JOURNAL OF GENERAL VIROLOGY

RESEARCH ARTICLE

Van Eynde *et al.*, *Journal of General Virology* 2020;101:651–666

DOI 10.1099/jgv.0.001412

MICROBIOLOGY SOCIETY

Exploration of the virome of the European brown shrimp (*Crangon crangon*)

Benigna Van Eynde[1,2], Olivier Christiaens[1], Daan Delbare[2], Chenyan Shi[3], Emiel Vanhulle[3], Claude Kwe Yinda[3], Jelle Matthijnssens[3] and Guy Smagghe[1,*]

Abstract

Crangon crangon is economically a very important species. Recently, promising culture attempts have been made, but a major problem is the uncontrollable mortality during the grow-out phase. As of yet, the life cycle of *C. crangon* is not closed in captivity so wild-caught individuals are used for further rearing. Therefore, it is important to investigate the virome of *C. crangon* both in wild-caught animals as in cultured animals. In recent years, next-generation-sequencing (NGS) technologies have been very important in the unravelling of the virome of a wide range of environments and matrices, such as soil, sea, potable water, but also of a wide range of animal species. This will be the first report of a virome study in *C. crangon* using NGS in combination with the NetoVIR protocol. The near complete genomes of 16 novel viruses were described, most of which were rather distantly related to unclassified viruses or viruses belonging to the *Picornavirales*, *Bunyavirales Nudiviridae*, *Parvoviridae*, *Flaviviridae*, *Hepeviridae*, *Tombusviridae*, *Narnaviridae*, *Nodaviridae*, *Sobemovirus*. A difference in virome composition was observed between muscle and hepatopancreatic tissue, suggesting a distinct tissue tropism of several of these viruses. Some differences in the viral composition were noted between the cultured and wild shrimp, which could indicate that in sub-optimal aquaculture conditions some viruses become more abundant. This research showed that a plethora of unknown viruses is present in *C. crangon* and that more research is needed to determine which virus is potentially dangerous for the culture of *C. crangon*.

INTRODUCTION

Since the discovery of the first marine viral pathogen in 1966 [1], a great deal of viral pathogens have been described from wild and cultured crustaceans [2]. Traditionally, identification and discovery of novel viral pathogens required propagation of these viruses in cell cultures [3]. Unfortunately, most of the viral species are not readily culturable, thereby hindering the discovery of novel pathogens. Furthermore, no crustacean cell lines are available yet. Identification of unculturable infectious agents became possible with successful implementation of molecular-based methods such as PCR [4, 5]. Unfortunately, characterization and classification using such molecular-based methods is only possible when nucleotide sequences of viruses are known [3].

To circumvent the need for sequence information, a sequence-independent metagenomics approach proved to be a powerful tool for the genomic analysis of a population of micro-organisms [6]. More specifically, viral metagenomics or viromics embodies the process to study uncultured viral populations in any given sample [7]. Furthermore, this approach makes it possible to identify and discover a large number of novel viral genomes without prior knowledge of the viruses in the samples [8, 9]. In recent years, next-generation sequencing technologies or NGS have profoundly facilitated the study of viromes of both a wide range of environments

Received 05 December 2019; Accepted 15 March 2020; Published 11 May 2020

Author affiliations: [1]Department of Plant and Crops, Faculty of Bioscience Engineering, Ghent University, 9000 Ghent, Belgium; [2]Animal Sciences Unit-Fisheries, Flanders research institute for agriculture, fisheries and food (ILVO), 8400 Ostend, Belgium; [3]KU Leuven, Department of Microbiology, Immunology and Transplantation, Rega Institute, Laboratory of Viral Metagenomics, Leuven, Belgium.

***Correspondence:** Guy Smagghe, guy.smagghe@ugent.be

Keywords: aquaculture; *Crangon crangon*; North Sea; virome; viruses.

Abbreviations: CcBV, Crangon crangon bacilliform virus; CcDLV, Crangon crangon dicistro-like virus; CcFLV, Crangon crangon flavi-like virus; CcHLV, Crangon crangon hepe-like virus; CcKLV, Crangon crangon kelpfly-like virus; CcNaLV, Crangon crangon narna-like virus; CcNoLV, Crangon crangon noda-like virus; CcNV, Crangon crangon nudivirus; CcPaLV, Crangon crangon parvo-like virus; CcSLV, Crangon crangon sobemo-like virus; CcTLV, Crangon crangon tombus-like virus; HP, Hepatopancreas; HPV, Hepatopancreatic parvovirus; IHHNV, Infectious hypodermal and haematopoietic necrosis virus; MBV, Monodon baculuvirus; NGS, Next generation sequencing; Nr, Non-redundant; RPM, Reads per million; TSV, Taura syndrome virus; WSSV, White spot syndrome virus; YHV, Yellow head virus.

Nine supplementary tables are available with the online version of this article.

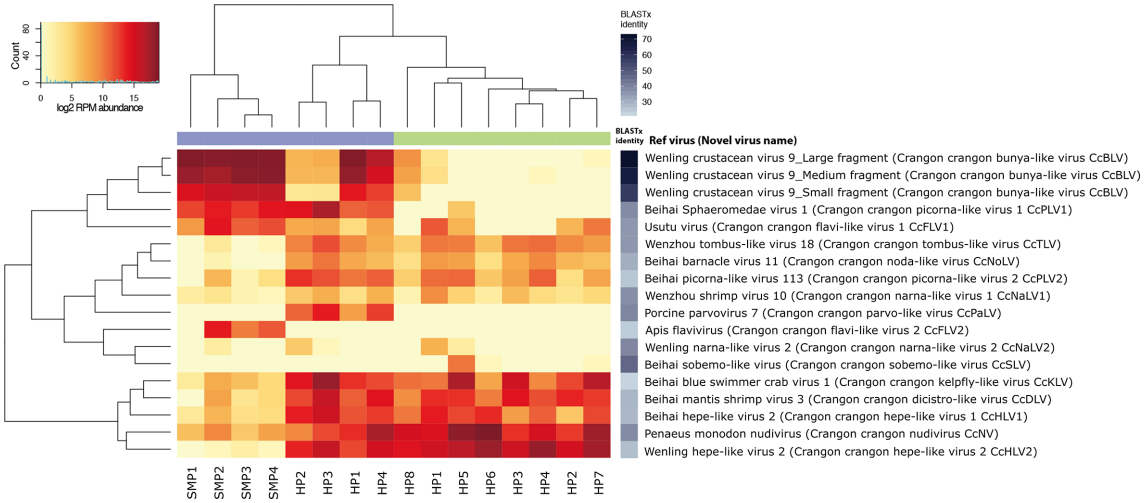

Fig. 1. Heat map of the putative viruses in both cultured (blue) and wild-caught (green) samples. The Ref virus names are from the taxonomic annotation by DIAMOND and KronaTools. The novel virus names are in the bracket after the Ref virus. The hierarchical clustering above shows the clustering per tissue type (Muscle vs hepatopancreas) or per origin (cultured vs wild samples). The hierarchical clustering on the left, shows the clustering in abundance per virus across the different samples. SMP: shrimp muscle tissue. HP: hepatopancreatic tissue.

