Ticks and Tick-borne Diseases 10 (2019) 269–279

Contents lists available at ScienceDirect

Ticks and Tick-borne Diseases

journal homepage: www.elsevier.com/locate/ttbdis

Sequencing and genetic characterization of two strains Paramushir virus obtained from the Tyuleniy Island in the Okhotsk Sea (2015)

Marina V. Safonova[a], Mikhail Yu. Shchelkanov[b,c,d], Kamil Khafizov[e], Alina D. Matsvay[e], Andrey A. Ayginin[e], Anna S. Dolgova[e], Egor M. Shchelkanov[b], Ekaterina V. Pimkina[e], Anna S. Speranskaya[e,f], Irina V. Galkina[b], Vladimir G. Dedkov[g,h,*]

[a] *Anti-Plague Center, Federal Service on Consumers' Rights Protection and Human Well-Being Surveillance, Moscow, Russia*
[b] *Far Eastern Federal University, School of Biomedicine, Vladivostok, Russia*
[c] *Federal Scientific Center of Eastern Asia Terrestrial Biodiversity, Vladivostok, Russia*
[d] *National Scientific Center of Marine Biology, Vladivostok, Russia*
[e] *Central Research Institute for Epidemiology, Federal Service on Consumers' Rights Protection and Human Well-Being Surveillance, Moscow, Russia*
[f] *Lomonosov State University, Moscow, Russia*
[g] *Saint-Petersburg Pasteur Institute, Center, Federal Service on Consumers' Rights Protection and Human Well-Being Surveillance, Saint-Petersburg, Russia*
[h] *Martsinovsky Institute of Medical Parasitology, Tropical and Vector Borne Diseases, Sechenov First Moscow State Medical University, Moscow, Russia*

ARTICLE INFO

Keywords:
Broad-range PCR
High throughput sequencing
Paramushir virus
Orthonairoviruses
Russia

ABSTRACT

Paramushir virus belongs to Sakhalin virus genogroup within *Orthonairovirus* genus and is one of the poorly studied viruses with unknown pathogenicity. At the moment, only one nearly complete sequence of Paramushir virus genome, isolated in 1972, is available. Two new strains of PARV were isolated in 2015 from a sample collected at the Tyuleniy Island in the Okhotsk Sea and sequenced using a combination of high throughput sequencing and specific multiplex PCR. Both strains are closely related to the early sequenced PARV strain LEIV–1149 K. The signs of intersegment reassortment and probable recombination were revealed, which point to a high variability potential of Paramushir virus and may lead to the formation of strains with novel properties, different from those of the predecessors.

The new data regarding Paramushir virus can promote a better understanding of the diversity and relations within *Orthonairovirus* genus and help define intragenic demarcation criteria, which have not yet been established.

1. Introduction

At present, *Orthonairovirus* genus (formerly *Nairovirus* genus) contains 34 species; of them, 22 have not yet been classified (ICTV, 2016). All members of the *Orthonairovirus* genus have a negative-sense tripartite RNA genome (small [S], medium [M], and large [L] segments) with a high genetic diversity. Some of the orthonairoviruses were isolated from Eurasia, including the Crimean-Congo hemorrhagic fever virus (CCHFV), which is the causative agent of severe fever illness in humans (Deyde et al., 2006). Despite this fact, the information regarding biodiversity, spreading, and epidemiological potential of orthonairoviruses (excluding CCHFV, due to its medical significance) is insufficient. Paramushir virus (PARV) is one of the orthonairoviruses, for which the taxonomy has not yet been defined. After Dr. D. K. Lvov found two strains of PARV in 1972, no new strain has been identified until now (Lvov et al., 1976). Those strains (LEIV–2238 and LEIV–1149 K) were isolated from *Ixodes signatus* ticks. The ticks were collected from seabird nests at Paramushir Island (the Kuril Islands in the Okhotsk Sea) (Fig. 1a, c). At the moment, PARV, along with Avalon virus (AVAV), Tillamook virus (TILV), Sakhalin virus (SAKV), Taggert virus (TAGV), and Clo Mor virus (CMV) belong to the Sakhalin genogroup within *Orthonairovirus* genus (Walker et al., 2016). A partial sequence of the PARV L gene, strain LEIV–2238, is available in the NCBI GenBank under accession no KF801657 from March 2014, whereas the whole CDS sequence of PARV strain LEIV–1149K was submitted in 2017 and is now available in the NCBI GenBank under the accession nos. KP792717 – KP792719.

In this work, we describe two new strains of Paramushir orthonairovirus, which were isolated in Tyuleniy Island, located in the Far Eastern part of Russia in the Okhotsk Sea (see Fig. 1a and b). Both

---

[*] Corresponding author at: 197101, Saint-Petersburg Pasteur Institute, 14 Mira Street, Saint-Petersburg, Russian Federation.
E-mail address: vgdedkov@yandex.ru (V.G. Dedkov).

https://doi.org/10.1016/j.ttbdis.2018.11.004
Received 16 August 2018; Received in revised form 14 October 2018; Accepted 8 November 2018

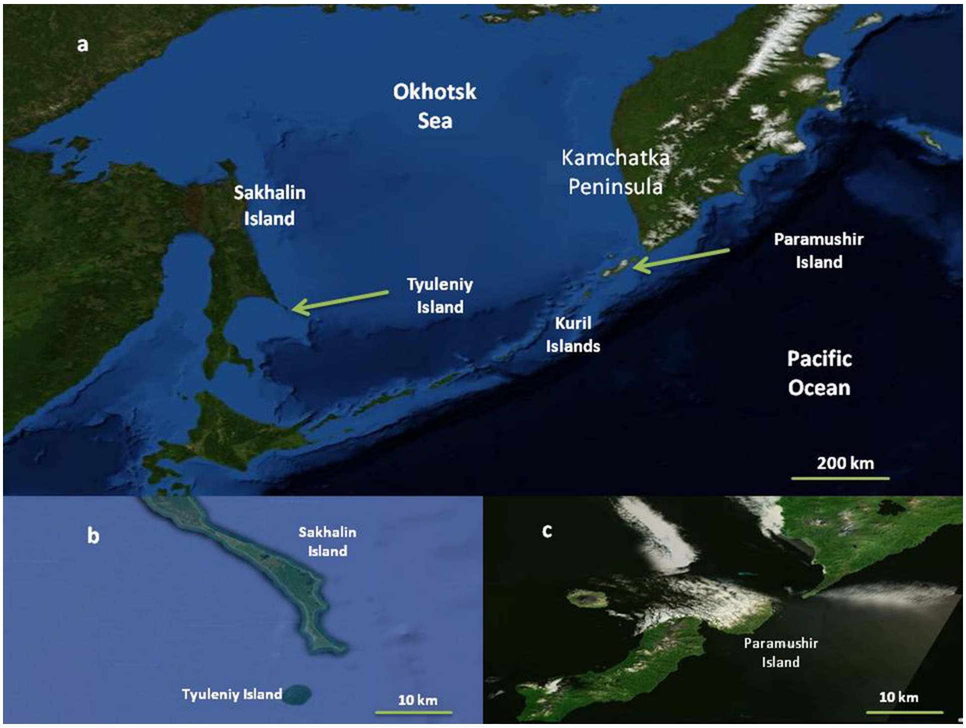


Fig. 1. Island territories of the Russian Far East (courtesy Google maps).
a– Sakhalin Island and Kuril Ridge.
b– Tyuleniy Island – the territory where PARV strain 133 and PARV strain 137 were isolated.
c– Paramushir Island – the territory where the first isolation of PARV occurred.

strains were isolated in 2015 from *Ixodes uriae* White, 1852 (Acari: Ixodidae) ticks, and this was the first case of isolation of PARV after 43 years of the PARV discovery in the Far East of Russia.

Tyuleniy Island is known to be the source of many viruses isolated from *I. uriae* ticks. These viruses include Tyuleniy virus (TYUV), belonging to *Flavivirus* genus in the *Flaviviridae* family (Lvov et al., 2001, 2014a, 2015); Sakhalin virus, (SAKV), belonging to *Orthonairovirus* genus in the *Nairoviridae* family (Lvov et al., 2001, 2014b, 2015); Zaliv Terpeniya virus (ZTV), Komandory virus (KOMV) and Rukutama virus (RUKV), which belong to *Phlebovirus* genus in the *Phenuiviridae* family (Lvov et al., 2001, 2014b, 2014c, 2015, Alkhovsky et al., 2013); Okhotskiy virus (OKHV) and Aniva virus (ANIV), which belong to Orbivirus genus within the *Reoviridae* family (Lvov et al., 2001, 2014d, 2015). Thereby, this island is of great interest to the scientists who study the biodiversity and ecology of arboviruses. However, due to a large distance from the mainland and harsh climate, the accessibility to this area is limited. Nevertheless, in recent years, the interest in studying the island territories of the far Eastern part of Russia has increased, and some expeditions have been organized to Tyuleniy Island, including the one in August 2015 after a 25-year break. (Shchelkanov et al., 2017). In the course of the expedition, two strains of PARV were isolated, then sequenced using high throughput sequencing techniques, and subsequently characterized.

2. Materials and methods

2.1. Sample preparation and RNA extraction

During an arbovirus surveillance and control program, *I. uriae* ticks were collected from the nests of birds located in Tyuleniy Island (48°29′ N., 144° 38′ E) in the Okhotsk Sea near Sakhalin island (Fig. 1a–c).

The identity and sex of the ticks were determined based on the morphology following earlier recommendations (Filippova, 1977). The ticks were classified according to species, sex, and developmental stages, and later processed in groups of 10 (immature ticks) and 25 individuals (adults). All the ticks were stored alive in a wet chamber until separation. The grouped tick samples ($n = 30$) were homogenized using Tissue Lyser LT (Qiagen, Germany) for 5 min in 100 µL of phosphate buffered saline (pH 7.0, Dako, Denmark) before extraction, centrifuged at 10,000 *g* for 10 min, and the supernatant was used for nucleic acid extraction.

Total nucleic acids were extracted and purified using the RIBO-prep DNA/RNA extraction kit (k2-9-Et-100CE, AmpliSens, Russia), according to the recommendations of the manufacturer. DNA/RNA was eluted with 50 µL of the elution buffer (AmpliSens, Russia) and stored at −70 °C until evaluation.

