus* and *Anopheles yatsushiroensis* mosquitoes in China (5) and in *I. ricinus* ticks in Finland (11). It was shown that ALSV is able to infect mammals and cause illness in humans (with symptoms including headache, fever, fatigue and depression) (5, 18). Thus, additional isolates of this virus from new geographical locations will be important for understanding its distribution, as well as for the local healthcare system. New genome data could also provide additional information on the structural organization of ALSV genome.

We isolated three strains of ALSV from *I. persulcatus* ticks collected in Chelyabinsk region, in the south of European Russia and in the Republic of Karelia in north-west Russia. We sequenced two full genomes of the new ALSV strains Miass527 and Miass519 from Chelyabinsk region and obtained partial sequences of segment 2 for the ALSV strain Galozero-14-T20426

from the Republic of Karelia. Although these three strains were isolated from territories distant from each other, phylogenetic analysis of the segment 2 sequences showed that they form a single monophyletic group with the strain H3 isolated in China.

The segments 1, 3 and 4 of the strains Miass519 and Miass527 were more similar to each other than to previously-described ALSV strains. Segment 2 of the strain Miass527 was more similar to the strain H3 than to Miass519, while segment 2 of the strain Miass519 was closer to the isolate JMTV/I.ricinus/France then to Miass527. This data suggests that reassortment of genome segments within ALSV could occur. Similar observations were made earlier for JMTV (4).

Chinese strains of ALSV were isolated from mammalian blood samples using Vero (African green monkey) cells (5, 18) while the isolation of the virus detected in Finnish ticks failed in Vero and two human cell lines (11). The Chinese JMTV strain isolated from *A. javanense* ticks replicated successfully in the *R. microplus* embryo-derived tick cell line BME/CTVM23 but failed to grow in Vero, DH82 (canine monocyte-macrophage-derived cells) and C6/36 (*Aedes albopictus* mosquito) cells (8). We isolated three tick-derived strains of ALSV using the tick cell line IRE/CTVM19; our results, together with those of the previous studies reporting successful *in vitro* isolation and continuous propagation of JMTV group viruses (5, 8, 18), suggest the possibility that mammal-derived viruses will more readily infect mammalian cells, while tick-derived viruses may be more infective for tick cells. However, additional isolations of JMTV group viruses will be required to test this theory and the possible underlying genetic causes. For the first time we showed that the ALSV strain Miass527 can persist in tick cells for at least three years. It was previously shown that the tick cell line IRE/CTVM19 may harbour an unknown rhabdovirus (31). Our experiment confirmed the presence of a rhabdovirus and suggests that it does not interfere with the persistence and active reproduction of Alongshan virus in IRE/CTVM19 cells. Persistent infection in tick cells and ticks could affect virus properties, as was previously shown for tick-borne encephalitis virus (34).

Previously, it was shown that the virion sizes of JMTV group representatives ranged from 30-35 nm to 70-100 nm (2, 4, 5). ALSV virions were reported to be enveloped spherical (or nearly spherical) particles with a diameter of approximately 80-100 nm (5). The size of virions of our strain Miass527 was 40.5±3.7 nm. In transmission electron microscopy, some of the viral particles had electron-dense cores, and could hypothetically be defective virions without genomic RNA (35). In addition, we found smaller particles with a diameter of 13.1±2.1 nm, which had either electron-translucent or electron-dense cores. Presumably, they could represent either: a) glycoprotein aggregates, similar to the “classic” flavivirus nonstructural protein 1, which is secreted from infected cells (36, 37), b) protein structures appearing after long-term

persistence, c) virions with incomplete genome, or d) other unknown structures present in IRE/CTVM19 cells. Further investigations are needed to clarify this point.

Although the strain Miass527 of ALSV did not differ in overall genome structure from the previously isolated ALSV (5), the newly obtained data allowed us to propose a novel protein encoded by both ALSV and JMTV and, probably, by some other members of the JMTV group. The nuORF, proposed here, is conserved in all the full segment 2 sequences of ALSV, JMTV and Yanngou virus available at the time of writing. Considering that ALSV and JMTV are somewhat distantly related (around 70% nucleotide identity and around 80% amino acid identity in the NS5-like protein), it is likely that the nuORF indeed encodes a functional virus protein. Nevertheless, further studies are needed to confirm its expression and function during the virus life cycle.

Interestingly, an ORF of a more distant member of the JMTV group, the mosquito-borne Guaico Culex virus (2), was discovered in the same genome region. It also encodes a small protein, but with no amino acid identity to the nuORF. A mass spectrometry analysis of the Guaico Culex virus virion proteins did not identify any peptides that belong to this putative ORF, thus making its product (if it is indeed expressed) a nonstructural protein (2).