and matrices, such as soil, sea, ballast water, potable water and activated sludge, and a wide range of animal species [7, 10–13].

Little effort has been given to the study of viromes associated with marine invertebrates important for aquaculture. Only a few studies have reported the metagenomic identification of novel viruses in *P. monodon*, in *Farfantepenaeus duorarum*, in copepods and in bivalves [7, 14–17]. Apart from these economically important species, few studies were published concerning the viromes of ecological important marine invertebrates, such as ghost and volcano shrimps, tunicates, non-scleractinian corals, amphipods and echinoderms [18–22]. However, recent comprehensive surveys have been conducted to discover the genomic diversity of multiple DNA and RNA viruses in arthropods, including marine species [23–25].

In most cases, identification of pathogens is initiated after major events of disease outbreaks and subsequent substantial economic losses. Unfortunately, these late follow-ups have hampered the development of diagnostic tools, research on the understanding of the mechanism of action of disease epidemiology and the development of treatments to control or prevent diseases [15, 26]. The possibility that numerous of these viruses are likely opportunistic pathogens arising from wild hosts has long been neglected. These pathogens seem to be latent in favourable environmental conditions but may become virulent in less favourable circumstances such as in aquaculture settings [26, 27]. Therefore, a proactive characterization of putative viral pathogens is useful to rapidly diagnose valuable aquaculture species, allowing appropriate measures to be taken to avoid further disease outbreaks with major economic losses [11].

The evaluation of *Crangon crangon* as an aquaculture target has been investigated in promising culture attempts [28]. The major challenge remains the high mortality during the grow-out phase. One possible cause of this uncontrollable mortality might be *C. crangon* bacilliform virus (CcBV). This virus is highly prevalent in wild *C. crangon* and it has been shown via histology in previous studies that there is a loss of structural integrity of the hepatopancreas causing possible dysfunction [2, 29, 30]. Besides CcBV, *C. crangon* possibly harbours multiple other viruses, which might become virulent in suboptimal aquaculture conditions [2].

The goal of this study was to characterize the virome of *C. crangon* in both cultured and wild individuals and to describe and highlight potential viral infections. The knowledge gained from this virome study, could enable us to develop detection and prevention (e.g. vaccination) tools in the near future, to accurately identify potential danger of pathogens present in aquaculture settings, especially when wild populations are needed as source for stocking.

METHODS

Sampling *C. crangon*

C. crangon were sampled using a 3 m beam trawler with a mesh size of 11 mm, which was towed by hand along the coastline of Ostend, Belgium (51°14′20.3″N, 2°55′47.5″E) [31]. In May 2016, shrimp were sampled and the animals were kept in natural seawater prior transport to the rearing facility. These shrimp were stocked in rearing tanks for 5 months (9 °C, salinity: 32 p.p.t.). All the tanks were connected to a recirculation system. They were fed daily *ad libitum* with mussels, white fish and/or ragworms. The feed was not treated or

Table 1. Overview of the closest relative of each virus based on NCBI BLAST results of the whole genome. The genome size (in bp) is the length that was found by assembly of the contigs

Virus	Abbr.	NCBI BLAST results	E-value	Identity	Genome size (bp)	Poly A tail
Crangon crangon bunya-like virus	CcBLV	Wenling crustacean virus 9 L	0.0	72%	6692	No
		Wenling crustacean virus 9 M	0.0	65%	3322	
		Wenling crustacean virus 9 s	2e-78	58%	1699	
Crangon crangon picorna-like virus1	CcPLV1	Beihai Sphaeromedae virus 1	0.0	39%	9694	Yes
Crangon crangon dicistro-like virus	CcDLV	Beihai mantis shrimp virus 3	0.0	36%	9044	Yes
Crangon crangon kelpfly-like virus	CcKLV	Beihai blue swimmer crab virus 1	2e-84	27%	12472	Yes
Crangon crangon picorna-like virus 2	CcPLV2	Beihai picorna-like virus 113	2e-67	26%	10852	Yes
Crangon crangon hepe-like virus 1	CcHLV1	Beihai hepe-like virus 2	1e-73	29%	12722	No
Crangon crangon hepe-like virus 2	CcHLV2	Wenlinghepe-like virus 2	1e-165	34%	12055	No
Crangon crangon tombus-like virus	CcTLV	Wenzhou tombus-like virus 18	3e-125	45%	3346	Yes
Crangon crangon narna-like virus 1	CcNaLV1	Wenzhou shrimp virus 10	3e-154	38%	2938	No
Crangon crangon sobemo-like virus	CcSLV	Beihai sobemo-like virus 13	1e-108	46%	3540	No
Crangon crangon flavi-like virus 1	CcFLV1	Usutu virus	0.0	35%	11111	Yes
Crangon crangon flavi-like virus 2	CcFLV2	Apis flavivirus	3e-83	31%	21799	No
Crangon crangon parvo-like virus	CcPaLV	Porcine parvovirus 7	2e-21	41%	4083	No
Crangon crangon narna-like virus 2	CcNaLV2	Wenlingnarna-like virus 2	9e-122	39%	2185	No
Crangon crangon nudivirus	CcNV	Penaeus monodon nudivirus	0.0	36%	119291	No
Crangon crangon noda-like virus	CcNoLV	Beihai Barnacle virus 11	1e-18	31%	1800	No

Abbr., abbreviation; Bp, basepairs.

sterilized. After 5 months, the shrimp were sacrificed for NGS. The hepatopancreases (HP) of 20 shrimp were dissected with sterile equipment and randomly divided in four Eppendorf tubes (four pools with five HP). Next, muscle tissue samples of 20 different shrimp were dissected and randomly divided in four Eppendorf tubes (four pools with five muscles). All the pools were stored at −80 °C prior to analysis.