2.2. Broad-range reverse transcriptase-polymerase chain reaction (RT-PCR)

Broad-range RT-PCR was performed using a set of broadly reactive degenerate oligonucleotides designed to target each viral species that could be transmitted by ticks within the following genera: *Flavivirus*, *Orthonairovirus*, *Phlebovirus*, *Orthobunyavirus*, *Orthoreovirus,* and *Orbivirus* (Table 1).

Reverse transcription and touch-down amplification were performed using the SuperScript III One-Step RT-PCR System with Platinum Taq DNA Polymerase (Invitrogen, USA) in a reaction mixture containing 2 µL of RNA, 0.4 µM sense primer, 0.4 µM antisense primer, 12.5 µL of 2X reaction buffer, 1 µL of enzyme mix, 0.2 µL of 50 mM Mg_2SO_4, 1 µL of bovine serum albumin (BSA; 1 ng/µL), and H_2O in a total volume of 25 µL.

The following thermal cycling parameters were used: 45 °C for 20 min and 94 °C for 3 min, followed by 5 cycles of 94 °C for 20 s, 55 °C to 45 °C (−2 °C per cycle) for 20 s, and 72 °C for 30 s; and then 40 cycles of 94 °C for 20 s, 45 °C for 20 s, 72 °C for 30 s; and finally 72 °C for 3 min. All amplification steps were performed using Maxy Gene gradient thermocycler (Axygen, USA). The products of the amplification

Table 1
Primers used for broad-range PCR.

Primer set no	Genera	Primer name	Sequence 5’-3’	Expected Product Length, bp	References
1	Flavivirus	PhlP2	GGCTACTTCAARAAYAARGANGA	507	Klimentov et al., 2016
		PhlM2	CTCTCTCAGICCICCRTGYTG		
2	Orthonairo-virus	Nairo-F	TCTCAAAGAAACACGTGCCGC	400	Lambert and Lanciotti, 2009
		Nairo-R	GTCCTTCCTCCACTTGWGRGCAGCCTGCTGGTA		
3	Phlebovirus	Phlebo F1	TTTGCTTATCAAGGATTTGACC	370	Lambert and Lanciotti, 2009
		Phlebo R1	TCAATCAGTCCAGCAAAGCTGGGATGCATCAT		
4	Ortho-bunyavirus	Bun F	CTGCTAACACCAGCAGTACTTTTGAC	222-286	Lambert and Lanciotti, 2009
		Cal/Bwa F	GCAAATGGATTTGATCCTGATGCAG		This study
		Oropouche F	GGCCCATGGTTGACCTTACTTT		
		Wyeomyia-F	ATGTCTGAAATTGTATTTGATGATATTGG		
		Bun R	TGGAGGGTAAGACCATCGTCAGGAACTG		
		Cal/Bwa R	TTGTTCCTGTTTGCTGGAAAATGAT		
		Oropouche R	ACCAAAGGGAAGAAAGTGAAT		
		Wyeomyia-R	TATTTCGATTCCCCGGAAAGT		
5	Orbivirus	OrbiVP1 F2295a	CGGAGCAGTATGTGGGNGATGAYA	228	Dedkov et al., 2016
		OrbiVP1 F2295b	CGGAGATGTATGTGGGNGATGAYA		
		OrbiVP1-R2523a	TCWGAWGAKATKATCATCATTCGRTCTTG		
		OrbiVP1 R2523b	TCAGACGAGATCAACATCATTCGRTCTTG		
		OrbiVP1 R2523c	TCAGAGGAGACCATCATCATTCGRTCTTG		
		OrbiVP1 R2523d	TCCGATGAAATGAACATCATTCGRTCTTG		
		OrbiVP1 R2523e	TCGGACGAAACTAGCATCATTCGRTCTTG		
		OrbiVP1 R2523f	TCAGAACTGACAATCATCATTCGRTCTTG		
6	Orthoreovirus	Orthoreo -2090F	GGBTCMACNGCYACYTCBACYGAGCA	244	Landolfi et al., 2010
		Orthoreo -2334R	CDATGTCRTAHWYCCANCCRAA		

reactions were analyzed by 1.2% agarose gel electrophoresis.

All PCR products of the expected size were cut out from the gel under UV illumination, extracted using a MinElute gel extraction kit (Qiagen, Germany) and ligated into the pGEM-T plasmid vector (Promega, USA). These plasmids were transformed into *Escherichia coli* (XL1 blue strain) (Maniatis et al., 1989) and 10 white colonies for each construct were picked and sequenced using standard M13R primers and ABI Prism 3500 XL sequencer (Applied Biosystems, USA). The obtained sequences were examined using the Basic Local Alignment Search Tool (BLAST, Altschul et al., 1990) NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.3. Primer design for the complete genome sequencing

In order to obtain the complete genome sequences of two PARV strains (excluding the 5′ and 3′ ends), a total of 60 primer pairs were designed (Table 2) using a web-based primer design tool – Primal Scheme (http://primal.zibraproject.org). This tool offers a complete pipeline for the development of efficient multiplex primer schemes for generation of the overlapping products; the size of which is determined by the target genome length, amplicon length, and required overlap (Quick et al., 2017). For PARV, we used the amplicon lengths of about 400 nts with a 50-nt overlap. The sequences of PARV strain LEIV–1149 K (GenBank NCBI ac. nos. KP792717 – KP792719) were used as reference genomes.

2.4. Library preparation and near complete genome sequencing

Reverse transcription reaction was performed using random hexanucleotide primers and Reverta-L kit (AmpliSens, Russia) according to the manufacturer’s instructions. The cDNA was stored at −70 °C and used subsequently as a template for amplification. The designed primers were assorted into three groups, each containing twenty primer pairs. As a result, the three groups of 200–400 bp DNA fragments were amplified that were suitable for the subsequent 400-bp sequencing on an Ion S5 System (Table 2).

Hot-start multiplex PCR amplification reactions were performed in a 25-µL total volume containing 2 µL of the template cDNA, 0.1 µM of each sense primer, 0.1 µM of each antisense primer, 2.5 µL of dNTPs (1.76 mM, AmpliSens, Russia), and 10 µL of PCR buffer blue-2 with 7.5 mM $MgCl_2$ (AmpliSens, Russia). The following thermal cycling parameters were employed: 94 °C for 2 min; 40 cycles at 94 °C for 10 s, 55 °C for 15 s, 72 °C for 60 s; and a final extension at 72 °C for 5 min. The reactions were performed in a MaxyGene gradient thermocycler (Axygen, USA). The products were analyzed by 1.2% agarose gel electrophoresis (Fig. 2a).

The library preparation included 5′ phosphorylation and incorporation of barcoded adapters followed by the amplification of the final library. For this purpose, T4 Polynucleotide Kinase and T4 DNA Ligase (New England Biolabs, USA) were used with a slightly modified manufacturer's protocol. Amplification was performed using the PCR-mix 2 FEP/FRT kit (AmpliSens, Russia).

The concentrations of the fragments were measured with a Qubit 2.0 fluorimeter (Invitrogen, USA) using the Qubit dsDNA H Assay Kit (Invitrogen, USA). The quality assessment of the final libraries was carried out on an Agilent 2100 Bioanalyzer instrument (Agilent Genomics, USA), employing the Agilent High Sensitivity DNA Kit (Agilent Genomics, USA) (Fig. 2b). The size selection of the final library was carried out on 2% E-Gel™ SizeSelect™ II Agarose Gels (Thermo Fisher Scientific, USA) on an E-Gel electrophoresis system (Thermo Fisher Scientific, USA).

Sequencing was performed using an Ion S5 System (Thermo Fisher Scientific, USA) with Ion 530 chips (Thermo Fisher Scientific, USA).

2.5. Genetic analysis

The overall quality of the Ion S5 (Thermo Fisher Scientific) reads was confirmed using the FastQC software (Andrews, 2010), and low-quality sequences were filtered by the PRINSEQ program (Schmieder and Edwards, 2011). The filtered reads were aligned to the reference viral protein sequences retrieved from the NCBI database (NCBI Resource Coordinators, 2017) using the BLASTx program (Altschul et al., 1990) to remove all non-target sequences. All the remaining reads were then mapped to the reference genome of Paramushir virus (GenBank NCBI ac. nos. KP792717 – KP792719) to confirm the coverage of the target regions using bwa mem tool (Li and Durbin, 2009). The

Table 2
Primers used for near complete genome sequencing of PMRV.