If the nuORF product protein is indeed functional at some point during ALSV replication, it would most likely be a minor membrane protein. Such proteins are present in a variety of RNA viruses and perform different functions. For example, in flaviviruses, small membrane proteins NS2B and NS4A (with sizes similar to the nuORF) perform multiple functions, ranging from enzyme cofactors (38, 39) to participation in particle formation (40, 41) and protection against the cellular antiviral response. Influenza A M2 protein acts as a proton-selective ion channel (42), whereas Coronavirus E protein acts as a viroporin and plays a role in the induction of the regulation of the unfolded protein response (43). Nevertheless, confirmation of the nuORF expression and understanding of its possible functions should be subjects for further studies.

Conclusion

We isolated three strains of ALSV from *I. persulcatus* ticks in Russia. Although these strains were isolated from territories far from each other, in phylogenetic analysis they formed a single monophyletic group with the virus isolated in China. Our study confirmed the value of tick cell lines in virus isolation and maintenance of persistent infection. Availability of these novel ALSV strains has enabled better understanding of the phylogenetic relationships within the JMTV group, and will facilitate further study of the origins, structure, replication, epidemiology,

transmission, vertebrate hosts and medical and veterinary significance of these enigmatic segmented RNA viruses.

Acknowledgments

Tick collection by Sergey V. Bugmyrin and Liubov A. Bespyatova is supported by the state order (project № 0218-2019-0075). Lesley Bell-Sakyi is supported by the United Kingdom Biotechnology and Biological Sciences Research Council grant number BB/P024270/1.

References

1. M. Shi, *et al.*, Divergent Viruses Discovered in Arthropods and Vertebrates Revise the Evolutionary History of the Flaviviridae and Related Viruses. *J. Virol.* **90**, 659–69 (2016).
2. J. T. Ladner, *et al.*, A Multicomponent Animal Virus Isolated from Mosquitoes. *Cell Host Microbe* **20**, 357–367 (2016).
3. S. R. Maruyama, *et al.*, Characterisation of divergent flavivirus NS3 and NS5 protein sequences detected in Rhipicephalus microplus ticks from Brazil. *Mem. Inst. Oswaldo Cruz* **109**, 38–50 (2014).
4. X.-C. Qin, *et al.*, A tick-borne segmented RNA virus contains genome segments derived from unsegmented viral ancestors. *Proc. Natl. Acad. Sci.* **111**, 6744–6749 (2014).
5. Z.-D. Wang, *et al.*, A New Segmented Virus Associated with Human Febrile Illness in China. *N. Engl. J. Med.* **380**, 2116–2125 (2019).
6. E. J. Lefkowitz, *et al.*, Virus taxonomy: The database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res.* **46**, D708–D717 (2018).
7. C. Y. Yong, S. K. Yeap, A. R. Omar, W. S. Tan, Advances in the study of nodavirus. *PeerJ* **2017**, 1–31 (2017).
8. N. Jia, *et al.*, Emergence of human infection with Jingmen tick virus in China: A retrospective study. *EBioMedicine* **43**, 317–324 (2019).
9. E. Dinçer, *et al.*, Survey and Characterization of Jingmen Tick Virus Variants. *Viruses* **11**, 1071 (2019).
10. P. Emmerich, *et al.*, Viral metagenomics, genetic and evolutionary characteristics of Crimean-Congo hemorrhagic fever orthonairovirus in humans, Kosovo. *Infect. Genet. Evol.* **65**, 6–11 (2018).
11. S. Kuivanen, *et al.*, Detection of novel tick-borne pathogen , Alongshan virus , in Ixodes ricinus ticks , south-eastern Finland ,. *Eurosurveillance* **24**, pii=1900394 (2019).
12. S. Temmam, *et al.*, Insights into the Host Range, Genetic Diversity, and Geographical Distribution of Jingmenviruses. *mSphere* **4**, 1–13 (2019).
13. J. de O. Pascoal, S. M. de Siqueira, R. da C. Maia, M. P. Juan Szabó, J. Yokosawa, Detection and molecular characterization of Mogiana tick virus (MGTV) in Rhipicephalus microplus collected from cattle in a savannah area, Uberlândia, Brazil. *Ticks Tick. Borne. Dis.* **10**, 162–165 (2019).
14. E. C. Villa, S. R. Maruyama, I. K. F. de Miranda-Santos, G. Palacios, J. T. Ladner, Complete Coding Genome Sequence for Mogiana Tick Virus, a Jingmenvirus Isolated from Ticks in Brazil. *Genome Announc.* **5**, 17–18 (2017).
15. C. L. Webster, *et al.*, The discovery, distribution, and evolution of viruses associated with

drosophila melanogaster. *PLoS Biol.* **13**, 1–33 (2015).