In February 2018, shrimp were again sampled and transported to the rearing facility. Once in the rearing facility, the hepatopancreases of 40 shrimp were directly dissected and randomly divided in eight Eppendorf tubes (eight pools with five HP). All the pools were stored at −80 °C prior to analysis. None of the shrimp were checked with TEM or histology to confirm infection. Furthermore, none of the sampled shrimp (wild and cultured) showed symptoms of disease.

Isolation of viral particles and sequencing

The NetoVIR protocol was followed to isolate, extract and further sequence viral particles [32]. Sterile PBS (10 % m/v) was added to each pool and the samples were homogenized in a Minilys tissue homogenizer (3000 r.p.m. for 60′) together with beads. The samples were then centrifugated at 13000 ***g*** for 3′. The obtained supernatant was then filtered through a PES filter (0.8 µm) at 13000 ***g*** for 1′.

Next, a nuclease treatment was performed to remove host DNA and other free-floating nucleic acids by adding 7 µl of 20× homemade buffer (1 M Tris, 100 mM $CaCl_2$ and 30 mM $MgCl_2$) to 130 µl of sample filtrate. Also, 2 µl of benzonase and 1 µl microccocal nuclease was added and the mixture was gently inverted three times before incubation for 2 h at 37 °C. The reaction was stopped by adding 7 µl of 10 nM EDTA.

Lastly, DNA and RNA was extracted using the QIAamp Viral RNA Mini kit (Qiagen), reverse transcribed and randomly amplified using a modified WTA2 or Complete Whole Transcriptome Amplification kit protocol. WTA2 products were purified, and the libraries were prepared for Illumina sequencing using the Nextera XT DNA sample preparation kit (Illumina). After library synthesis a cleanup was performed using a 1.8 ratio of Agencourt AMPure XP beads (Beckman Coulter). Samples were pooled at equimolar rations, and stored immediately at −80 °C to avoid primer hopping. Sequencing was performed on a NextSeq platform (Illumina) for 300 cycles (2×150 bp paired ends).

