FULL PAPER

Parasitology

Viral population analysis of the taiga tick, *Ixodes persulcatus,* by using Batch Learning Self-Organizing Maps and BLAST search

Yongjin QIU[1,2)], Takashi ABE[3)], Ryo NAKAO[4,5)], Kenro SATOH[3)] and Chihiro SUGIMOTO[1,6)]*

[1)]Division of Collaboration and Education, Hokkaido University Research Center for Zoonosis Control, Kita 20 Nishi 10, Kita-ku, Sapporo, Hokkaido 001-0020 Japan
[2)]Hokudai Center for Zoonosis Control in Zambia, the University of Zambia, Lusaka, 10101 Zambia
[3)]Graduate School of Science and Technology, Niigata University, Ikarashi 2 no-cho 8050, Nishi-ku, Niigata, Niigata 950-2181 Japan
[4)]Unit of Risk Analysis and Management, Hokkaido University Research Center for Zoonosis Control, Kita 20 Nishi 10, Kita-ku, Sapporo, Hokkaido 001-0020 Japan
[5)]Laboratory of Parasitology, Faculty of Veterinary Medicine, Graduate School of Infectious Diseases, Hokkaido University, Kita 18 Nishi 9, Kita-ku, Sapporo, Hokkaido 060-0818 Japan
[6)]Global Institution for Collaborative Research and Education (GI-CoRE), Hokkaido University, Kita 20 Nishi 10, Kita-ku, Sapporo, Hokkaido 001-0020 Japan

ABSTRACT. Ticks transmit a wide range of viral, bacterial, and protozoal pathogens, which are often zoonotic. Several novel tick-borne viral pathogens have been reported during the past few years. The aim of this study was to investigate a diversity of tick viral populations, which may contain as-yet unidentified viruses, using a combination of high throughput pyrosequencing and Batch Learning Self-Organizing Map (BLSOM) program, which enables phylogenetic estimation based on the similarity of oligonucleotide frequencies. DNA/cDNA prepared from virus-enriched fractions obtained from *Ixodes persulcatus* ticks was pyrosequenced. After *de novo* assembly, contigs were cataloged by the BLSOM program. In total 41 different viral families and order including those previously associated with human and animal diseases such as *Bunyavirales*, *Flaviviridae*, and *Reoviridae*, were detected. Therefore, our strategy is applicable for viral population analysis of other arthropods of medical and veterinary importance, such as mosquitos and lice. The results lead to the contribution to the prediction of emerging tick-borne viral diseases. A sufficient understanding of tick viral populations will also empower to analyze and understand tick biology including vector competency and interactions with other pathogens.

KEY WORDS: Batch Learning Self-Organizing Map, high throughput pyrosequencing, *Ixodes persulcatus*, viral population analysis

Ticks (ixodida) are haematophagous parasitic arthropods that feed on the blood of vertebrates ranging from mammals and birds to reptiles. They can harbor and transmit a wide range of viral, bacterial, and protozoal pathogens, which are zoonotic frequently [4, 15, 16, 52]. A change in the world's climate and environments and developing international exchanges can cause changes or extension of tick distributions, because of which contacts with ticks are increasing [5, 17, 35, 36]. Furthermore, recognitions of new tick-borne diseases are in on the rise. For example, new tick-borne viral diseases, such as severe fever with thrombocytopenia syndrome (SFTS) and Heartland virus infection have appeared as emerging diseases, recently [31, 52].

Currently the majority of tick-borne virus studies have heavily concentrated on the well described RNA arboviruses containing the families *Nairoviridae, Phenuiviridae, Peribunyaviridae, Flaviviridae,* and *Reoviridae*, including the causative agents of Crimean-Congo hemorrhagic fever, tick-borne encephalitis, and Colorado tick fever in humans [3, 34]. Little efforts focus on potentially emergent viruses that have not yet made their presence known on ticks. These undiscovered viruses can also provide a gene pools serving as sources of biodiversity by the mechanisms of recombination or reassortment with existing pathogenic tick-borne viruses. Understanding the total diversity of tick-borne viruses is important for prediction and preparation of emerging and re-emerging tick-borne viral diseases; however, characterizing new viruses is difficult due to limitation of currently available detection methods, such as culture- or experimental animal-based isolation, conventional PCRs, and pan-viral microarrays.

Traditional molecular methods have technical limitations in finding and characterizing novel viruses that share limited homology to know viral sequences in the database.

An effective strategy for exploiting "virome" would need to simultaneously identify a wide range of viral taxa. To elude these problems, recent studies have employed the effectiveness of viral particle purification and shotgun sequencing for finding of novel viruses. Using this combination of techniques, novel viruses have been characterized or viral populations were detected from several biological and environmental samples [19, 20, 33, 51]. These studies employed BLAST analysis to classify sequence reads, with which only a few percentage of the reads were assigned to viruses and over half of the reads were not able to assign to any biological groups [6, 12, 14, 28, 41, 51]. Thus, there is a positivity of the presence of the viral world which has not been explored previously.

Batch Learning Self-Organizing Map (BLSOM) is a novel bioinformatics tool for large-scale comprehensive, phylogenetic studies on big sequence data, which is applicable to any environmental sample: a tool that can overview all available sequences from prokaryotic, eukaryotic, organelle and viral genomes at once [1, 23]. Suppose only a massive amount of fragmental sequences (e.g., 1 kb sequences) derived from mixed genomes of multiple organisms in an environmental sample are available, it appears impossible to identify how many and what types of genomes are present in the sample. However, we found that BLSOM can classify the genomic fragments according to phylotype without any information other than oligonucleotide composition; BLSOM can properly recognize species-specific characteristics of oligonucleotide composition in most genomic fragments, permitting phylotype-specific clustering (self-organization) of sequences and unveiling diagnostic oligonucleotides responsible for the phylotype-specific clustering [2, 47]. In our previous study, this approach was successfully applied to analyze the bacterial flora of ticks, which resulted in the identification of over a hundred different genera, including novel Chlamydiae, that had previously not been found in ticks [32]. Additionally, phylotype-specific classification methods theoretically similar to BLSOM were applied to metagenomic studies [10, 13, 29, 50].

The aim of this study was to reveal virome in *Ixodes persulcatus* tick by using a combination of a high throughput sequencing technique and a bioinformatics tool based on BLSOM. This tick species is a vector of tick-borne encephalitis (TBE) virus in the Far East [18]. In Hokkaido, first human case of TBE has been reported in 1993 [44]. However, the responsible vector of TBE in the case was supposed to be *I. ovatus* [45]. After 23 years, a second TBE case was reported in 2016 [43]. Recently, new tick-borne *Phlebovirus* was detected from *I. persulcatus* in Hokkaido [30]. Although, the potency of *I. persulcatus* as a vector of viral pathogens is still unknown.

MATERIALS AND METHODS

Sample collection site

Adult *I. persulcatus* ticks were collected in Hidaka (42.97 N 142.68 E) in Hokkaido, Japan by flagging method. No specific permissions were required for the location and activities. Our field activities did not involve endangered or protected species. Live adult ticks were separated by sexes and stored in an incubator at 4°C with over 80% humidity until used for the next sample preparation step.