16. D. M. Callister, A. D. Winter, A. P. Page, R. M. Maizels, Four abundant novel transcript genes from Toxocara canis with unrelated coding sequences share untranslated region tracts implicated in the control of gene expression. *Mol. Biochem. Parasitol.* **162**, 60–70 (2008).
17. K. K. A. Tetteh, A. Loukas, C. Tripp, R. M. Maizels, Identification of abundantly expressed novel and conserved genes from the infective larval stage of Toxocara canis by an expressed sequence tag strategy. *Infect. Immun.* **67**, 4771–4779 (1999).
18. Z.-D. Wang, *et al.*, Prevalence of the emerging novel Alongshan virus infection in sheep and cattle in Inner Mongolia , northeastern China. *Parasit. Vectors* **12**, 1–7 (2019).
19. N. A. Filippova, *Ixodid Ticks of the Subfamily Ixodinae. Fauna of the USSR: Arachnoides. Vol. 4(4)* (Nauka, 1977).
20. N. A. Filippova, *Fauna of Russia and neighbouring countries. Ixodid ticks of subfamily Amblyomminae. Arachnoidea. Vol.4(5)* (Nauka, 1997).
21. N. Scaramozzino, *et al.*, Comparison of Flavivirus universal primer pairs and development of a rapid, highly sensitive heminested reverse transcription-PCR assay for detection of flaviviruses targeted to a conserved region of the NS5 gene sequences. *J. Clin. Microbiol.* **39**, 1922–1927 (2001).
22. L. Bell-Sakyi, E. Zweygarth, E. F. Blouin, E. A. Gould, F. Jongejan, Tick cell lines: tools for tick and tick-borne disease research. *Trends Parasitol.* **23**, 450–457 (2007).
23. S. Weisheit, *et al.*, Ixodes scapularis and Ixodes ricinus tick cell lines respond to infection with tick-borne encephalitis virus: Transcriptomic and proteomic analysis. *Parasites and Vectors* **8**, 1–26 (2015).
24. N. Nikitin, *et al.*, Comparative study of non-enveloped icosahedral viruses size. *PLoS One* **10**, 1–11 (2015).
25. C. C. Lo, P. S. G. Chain, Rapid evaluation and quality control of next generation sequencing data with FaQCs. *BMC Bioinformatics* **15**, 1–8 (2014).
26. A. Bankevich, *et al.*, SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
27. A. E. Firth, Mapping overlapping functional elements embedded within the protein-coding regions of RNA viruses. *Nucleic Acids Res.* **42**, 12425–12439 (2014).
28. R. Gupta, E. Jung, S. Brunak, Prediction of N-glycosylation sites in human proteins. **46**, 203–206 (2004).
29. A. Krogh, È. Larsson, G. Von Heijne, E. L. L. Sonnhammer, Predicting Transmembrane Protein Topology with a Hidden Markov Model□: Application to Complete Genomes. *J.*

Mol. Biol. **305**, 567–580 (2001).

30. S. Janssen, R. Giegerich, The RNA shapes studio. *Bioinformatics* **31**, 423–425 (2015).
31. L. Bell-Sakyi, H. Attoui, Virus discovery using tick cell lines. *Evol. Bioinforma.* **12**, 31–34 (2016).
32. E. Huhtamo, *et al.*, Characterization of a Novel Flavivirus from Mosquitoes in Northern Europe That Is Related to Mosquito-Borne Flaviviruses of the Tropics. *J. Virol.* **83**, 9532–9540 (2009).
33. E. Huhtamo, *et al.*, Novel flaviviruses from mosquitoes: Mosquito-specific evolutionary lineages within the phylogenetic group of mosquito-borne flaviviruses. *Virology* **464**–**465**, 320–329 (2014).
34. O. A. Belova, *et al.*, Properties of the tick-borne encephalitis virus population during persistent infection of ixodid ticks and tick cell lines. *Ticks Tick. Borne. Dis.* **8**, 895–906 (2017).
35. V. K. Korboukh, *et al.*, RNA Virus Population Diversity: An Optimum for Maximal Fitness and Virulence. *J. Biol. Chem.* **289**, 29531–29544 (2014).
36. A. C. Alcalá, *et al.*, The dengue virus non-structural protein 1 (NS1) is secreted efficiently from infected mosquito cells. *Virology* **488**, 278–287 (2016).
37. D. A. Muller, *et al.*, Structure of the dengue virus glycoprotein non-structural protein 1 by electron microscopy and single-particle analysis. *J. Gen. Virol.* **93**, 771–779 (2012).
38. B. D. Lindenbach, C. M. Rice, Molecular biology of flaviviruses. *Adv. Virus Res.* **59**, 23–61 (2003).
39. S. a Shiryaev, A. V Chernov, A. E. Aleshin, T. N. Shiryaeva, A. Y. Strongin, NS4A regulates the ATPase activity of the NS3 helicase: a novel cofactor role of the non-structural protein NS4A from West Nile virus. *J. Gen. Virol.* **90**, 2081–5 (2009).
40. X. Li, *et al.*, Transmembrane Domains of NS2B Contribute to both Viral RNA Replication and Particle Formation in Japanese Encephalitis Virus. *J. Virol.* **90**, 5735–5749 (2016).
41. J. L. Muñoz-Jordan, G. G. Sánchez-Burgos, M. Laurent-Rolle, A. García-Sastre, Inhibition of interferon signaling by dengue virus. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 14333–8 (2003).
42. D. DiMaio, Viral Miniproteins. *Annu. Rev. Microbiol.* **68**, 21–43 (2014).
43. D. Schoeman, B. C. Fielding, Coronavirus envelope protein□: current knowledge. *Virol. J.* **16**, 1–22 (2019).