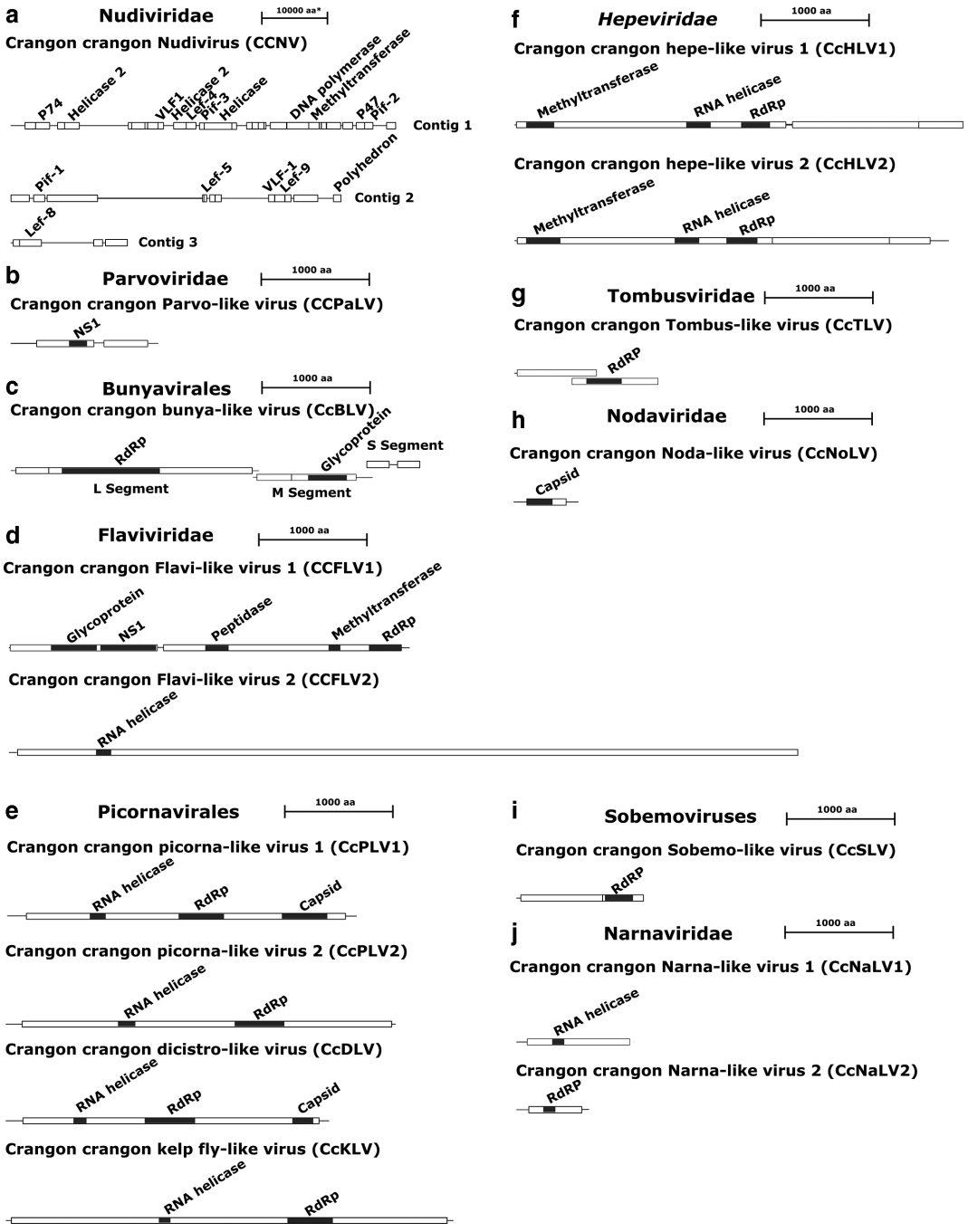

Fig. 2. Genome structures of novel putative viruses. All the genomes or contigs were analysed with ORF Finder (NCBI) and putative gene motifs were analysed with HMMER. The most important gene motifs are shown in the figure. (a) Crangon crangon nudivirus (CcNV). The length of the contigs was: 55020, 32659, 16687 bp, respectively. (b) Putative genome structure of Crangon crangon parvo-like virus (CcPaLV). (c) Putative genome structure of Crangon crangon bunya-like virus (CcBLV). The genome consists of three segments: L segment, M segment and S segment. (d) Putative genome structures of Crangon crangon flavi-like virus 1 and 2 (CcFLV1 and 2). (e) Putative genome structures of Crangon crangon picorna-like virus 1 and 2 (CcPLV1 and CcPLV2), Crangon crangon dicistro-like virus (CcDLV) and Crangon crangon kelp fly-like virus (CcKLV). (f) Putative genome structures of Crangon crangon hepe-like virus 1 and 2 (CcHLV1 and 2). (g) Putative genome structures of Crangon crangon tombus-like virus (CcTLV). (h) Putative genome structures of Crangon crangon noda-like virus (CcNoLV). (i) Putative genome structures of Crangon crangon sobemo-like virus (CcSLV). (j) Putative genome structures of Crangon crangon narna-like virus 1 and –2 (CcNaLV1 and 2).

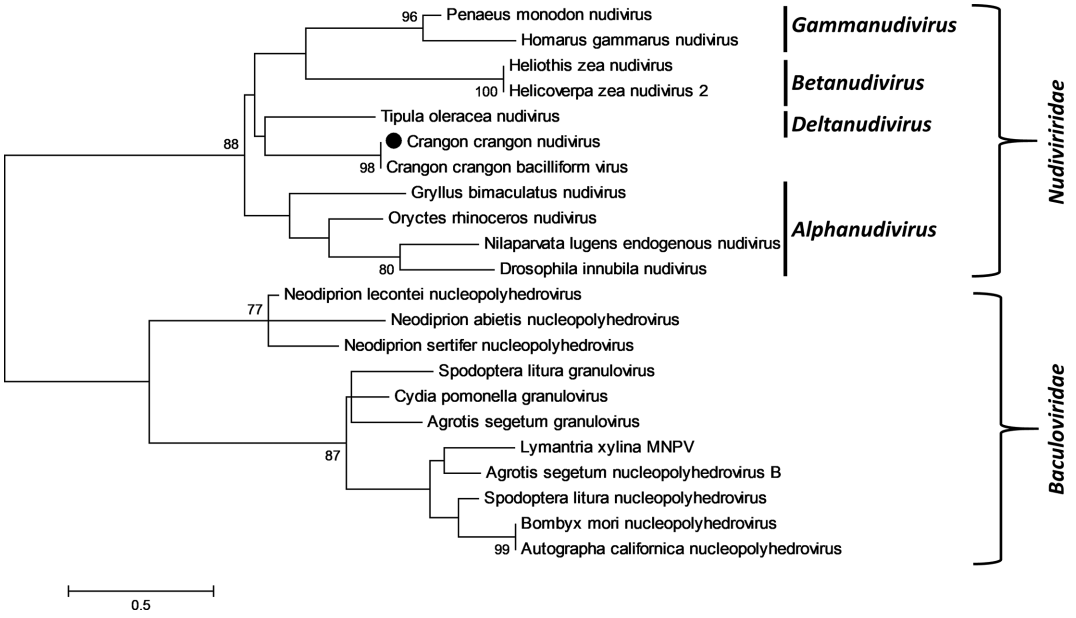

Fig. 3. *Baculo-* and *Nudiviridae* phylogeny obtained from maximum-likelihood analysis with MEGA7 software, based on *lef-8* aa sequences of selected dsDNA viruses. The robustness of the tree was tested using bootstrap (1000) analysis and the percent values are indicated at the nodes. In the case of *Nudiviridae*, viruses of four genera were used to generate the tree. GenBank accession numbers from these viral *Lef-8* aa sequences are listed in Table S1.

Genomic and phylogenetic analysis

Raw pair-end reads were trimmed for quality and adapters using Trimmomatic and were *de novo* assembled into contigs using SPAdes. Contigs were then annotated using DIAMOND BLAST with the sensitive option using non-redundant (nr) database of GenBank [33–37]. The data was visualized using Krona Tools [38]. ORFs were identified with ORFfinder (NCBI) and the translated aa sequences were further analysed with HMMER (EMBL-EBI) to identify conserved motifs. Next, phylogenetic trees were generated for each novel virus based on the nucleotide sequences of the *RdRp*, *Lef-8* or *NS1* gene. The novel sequences were aligned using Muscle in MEGA7, with relevant reference sequences. Then, phylogenetic trees were generated using the maximum-likelihood method [25]. Sequences used in the phylogenetic analysis were representatives of the different virus families and their accession numbers are listed in Tables S1–S9 (available in the online version of this article).

Heat map

The magnitudes (number of reads) of each virus in each sample were obtained by mapping the raw reads of each sample to the novel viral genome contigs using BBMap [39]. Next, the mapped read numbers were normalized for the reads per million or RPM. Here, the total number of reads in a sample was divided by 1000000 resulting in a 'per million' scaling factor. Then, to get the RPM counts, the mapped read numbers were divided by the 'per million' scaling factor. The heat map in Fig. 1 shows the normalized RPM read counts on log2 scale using the metagenomeSeq [40]. The hierarchical clustering is based on the Euclidean distance matrix calculated from the normalized read counts.

RESULTS

Sixteen highly divergent putative novel viral genomes identified belonging to various viral families

Based on the DIAMOND BLAST results, contigs were assigned to particular virus species, genera or families. The putative viral genomes were confirmed on NCBI with the BLASTX tool to determine the closest relative. The closest relative for each of the putative genomes found in both reared shrimp samples and wild-caught shrimp samples is shown in Table 1. Furthermore, the putative genome length is described and the presence or absence of a poly-A tail.

In total, 15 viruses were identified to be nearly complete when compared to the expected genome length of related viruses. For one virus, Crangon crangon nudivirus, three large contigs (55020, 32659, 16687 bp) were found in the samples, but it was not possible to obtain the complete genome of this virus, most likely due to repeated elements in the genome.