Region No.	Pool	Left Primer Name	Left Primer Sequence	Right Primer Name	Right Primer Sequence	Product Length
1	1	ParamL_1_LEFT	TCCTTTAACTGGCATCCCTCTCT	ParamL_1_RIGHT	CGTAACGCATGACATGGTCACA	386
2	2	ParamL_2_LEFT	AACGAATGGGGCTCAACACTTG	ParamL_2_RIGHT	TGGGGATCAGCTTTCGAACTCT	407
3	3	ParamL_3_LEFT	GTTGGTGAACGAGGAAGAGCTG	ParamL_3_RIGHT	GTCCTTAGCAACCCAGGTAGGT	373
4	1	ParamL_4_LEFT	CTCTGACTCTAAGAACACTTGGTCAT	ParamL_4_RIGHT	GGATGCTCGTACAGGTCCTTCT	365
5	2	ParamL_5_LEFT	AGAAGAAGTTCATAACCAACTGCCT	ParamL_5_RIGHT	GCTCTGTCAACTGCTCCTTGTC	382
6	3	ParamL_6_LEFT	TGAAACAGAAGACTTTGGTCACCT	ParamL_6_RIGHT	AGGAAGGATTGCCTGTGTTCCT	404
7	1	ParamL_7_LEFT	AGGAGAGTCTGAAGCTGAGCAC	ParamL_7_RIGHT	TGAACTCTGGGTGAAGCTGAGT	400
8	2	ParamL_8_LEFT	ATGGAGGTATGTGCCAGAGGAC	ParamL_8_RIGHT	TTGATGAGGAGGCCGTCTTCAG	406
9	3	ParamL_9_LEFT	CGGTGCCATTGAAGTGTTCGAA	ParamL_9_RIGHT	TTAACCACTGACCCTGCCTTGA	427
10	1	ParamL_10_LEFT	TCCTCAAATCTTCAATCAGCTTTCTCT	ParamL_10_RIGHT	GGTCTTCCTTTAGCCAGGCAAC	424
11	2	ParamL_11_LEFT	TCATTGACGCCTTTGAGAGAACA	ParamL_11_RIGHT	GTGAGCCTTACCAGCCTATCCA	368
12	3	ParamL_12_LEFT	CCCACAGTGAGACGAAGAGTTG	ParamL_12_RIGHT	CTGGCACTTTGAGTTGGCTTGA	403
13	1	ParamL_13_LEFT	AGTTGGCAGAGGAGAAGCTGAA	ParamL_13_RIGHT	GCCTCAACATCAATTCCGCAGA	361
14	2	ParamL_14_LEFT	AGGAGACTACATCAGCAACAAGC	ParamL_14_RIGHT	CAGGATGTAGCTGTAGGAGGCT	431
15	3	ParamL_15_LEFT	TGACCTGCTGTTTATACAATAGCAGT	ParamL_15_RIGHT	GGAGCTGTTTGTTTTGCTGCTG	424
16	1	ParamL_16_LEFT	ACACTCATGACCACTTTATAAGGGTT	ParamL_16_RIGHT	CGATGCAGCCCTCATCAAAGTT	430
17	2	ParamL_17_LEFT	TCTATGGAGCCTGGGTCAACAG	ParamL_17_RIGHT	TCCCTTTCTACCTTGAACCCTGA	428
18	3	ParamL_18_LEFT	AGGTCGGTTGGTTCCAGTAACA	ParamL_18_RIGHT	AGCATGTGTGTCTTTGGGACTG	387
19	1	ParamL_19_LEFT	TCCCACCTGATGCAATAGAAAAGT	ParamL_19_RIGHT	GCTGCTAAGGTCTTAAGCCAGC	419
20	2	ParamL_20_LEFT	TTCTACCACTGGAAGGAGCTGG	ParamL_20_RIGHT	AGTTAGAGTTGGGAGTGCAGCT	370
21	3	ParamL_21_LEFT	TTCAGAGTGCCTGGGAGAAGAC	ParamL_21_RIGHT	CCAAGAGTTGGCAGTCAGGTTC	376
22	1	ParamL_22_LEFT	CCCTGGTGTGTCCACTACAAGT	ParamL_22_RIGHT	CCATAGAAACTCATGCCTGCCG	370
23	2	ParamL_23_LEFT	GCGTGAGAAGGTTGTTCAAGGT	ParamL_23_RIGHT	CCATCATCGTTCGTTGTGCTCA	427
24	3	ParamL_24_LEFT	AGGCTGTCGAGATCTACTGGTG	ParamL_24_RIGHT	TGCAGAGGGTATTTCCACCTGT	389
25	1	ParamL_25_LEFT	GTGGAATGATGCAGCAGCTACT	ParamL_25_RIGHT	TGAAAAGAAGTCTCGCGCTAGC	387
26	2	ParamL_26_LEFT	AGCTACAATCACATGGGTCAAGG	ParamL_26_RIGHT	TGATCACTGCCGGTGTAACTCT	377
27	3	ParamL_27_LEFT	GCCCGATGCTGTCAGATGAAAG	ParamL_27_RIGHT	AGCCGAGCAAGAGATAACCTCA	409
28	1	ParamL_28_LEFT	TGACGCTTGGGTCATTATCACC	ParamL_28_RIGHT	TGGACTTTGGATCCAGAGGTCC	409
29	2	ParamL_29_LEFT	GCCAGAACCATCCAAGATTACTCC	ParamL_29_RIGHT	GTTGCAAACCTGGTAGGAGCTC	407
30	3	ParamL_30_LEFT	GAAGACCCAATGATACAACTCATCCC	ParamL_30_RIGHT	GTCTCTGTAGTTCGTGCTGTGC	361
31	1	ParamL_31_LEFT	ACTCTCAGTGCCTCTGTACACA	ParamL_31_RIGHT	CAAGTTCTGCTTCTGGAGGCTG	414
32	2	ParamL_32_LEFT	CAATGACCTAGACATACTTCACAAAGAATT	ParamL_32_RIGHT	ACTGCAGTCCAGATCCTGTTGT	371
33	3	ParamL_33_LEFT	CCACAAGGATAGATCTATCAAACCTGC	ParamL_33_RIGHT	TGGCAGAGATGCAGTGGAGAAT	382
34	1	ParamL_34_LEFT	TCGTATGGGATCAAAGAAGGAAGGT	ParamL_34_RIGHT	GGCCCTCTCCCCATGTCTTAAA	413
35	2	ParamL_35_LEFT	AAGGAAAAGTATAAGTGACCTTATCGACT	ParamL_35_RIGHT	AGAGAGACGGAGGCTATCTTGC	370
36	3	ParamL_36_LEFT	CGGACTTAGTCTCAAAGGAACAGC	ParamL_36_RIGHT	AGTTTCGGAGAAAGTTGCATTTGC	384
37	1	ParamL_37_LEFT	AGAGACTCAGACCACATACCATGA	ParamL_37_RIGHT	TCCTTGTCAGTGATGTTGCTGT	367
38	2	ParamL_38_LEFT	TCCTGTCAAGTGTTCTCTGGCA	ParamL_38_RIGHT	ACTCATAAACGCATCAATGGTGCT	382
39	3	ParamL_39_LEFT	GCTTTCCGCTGTTTGGAACTTC	ParamL_39_RIGHT	AGCATTAGGACGGACACAGTGA	361
40	1	ParamL_40_LEFT	TGTCGATTGTCAGAAGCAGAAGC	ParamL_40_RIGHT	CTCTCGGTTGTTGCTGCTGATG	429
41	2	ParamL_41_LEFT	CCACCATTGGCAACCTCTTCAT	ParamL_41_RIGHT	TCTGGCTCAGGAACCACTTCTT	368
42	3	ParamM_1_LEFT	AGACCCACCTTCACATGCTCTG	ParamM_1_RIGHT	ACAAAATTAAGGCCGGTCCTGC	408
43	1	ParamM_2_LEFT	AGTCTGTTTACCACTTTCATTGAAGGT	ParamM_2_RIGHT	TTGTCACAGGAGGCCTTCTTGA	381
44	2	ParamM_3_LEFT	CAAAGGAAGATAGAGCACCTTGGT	ParamM_3_RIGHT	AATTGTGCACTCCTCGTTGTGG	375
45	3	ParamM_4_LEFT	TGGTACCATGTCGAATTGGTGG	ParamM_4_RIGHT	GTAAGCTTGTCACCCTTGTGCC	378
46	1	ParamM_5_LEFT	GCATCCAACTCCACTTTGATTCCA	ParamM_5_RIGHT	GCATTGGGCAGCTTATGTCGAT	431
47	2	ParamM_6_LEFT	CTTCAGTGATGCAGTGGGTGAC	ParamM_6_RIGHT	GAACTTGGTGCTGCAGTAAGGG	395
48	3	ParamM_7_LEFT	TGCCCTGAATGCCTAGACAGAT	ParamM_7_RIGHT	CTGGTTGTCACACTGGCTAACC	385
49	1	ParamM_8_LEFT	TCATGCTTGTCTCTCCGGTGAA	ParamM_8_RIGHT	CTTTGCTGCTGGAAAGTGACCA	424
50	2	ParamM_9_LEFT	TGGGACTCAGAAGAAGAAGTTGGT	ParamM_9_RIGHT	CATGAGAACTGCCAGGTGTTCC	406
51	3	ParamM_10_LEFT	GCAAGAAATTGGGGCTGCAATC	ParamM_10_RIGHT	CGCTCTTGACAAGGTGATCCAC	431
52	1	ParamM_11_LEFT	ACATCCTGACATGGTCACGACT	ParamM_11_RIGHT	GCTTTGCCTTGACATCCCAAGT	375
53	2	ParamM_12_LEFT	ACACAAAGTCGAAAGAGTGCTCT	ParamM_12_RIGHT	ACTGACAGGGGAGAACAGCTTT	398
54	3	ParamM_13_LEFT	GCCTGACAGATTTGCAGTGCAT	ParamM_13_RIGHT	ATGACAAGCCCAGCGCAAATAA	424
55	1	ParamM_14_LEFT	TGCTAGAAAACAGAGGGACCATCA	ParamM_14_RIGHT	TCAGTGAGGTGGGAGAATGCTT	383
56	2	ParamS_1_LEFT	GCAAGCTGGACTTTGAGGACAA	ParamS_1_RIGHT	CACTTGCAGTCTCGGTGAACTC	436
57	3	ParamS_2_LEFT	TCAAGAAGCGTGTCCCTCTTCT	ParamS_2_RIGHT	TACAAGTTTGCACAGGCCAGTG	372
58	1	ParamS_3_LEFT	TTAGGCGTAGGAACCTTGTGCT	ParamS_3_RIGHT	ATCATTGCACCGCTTGACTCTG	368
59	2	ParamS_4_LEFT	AGATGGTTGAGGAGCCTGACAA	ParamS_4_RIGHT	CATGCAGCCATATCTGACAGCC	401
60	3	ParamS_5_LEFT	CCTTCAAGTGGGGTAAGGGGAT	ParamS_5_RIGHT	GCCCTTCACTCTTGAGACGTTG	366