Legends to Figures

Figure 1. Phylogenetic tree of a 1732 bp fragment of segment 2 of Alongshan virus and other members of the Jingmen tick virus group.

Phylogenetic trees were constructed using MEGA 6.0 with the neighbour-joining method (1000 bootstrap replications). Bootstrap values (>70%) are shown at the branches. GenBank accession numbers are listed for each strain. ▼ Strains from the present study.

Figure 2. Transmission electron microscopy. Electron micrographs of purified viral particles of the Alongshan virus strain Miass527 propagated in IRE/CTVM19 cells. Small particles with electron-translucent (c) or electron-dense (d) core. Samples were stained with 2% uranyl acetate. Scale bars, 100 nm.

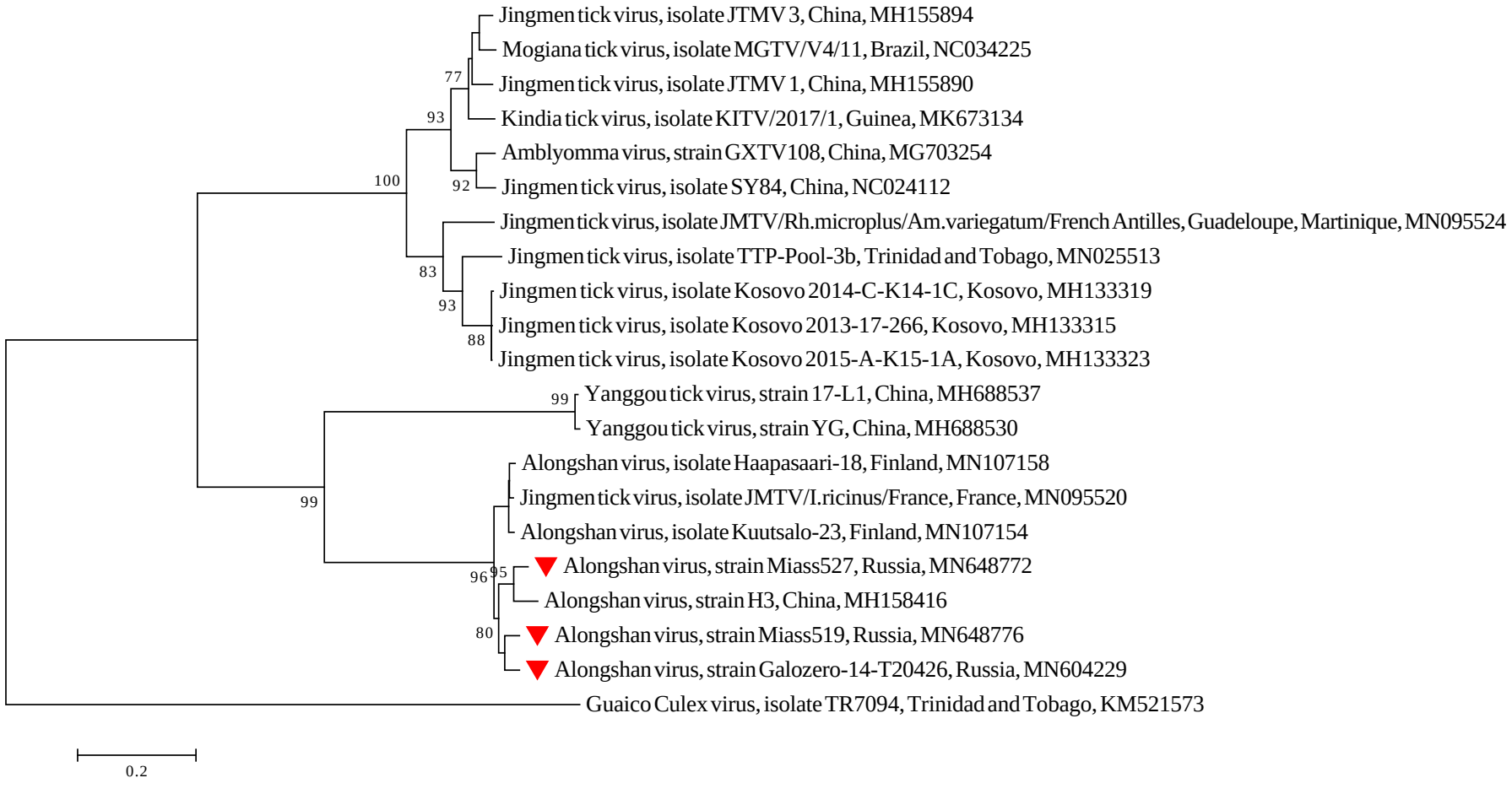
Figure 3. Genome of the Alongshan virus strain Miass527.