Several nearly identical viruses were identified in several or most of the investigated pools. In the case of Crangon crangon kelpfly-like virus (CcKLV), Crangon crangon picorna-like virus 2 (CcPLV2), Crangon crangon hepe-like virus 1 (CcHLV1) and Crangon crangon hepe-like virus 2 (CcHLV1), the genome sequences found in both reared and wild-caught

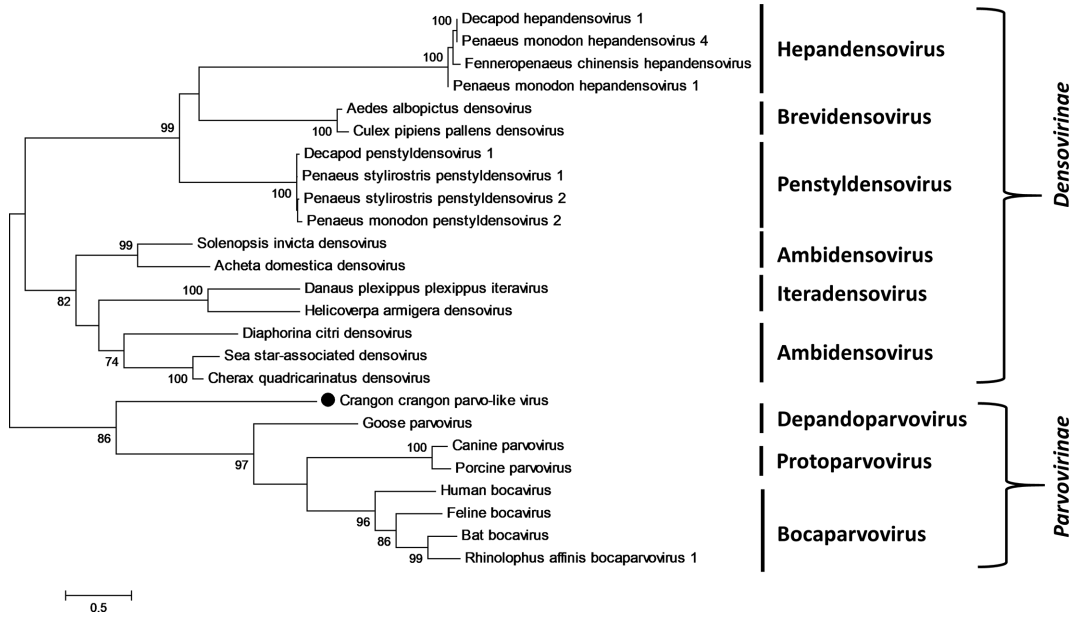

Fig. 4. *Parvoviridae* phylogeny obtained from maximum-likelihood analysis with MEGA7 software, based on the *NS1* aa sequences of selected *Parvoviridae* viruses. The viruses clustered into two subfamilies, namely the *Densovirinae* (above) and the *Parvovirinae* (under). Crangon crangon parvo-like virus formed an outgroup between *Densovirinae* and *Parvovirinae*. The robustness of the tree was tested using bootstrap (1000) analysis and the percent values are indicated at the nodes. GenBank accession numbers from these viral aa sequences are shown in Table S2.

samples were aligned with MEGA7 and the pairwise distance was determined with the p-distance method. The genomes found for CcKLV (99.3%), CcPLV2 (96.3%), CcHLV1 (97.7%) and CcHLV2 (95.7%) in the reared and wild-caught samples were very similar on the nucleotide level. Therefore, the largest contig per virus was selected and treated as the representative of that novel virus.

Multiple novel viruses are highly prevalent in multiple of the analysed *C. crangon* pools, but show distinct tissue tropism

A heat map was generated based on the obtained viral genomes to visualize the number of reads of each of the identified viruses per pool as shown in Fig. 1. The viral species names shown in the heat map are from the taxonomic annotation by DIAMOND and the novel virus names are in the bracket after the annotated name. The aa identity of the ORF(s) in the contig of each viral species with the highest BLASTX score to a reference sequence is shown in the shaded blue boxes.

The hierarchical clustering based on the Euclidean distance matrix (Fig. 1) showed that samples clustered per tissue type (muscle vs hepatopancreas), rather than origin (cultured vs wild samples). In particular, CcBLV, CcPLV1, CcFLV1 and CcFLV2 are very abundant in the muscle samples, whereas CcPaLV, CcSLV and CcNoLV are almost completely absent in the muscle samples. On the other hand, several virus are highly (CcKLV, CcDLV, CcHLV1, CcNV, CcHLV2) or medium (CcTLV, CcNoLV, CcPLV2, CcNaLV1) abundant in nearly all of the hepatopancreas samples, irrespective of the origin. The abundance of these viruses is much lower in the respective muscle samples. Furthermore, the abundance of some viruses such as CcBLV, CcPLV1 and CcPaLV has increased in the hepatopancreas of cultured shrimp samples compared to the wild-caught shrimp samples. Viruses such as CcNaLV2 and CcSLV are only present with lower abundances in a limited number of pools.

During the sampling of the wild shrimp no muscle tissue was analysed. One reason was that the hepatopancreas is in direct contact with the environment and the feed. When these wild shrimps are then stored under aquaculture conditions they can easily spread the viruses through the faeces and the water to the rest of the system. Furthermore, we were mainly interested if the virome would shift between wild and cultured shrimp.

ORF description and phylogenetic analysis

Based on the putative genomes, the ORFs were determined with the help of NCBI ORF Finder. Next, HMMSCAN was used to determine conserved motifs and assign putative functions to the identified ORFs. The putative genome structures are shown in Fig. 2. The genome structures of the (nearly) complete viral genomes are further discussed in the sections below. Furthermore, phylogenetic trees were generated for each novel viral genome based on the *RdRp*, *Lef-8* or *NS1* gene. The accession numbers of the used sequences are listed in Tables S1–S9.