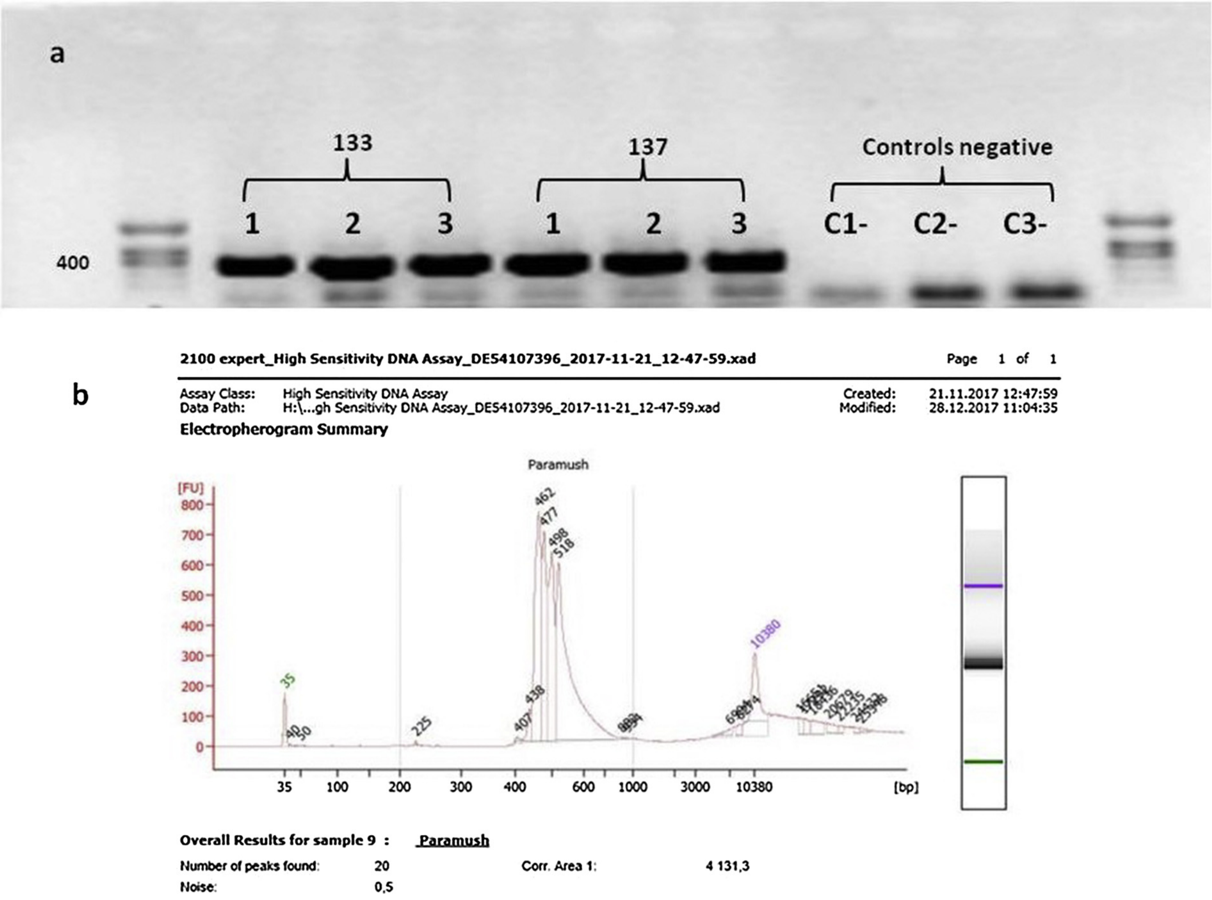
individual sets of contigs were assembled for both samples using SPAdes (Bankevich et al., 2012) in a single cell mode and the resulting sequences were mapped to the reference Paramushir virus genome.

A dendrogram analysis was performed using the Jukes-Cantor substitution model and the trees were reconstructed using the Neighbor-Joining (NJ) tree algorithm in MEGA 6.0 software (Hall, 2013). Definition of SNV and estimation of identity were performed using Vector NTI v.7.0 software (Invitrogen, USA).

A test for probable recombinations was performed with SimPlot 3.5.1.0 software (Lole et al., 1999) and with the Recombination Detection Program (RDP) 4 beta 80 using eight methods provided by this software.

The availability and the location of signal peptides in translated nucleotide sequences were predicted using SignalP 4.1 software (http://www.cbs.dtu.dk/services/SignalP/) (Nielson, 2017). The secondary structures of glycoproteins were predicted using PSIPRED v3.3 software (Buchan et al., 2013).

All molecular graphics were produced using PyMOL software (PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC).



2100 expert_High Sensitivity DNA Assay_DE54107396_2017-11-21_12-47-59.xad Page 1 of 1

Assay Class: High Sensitivity DNA Assay Created: 21.11.2017 12:47:59
Data Path: H:\...gh Sensitivity DNA Assay_DE54107396_2017-11-21_12-47-59.xad Modified: 28.12.2017 11:04:35

Electropherogram Summary

Overall Results for sample 9 : Paramush

Number of peaks found: 20 Corr. Area 1: 4 131,3
Noise: 0,5

Fig. 2. Library preparation for near complete genome sequencing using the Ion S5 system.
a – PCR products after multiplex PCR.
b – Mix of final RARMV libraries for Ion S5 system after size selection.

3. Results

3.1. Sequencing

In two out of thirty analyzed samples (IDs 133 and 137), single bands of expected length for the Orthonairovirus genera were found using a broad range PCR. These bands were purified, sequenced by the Sanger method and examined using BLAST. Both sequences were of 400 bp in length and shared 97% nucleotide identity with the L segment of PARV (strain LEIV–1149 K, ac. no KP792717). Thereby, both strains were identified as PARV. The strains were designated as Paramushir virus/Ixodes uriae/Russia/Tyuleniy Island/133/2015 and Paramushir virus/Ixodes uriae/Russia/Tyuleniy Island/137/2015 (briefly PARV strain 133 and PARV strain 137).

A total of 60 primer pairs were newly designed and used for the high-throughput sequencing library preparation for genomes of the two PARV strains. As a result, in total 489,436 and 358,075 sequencing reads were obtained for PARV strain 133 and PARV strain 137, respectively, and the nearly complete genomes of both strains were assembled (except for the 5′ and 3′ ends). The obtained reads were distributed unequally along the reference genome sequences. The S segments were 100% covered for both strains, the L segments were 98.4% and 95.7% covered for PARV strain 137 and PARV strain 133, respectively, and the M segments were 72.8% and 23.0% covered (for PARV strain 137 and PARV strain 133, respectively). The gaps were additionally sequenced using Sanger's sequencing technique by means of the ABI Prism 3500 XL (Applied Biosystems, USA); PARV strain 133 and PARV strain 137 were submitted to the GenBank NCBI under accession numbers MH124637, MH638289, MH124638, and MH124634-MH124636, respectively.

3.2. Genetic analysis

The comparative analysis of PARV sequences showed the existence of SNVs in the sequence of the PARV strain 133 (L segment: n = 260, including 19 nonsynonymous substitutions; M segment: n = 861, 49 nonsynonymous; S segment: n = 37, one nonsynonymous) as well as in the sequence of PARV strain 137 (L segment: n = 232, 25 nonsynonymous substitutions; M segment: n = 367, 49 nonsynonymous; S segment: n = 35, 3 nonsynonymous). In addition, the PARV strain 133 had 114 unique SNVs in the L segment, 174 unique SNVs in the M segment and 6 unique SNVs in the S segment, whereas PARV strain 137 had 86 unique SNVs in the L segment, 11 unique SNVs in the M segment, and 4 unique SNVs in the S segment.

The two novel strains of PARV shared 98% and 97–98% identities with the L and S segment sequences of PARV LEIV–1149 K (ac. nos. KP792717, KP792719), and differed from each other only by 1–2% within the L and S segments' nucleotide sequences. The comparisons of the M segment sequences demonstrated a 90%-nucleotide identity between PARV LEIV–1149 K (ac. no KP792718) and PARV strain 137, whereas the nucleotide identity between strains PARV LEIV–1149 K and PARV strain 133 was much lower (67%). Two newly sequenced strains shared 68% of nucleotide identity in the sequence of the M segment (Table 3). Thus, based on the nucleotide sequences of the new PARV strains, the M segment was found to be the most divergent.

The dendrogram analysis of nucleotide sequences of the L, M and S segments of PARV strain 133 and PARV strain 137, along with the appropriate sequences of other orthonairoviruses, supports their

Table 3
Nucleotide Identities of the L, M and S segment sequences within the Sakhalin genogroup.

strain	PARV LEIV-1148 K identity, nucleotide level, %			PARV 133, identity, nucleotide level, %			PARV 137, identity, nucleotide, %		
	L	M	S	L	M	S	L	M	S
PARV LEIV-1148 K	100	100	100	98	64	97	98	90	98
PARV 133	98	64	97	100	100	100	98	64	99
PARV 137	98	90	98	98	64	99	100	100	100
Avalon CanAr173	78	75	82	78	57	83	78	74	83
Avalon Bres T/Ar/T-261	78	74	82	78	56	82	78	74	82
Sakhalin LEIV-71C	77	67	66	77	47	66	77	67	66
Sakhalin LEIV-71S	76	68	67	76	52	66	76	68	66
Tillamook RML 86	76	67	67	76	47	66	76	67	66
Taggert Ml 14850	73	68	67	73	48	67	73	67	67
Clo Mor SCOT Ar7	75	69	69	75	49	68	76	69	69

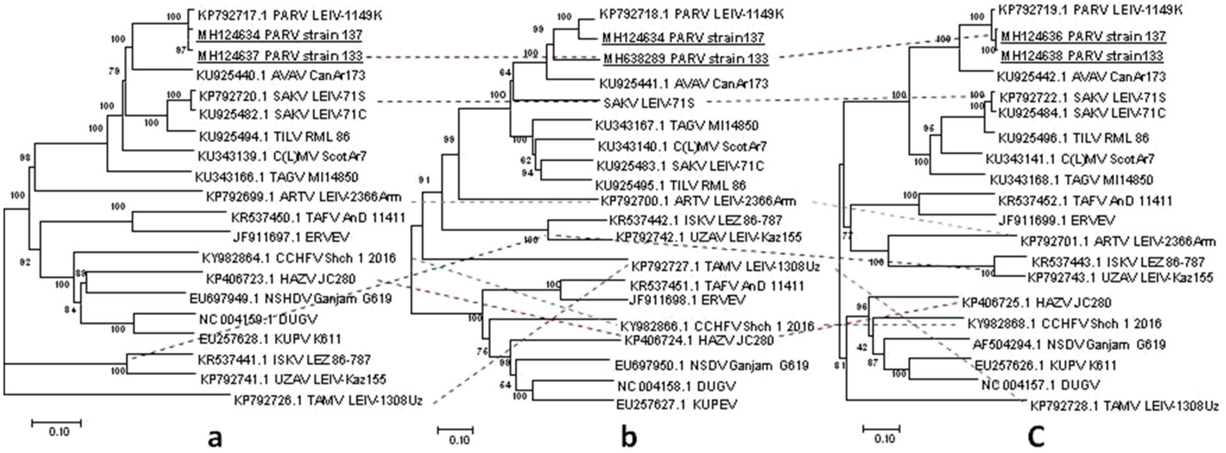


Fig. 3. Dendrograms for near complete L, M and S sequences of PARV (at the nucleotide level).
Phylogenetic analysis was performed using the Jukes-Cantor substitution model and the trees were reconstructed using the Neighbor-Joining (NJ) tree algorithm in MEGA 6.0 software. The robustness of the trees was tested using 1000 bootstrap replicates. The locations on the trees of the Paramushir virus/Ixodes uriae/Russia/Tyuleniy Island/133/2015 and Paramushir virus/Ixodes uriae/Russia/Tyuleniy Island/137/2015strains are marked with black dots.
A – Dendrogram for near complete L sequences of PARV. Other orthonairoviruses used for tree building.
B – Dendrogram for near complete M sequences of PARV. Other orthonairoviruses used for tree building.
c – Dendrogram for near complete S sequences of PARV. Other orthonairoviruses used for tree building.

affiliation with *Orthonairovirus* genus (Fig. 3). The location of PARV strain 133 and PARV strain 137 on the trees indicates their membership with the Sakhalin virus subgroup and a close relationship with the PARV strain LEIV–1149 K, which was sequenced previously.