Enrichment of virion populations from ticks

The ticks were washed with 70% of ethanol and rinsed with SM buffer with gelatin (1 *l* of SM buffer containing 5.8 g NaCl, 1.2 g MgSO$_4$, 50 m*l* of 1 M Tris-HCl pH 7.5, and 0.1 g gelatin) several times. Homogenates were prepared from 10 whole ticks of the same gender by placing them in a 2 m*l* microtube containing two stainless beads with 100 *μl* of SM buffer and then shaking the tubes in a beads homogenizer (Tomy, Tokyo, Japan) at 3,000 rpm for 30 sec. After homogenization, 300 *μl* of SM buffer was added into each tube and the homogenates were remixed with fresh SM buffer. To remove tick debris and intact cells, the homogenates were centrifuged at 5,800 g for 30 min. Supernatants obtained from 10 tubes of the same gender were pooled and filtered through 0.45 and 0.22 *μm* pore-size polyether membranes (Whatman, Maidstone, U.K.). Small particles in the filtrates were concentrated with a tangential flow filtration cassette with a 30 kDa nominal molecular weight cut-off (NMWCO) regenerated cellulose membrane (Millipore, Burlington, MA, U.S.A.). The virus-enriched fraction was recovered from the filter-retained part and resuspended in DNase buffer. Then the virus-enriched fraction was treated with 2.5 U/*μl* DNase I (Nippon gene, Tokyo, Japan) and 2 U/*μl* RNase I$_f$ (New England Biolabs, Ipswich, MA, U.S.A.) at 37°C for 60 min. After enzyme inactivation at 75°C for 20 min, viral nucleic acids were extracted using NucleoSpin RNA XS (Takara, Kusatsu, Japan) according to the manufacturer's instructions.

Reverse transcriptase priming and amplification

Single-stranded DNA synthesis was performed on the extracted viral nucleic acids with SuperScript III reverse transcriptase (Invitrogen, Waltham, MA, U.S.A.) using previously published methods [9, 21, 22]. Viral nucleic acids were mixed with 100 *p*mol of random primer A was heated at 65°C for 5 min and cooled on ice for 2 min to denature secondary structure. The random primer A (GTTTCCCAGTCACGATCNNNNNNNNN) consists of two parts; GTTTCCCAGTCACGATC corresponding to the sequence of primer B used in the following step, and randomly arranged 9 nucleotides (NNNNNNNNN). Subsequently, the following components were added: 4 *μl* of 5× first strand buffer, 1 *μl* of 10 mM deoxynucleoside triphosphate (dNTP) mix, 40 units of RNase OUT, 1 *μl* of 0.1 M dithiothreitol (DTT), and 200 units of SuperScript III reverse transcriptase (Invitrogen). The reaction mixture was incubated at 25°C for 5 min and 50°C for 60 min, followed by 75°C for 15 min to inactivate the transcriptase. To synthesize

double-stranded DNA, 2.5 units of Klenow Fragment (3′-5′ exo-) (New England Biolabs) was added to the cDNA mixture containing random primer A. After incubation at 37°C for 60 min, the enzyme was inactivated at 75°C for 10 min.

The synthesized double-stranded DNA was amplified by employing sequence-independent single-primer amplification (SISPA) method established in a previous study [40]. PCR of the first step products was performed in a total volume of 50 μl containing 5 μl of 10× EX taq buffer, 4 μl of 10 mM dNTPs, 5 units of EX taq, and 500 $p$mol of primer B. The reaction mixture was incubated under the following conditions: 40 cycles of 94°C for 21 sec, 40°C for 30 sec, 50°C for 30 sec, and 72°C for 1 min. Amplified products were purified using Wizard® SV Gel and PCR Clean-Up System (Promega Corp., Madison, WI, U.S.A.) The concentrations and quality of the amplified products were assessed on an Agilent 2100 Bioanalyzer using a DNA1000 Lab Chip kit (Agilient, Palo Alto, CA, U.S.A.).

Pyrosequencing and data analysis

Sequencing of the amplified products was performed on a 454 pyrosequencing Genome Sequencer Junior (GS Junior) (Roche, Basel, Switzerland) according to the manufacturer's protocol. The raw sequencing data file in a standard fogram format (.sff) was converted into a FASTA file, the primer sequence was trimmed, and low quality and short (<150 bp) reads were removed using CLC Genomics Workbench. *De novo* assembly was also performed using this software with default settings. The contigs were subjected to a homology search using BLASTx with the GenBank database, and a phylogenic tree was constructed with Mega 6.05 software. Furthermore, the contigs consisting of more than 300 bases were used to identify viruses using BLSOM analysis.

BLSOM analysis

Self-Organizing Map (SOM) is a neural network algorithm based on unsupervised learning that carries out a characteristic nonlinear projection from the high-dimensional space of input data onto a two-dimensional array of weight vectors [24]. We used the 'Batch Learning SOM' (BLSOM), which is a modified version of the conventional SOM for genome informatics that makes the learning process and creation of the resulting map independent of the order of data input [1, 23]. Instead of random values, they defined the initial weight vectors by Principal Component Analysis (PCA). BLSOM learning was conducted as described previously [1]. We developed a software, with which researchers could easily predict the phylotype for each of a large number of metagenomics sequences using BLSOM. The software, which is named PEMS (Phylogenetic Estimation of Metagenomic sequences on the basis of batch-learning Selforganizing map), can be downloaded freely at http://bioinfo.ie.niigata-u.ac.jp/?PEMS_Soft_e.

In advance, two types of large-scale BLSOMs, namely Kingdom- and Virus group-BLSOM, were constructed to identify viruses from metagenomic sequences using sequences deposited in DDBJ/EMBL/GenBank as previously described [1]. Kingdom-BLSOM was constructed with tetranucleotide frequencies in all 5-kb sequences derived from the whole-genome sequences of 111 eukaryotes, 2,813 prokaryotes, 1,728 mitochondria, 110 chloroplasts, and 31,486 viruses. Virus group-BLSOM was constructed with a total of 602,951 1-kb sequences from 97 families.

After *de novo* assembly, contigs longer than 300 bp were mapped using Virus group-BLSOM. The mapping was conducted by finding the lattice point with the minimum Euclidean distance in the multidimensional space and was assigned to Virus group-BLSOM on the basis of statistical tests. To taxonomically classify the contigs that could not be assigned using Virus group-BLSOM, Kingdom-BLSOM was employed.

Nucleotide sequence accession number

Pyrosequencing data have been submitted to DDBJ under accession number of DRA002848.

RESULTS

Results of pyrosequencing

A total of 133,932 and 175,545 sequence reads were obtained from female and male ticks, respectively. After trimming tag sequences and removing short (<150 bp) and low quality reads, 100,634 and 156,837 sequence reads with average lengths of 351 and 429 bp from female and male samples, respectively, were finally used for further analyses.

De novo assembly and taxonomic classification by BLASTx

Sequence reads were assembled by using CLC Genomics Workbench version 7.5.1 (Qiagen Inc., Valencia, CA, U.S.A.). After *de novo* assembly, 577 and 386 contigs were obtained from female and male samples, respectively. The sizes of longest contigs were 4,291 and 8,972 bp in length from female and male samples, respectively. Taxonomic classifications of contigs were firstly performed with BLASTx analysis (cut-off e-value: $<10^{-5}$). In female sample, 7.6% (44/577) of contigs were assigned to viruses (Fig. 1A). In male sample, 11.9% (46/386) of contigs were assigned to viruses (Fig. 1B). At the lower taxonomic levels, these viral contigs were classified as sequences derived from members of the order Mononegavirales, or the families *Rhabdoviridae*, *Nairoviridae*, *Peribunyaviridae*, *Phenuiviridae*, and *Totiviridae* (Fig. 1C and 1D).