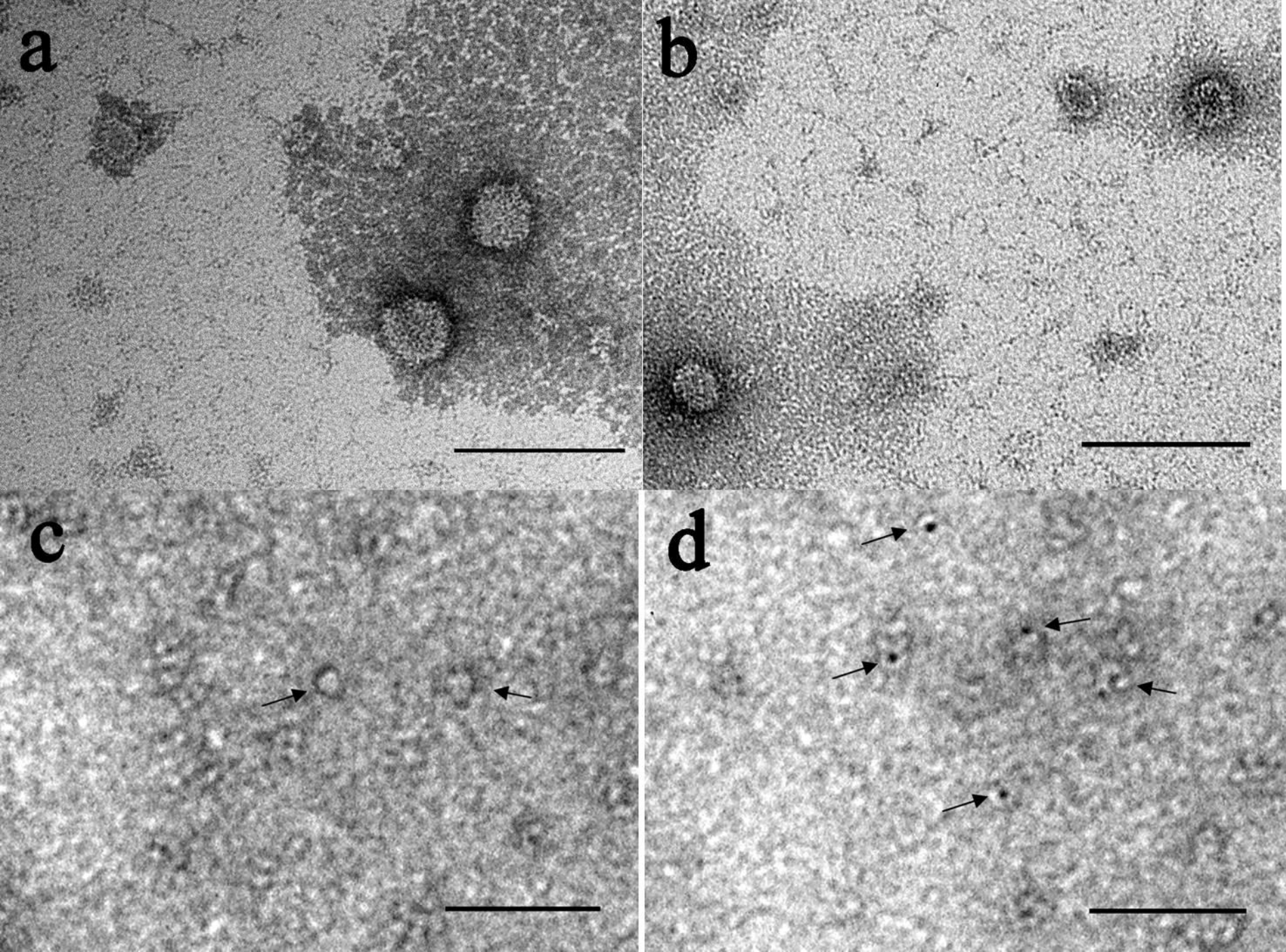
Transmembrane regions are shown in yellow, the part of the VP1b protein derived from the proposed ribosomal frameshift is shown in blue, and the proposed novel protein in segment 2 is in green. Blue stars mark the putative N-glycosylation sites. Proposed ribosomal frameshift sites are in blue.

Figure 4. Assosiation of the nuORF with the region of high conservation in the Alongshan VP1a and Jingmen tick virus VP1 open reading frames.

(A) Codon alignment of 7 full sequences of the Alongshan virus VP1a ORF was used in the analysis (Supplementary file 3). (B) Codon alignment of all 21 full sequences of the Jingmen tick virus VP1 ORF was used in the analysis (Supplementary file 3). In each section the top panel represents the virus genome map of the tested region. The middle panel depicts probability that the degree of ORF conservation within a 25-codon sliding window could be obtained under neutral evolution. The grey dashed line indicates p=0.005 significance (after correcting for multiple tests, where the number of tests is the length of a coding sequence divided by the window size). The bottom panel displays the relative amount of synonymous-site conservation at a 25-codon sliding window by showing the ratio of the observed number of synonymous substitutions to the expected number. The analysis was done using the Synplot2 program (27).