However, the dendrograms of different genomic segments revealed distinct genetic relationships among PARV LEIV–1149 K, PARV strain 133, and PARV strain 137 (Fig. 3). Regarding segments L and S, PARV strain 133 and PARV strain 137 form the reliably supported clade related to PARV strain LEIV–1149 K. In the segment M, PARV strain LEIV–1149 K and PARV strain 137 form the reliably supported clade related to PARV strain 133. Moreover, the different phylogenetic positions of SAKV strain LEIV–71S, ARTV strain LEIV–2366Arm, HAZV strain JC280, and TAMV strain LEIV–1308Uz sequences were observed. Regarding the L and S segments, the SAKV strain LEIV–71S grouped reliably with the SAKV strain LEIV–71C and TILV strain RML 86, but grouped with the AVAV strain CanAr173 and all known PARV strains in the M segment. In the L and M segments, the ARTV strain LEIV–2366Arm was an outgroup for the strains of PARV, SAKV, AVAV, TILV, TAGV and C (L)MV, whereas, in the S segment, TAFV strains AnD 11411 and ERVEV were reliably clubbed in the group, but the ARTV strain LEIV–2366Arm was an outgroup as well as the strains: ISKV LEZ 86–787 and UZAV LEIV-Kaz155. The CCHFV strain Shch 1 2016 was an outgroup for the reliably grouped HAZV strain JC280, NSHDV strain Ganjam G619, DUGV, KUPV strain K611 in the L and M segments, but in the S segment, it changed its place with HAZV strain JC280. In addition the TAMV strain LEIV–1308Uz was an outgroup for all analyzed strains in the L and S segments, but it reliably formed an internal group with the strains of PARV, SAKV, AVAV, TILV, TAGV, C (L)MV, ARTV strain LEIV–2366Arm, ISKV strain LEZ 86–787, and UZAV strain LEIV-Kaz155.

A Bootscan analysis with Simplot version 3.5.1 software showed the absence of probable recombination events in the L and S segments but revealed one potential recombination event at the nucleotide position 1366–2897 in the M segment (Fig. 4a). An analysis provided by the RDP 4 beta 80 software showed two probable recombination events with recombinant fragments from the nucleotide position 1523–1910 and recombinant fragments from the nucleotide position 2210–2722 (according to genome alignment of Sakhalin virus genogroup); the PARV strain 133 is a recombinant strain, whose probable major parent is an unknown strain closely related to the AVAV. The PARV strain LEIV–1149 K is a potential minor parent in the first probable recombination event, and the PARV strain 137 is a potential minor parent in the second probable recombination event (Fig. 4b).

3.3. Analysis of putative proteins in the M segment

The glycoprotein precursors of orthonairoviruses form a number of smaller proteins after post-translational modifications (Kuhn et al., 2016). The glycoprotein precursor of PARV strain 137 has a molecular weight of 153.4 kDa and contains 1376 aa, which is the same as that of

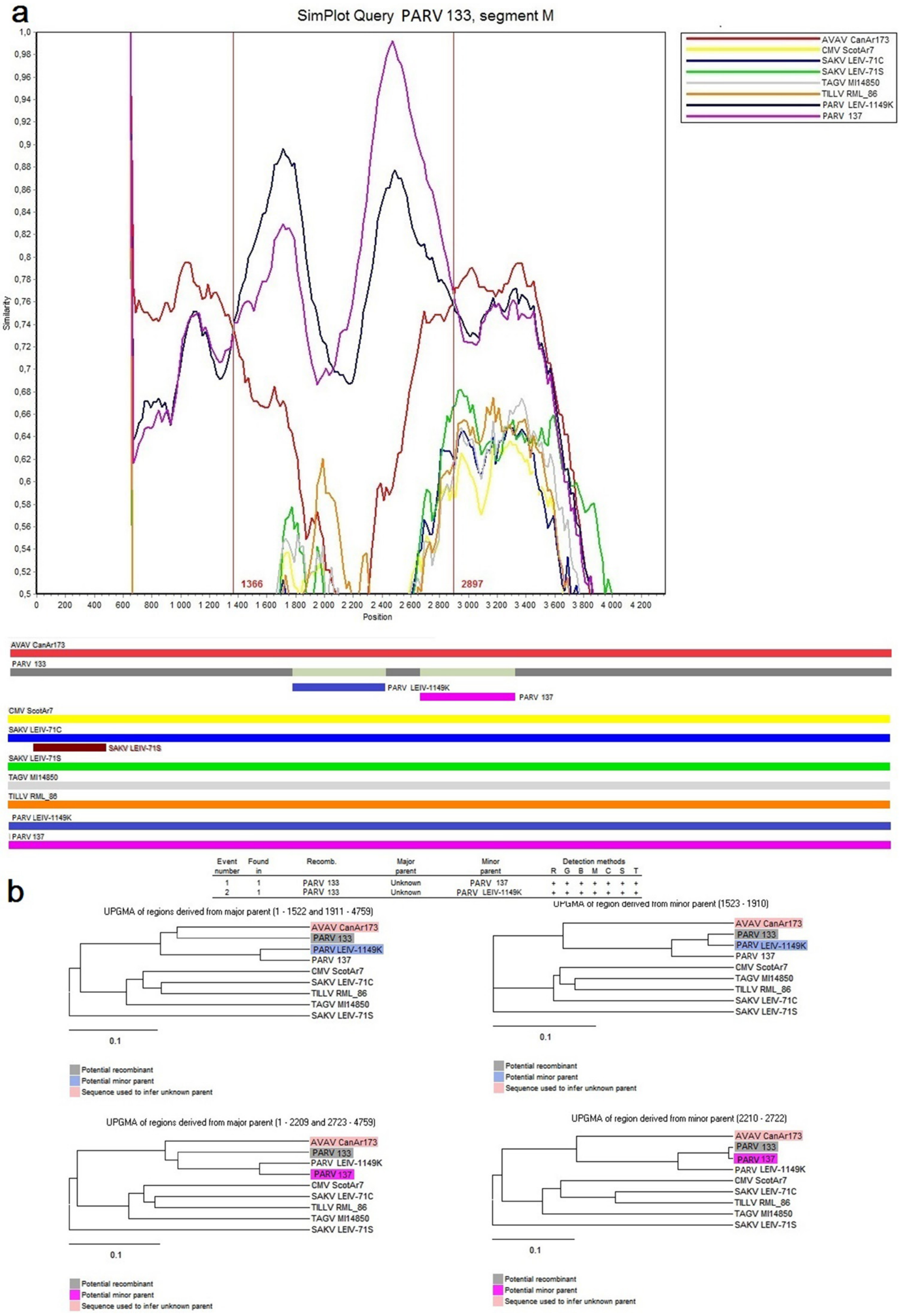


Fig. 4. Results of recombination analysis.

a– Bootscan analysis: A Bootscan analysis of the M segment was performed using Simplot version 3.5.1 software. Probable recombination fragments are located in the region from 1366 to 2897 nt.

b- Analysis RDP 4 beta 80 software. Two probable recombination fragments were revealed (first was located in the region from 1523 to 1910 nt, second was located in the region from 2210 to 2722 according to genome alignment of Sakhalin virus genogroup).

PARV strain LEIV–1149 K (AKC89338.1) (Fig. 5a). The cleavage site in the signal peptide was predicted at the N-terminal of the amino acid sequence of PARV strain 137 at position 28/29 aa, whereas in the amino acid sequences of the closely related viruses, its positions were 23/24 aa for PARV strain LEIV–1149 K (AKC89338.1), 21/22 aa for AVAV strain CanAr 173 (AMT75378.1), 30/31 aa for SAKV strain LEIV–71C (AKC89335.1). This peculiarity has been described for CCHFV strains, whose cleavage sites are located from 20/21 aa to 27/28 aa of amino acid sequences (Goedhals et al., 2015). However, in contrast to CCHFV, two newly sequenced strains of PARV have no furin-dependent cleavage site R-X-[K/R]-R↓. Therefore, instead of two proteins (Mucin-like protein and GP38), only one protein is formed after proteolysis from the N-terminal of the polypeptide (G1) with a molecular weight of 37 kDa. However, this peculiarity is common for a number of species belonging to Orthonairovirus genus (Lenz et al., 2001). The N-terminus of glycoprotein (Gn), for both PARV strain 133 and PARV strain 137, is cleaved off from the G1 by the means of subtilisin kexin isozyme 1 (SKI–1)/site 1 protease (S1P), which recognizes