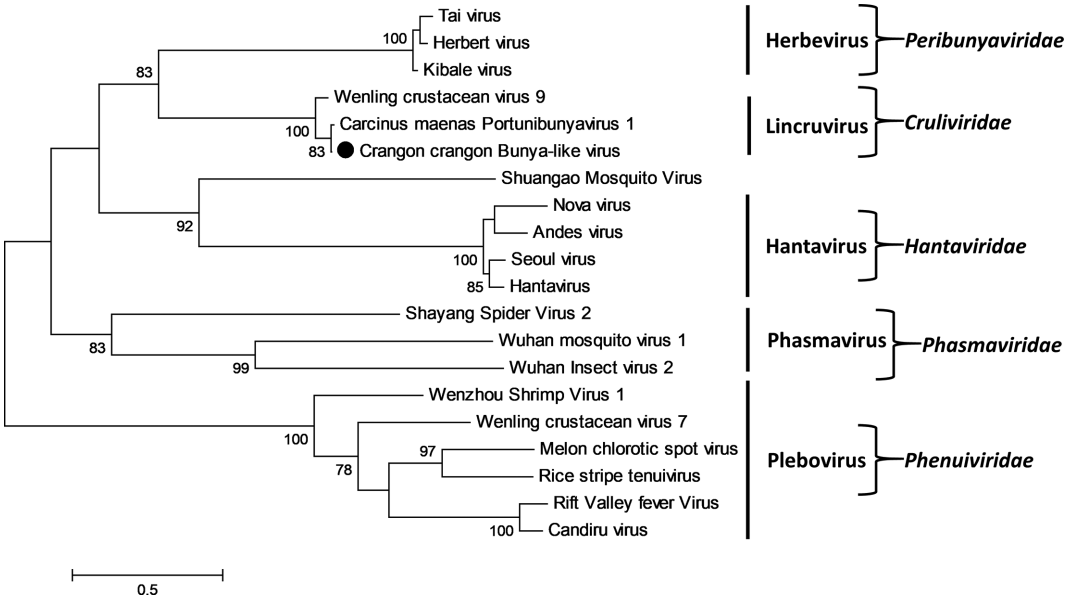

Fig. 5. *Bunyavirales* phylogeny obtained from maximum-likelihood analysis with MEGA7 software, based on the *RdRp* aa sequences of selected *Bunyavirales* viruses. The viruses clustered into five families, namely the *Hantaviridae, Peribunyaviridae, Cruliviridae, Phasmaviridae* and *Phenuiviridae*. Crangon crangon bunya-like virus resides withing the *Cruliviridae* and is probably a member of the genus *Lincruvirus* together with two other crustacean viruses. The robustness of the tree was tested using bootstrap (1000) analysis and the percent values are indicated at the nodes. GenBank accession numbers from these viral aa sequences are shown in Table S3.

dsDNA viruses

Nudiviridae

Based on the genomic information obtained during this study, we found a virus (CcNV) related to members of the *Nudiviridae*. Our analyses resulted in three contigs related to known nudiviruses and their identified ORF structure is shown in Fig. 2a. Unfortunately, it was not possible to connect the three contigs into a single complete genome, probably due to repeat regions in the genomes, which could not be crossed with (short) Illumina reads. A phylogenetic tree was constructed based on the *lef-8* gene of the known members of the *Nudiviridae* (Fig. 3). Furthermore, the *lef-8* gene (CcBV) provided by K. Bateman was also added in the phylogenetic analysis (Fig. 3). Based on the phylogenetic tree, it is clear that CcNV and CcBV are (nearly) identical and therefore cluster together. Together with Tipula oleracea nudivirus, they form an outgroup inside the *Nudiviridae* family and potentially represent the founding member of a novel genus, namely the *Deltanudivirus*.

ssDNA viruses

Parvoviridae

The genome of Crangon crangon parvo-like virus (CcPaLV) is nearly complete with 4083 bp, but highly divergent from known members of the *Parvovirinae* (Table 1). The genome structure consists of a non-structural ORF (*NS1* and *NS2*) and a structural ORF (*VP1*, *VP2* and *VP3*). However, for CcPaLV, only the *NS1* motif could be identified in the first ORF as shown in Fig. 2b. A phylogenetic tree was generated using the putative *NS1* gene and is shown in Fig. 4. Based on the phylogenetic tree, two subfamilies can be distinguished, e.g. *Parvovirinae* and *Densovirinae*. In the latter subfamily, most of the members are known to infect insects or crustaceans, while the members of the *Parvovirinae* are known to infect vertebrates. The novel virus CcPaLV clusters between both of these families, although it is slightly more related to the subfamily *Parvovirinae*. It remains to be seen by taxonomists, whether this novel virus represent a novel subfamily in the *Parvoviridae* or rather a very distinct genus in the subfamily *Parvovirinae*.

Negative single-stranded RNA viruses

Bunyavirales

The order of *Bunyavirales* contains ten families known to infect a large range of species (humans, mammals, plants and invertebrates) [41]. Most viruses of this order have a tripartite RNA genome, possessing large (L – 6.8 to 12 kb), medium (M – 3.2 to 4.9 kb) and small (S – 1 to 3 kb) segments [42–45]. In this study, the size of the near-complete segments of CcBLV, was 6516, 2745 and 1699 nt, respectively (Fig. 2c), with aa similarities ranging from 58 to 72% with Wenling crustacean virus 9 (Table 1). The L segment contained an ORF with a putative *RdRp* gene motif, while the M segment contained an ORF with a putative glycoprotein gene function. The S segment had two ORFs, but HMMER analysis did not result in the