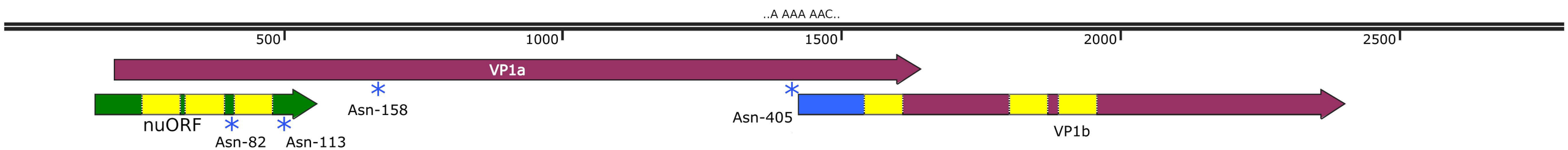
Figure 5. Transmembrane domains of the proposed nuORF product of (A) Alongshan virus strain Miass527, (B) Jingmen tick virus strain SY84. The analysis was done using TMHMM Server v. 2.0 program (29).


Jingmen tick virus, isolate JTMV 3, China, MH155894
Mogiana tick virus, isolate MGTV/V4/11, Brazil, NC034225
Jingmen tick virus, isolate JTMV 1, China, MH155890
Kindia tick virus, isolate KITV/2017/1, Guinea, MK673134
Amblyomma virus, strain GXTV108, China, MG703254
Jingmen tick virus, isolate SY84, China, NC024112
Jingmen tick virus, isolate JMTV/Rh.microplus/Am.variegatum/French Antilles, Guadeloupe, Martinique, MN095524
Jingmen tick virus, isolate TTP-Pool-3b, Trinidad and Tobago, MN025513
Jingmen tick virus, isolate Kosovo 2014-C-K14-1C, Kosovo, MH133319
Jingmen tick virus, isolate Kosovo 2013-17-266, Kosovo, MH133315
Jingmen tick virus, isolate Kosovo 2015-A-K15-1A, Kosovo, MH133323
Yanggou tick virus, strain 17-L1, China, MH688537
Yanggou tick virus, strain YG, China, MH688530
Alongshan virus, isolate Haapasaari-18, Finland, MN107158
Jingmen tick virus, isolate JMTV/I.ricinus/France, France, MN095520
Alongshan virus, isolate Kuutsalo-23, Finland, MN107154
Alongshan virus, strain Miass527, Russia, MN648772
Alongshan virus, strain H3, China, MH158416
Alongshan virus, strain Miass519, Russia, MN648776
Alongshan virus, strain Galozero-14-T20426, Russia, MN604229
Guaico Culex virus, isolate TR7094, Trinidad and Tobago, KM521573
77
93
100
92
83
93
88
99
99
96
95
80
0.2

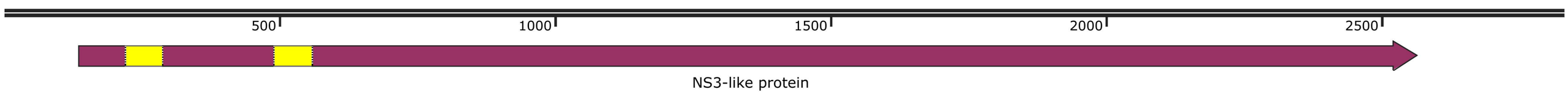

a
b
c
d

Created with SnapGene®

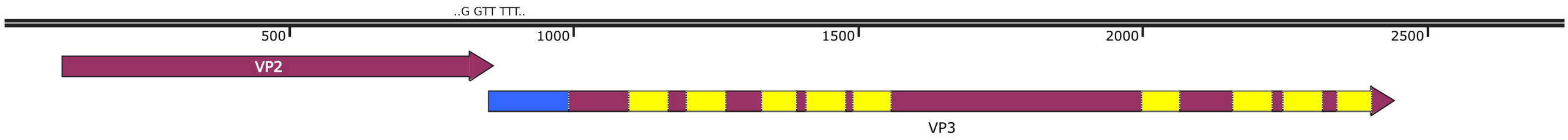
Alongshan virus strain Miass527 segment 1
3006 bp