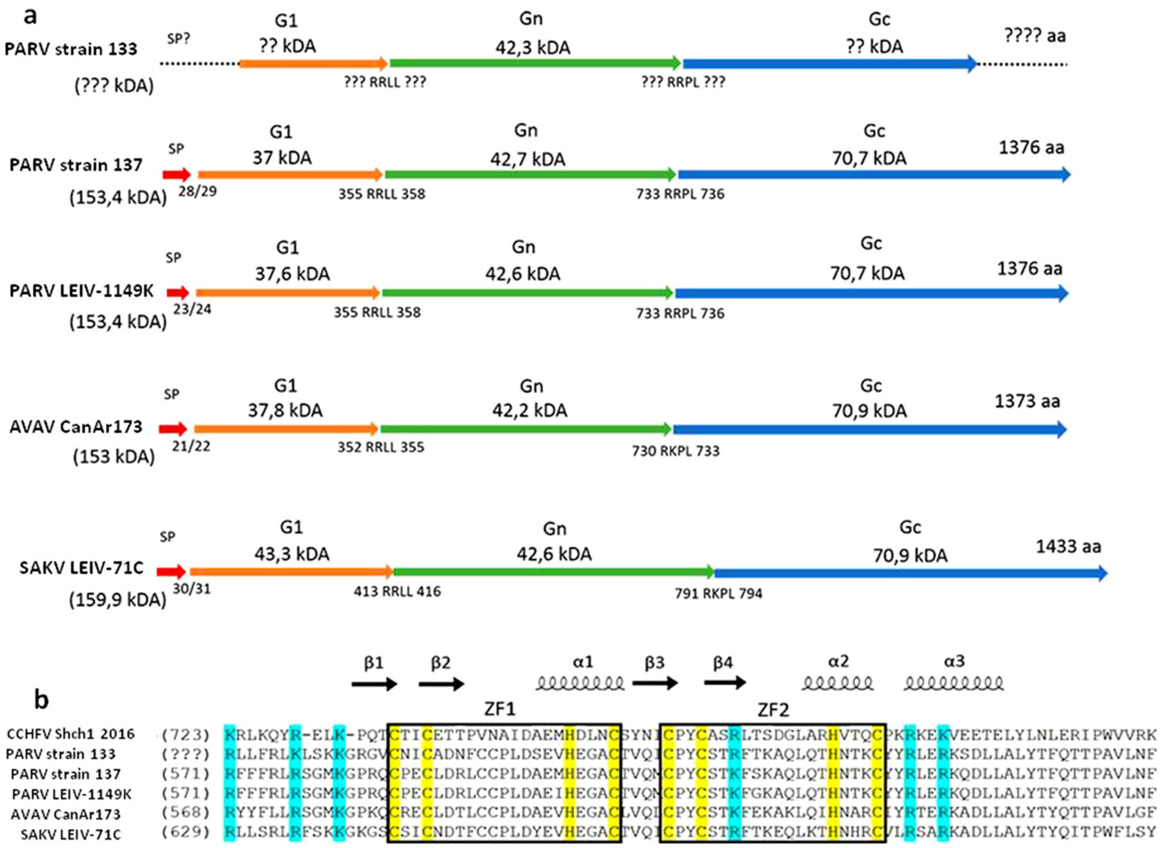


Fig. 5. Putative proteins encoded in M segment of PARV.

a – Schematic illustration of the structures of the polyglycoproteins encoded in the M segments of Sakhalin orthonairoviruses. Regions corresponding to the G1 (orange), Gn (green) and Gc (blue) are shaded. Dotted lines represent unsequenced regions. Predicted signal peptidase cleavage sites (SP) and potential SKI-I cleavage sites are shown. The predicted molecular weights (kDa) of GPCs are annotated.

b. Sequence alignment of the Gn tails of the members from the Sakhalin genogroup with the known structure of zinc fingers of CCHF. The conserved zinc finger motifs (boxed) and conserved basic residues (in blue) are mapped.

the conservative domain RRLL↓. As shown above, this protease is crucial in the surface glycoprotein processing for some pathogenic viruses, such as Lassa virus (LASV) (Lenz et al., 2001), lymphocytic choriomeningitis virus (LCMV) (Beyer et al., 2003), and CCHFV (Vincent et al., 2003).

The two smaller glycoproteins (Gn) contain 378 aa for the both PARV strain 133 and PARV strain 137, and have molecular weights of 42.3 kDa and 42.7 kDa, respectively. The C-terminal end of Gc is cleaved at site RRPL↓ by means of SKI–1/S1P-like protease. The 733RRPL736 motif is identical to that of the previously described PARV strain LEIV–1149 K, but differs from that of the closely related species AVAV strain CanAr 173 (730RKPL733) and SAKV strain LEIV–71C (730RKPL733). The C-terminal Gc protein of PARV strain 137 highly conserved, with only 7 aa substitutions (98.9% aa identity) in comparison with PARV strain LEIV–1149 K (with 8 substitutions in Gn (95.2%) and 24 in G1 (91.3%)). Moreover, in comparison to the closely related viruses, this protein is also highly conserved with 87.8% aa identity with Gc of AVAV strain CanAr 173 and 74.7% aa identity with Gc of SAKV strain LEIV–71C. Sixty-six amino acid substitutions in the sequenced fragment of PARV strain 133 Gc were found in comparison with PARV strain LEIV–1149 K (86.4% aa identity) and 55 (85.4%) substitutions in Gn. The N-terminus of the G1 protein was the most variable fragment. The sequenced fragment of PARV strain 133 G1 had 41 substitutions in comparison with PARV strain LEIV–1149 K, but only 30 substitutions in comparison with AVAV strain CanAr 173. The four transmembrane domains of the polyprotein M were predicted by the TMHMM server, which form two regions outside of the membrane, and two regions in the cytoplasm. This is most likely similar to the structure observed in Qalyub virus (QYBV) (Walker et al., 2013).

The M segment of PARV strain 137 contains 63 cysteine residues, whereas the sequenced fragment of M segment in PARV strain 133 contains 50 cysteine residues located mainly in the conservative positions in comparison with PARV strain LEIV–1148 K (AKC89338.1), AVAV strain CanAr 173 (AMT75378.1), and SAKV strain LEIV–71C (AMT75420.1).This suggests the presence of a large number of disulfide bonds and a complex spatial structure of proteins similar to the related viruses. For instance, by comparing the C-terminus of Gn with CCHFV, which is the only virus with an identified molecular structure among orthonairoviruses, the presence of the conservative motifs CXCXHXC of cysteine and histidine (where X is any amino acid) was characteristic of the ββα-type zinc fingers of bunyaviruses (Fig. 5b). In some *in vitro* experiments, the possibility of RNA binding has been shown for this structure, which may indicate the participation of Gn-tail in interaction with viral RNA (Estrada and De Guzman, 2011). In general, the molecular structure of orthonairoviruses is poorly studied, even for the CCHFV proteins, and nor the spatial structure, nor the functional domain have been completely described.

The secondary structures of G1, Gn, and Gc, deduced from PSIPRED server and identified as the distribution of helices, strands, and coils, are shown in Supplementary Figs. 1, 2 and 3, respectively. It can be seen that the secondary protein structures are highly conservative for all compared viruses. A significant difference can be seen only at the N-terminus of G1 of the SAKV virus. The experimental high-resolution structural data have been determined only for Gc protein of several othounyaviruses: Rift Valley virus (RFV), Puumala virus (PUUV) and Heartland virus (HRTV) (Dessau and Modis, 2013; Willensky et al.,

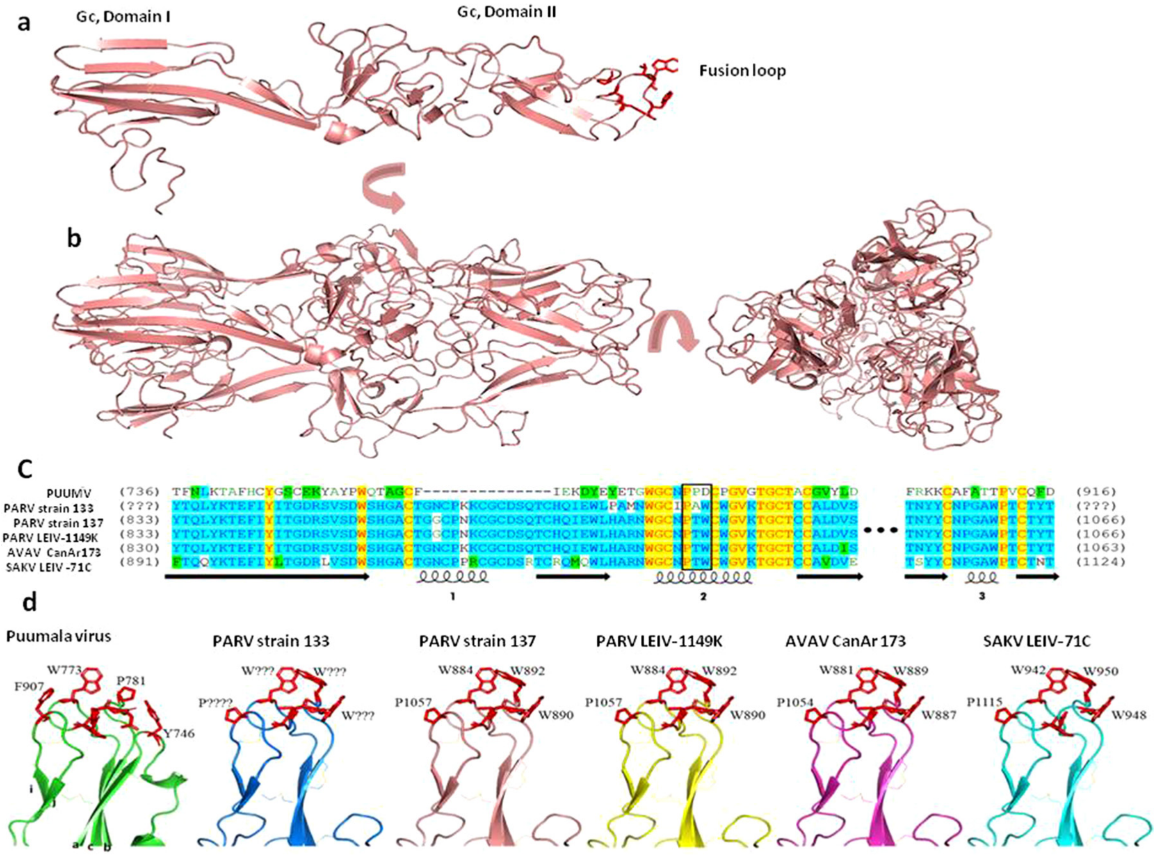


Fig. 6. Schematic representation of PARV strain 137 Gc structure.

a – Schematic 3D diagram of predicted PARV strain 137 Gc structure (domain I and II). A monomeric form with fusion loop residues is displayed in red.

b – Trimer form in different projections.

C – Structure-based multiple sequence alignment of the fusion loop regions. The PXX PUUV-like insertion is marked by a black box. The secondary structure elements are displayed below the sequences. Strictly conserved, identical, and highly similar residues are highlighted in yellow, blue and green, respectively. The residues correspond to polyprotein precursor numbering.

d – The putative fusion loop of Puumala GC and the fusion loops of the Sakhalin group viruses. The hydrophobic residues that anchor the protein to the cellular membrane are shown in red and the disulfide bonds in stick representation.

*All molecular graphics were produced using PyMOL (PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC).