The contigs assigned to viruses by BLASTx analysis are listed in Table 1. Some contigs were identified as a part of the RNA-dependent RNA polymerase gene of the family *Peribunyaviridae* (Table 1). Also, some contigs showed homology to the sequences of South Bay virus. The longest L segment-like contigs (IPf_95 and IPm_68) from both samples were clustered with sequences from viruses belonging to the genus *Nairovirus* in the phylogenetic tree (Fig. 2). The similar sequences of South Bay virus S
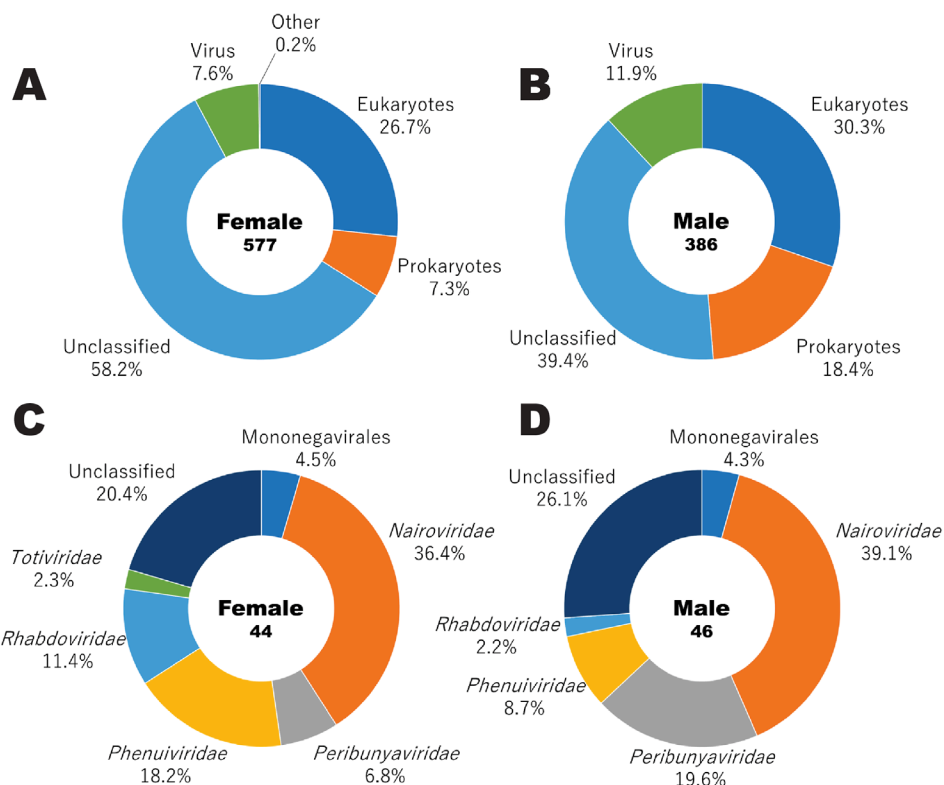
Fig. 1. Classification of the contigs from female and male samples using BLASTx analysis. A, Kingdom classification of contigs in female; B, Kingdom classification of contigs in male; C, Order and family classification of the viral contigs in female; D, Order and family classification of the viral contigs in male.

segments were found in both samples (Fig. 3). The contigs similar to deer tick mononegavirales-like virus and blacklegged tick phlebovirus were also detected in both male and female samples (Table 1). *Nairoviridae*, *Peribunyaviridae*, and *Phenuiviridae* have segmented genome, however, none of their M segments were detected in the samples.

A total of 507 and 324 contigs were yielded from female and male samples, respectively, with a length of over 300 bp. BLSOM analysis of these contigs showed that about half of the contigs fell on the clusters of viruses (Fig. 4). Only 2.0% (10/507) and 2.5% (8/324), from female and male samples, respectively, were not locatable in the BLSOM map (Fig. 4). The contigs assigned to virus by BLSOMs were several viral groups (Fig. 5). Sequences of double strand (ds) DNA viruses were occupied nearly 50% of viral contigs from female and male samples (Fig. 5). Female and male ticks respectively carried 36 and 29 different viral order and families, and the total viral taxa detected by this analysis covered 1 order and 40 families (Table 2). One order and 23 families were detected in both female and male ticks (Table 2). These families and order were containing viruses infecting vertebrates, insects, plants and bacteria. Over 10 contigs were assigned *Bunyavirales* and each family of *Herpesviridae* and *Myoviridae* (Table 2). Among the contigs (>300 bp) assigned to viruses by BLASTx analysis, 76.2% (32/42) and 76.7% (33/43) of them, respectively, from females and males were also assigned to viruses by BLSOM analysis.

Assessment of BLSOM analysis

To validate accuracy of BLSOM analysis, mock data sets were created from three datasets (A, B, and C) were prepared from the viral and non-viral (eukaryotic and prokaryotic) sequences deposited in DDBJ/EMBL/GenBank. The datasets A, B, and C contained BLAST-identified viral sequences with lengths ranging between 300–1,000 bp, 500–1,000 bp, 750–1,000 bp, respectively. When the dataset C was tested, approximately 80% of the fragments were correctly detected and approximately 95% of the fragments were correctly classified into the corresponding taxa by Kingdom-BLSOM at kingdom level. Furthermore, about 80% of these viral sequences were assigned to the corresponding taxa at the family level with accuracy (Supplemental Table S1).

DISCUSSION

Tick viromes might be constituted of a variety of viral families including those containing human and animal pathogens, such as order of *Bunyavirales* and members of the families *Flaviviridae*, *Reoviridae*, and *Orthomyxoviridae*. It is interesting to understand the relationships between viruses found in ticks and known pathogenic viruses from the viewpoint of viral evolution, especially acquirement of pathogenicity against mammals. It is also urgent to analyze whether ticks can transmit those yet-unknown viruses to animals and cause emerging diseases.

Table 1. Contigs from female (A) and male (B) sample assigned to viruses by BLASTx analysis

(A) Female

Contig ID	Length (bp)	Average Covarege	BLASTx	Identity (%)	Mach with BLSOM
IPf_2	766	32.9	RNA-dependent RNA polymerase [Tacheng Tick Virus 3]	36.7 [a]	
IPf_3	766	122.7	RNA-dependent RNA polymerase [Wuhan Tick Virus 1]	31.9 [a]	
IPf_8	462	240.9	N protein [South Bay virus]	51.0 [a]	
IPf_10	1,628	106.0	L protein [South Bay virus]	63.8 [b]	
IPf_15	1,544	3,032.5	capsid precursor, partial [Drosophila A virus]	33.6 [a]	
IPf_19	2,038	36.9	polymerase [Deer tick mononegavirales-like virus]	83.1	
IPf_20	2,626	2,705.1	putative RNA dependent RNA polymerase [Ourmia melon virus]	36.3 [a]	
IPf_29	261	15.9	N protein [South Bay virus]	52.8 -	
IPf_32	492	1,237.1	L protein [South Bay virus]	39.6 [a]	
IPf_33	339	84.5	L protein [South Bay virus]	49.4	
IPf_34	344	65.1	L protein [South Bay virus]	47.2	
IPf_35	2,303	134.4	L protein [South Bay virus]	32.4	
IPf_47	718	75.8	L protein [South Bay virus]	76.9 [a]	
IPf_48	718	34.1	L protein [South Bay virus]	76.9 [a]	
IPf_52	3,463	94.5	RNA-dependent RNA polymerase [Huangpi Tick Virus 3]	48.5	
IPf_77	4,152	92.7	glycoprotein precursor [Leopards Hill virus]	36.7 [b]	
IPf_81	1,234	30.8	RNA-dependent RNA polymerase [Wenzhou Tick Virus]	28.4 [b]	
IPf_84	2,657	23.4	putative glycoprotein [Suffolk virus]	80.1 [a]	
IPf_86	3,079	97.9	L protein [South Bay virus]	68.6 [b]	
IPf_95	4,291	105.4	L protein [South Bay virus]	65.2 [b]	
IPf_96	1,513	59.1	RNA-dependent RNA polymerase [Wenzhou Tick Virus]	34.5 [a]	
IPf_99	1,315	71.5	RNA-dependent RNA polymerase [Camponotus yamaokai virus]	24.3	
IPf_103	2,553	67.0	hypothetical protein, partial [Ixodes scapularis associated virus 1]	78.1 [a]	
IPf_119	1,398	60.0	L protein [Blacklegged tick phlebovirus-2]	32.4 [b]	
IPf_120	1,543	101.0	nucleocapside [Leopards Hill virus]	35.3	
IPf_125	3,576	41.4	polymerase [Deer tick mononegavirales-like virus]	74.2 [b]	
IPf_126	331	56.7	polymerase [Deer tick mononegavirales-like virus]	66.4	
IPf_127	1,331	66.2	L protein [South Bay virus]	50.9 [a]	
IPf_128	629	21.5	polyprotein [Salanga virus]	37.1 [a]	
IPf_150	1,296	43.5	RNA polymerase [Ingwavuma virus]	27.9 [a]	
IPf_152	1,012	28.0	N protein [Blacklegged tick phlebovirus-1]	38.5 [a]	
IPf_169	3,619	40.7	RNA-dependent RNA polymerase [Kasokero virus]	42.2 [b]	
IPf_219	1,441	30.4	ORF3 [Suffolk virus]	66.8	
IPf_223	1,073	10.8	RNA-dependent RNA polymerase [Huangpi Tick Virus 1]	58.3 [a]	
IPf_230	559	28.4	L protein [South Bay virus]	87.7 [b]	
IPf_248	1,770	13.7	RNA-dependent RNA polymerase [Huangpi Tick Virus 1]	59.3 [b]	
IPf_257	595	13.1	RNA-dependent RNA polymerase [Huangpi Tick Virus 1]	55.1 [a]	
IPf_307	630	56.3	polymerase [Deer tick mononegavirales-like virus]	67.7 [a]	
IPf_308	201	9.2	polymerase [Deer tick mononegavirales-like virus]	59.7 -	
IPf_322	1,492	8.5	L protein [Blacklegged tick phlebovirus-2]	62.9 [b]	
IPf_338	555	15.9	L protein [Blacklegged tick phlebovirus-2]	42.6	
IPf_418	472	3.1	L protein [Blacklegged tick phlebovirus-1]	42.5 [a]	
IPf_449	597	2.8	RNA-dependent RNA polymerase [Whenzhou Shrimp Virus 2]	32.1 [a]	
IPf_487	760	2.3	L protein [Blacklegged tick phlebovirus-1]	66.3 [a]	