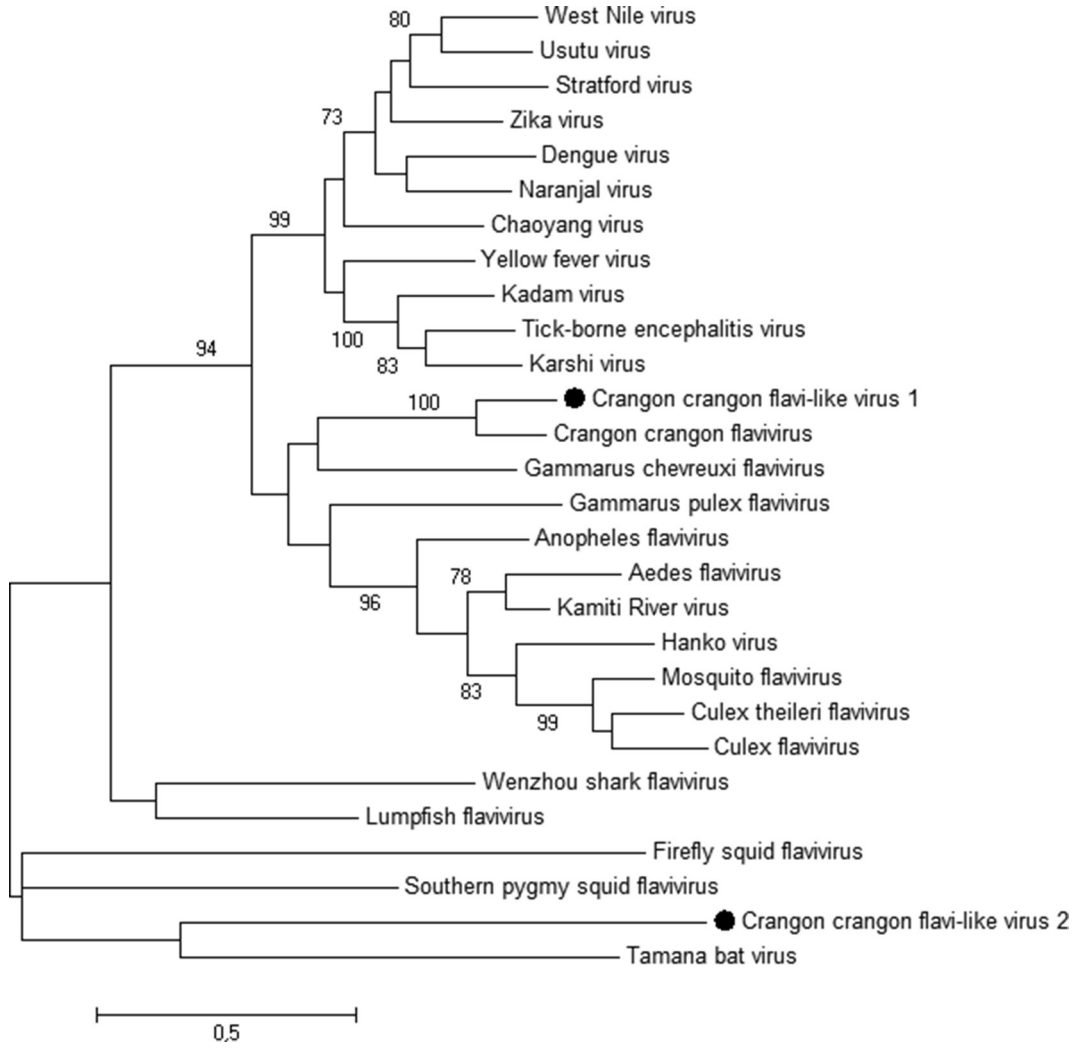

Fig. 6. *Flaviviridae* phylogeny obtained from from maximum-likelihood analysis with MEGA7 software, based on *RdRp* aa sequences of selected *Flaviviridae* viruses. The robustness of the tree was tested using bootstrap (1000) analysis and the percent values are indicated at the nodes. GenBank accession numbers are listed in Table S4.

characterization of a putative gene function. A phylogenetic tree was constructed with representatives of five families (e.g. *Peribunya-, Cruli-, Hanta-, Phasma-* and *Phenuiviridae*), showing that Carcinus maenas portunibunyavirus 1, belonging to *Cruliviridae* was the closest relative of CcBLV (Fig. 5).

Positive single-stranded RNA viruses

Flaviviridae

Flaviviridae is a family that is comprised of four genera: *flavivirus, hepacivirus, pestivirus* and a newly proposed *pegivirus*. Their genomes are generally unsegmented (+)ssRNAs of 9.6 to 12.3 kb in length [46]. To our knowledge, a small number of viruses belonging to the *Flaviviridae,* are described in crustaceans, except for the ones described by Shi *et al.* [46] and more recently by Parry and Asgari. The flavi-like viruses described in those studies still need further characterization in order to classify them in the correct family and in the correct genus [47]. The two viruses found during this study were approximately 11 and 22 kb (CcFLV1 and CcFLV2, respectively) long. The latter being much larger than other viruses within the *Flaviviridae*. CcFLV1 possessed two ORFs as shown in Fig. 2d. In the first ORF, putative motifs were found of a glycoprotein and the *NS1* gene, while in the second ORF putative motifs were found of peptidase, methyltransferase and RdRp. In the case of CcFLV2, only one ORF was found with RNA helicase and RdRp motifs (Fig. 2d). Phylogenetically, CcFLV1 clustered together with the recently discovered Crangon crangon flavivirus and Gammarus chevreuxi flavivirus [47]. These viruses are distantly related to a group of insect-only infecting viruses, (e.g. aedes flavivirus, mosquito flavivirus, culex theileri flavivirus and culex flavivirus), inside the *flavivirus* genus. CcFLV2, on the other hand forms a distant outgroup and is