2016; Zhu et al., 2018). The structural studies have revealed that both hantaviral and phenuiviral Gc fold as a class II membrane fusion proteins, which is induced by a low pH during endocytosis and undergo a conformational change by exposing a hydrophobic fusion peptide or fusion loop. This is highly possible that the tertiary structure of Gc protein from strain 137 has the same post-fusion trimeric conformation and is also a member of the class II fusion protein. The bioinformatical methods, considering the amino acid sequences of most closely related PUUV, have partially predicted the tertiary structure of the PARV strain 137 Gc protein (Fig. 6a) (Biasini et al., 2014). The sequences of domains I and II match well with the sequence of the Puumala virus, but domain III differs to a large extent, and no spatial structure could be determined, same being true for the particularly sequenced M segment of PARV strain 133. However, a conserved block of the sequences forming putative fusion loops and the endosomal membrane anchors has been seen as a part of the domain II, which is distal to the domain I (Fig. 6b). The fusion loop, which is responsible for the initial insertion into the target membrane, is the key element in the fusogenic process (White et al., 2008)

PARV strain 133 and PARV strain 137, as well as other bunyaviruses, have more than one fusion loop (Fig. 6c) (Zhu et al., 2018). In comparison with the PUUV, on which the model was built, there is a significant difference in the first loop: for PARV strain 137 and PARV strain 133 they are outside of the bonding plane with the membrane (Fig. 6d). However, this could be due to a flaw in the constructed model. The two remaining loops corresponding to the bc and ij loop of PUUV have all the conserved residues and two disulfide bonds fixing the fusion loop conformation. At the same position, PARV strain 137 Gc fusion loop contains a three-residue insertion 888P-T-W890, which differs from 777P-X-D779, mostly conserved among hantaviruses (Willensky et al., 2016). Instead of the conserved F907 of hantaviruses, a proline amino acid (P1057 in PARV strain 137 isolate) is present in the Sakhalin genogroup. As previously shown, Gc of Andes virus (ANDV) with single mutations in the conserved residues W773, N776, and D779, located at the fusion loop, eliminated the cell-cell fusion activity and infectivity of the ANDV pseudotyped particles (Cifuentes-Muñoz et al., 2011). The third amino acid (D) of PARV strain 137 and PARV strain 133 is replaced by W (W884, N887, and W890) and for PARV strain 133 N is changed to I. This suggests that, despite the fact that the structure of the glycoprotein C in the Sakhalin genogroup has a high homology with other representatives of bunyaviruses, the substitutions in the conserved residues most likely result in the differences in pathogenesis processes and pathways and pathogenicity to humans.

4. Discussion

In general, within the Sakhalin genogroup, the nucleotide identities ranged from 78% to 98% in the L segment, from 69% to 90% in the M segment, and from 66% to 99% in the S segment. Therefore, the S segment is known as the most variable region within this genogroup (see Table 3). However, based on our data, the sequences of the M segment among PARV strains share a lower identity in comparison to the average volume within the genogroup. This peculiarity likely explains why the number of sequencing reads belonging to the segment M was the lowest and caused significant difficulties in the primer design and sequencing.

By using phylogenetic analysis, we observed close relations between the novel strains and PARV strain LEIV–1149 K, also confirmed their association with the Sakhalin genogroup. Moreover, the phylogenetic analysis identified distinct and reliable grouping patterns among PARV LEIV–1149 K and PARV strains 133, 137 in the different genome segments. This indicates that the M segment reassortment is common in PARV and also supports the assignment of all these strains to the same species, because their genome segments are genetically compatible and could be involved in the reassortment. Besides PARV, other orthonairoviruses that were analyzed also demonstrated signs of multiple reassortment, in which both M and S segments were involved. Such peculiarity is probably caused by genome structure and comparability and allows for the reassortment process.

In view of the change in tree topology at the M segment coding the region of PARV strain 133, recombination analysis was performed to detect the probable recombination events.

Using RDP 4 beta 80 software, two probable recombination events in a fragment from nucleotide positions 1523–2722 were revealed. The unknown strain that belongs to the Sakhalin genogroup, which is most closely related to the AVAV strain CanAr173, was the major parent in both events. The minor parent in the first event was the strain closely related to PARV LEIV–1149 K, while in the second event it was the strain that is most closely related to PARV strain 137. Given the close location of the probable recombination events and a high nucleic acid identity between both minor parents, one expanded recombination event could be assumed in a fragment from nucleotide positions 1523–2722 bp (according to genome alignment of Sakhalin virus genogroup). In the M segment of PARV strain 133, this fragment is located from nucleotide positions 1100 to 2317.

A single probable recombination event was confirmed by SimPlot analysis. However, RDP 4 beta software indicated a narrower range of recombination site due to the use of several statistical models. The minor parent in this event was, probably, an unknown strain of PARV, sharing a number of features with PARV LEIV–1149 K and PARV 137.

The pairwise analysis of the M segment's nucleic acid identity within the Sakhalin genogroup particularly confirmed this assumption. The M segment identity of PRMV strain 133 with other known PARV strains in the range 1523–2722 bp was significantly higher than that for the whole M segment. The inherited fragment of the major parent was closer to the strains of PARV, rather than to the strains of AVAV. However, the nucleic acid identity did not exceed 61%.

The recombination phenomenon as a source of genetic diversity of orthonairoviruses was previously described for the CCHFV (Lukashev, 2005). The recombination event along with the intersegment reassortation points to a high variability potential of the virus, which may lead to the formation of strains with new properties.

5. Conclusions

In this study, we report the identification and nearly complete genome sequencing of two strains of Paramushir virus using a combination of high throughput sequencing and specific multiplex PCR following an earlier approach (Quick et al., 2017). The results of the sequence analysis confirmed the applicability of this approach for resequencing of known viral genomes, especially in the case of a lack of viral RNA.

Both strains of PARV were isolated in 2015 on the Tyuleniy Island in the Okhotsk Sea after a 43-years break. The phylogenetic analysis was used to determine their relation with the Sakhalin genogroup in the *Orthonairovirus* genus and with PARV strain LEIV–1149 K. The signs of multiple reassortment among the orthonairoviruses, including PARV, and of probable recombination events within the Sakhalin genogroup were revealed. However, these phenomena need to be studied in more detail.

We hope our study will contribute to the knowledge of orthonairoviruses, which is still poor, especially regarding the Sakhalin genogroup of unknown pathogenicity. For a more detailed understanding of the complicated phylogeny within the Sakhalin genogroup, further ecological studies are needed, and additional isolates belonging to the genogroup should be analyzed.

More data regarding Paramushir virus could provide a better understanding of its diversity and relation within *Orthonairovirus* genus and could help define intragenic demarcation criteria which have not yet been established.

Foundation

This work was supported by RSF grant N. 17-74-20096

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ttbdis.2018.11.004.

References

Alkhovsky, S.V., Lvov, D.K., Shchelkanov, M.Yu., Shchetinin, A.M., Deryabin, P.G., Botikov, A.G., Gitelman, A.K., Samokhvalov, E.I., 2013. Genetic characterisation of new Komandory virus (KOMV; Bunyaviridae, Phlebovirus) isolated from ticks Ixodes uriae collected in guillemot (Uria aalge) nesting sites on Komandorsky islands, Bering sea. Vopr. Virusol. 58 (6), 18–22. [In Russian]. http://elibrary.ru/download/52643105.pdf.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215 (3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Andrews, S., 2010. Fast QC: a Quality Control Tool for High Throughput Sequence Data. Available online at:. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19 (5), 455–477. https://doi.org/10.1089/cmb.2012.0021.

Beyer, W.R., Pöpplau, D., Garten, W., Von Laer, D., Lenz, O., 2003. Endoproteolytic processing of the lymphocytic choriomeningitis virus glycoprotein by the subtilase SKI-1/S1P. J. Virol. 77 (5), 2866–2872. https://doi.org/10.1128/JVI.77.5.2866-2872.2003.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., Schwede, T., 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 2014https://doi.org/10.1093/nar/gku340. Jul 1; 42 (Web Server issue): W252–W258. Published online.

Buchan, D.W.A., Minneci, F., Nugent, T.C.O., Bryson, K., Jones, D.T., 2013. Scalable web services for the PSIPRED protein analysis workbench. Nucleic Acids Res. 41, W349–W357. https://doi.org/10.1093/nar/gkt381.

Cifuentes-Muñoz, N., Barriga, G.P., Valenzuela, P.D., Tischler, N.D., 2011. Aromatic and polar residues spanning the candidate fusion peptide of the Andes virus Gc proteinare essential for membrane fusion and infection. J. Gen. Virol. 92 (Pt3), 552–563. https://doi.org/10.1099/vir.0.027235-0.

Dedkov, V.G., Lukashev, A.N., Deviatkin, A.A., Kuleshov, K.V., Safonova, M.V., Poleshchuk, E.M., Drexler, J.F., Shipulin, G.A., 2016. Retrospective diagnosis of two rabies cases in humans by high throughput sequencing. J. Clin. Virol. 78, 74–81. https://doi.org/10.1016/j.jcv.2016.03.012.

Dessau, M., Modis, Y., 2013. Crystal structure of glycoprotein C from Rift Valley fever virus. Proc. Natl. Acad. Sci. U. S. A. 110 (January (5)), 1696–1701. https://doi.org/10.1073/pnas.1217780110. Published online.

Deyde, V.M., Khristova, M.L., Rollin, P.E., Ksiazek, T.G., Nichol, S.T., 2006. Crimean-Congo hemorrhagic fever virus genomics and global diversity. J. Virol. 80 (17), 8834e8842. https://doi.org/10.1128/JVI.00752-06).

Estrada, D.F., De Guzman, R.N., 2011. Structural characterization of the crimean-congo hemorrhagic fever virus gn tail provides insight into virus assembly. J. Biol. Chem. 286 (24), 21678–21686. https://doi.org/10.1074/jbc.M110.216515. Published online.

Filippova, N.A., 1977. Ixodid ticks of the subfamily ixodinae. Fauna USSR: Arachnoides 4 (4) Nauka Leningrad (in Russian).

Goedhals, D., Bester, P.A., Paweska, J.T., Swanepoel, R., Burt, F.J., 2015. Comparative analysis of the L, M, and S RNA segments of Crimean-Congo haemorrhagic fever virus isolates from southern Africa. J. Med. Virol. 87 (5), 717–724. https://doi.org/10.1002/jmv.24079. Epub 2015 Feb 18.

Hall, B.G., 2013. Building phylogenetic trees from molecular data with MEGA. Mol. Biol. Evol. 30 (5), 1229–1235. https://doi.org/10.1093/molbev/mst012.