In BLASTx analysis, some contigs were revealed to be related to South Bay virus (Fig. 2). This virus was originally detected in *I. scapularis* using a high throughput sequencing technique [43], but its pathogenic potential in animals is still unknown. Contigs similar to *I. scapularis*-associated virus, deer tick mononega-like virus, and blacklegged tick phlebovirus were also identified (Table 1). Though these viruses were firstly detected in *I. scapularis* collected in New York [46], the present study demonstrated the prevalence of similar viruses in tick populations in Japan. M segments of South Bay virus and blacklegged tick phlebovirus were not found in the previous studies [46]. Similarly, the sequence of M segment was not detected in the present study. Another virome study on a wide range of arthropods including ticks reported the detection of a number of *Bunyavirales* viruses [27]. Despite the fact that the authors used the consistent sample and data processing protocols, M segments were not discovered in all viruses detected in the study. The M segment of prototypical *Bunyavirales* encodes a glycoprotein responsible for cellular attachment. The sequences of M segments are more variable than those of L and S segments [37]. These facts may indicate that some *Bunyavirales* viruses lack M segments in their genomes or that the sequences of M segments are quite different from those

(B) Male

| Contig ID	Length (bp)	Average Covarege	BLASTx	Identity (%)	Mach with BLSOM
IPm_10	3,528	211.9	RNA-dependent RNA polymerase [Huangpi Tick Virus 3]	54.2	
IPm_11	3,989	1,420.6	polymerase [Brazoran virus]	25.8	a)
IPm_45	1,034	653.7	L protein [South Bay virus]	56.9	a)
IPm_46	402	560.6	RNA-dependent RNA polymerase [Whenzhou Shrimp Virus 2]	34.9	
IPm_50	357	54.3	N protein [South Bay virus]	53.4	a)
IPm_51	377	35.3	N protein [South Bay virus]	54.4	a)
IPm_55	533	128.0	N protein [South Bay virus]	50.7	a)
IPm_56	513	167.4	N protein [South Bay virus]	52.7	
IPm_59	206	4.0	N protein [South Bay virus]	54.6	-
IPm_61	468	922.3	RNA-dependent RNA polymerase [Whenzhou Shrimp Virus 2]	34.0	a)
IPm_68	8,972	242.4	L protein [South Bay virus]	68.1	b)
IPm_82	616	47.9	L protein [South Bay virus]	33.3	a)
IPm_90	3,039	629.7	L protein [South Bay virus]	31.7	
IPm_92	1,385	111.5	glycoprotein precursor [Whenzhou Shrimp Virus 2]	22.6	a)
IPm_105	466	72.6	L protein [South Bay virus]	48.2	a)
IPm_106	474	113.5	L protein [South Bay virus]	48.2	a)
IPm_111	4,206	123.2	polymerase [Deer tick mononegavirales-like virus]	72.8	b)
IPm_116	2,480	22.4	polymerase [Lishi Spider Virus 1]	39.5	
IPm_122	442	125.1	L protein [South Bay virus]	65.7	b)
IPm_123	442	118.5	L protein [South Bay virus]	65.7	b)
IPm_134	742	89.3	polymerase [Deer tick mononegavirales-like virus]	87.8	
IPm_135	502	6.4	polymerase [Deer tick mononegavirales-like virus]	92.2	
IPm_141	2,375	44.4	hypothetical protein, partial [Ixodes scapularis associated virus 2]	58.4	a)
IPm_143	1,873	690.1	putative glycoprotein [Suffolk virus]	74.4	a)
IPm_147	2,563	126.0	ORF3 [Suffolk virus]	63.7	b)
IPm_152	702	9.1	RNA-dependent RNA polymerase [Huangpi Tick Virus 1]	38.5	a)
IPm_153	1,266	56.3	putative nucleoprotein [Wuhan Mosquito Virus 8]	43.2	a)
IPm_161	1,696	55.9	RNA-dependent RNA polymerase [Wenzhou Tick Virus]	48.3	a)
IPm_173	1,092	23.0	polymerase [Wenzhou Crab Virus 2]	59.4	a)
IPm_176	535	10.3	polymerase [Deer tick mononegavirales-like virus]	66.7	a)
IPm_178	956	39.7	L protein [Blacklegged tick phlebovirus-2]	50.8	b)
IPm_194	1,111	12.9	RNA-dependent RNA polymerase [Tacheng Tick Virus 1]	52.7	a)
IPm_197	524	54.3	RNA-dependent RNA polymerase [Huangpi Tick Virus 1]	57.7	b)
IPm_207	450	3.3	L protein [Blacklegged tick phlebovirus-1]	63.8	a)
IPm_213	891	20.3	RNA-dependent RNA polymerase [Tacheng Tick Virus 1]	33.0	
IPm_218	2,075	14.3	glycoprotein precursor [Tacheng Tick Virus 1]	42.5	a)
IPm_223	999	25.3	RNA-dependent RNA polymerase [Huangpi Tick Virus 1]	63.7	a)
IPm_235	1,177	6.6	RNA-dependent RNA polymerase [Huangpi Tick Virus 1]	65.9	a)
IPm_249	400	2.3	RNA-dependent RNA polymerase [Wenzhou Tick Virus]	52.3	
IPm_274	541	15.8	RNA-dependent RNA polymerase [Huangpi Tick Virus 1]	57.7	
IPm_286	762	7.8	L protein [Blacklegged tick phlebovirus-1]	52.8	a)
IPm_319	703	4.1	L protein [Blacklegged tick phlebovirus-2]	25.8	b)
IPm_332	260	15.2	L protein [South Bay virus]	53.1	-
IPm_336	482	2.7	RNA-dependent RNA polymerase [Huangpi Tick Virus 1]	59.0	a)
IPm_362	299	1.9	hypothetical protein, partial [Ixodes scapularis associated virus 1]	88.9	-
IPm_383	630	1.3	RNA-dependent RNA polymerase [Tacheng Tick Virus 1]	55.1	a)

a) Contig correctly assigned to virus at kingdom level by BLSOM. b) Contig correctly assigned to virus at family and/or order level by BLSOM. -: Less than 300 bp contig could not be used for BLSOM analysis.

deposited in the database which leads to false negative results using BLAST analysis. Further studies are warranted to investigate the pathogenic potential of these newly identified viruses. The results obtained by BLASTx analysis indicate that metagenomic approach coupled with viral purification steps is robust for the detection and characterization of a wide range of viruses, especially previously unknown viruses from arthropods.