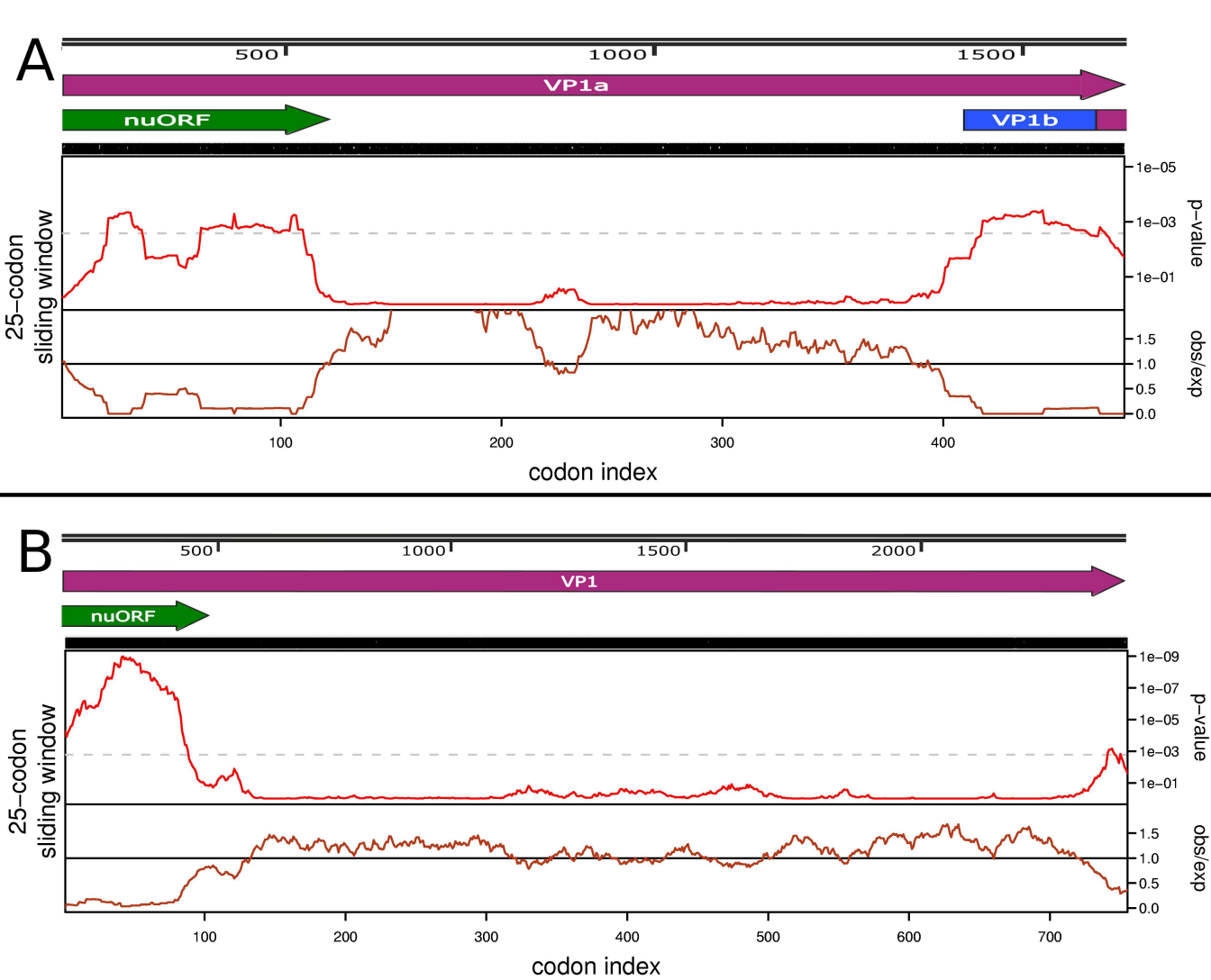
Alongshan virus strain Miass527 segment 2
2790 bp



Alongshan virus strain Miass527 segment 3
2828 bp



Alongshan virus strain Miass527 segment 4
2737 bp


A
500
1000
1500
VP1a
nuORF
VP1b
25−codon
sliding window
1e−05
1e−03
1e−01
p−value
1.5
1.0
0.5
0.0
obs/exp
100
200
300
400
codon index
B
500
1000
1500
2000
VP1
nuORF
25−codon
sliding window
1e−09
1e−07
1e−05
1e−03
1e−01
p−value
1.5
1.0
0.5
0.0
obs/exp
100
200
300
400
500
600
700
codon index