ICTV, 2016. Virus Taxonomy: the Classification and Nomenclature of Viruses the Online (10th) Report of the. Available at. . https://talk.ictvonline.org/ictv-reports/ictv_online_report/.

Klimentov, A.S., Butenko, A.M., Khutoretskaya, N.V., Shustova, E.Yu., Larichev, V.F., Isaeva, O.V., Karganova, G.G., Lukashev, A.N., Gmyl, A.P., 2016. Development of pan-phlebovirus RT-PCR assay. J. Virol. Methods 232, 29–32. https://doi.org/10.1016/j.jviromet.2016.02.009.

Kuhn, J.H., Wiley, M.R., Rodriguez, S.E., Bào, Y., Prieto, K., Travassos da Rosa, A.P., Guzman, H., Savji, N., Ladner, J.T., Tesh, R.B., Wada, J., Jahrling, P.B., Bente, D.A., Palacios, G., 2016. Genomic Characterization of the Genus Nairovirus (Family Bunyaviridae). Viruses 8 (6). https://doi.org/10.3390/v8060164.

Lambert, A.J., Lanciotti, R.S., 2009. Consensus amplification and novel multiplex sequencing method for S segment species identification of 47 viruses of the Orthobunyavirus, Phlebovirus, and Nairovirus genera of the family Bunyaviridae. J. Clin. Microbiol. 47, 2398–2404. https://doi.org/10.1128/JCM.00182-09.

Landolfi, J.A., Terio, K.A., Kinsel, M.J., Langan, J., Zachariah, T.T., Childress, A.L., Wellehan, J.F.Jr., 2010. Orthoreovirus infection and concurrent cryptosporidiosis in rough green snakes (Opheodrys aestivus): pathology and identification of a novel orthoreovirus strain via polymerase chain reaction and sequencing. J. Vet. Diagn. Invest. 22, 37–43. https://doi.org/10.1177/104063871002200106.

Lenz, O., ter Meulen, J., Klenk, H.D., Seidah, N.G., Garten, W., 2001. The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1yS1P. PNAS 98 (22), 12701–12705. https://doi.org/10.1073/pnas.221447598.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25 (14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., Ray, S.C., 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J. Virol. 73 (1), 152–160.

Lukashev, A.N., 2005. Evidence for recombination in Crimean-Congo hemorrhagic fever virus. J. Gen. Virol. 86 (Pt 8), 2333–2338. https://doi.org/10.1099/vir.0.80974-0.

Lvov, D.K., Sazonov, A.A., Gromashevsky, V.L., Skvortsova, T.M., Beresina, L.K., Aristova, V.A., Timofeeva, A.A., Zakharyan, V.A., 1976. “Paramushir” virus, a new arbovirus, isolated from ixodid ticks in nesting sites of birds on islands in north-western part of the Pacific Ocean basin. Arch. Virol. 51, 157–161.

Lvov, D.K., Deryabin, P.G., Aristova, V.A., Butenko, A.M., Galkina, I.V., Gromashevsky, V.L., Davidova, A.A., Kolobukhina, L.V., Lvov, S.D., Shchelkanov, M.Yu., 2001. Atlas of Distribution of Natural Foci Virus Infections on the Territory of Russian Federation. SMC MPH RF Publ., Moscow 2001 192 p. [in Russian].

Lvov, D.K., Alkhovskiy, S.V., Shchelkanov, M.Iu., Shchetinin, A.M., Deriabin, P.G., Samokhvalov, E.I., Gitel’man, A.K., Botikov, A.G., 2014a. Genetic characterization of viruses from antigenic complex Tyuleniy (*Flaviviridae*, Flavivirus): Tyuleniy virus (TYUV) isolated from ectoparasites of colonial seabirds – Ixodes (Ceratixodes) uriae White, 1852 ticks collected in the high latitudes of Northern Eurasia – and Kama virus (KAMV) isolated from Ixodes lividus Roch, 1844 collected in the digging colonies of the middle part of Russian Plane. Vopr. Virusol. 59 (1), 18–24. [In Russian]. http://elibrary.ru/download/37296578.pdf.

Lvov, D.K., Alkhovskiy, S.V., Shchelkanov, M.Iu., Shchetinin, A.M., Deriabin, P.G., Aristova, V.A., Gitel’man, A.K., Samokhvalov, E.I., Botikov, A.G., 2014b. Genetic characterization of Sakhalin virus (SAKV), Paramushir virus (PMRV) (Sakhalin group, Nairovirus, Bunyaviridae) and Rukutama virus (RUKV) (Uukuniemi group, Phlebovirus, Bunyaviridae), isolated from the obligate parasites of colonial sea-birds - ticks Ixodes (Ceratixodes) uriae, White 1852 and I. Signatus Birulya, 1895 in water area of sea of Okhotsk and Bering sea. Vopr. Virusol. 59 (3), 11–17. [In Russian]. http://elibrary.ru/download/32410409.pdf.

Lvov, D.K., Alkhovsky, S.V., Shchelkanov, M., Yu Shchetinin, A.M., Deriabin, P.G., Gitel’man, A.K., Samokhvalov, E.I., Botikov, A.G., 2014c. Genetic characterization of Zaliv Terpeniya virus (ZTV, Bunyaviridae, Phlebovirus, Uukuniemi serogroup) strains, isolated from ticks Ixodes (Ceratixodes) uriae White, 1852, obligate parasites of Alcidae birds, in high latitudes of Northern Eurasia and mosquitoes Culex modestus Ficalbi, 1889 in subtropics Transcaucasus. Vopr. Virusol. 59 (1), 12–18. [In Russian]. http://elibrary.ru/download/23250558.pdf.

Lvov, D.K., Alkhovskiy, S.V., Shchelkanov, M.Iu., Shchetinin, A.M., Deriabin, P.G., Gitel’man, A.K., Samokhvalov, E.I., Botikov, A.G., 2014d. Molecular-genetic characterization of Okhotskiy virus (OKHV) и Aniva virus (ANIV) (Orbivirus, Reoviridae), isolated from ticks Ixodes (Ceratixodes) uriae White, 1852 in high latitudes of Northern Eurasia. Vopr. Virusol. 59 (2), 22–28. [In Russian]. http://elibrary.ru/download/19327065.pdf.

Lvov, D.K., Shchelkanov, M.Yu., Alkhovsky, S.V., Deryabin, P.G., 2015. Zoonotic Viruses of Northern Eurasia. Taxonomy and Ecology. Academic Press ISBN: 978-0-12-801742-5.

Maniatis, T., Fritsch, E.F., Sambrook, J., 1989. Molecular Cloning, a Laboratory Manual. Cold Spring Harbor laboratory, N.Y.: Cold Spring Harbor.

NCBI Resource Coordinators, 2017. Resources of the national center for biotechnology information. Database. Nucleic Acids Res. 45https://doi.org/10.1093/nar/gkw1071. Database issue.

Nielson, H., 2017. Predicting secretory proteins with SignalP. Methods Mol. Biol. 1611, 59–73. https://doi.org/10.1007/978-1-4939-7015-5_6.

Quick, J., Grubaugh, N.D., Pullan, S.T., Claro, I.M., Smith, A.D., Gangavarapu, K., Oliveira, G., Robles-Sikisaka, R., Rogers, T.F., Beutler, N.A., Burton, D.R., Lewis-Ximenez, L.L., de Jesus, J.G., Giovanetti, M., Hill, S.C., Black, A., Bedford, T., Carroll, M.W., Nunes, M., Alcantara, L.C.Jr., Sabino, E.C., Baylis, S.A., Faria, N.R., Loose, M., Simpson, J.T., Pybus, O.G., Andersen, K.G., Loman, N.J., 2017. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat. Protoc. 12 (6), 1261–1276. https://doi.org/10.1038/nprot.2017.066.

Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27 (6), 863–864. https://doi.org/10.1093/bioinformatics/btr026.

Shchelkanov, M.Yu., Galkina, I.V., Ananiev, V., Samarsky, S.S., Lienho, V.Yu., Dedkov, V.G., Safonova, M.V., Orekhov, V.E., Shchelkanov, E.M., Alekseev, A.Yu., Shestopalov, A.M., Pitruk, D.L., Serkov, V.M., 2017. Ecological situation on the Tyuleniy Island in the Okhotsk Sea (2015): population interactions between pinnipeds, birds, ixodidae ticks and viruses. South of Russia: ecology. Development 12 (1), 30–43. https://doi.org/10.18470/1992-1098-2017-1-30-43. (In Russian).

Vincent, Mj, Sanchez, A.J., Erickson, B.R., Basak, A., Chretien, M., Seidah, N.G., Stuart, T., Nichol, S.T., 2003. Crimean-congo hemorrhagic fever virus glycoprotein proteolytic processing by subtilase SKI-1. J. Virol. 77 (16), 8640–8649. https://doi.org/10.1128/JVI.77.16.8640-8649.2003.

Walker, P.J., Widen, S.G., Wood, T.G., Guzman, H., Tesh, R.B., Vasilakis, N., 2016. A global genomic characterization of nairoviruses identifies nine discrete genogroups with distinctive structural characteristics and host-vector associations. Am. J. Trop. Med. Hyg. 94 (5), 1107–1122. https://doi.org/10.4269/ajtmh.15-0917.

White, J.M., Delos, S.E., Brecher, M., Schornberg, K., 2008. Structures and mechanisms of viral membrane fusion proteins multiple variations on a common theme crit. Rev. Biochem Mol. Biol. Author manuscript; available in PMC 2009 may 1. Published in final edited form as. Crit. Rev. Biochem. Mol. Biol. 43 (3), 189–219. https://doi.org/10.1080/10409230802058320.

Willensky, S., Bar-Rogovsky, H., Bignon, E.A., Tischler, N.D., Modis, Y., Dessau, M., 2016. Crystal structure of glycoprotein C from a hantavirus in the post-fusion conformation. PLoS Pathog., e1005948. https://doi.org/10.1371/journal.ppat.100594. Published online.

Zhu, Y., Wu, Y., Chai, Y., Qi, J., Peng, R., Feng, W.H., Gao, G.F., 2018. The postfusion structure of the heartland virus gc glycoprotein supports taxonomic separation of the bunyaviral families *phenuiviridae* and hantaviridae. J. Virol. 92 (1), e01558–17. https://doi.org/10.1128/JVI.01558-17. Prepublished online 2017 Oct 25. Published online 2017 Dec 14.