In BLSOM analysis, sequences related to several insect virus families were identified; that is, *Ascoviridae*, *Baculoviridae*, *Closteroviridae*, *Iridoviridae*, *Polydnaviridae* and *Poxviridae* (Table 2). Biological interactions between human or animal pathogens, their vector arthropods and their own viruses, have been reported [8, 11, 42], and such interactions can be utilized for disease control in the fields of agriculture and medicine. For example, mosquito-derived Densoviruses belonging to the
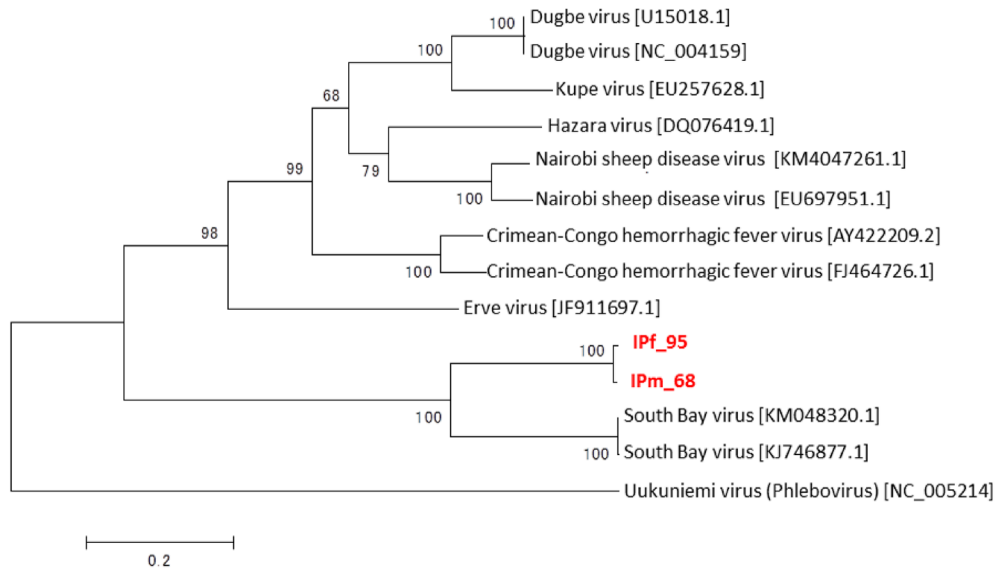
Fig. 2. Maximum-likelihood phylogenetic tree based on the nucleotide sequences of the longest contigs mapped to *Bunyavirales* L segment. All bootstrap values from 1,000 replications are shown on interior branch nodes.
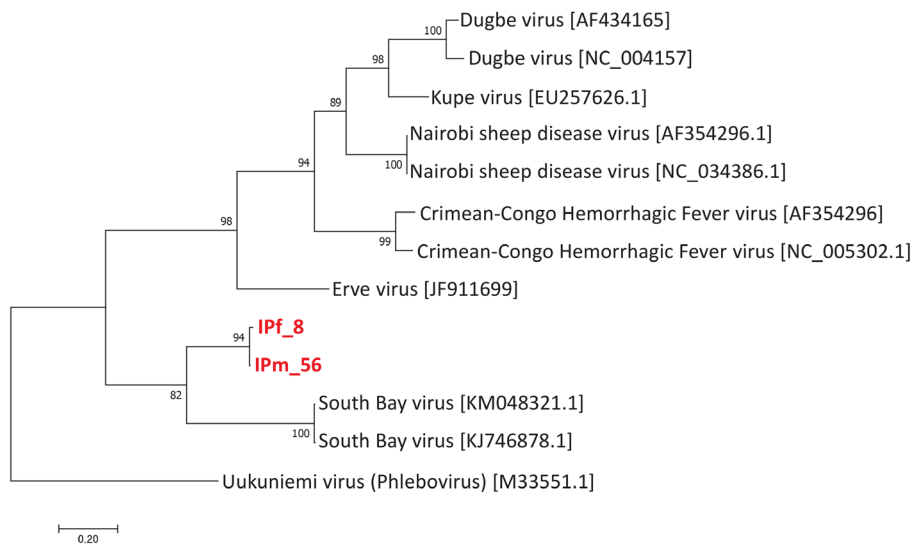
Fig. 3. Maximum-likelihood phylogenetic tree based on the nucleotide sequences of the longest contigs mapped to *Bunyavirales* S segment. All bootstrap values from 1,000 replications are shown on interior branch nodes.
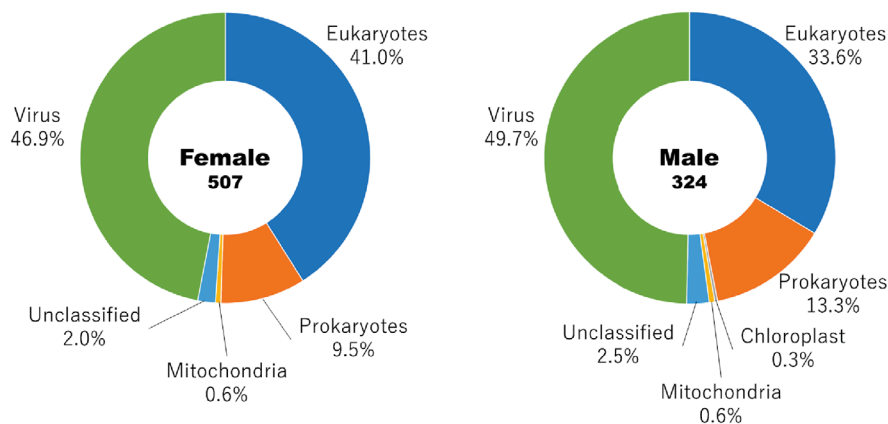
Fig. 4. Results of Kingdom-BLSOM of the contigs from female and male samples. Only contigs with a length of over 300 bp were used.
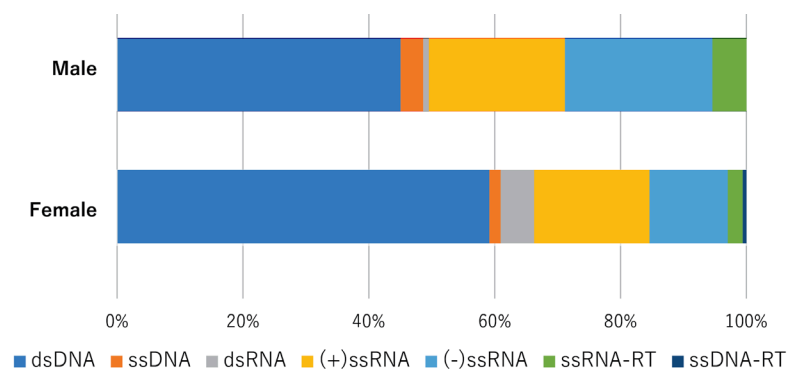
Fig. 5. Viral groups found in female and male samples. Contigs classified by using Virus group-BLSOM.