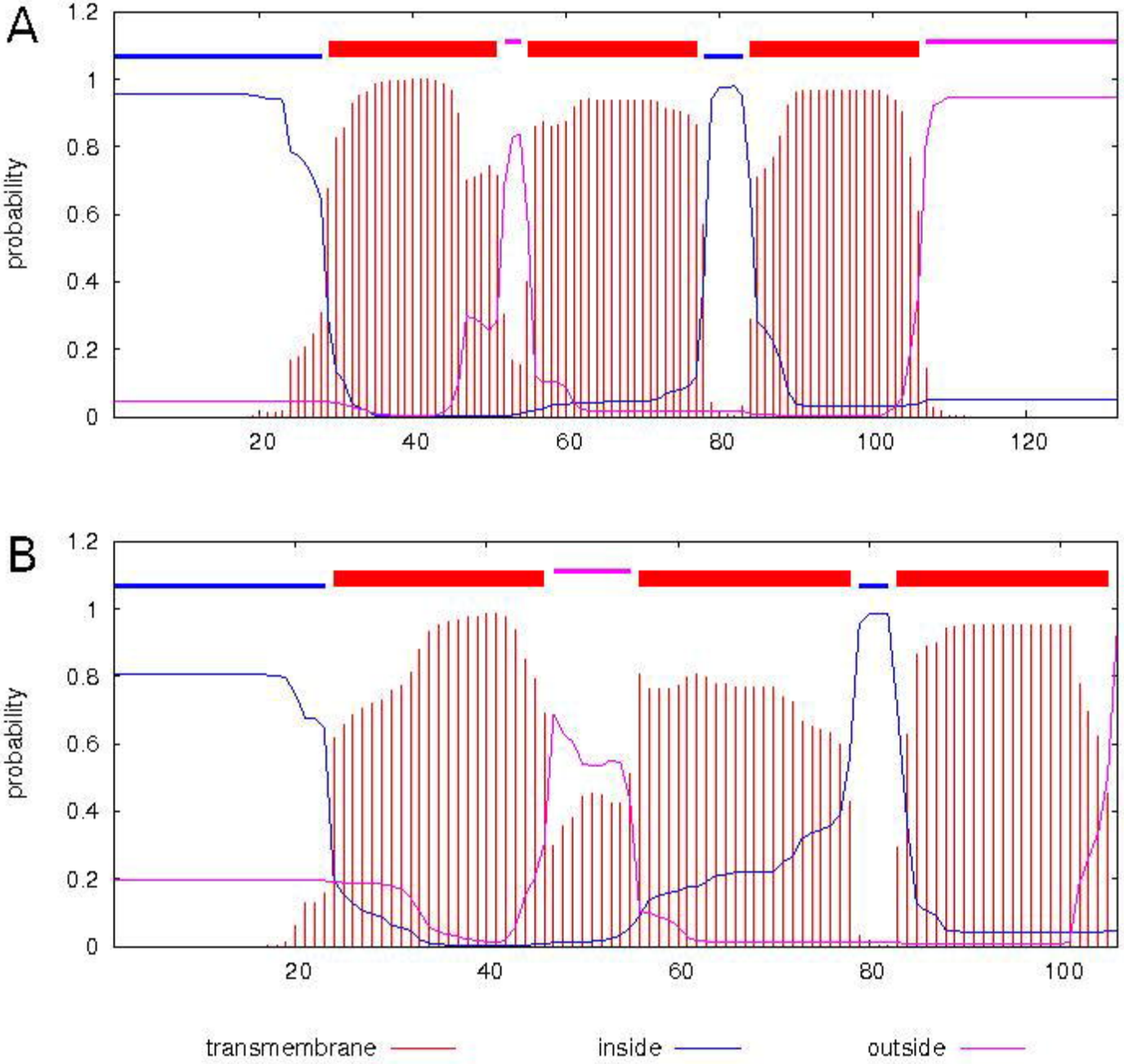

A
B
probability
1.2
1
0.8
0.6
0.4
0.2
0
20
40
60
80
100
120
transmembrane
inside
outside

Table 1. Alongshan virus strains isolated in unfed adult *Ixodes persulcatus* ticks from two geographically separate regions of Russia

Strain	Tick species	Year, region (GPS)	GenBank access. no.
Miass519	*I. persulcatus*	2014, Chelyabinsk region, Ilmen State Reserve (55°01.287´N 060°10.097´E)	MN648774 – MN648777
Miass527	*I. persulcatus*		MN648770 – MN648773
Galozero-14-T20426	*I. persulcatus*	2014, The Republic of Karelia (62°4.515´N 33°57.077´E)	MN604229

Table 2. Percentage nucleotide identity of the four genome segments of different strains of Alongshan virus for which full genomes were available at the time of writing. Genbank accession numbers for each strain and isolate are as follows: Miass519 (MN648774 – MN648777) (this study); Miass527 (MN648770 – MN648773) (this study); H3 (MH158415 – MH158418) (5); Kuutsalo-23 (MN107153 – MN107155) and Haapasaari-18 (MN107157 – MN107160) (11); JMTV/I.ricinus/France (MN095519 – MN095522) (12)

	Strain Miass527	Strain H3	Strain Kuutsalo-23	Strain Haapasaari-18	Isolate JMTV/I.ricinus /France
Segment 1. NS5-like protein					
Strain Miass519	90.6%	89.0%	89.2%	88.9%	89.1%
Strain Miass527		89.6%	89.2%	88.6%	89.7%
Strain H3			89.6%	89.2%	90.1%
Strain Kuutsalo-23				95.6%	96.0%
Strain Haapasaari-18					95.9%
Segment 2. VP1a and VP1b					
Strain Miass519	92.6%	92.0%	93.7%	93.6%	94.0%
Strain Miass527		94.7%	92.4%	92.2%	93.0%
Strain H3			91.7%	91.6%	92.1%
Strain Kuutsalo-23				98.3%	98.5%
Strain Haapasaari-18					98.6%
Segment 3. NS3-like protein					
Strain Miass519	96.7%	91.2%	91.3%	90.9%	91.2%
Strain Miass527		90.7%	91.1%	90.5%	91.0%
Strain H3			90.9%	91.3%	91.5%
Strain Kuutsalo-23				95.0%	97.4%
Strain Haapasaari-18					94.5%
Segment 4. VP2 and VP3					
Strain Miass519	98.6%	90.5%	90.8%	90.7%	91.2%
Strain Miass527		90.6%	90.7%	90.8%	91.3%
Strain H3			90.6%	90.3%	91.1%
Strain Kuutsalo-23				94.7%	98.0%
Strain Haapasaari-18					95.3%