Table 2. Number of contigs assigned to each family by BLSOM analysis

Viral group	Order / Family	No. of contigs	
		Female	Male
dsDNA	*Adenoviridae*	11	2
dsDNA	*Alloherpesviridae*	3	2
dsDNA	*Ascoviridae*	1	
dsDNA	*Asfarviridae*	2	
dsDNA	*Baculoviridae*	6	1
dsDNA	*Herpesviridae*	28	11
dsDNA	*Iridoviridae*	1	2
dsDNA	*Mimiviridae*	1	
dsDNA	*Myoviridae*	14	11
dsDNA	*Papillomaviridae*	3	2
dsDNA	*Phycodnaviridae*	3	2
dsDNA	*Podoviridae*	5	4
dsDNA	*Polyomaviridae*	1	
dsDNA	*Poxviridae*	2	
dsDNA	*Siphoviridae*	16	9
ssDNA	*Anelloviridae*		2
ssDNA	*Circoviridae*		1
ssDNA	*Geminiviridae*	2	1
ssDNA	*Parvoviridae*	1	
dsRNA	*Birnaviridae*	3	
dsRNA	*Reoviridae*	6	1
(+) ssRNA	*Alphaflexiviridae*	2	
(+) ssRNA	*Arteriviridae*		1
(+) ssRNA	*Betaflexiviridae*	4	1
(+) ssRNA	*Bromoviridae*	2	1
(+) ssRNA	*Closteroviridae*	1	4
(+) ssRNA	*Coronaviridae*	6	1
(+) ssRNA	*Flaviviridae*	4	5
(+) ssRNA	*Hepeviridae*	1	
(+) ssRNA	*Luteoviridae*	1	
(+) ssRNA	*Picornaviridae*	4	7
(+) ssRNA	*Potyviridae*	5	
(+) ssRNA	*Secoviridae*		2
(+) ssRNA	*Togaviridae*	1	2
(−) ssRNA	*Arenaviridae*		1
(−) ssRNA	*Bunyavirales*	12	13
(−) ssRNA	*Filoviridae*	1	1
(−) ssRNA	*Orthomyxoviridae*	4	8
(−) ssRNA	*Paramyxoviridae*	4	3
ssRNA-RT	*Retroviridae*	4	6
dsDNA-RT	*Caulimoviridae*	1	

family *Baculoviridae* are used as stable vectors for the transformation of mosquitoes, which has created interest in using these viruses for mosquito and malaria control, either directly as insect-killing agents or as carriers of transgenes whose products interfere with parasite development [39]. Members of the family *Baculoviridae* are utilized for the control of insect pests [7]. The insect viruses detected in this study may have potential as tools for the biological control of ticks and tick-borne diseases.

Myoviridae, *Podoviridae*, and *Siphoviridae*, bacteriophage families, were also detected (Table 2). Phages detected in the tick viromes may be infectious to bacterial hosts residing in ticks or to those derived from vertebrate hosts. Bacteria belonging to the genus *Bacillus* and *Pseudomonas*, the hosts of *Bacillus phage G* and *Pseudomonas phages* belonging to the *Myoviridae* family, are common bacterial species found in the microbiome of *I. persulcatus* [32]. A metagenomic study of bacterial communities associated with ticks has revealed broad bacterial diversity in ticks [32, 38]. It is of interest to investigate the interactions between bacteria and bacteriophages in ticks, as phages may affect the physiology and fitness of ticks indirectly through their interaction with host bacteria.

Herpesviridae was detected from female and male samples. Main host of this viral family is vertebrates except that a herpesvirus was found in pacific oyster as an invertebrate host [49]. Murid herpesvirus 4 strain 68 (genus *Rhadinovirus*, subfamily *Gammaherpesvirinae*) was detected from hard tick species, such as *I. ricinus*, in Europe [15, 25, 26, 48]. Thus, possible explanations are that the herpesvirus genome detected in this study was mechanically acquired by ticks through the blood feeding on vertebrate hosts and the virus cannot replicate in ticks, or that ticks are potential hosts of some uncharacterized members in *Herpesviridae*.

In BLSOM analysis, 24 out of 41 viral families and order were detected in both female and male ticks. Previous study suggested a difference of bacterial population between female and male ticks [38]. To reveal a difference of virome in female and male ticks, expanding sample number and/or tick species are required in the further study.

About 10% of the contigs were assigned to viruses using BLASTx approach, whereas about 40 to 60% of contigs could not be defined (Fig. 1A and 1B). The classification methods for microbes based on sequence similarity such as BLASTx have limitations for metagenomic analysis [6, 41, 51]. This is because current genome databases still cover the limited range of the genomes of organisms inhabiting the planet Earth, including viruses, despite the rapid increase in DNA entries. It is also possible that many of the sequence reads are too divergent from the sequence data deposited in reference databases, resulting in the difficulty of finding similar sequences. On the other hands, BLSOM does not require orthologous sequence data sets for phylogenetic classification of sequences [2, 47]. It is therefore possible to find taxonomical relationships of never-reported organisms to known, well-established organisms. This is one of the advantages of BLSOM, when it is applied to microbiomes composed of poorly characterized and highly diversified organisms [2]. About half of the contigs were assigned to viruses using BLSOM, and only a small percentage of the contigs could not be assigned to any organisms (Fig. 4), which supports that BLSOM is theoretically advantageous in detecting and classifying previously unknown viruses over the homology-based search.

It should be considered, however, that BLSOM-based classification has its own limitations especially in estimation accuracy as

demonstrated in the supplementary data (Supplemental Table S1). In fact, only about 76% of viral contigs (>300 bp) identified by BLASTx were allocated to viruses identified by BLSOM (Table 1). Moreover, about 70% of the pairwise comparisons between two methods were not in accordance at the family and/or order level (Table 1). These results indicate that some viral sequences might be overlooked in BLASTx analysis, while diversity of viral populations might be overestimated in BLSOM analysis. This discrepancy might be minimized if more entries are added to microbial sequence databases, especially those covering unexplored viral world. Therefore, we propose that BLSOM-based classification might be one of the complementing choices for taxonomic assignments of sequence reads. By combining with sequence homology-based methods, the results obtained from BLSOM analysis would be more valuable.

In conclusion, this study provides new and important information on tick virome. The nucleotide composition-based classification method, BLSOM, can be applied to medical and veterinary important vector arthropods, such as mosquito and fly. It might have great potential for mounting effective programs against vector-borne emerging infectious diseases. Both experimental and epidemiological studies are required to assess the risks of the identified viruses for human and animal health. Further identification of those viruses, at the species level or entire genome analyses could be achieved by using conventional methods, such as viral isolation from cell cultures or susceptible animals, or other molecular methods such as species-specific primer extension. Much deeper sequencing using larger amounts of nucleic acid materials than those used in this study might yield enough sequence data to assemble entire genomes of unknown viruses.

REFERENCES

1. Abe, T., Kanaya, S., Kinouchi, M., Ichiba, Y., Kozuki, T. and Ikemura, T. 2003. Informatics for unveiling hidden genome signatures. *Genome Res.* **13**: 693–702. [Medline] [CrossRef]
2. Abe, T., Sugawara, H., Kinouchi, M., Kanaya, S. and Ikemura, T. 2005. Novel phylogenetic studies of genomic sequence fragments derived from uncultured microbe mixtures in environmental and clinical samples. *DNA Res.* **12**: 281–290. [Medline] [CrossRef]
3. Belhouchet, M., Mohd Jaafar, F., Tesh, R., Grimes, J., Maan, S., Mertens, P. P. and Attoui, H. 2010. Complete sequence of Great Island virus and comparison with the T2 and outer-capsid proteins of Kemerovo, Lipovnik and Tribec viruses (genus *Orbivirus*, family *Reoviridae*). *J. Gen. Virol.* **91**: 2985–2993. [Medline] [CrossRef]
4. Berrada, Z. L. and Telford, S. R. 3rd. 2009. Burden of tick-borne infections on American companion animals. *Top. Companion Anim. Med.* **24**: 175–181. [Medline] [CrossRef]
5. Beugnet, F. and Marié, J. L. 2009. Emerging arthropod-borne diseases of companion animals in Europe. *Vet. Parasitol.* **163**: 298–305. [Medline] [CrossRef]
6. Bibby, K., Viau, E. and Peccia, J. 2011. Viral metagenome analysis to guide human pathogen monitoring in environmental samples. *Lett. Appl. Microbiol.* **52**: 386–392. [Medline] [CrossRef]
7. Bonning, B. C. and Hammock, B. D. 1996. Development of recombinant baculoviruses for insect control. *Annu. Rev. Entomol.* **41**: 191–210. [Medline] [CrossRef]
8. Carlson, J., Suchman, E. and Buchatsky, L. 2006. Densoviruses for control and genetic manipulation of mosquitoes. *Adv. Virus Res.* **68**: 361–392. [Medline] [CrossRef]
9. Carmona-Vicente, N., Buesa, J., Brown, P. A., Merga, J. Y., Darby, A. C., Stavisky, J., Sadler, L., Gaskell, R. M., Dawson, S. and Radford, A. D. 2013. Phylogeny and prevalence of kobuviruses in dogs and cats in the UK. *Vet. Microbiol.* **164**: 246–252. [Medline] [CrossRef]
10. Chan, C. K., Hsu, A. L., Tang, S. L. and Halgamuge, S. K. 2008. Using growing self-organising maps to improve the binning process in environmental whole-genome shotgun sequencing. *J. Biomed. Biotechnol.* **2008**: 513701. [Medline] [CrossRef]
11. Cheng, L., Chen, S., Zhou, Z. H. and Zhang, J. 2007. Structure comparisons of Aedes albopictus densovirus with other parvoviruses. *Sci. China C Life Sci.* **50**: 70–74. [Medline] [CrossRef]
12. Day, J. M., Ballard, L. L., Duke, M. V., Scheffler, B. E. and Zsak, L. 2010. Metagenomic analysis of the turkey gut RNA virus community. *Virol. J.* **7**: 313. [Medline] [CrossRef]
13. Dick, G. J., Andersson, A. F., Baker, B. J., Simmons, S. L., Thomas, B. C., Yelton, A. P. and Banfield, J. F. 2009. Community-wide analysis of microbial genome sequence signatures. *Genome Biol.* **10**: R85. [Medline] [CrossRef]
14. Fancello, L., Trape, S., Robert, C., Boyer, M., Popgeorgiev, N., Raoult, D. and Desnues, C. 2013. Viruses in the desert: a metagenomic survey of viral communities in four perennial ponds of the Mauritanian Sahara. *ISME J.* **7**: 359–369. [Medline] [CrossRef]
15. Ficová, M., Betáková, T., Pančík, P., Václav, R., Prokop, P., Halásová, Z. and Kúdelová, M. 2011. Molecular detection of murine herpesvirus 68 in ticks feeding on free-living reptiles. *Microb. Ecol.* **62**: 862–867. [Medline] [CrossRef]
16. Gorenflot, A., Moubri, K., Precigout, E., Carcy, B. and Schetters, T. P. 1998. Human babesiosis. *Ann. Trop. Med. Parasitol.* **92**: 489–501. [Medline] [CrossRef]
17. Gray, J. S., Dautel, H., Estrada-Peña, A., Kahl, O. and Lindgren, E. 2009. Effects of climate change on ticks and tick-borne diseases in europe. *Interdiscip. Perspect. Infect. Dis.* **2009**: 593232. [Medline] [CrossRef]
18. Hayasaka, D., Ivanov, L., Leonova, G. N., Goto, A., Yoshii, K., Mizutani, T., Kariwa, H. and Takashima, I. 2001. Distribution and characterization of tick-borne encephalitis viruses from Siberia and far-eastern Asia. *J. Gen. Virol.* **82**: 1319–1328. [Medline] [CrossRef]
19. He, B., Li, Z., Yang, F., Zheng, J., Feng, Y., Guo, H., Li, Y., Wang, Y., Su, N., Zhang, F., Fan, Q. and Tu, C. 2013. Virome profiling of bats from Myanmar by metagenomic analysis of tissue samples reveals more novel Mammalian viruses. *PLoS One* **8**: e61950. [Medline] [CrossRef]
20. Hurwitz, B. L. and Sullivan, M. B. 2013. The Pacific Ocean virome (POV): a marine viral metagenomic dataset and associated protein clusters for quantitative viral ecology. *PLoS One* **8**: e57355. [Medline] [CrossRef]

21. Jones, M. S. 2nd., Harrach, B., Ganac, R. D., Gozum, M. M., Dela Cruz, W. P., Riedel, B., Pan, C., Delwart, E. L. and Schnurr, D. P. 2007. New adenovirus species found in a patient presenting with gastroenteritis. *J. Virol.* **81**: 5978–5984. [Medline] [CrossRef]

22. Jones, M. S., Lukashov, V. V., Ganac, R. D. and Schnurr, D. P. 2007. Discovery of a novel human picornavirus in a stool sample from a pediatric patient presenting with fever of unknown origin. *J. Clin. Microbiol.* **45**: 2144–2150. [Medline] [CrossRef]

23. Kanaya, S., Kinouchi, M., Abe, T., Kudo, Y., Yamada, Y., Nishi, T., Mori, H. and Ikemura, T. 2001. Analysis of codon usage diversity of bacterial genes with a self-organizing map (SOM): characterization of horizontally transferred genes with emphasis on the *E. coli* O157 genome. *Gene* **276**: 89–99. [Medline] [CrossRef]

24. Kohonen, T. 1990. The self-organizing map. *Proceedings of the IEEE*: 17.

25. Kúdelová, M., Belvončíková, P., Vrbová, M., Kovaľová, A., Štibrániová, I., Kocáková, P., Slovák, M., Špitalská, E., Lapuníková, B., Matúšková, R. and Šupolíková, M. 2015. Detection of murine herpesvirus 68 (MHV-68) in *Dermacentor reticulatus* ticks. *Microb. Ecol.* **70**: 785–794. [Medline] [CrossRef]

26. Kúdelová, M., Jánošová, M. and Belvončíková, P. 2018. First detection of murine herpesvirus 68 in adult *Ixodes ricinus* ticks. *Folia Microbiol. (Praha)* **63**: 511–515. [Medline] [CrossRef]

27. Li, C. X., Shi, M., Tian, J. H., Lin, X. D., Kang, Y. J., Chen, L. J., Qin, X. C., Xu, J., Holmes, E. C. and Zhang, Y. Z. 2015. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *eLife* **4**: e0537827. [Medline] [CrossRef]

28. Li, L., Victoria, J. G., Wang, C., Jones, M., Fellers, G. M., Kunz, T. H. and Delwart, E. 2010. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. *J. Virol.* **84**: 6955–6965. [Medline] [CrossRef]

29. Martin, C., Diaz, N. N., Ontrup, J. and Nattkemper, T. W. 2008. Hyperbolic SOM-based clustering of DNA fragment features for taxonomic visualization and classification. *Bioinformatics* **24**: 1568–1574. [Medline] [CrossRef]

30. Matsuno, K., Kajihara, M., Nakao, R., Nao, N., Mori-Kajihara, A., Muramatsu, M., Qiu, Y., Torii, S., Igarashi, M., Kasajima, N., Mizuma, K., Yoshii, K., Sawa, H., Sugimoto, C., Takada, A. and Ebihara, H. 2018. The unique phylogenetic position of a novel tick-borne phlebovirus ensures an ixodid origin of the genus *Phlebovirus*. *MSphere* **3**: e0023918. [Medline] [CrossRef]

31. McMullan, L. K., Folk, S. M., Kelly, A. J., MacNeil, A., Goldsmith, C. S., Metcalfe, M. G., Batten, B. C., Albariño, C. G., Zaki, S. R., Rollin, P. E., Nicholson, W. L. and Nichol, S. T. 2012. A new phlebovirus associated with severe febrile illness in Missouri. *N. Engl. J. Med.* **367**: 834–841. [Medline] [CrossRef]

32. Nakao, R., Abe, T., Nijhof, A. M., Yamamoto, S., Jongejan, F., Ikemura, T. and Sugimoto, C. 2013. A novel approach, based on BLSOMs (Batch Learning Self-Organizing Maps), to the microbiome analysis of ticks. *ISME J.* **7**: 1003–1015. [Medline] [CrossRef]

33. Ng, T. F., Willner, D. L., Lim, Y. W., Schmieder, R., Chau, B., Nilsson, C., Anthony, S., Ruan, Y., Rohwer, F. and Breitbart, M. 2011. Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. *PLoS One* **6**: e20579. [Medline] [CrossRef]

34. Nuttall, P. A. 2009. Molecular characterization of tick-virus interactions. *Front. Biosci. (Landmark Ed.)* **14**: 2466–2483. [Medline] [CrossRef]

35. Ogden, N. H., Maarouf, A., Barker, I. K., Bigras-Poulin, M., Lindsay, L. R., Morshed, M. G., O'callaghan, C. J., Ramay, F., Waltner-Toews, D. and Charron, D. F. 2006. Climate change and the potential for range expansion of the Lyme disease vector *Ixodes scapularis* in Canada. *Int. J. Parasitol.* **36**: 63–70. [Medline] [CrossRef]

36. Parola, P. 2004. Tick-borne rickettsial diseases: emerging risks in Europe. *Comp. Immunol. Microbiol. Infect. Dis.* **27**: 297–304. [Medline] [CrossRef]

37. Prinz, M., Fuchs, J., Ehrmann, S., Scherer-Lorenzen, M., Kochs, G. and Panning, M. 2017. Molecular identification of novel phlebovirus sequences in European ticks. *Ticks Tick Borne Dis.* **8**: 795–798. [Medline] [CrossRef]

38. Qiu, Y., Nakao, R., Ohnuma, A., Kawamori, F. and Sugimoto, C. 2014. Microbial population analysis of the salivary glands of ticks; a possible strategy for the surveillance of bacterial pathogens. *PLoS One* **9**: e103961. [Medline] [CrossRef]

39. Ren, X., Hoiczyk, E. and Rasgon, J. L. 2008. Viral paratransgenesis in the malaria vector *Anopheles gambiae*. *PLoS Pathog.* **4**: e1000135. [Medline] [CrossRef]

40. Reyes, G. R. and Kim, J. P. 1991. Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. *Mol. Cell. Probes* **5**: 473–481. [Medline] [CrossRef]

41. Roux, S., Enault, F., Robin, A., Ravet, V., Personnic, S., Theil, S., Colombet, J., Sime-Ngando, T. and Debroas, D. 2012. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS One* **7**: e33641. [Medline] [CrossRef]

42. Rwegoshora, R. T. and Kittayapong, P. 2004. Pathogenicity and infectivity of the Thai-strain densovirus (AThDNV) in *Anopheles minimus* S.L. *Southeast Asian J. Trop. Med. Public Health* **35**: 630–634. [Medline]

43. Tajima, Y., Yaguchi, H. and Mito, Y. 2018. Fatal meningoencephalomyelitis due to the tick-borne encephalitis virus: The first detailed neurological observation in a Japanese patient from the central part of Hokkaido Island. *Intern. Med.* **57**: 873–876. [Medline] [CrossRef]

44. Takashima, I., Morita, K., Chiba, M., Hayasaka, D., Sato, T., Takezawa, C., Igarashi, A., Kariwa, H., Yoshimatsu, K., Arikawa, J. and Hashimoto, N. 1997. A case of tick-borne encephalitis in Japan and isolation of the the virus. *J. Clin. Microbiol.* **35**: 1943–1947. [Medline]

45. Takeda, T., Ito, T., Chiba, M., Takahashi, K., Niioka, T. and Takashima, I. 1998. Isolation of tick-borne encephalitis virus from *Ixodes ovatus* (Acari: Ixodidae) in Japan. *J. Med. Entomol.* **35**: 227–231. [Medline] [CrossRef]

46. Tokarz, R., Williams, S. H., Sameroff, S., Sanchez Leon, M., Jain, K. and Lipkin, W. I. 2014. Virome analysis of *Amblyomma americanum*, *Dermacentor variabilis*, and *Ixodes scapularis* ticks reveals novel highly divergent vertebrate and invertebrate viruses. *J. Virol.* **88**: 11480–11492. [Medline] [CrossRef]

47. Uehara, H., Iwasaki, Y., Wada, C., Ikemura, T. and Abe, T. 2011. A novel bioinformatics strategy for searching industrially useful genome resources from metagenomic sequence libraries. *Genes Genet. Syst.* **86**: 53–66. [Medline] [CrossRef]

48. Vrbová, M., Belvončíková, P., Kovaľová, A., Matúšková, R., Slovák, M. and Kúdelová, M. 2016. Molecular detection of murine gammaherpesvirus 68 (MHV-68) in *Haemaphysalis concinna* ticks collected in Slovakia. *Acta Virol.* **60**: 426–428. [Medline] [CrossRef]

49. Wang, N., Baldi, P. F. and Gaut, B. S. 2007. Phylogenetic analysis, genome evolution and the rate of gene gain in the *Herpesviridae*. *Mol. Phylogenet. Evol.* **43**: 1066–1075. [Medline] [CrossRef]

50. Weber, M., Teeling, H., Huang, S., Waldmann, J., Kassabgy, M., Fuchs, B. M., Klindworth, A., Klockow, C., Wichels, A., Gerdts, G., Amann, R. and Glöckner, F. O. 2011. Practical application of self-organizing maps to interrelate biodiversity and functional data in NGS-based metagenomics. *ISME J.* **5**: 918–928. [Medline] [CrossRef]

51. Winter, C., Garcia, J. A., Weinbauer, M. G., DuBow, M. S. and Herndl, G. J. 2014. Comparison of deep-water viromes from the atlantic ocean and the mediterranean sea. *PLoS One* **9**: e100600. [Medline] [CrossRef]

52. Yu, X. J., Liang, M. F., Zhang, S. Y., Liu, Y., Li, J. D., Sun, Y. L., Zhang, L., Zhang, Q. F., Popov, V. L., Li, C., Qu, J., Li, Q., Zhang, Y. P., Hai, R., Wu, W., Wang, Q., Zhan, F. X., Wang, X. J., Kan, B., Wang, S. W., Wan, K. L., Jing, H. Q., Lu, J. X., Yin, W. W., Zhou, H., Guan, X. H., Liu, J. F., Bi, Z. Q., Liu, G. H., Ren, J., Wang, H., Zhao, Z., Song, J. D., He, J. R., Wan, T., Zhang, J. S., Fu, X. P., Sun, L. N., Dong, X. P., Feng, Z. J., Yang, W. Z., Hong, T., Zhang, Y., Walker, D. H., Wang, Y. and Li, D. X. 2011. Fever with thrombocytopenia associated with a novel bunyavirus in China. *N. Engl. J. Med.* **364**: 1523–1532. [Medline] [CrossRef]