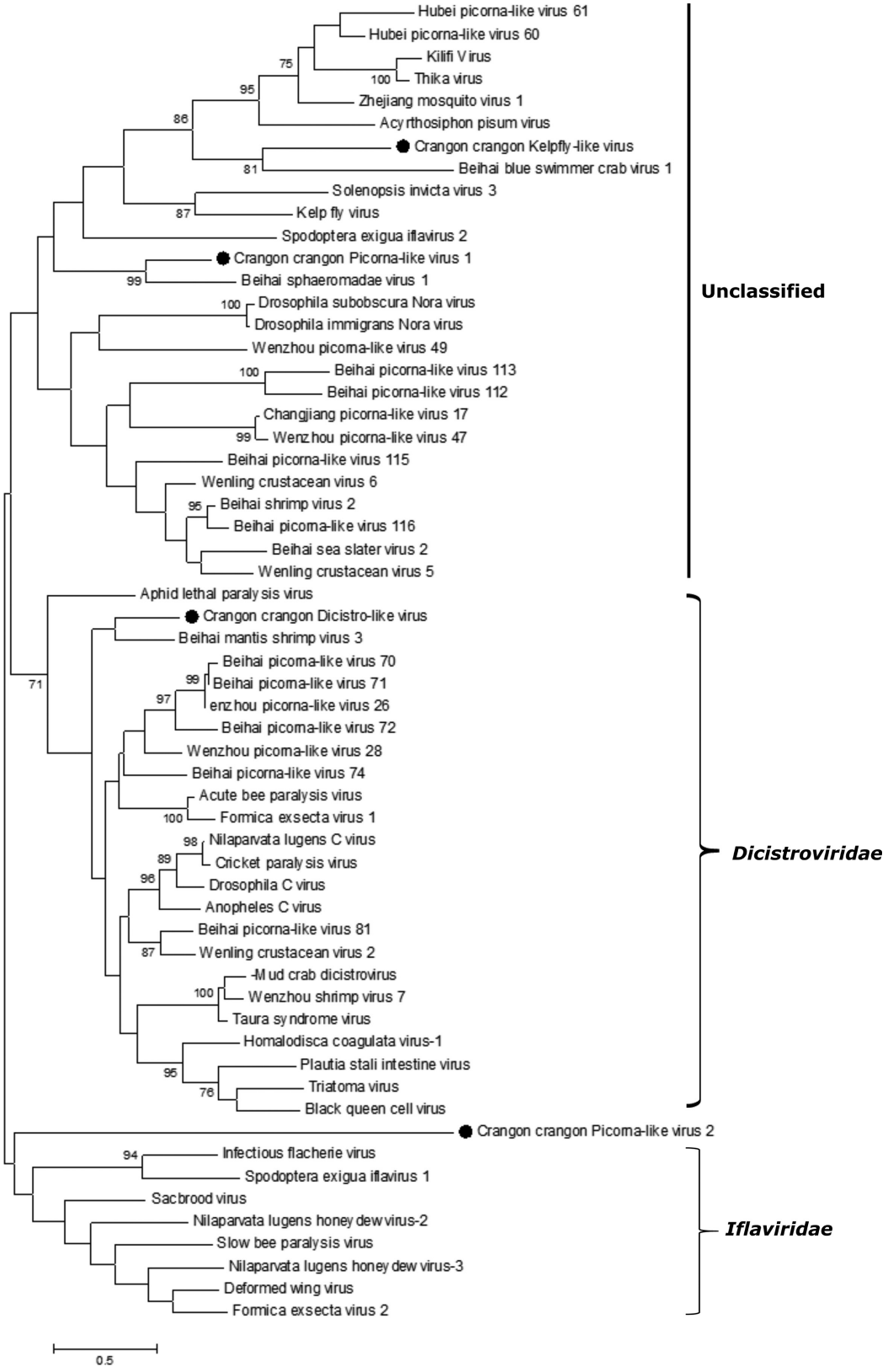

Fig. 7. *Picornavirales* phylogeny obtained from maximum-likelihood analysis with MEGA7 software, based on *RdRp* aa sequences of selected *Picornavirales* viruses. The robustness of the tree was tested using bootstrap (1000) analysis and the percent values are indicated at the nodes. Crangon crangon dicistro-like virus (CcDLV) clustered within the *Dicistroviridae*, while the other three novel viruses (CcPLV1, CcPLV2, CcKLV) clustered within unclassified groups. GenBank accession numbers are listed in Table S5.

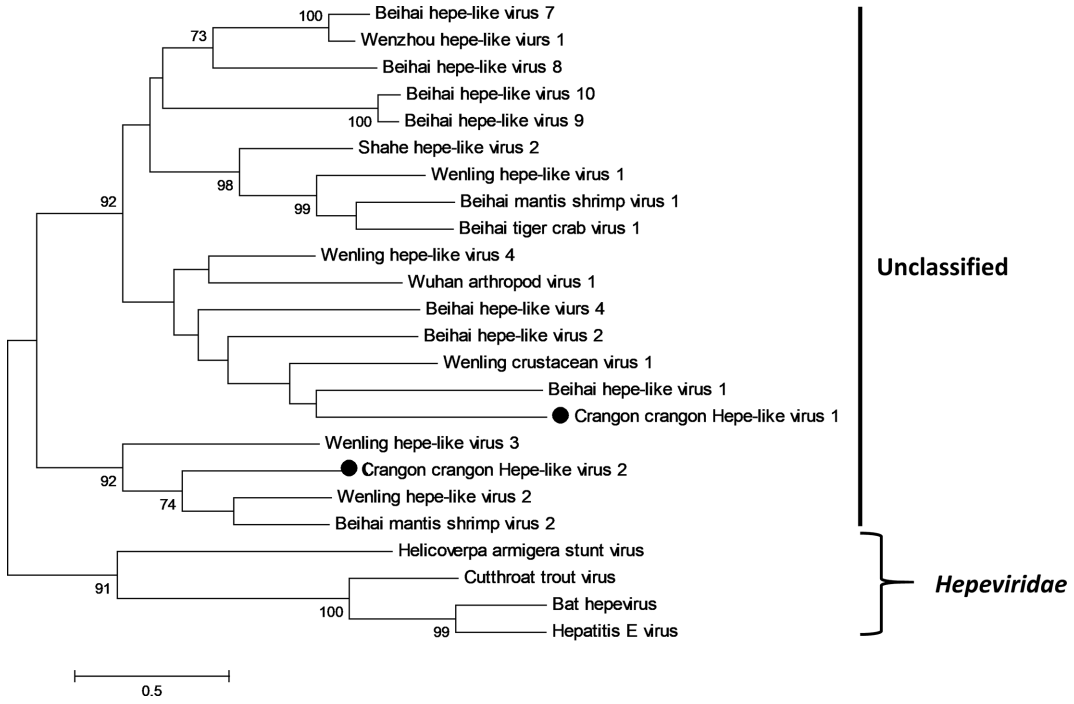

Fig. 8. *Hepeviridae* phylogeny obtained from maximum-likelihood analysis with MEGA7 software, based on RdRp aa sequences of selected *Hepeviridae* viruses. The robustness of the tree was tested using bootstrap (1000) analysis and the percent values are indicated at the nodes. GenBank accession numbers from these viral aa sequences are shown in Table S6.

most closely related to the tamana bat, and might represent a novel genus inside the *Flaviviridae* (Fig. 6).

Picornavirales

Four shrimp viruses were found to be distantly related (26–39% aa similarity, Table 1) to members of the order *Picornavirales*. The genome lengths of all four viruses are within the range of other viruses residing within the *Picornavirales* [8.5–10.2 kb; (+)ssRNA] (Fig. 2e) [48]. For CcPLV1 and CcDLV, the genome organization consisted of one ORF with motifs for an RNA helicase, RdRp and capsid proteins. In the case of CcPLV2 and CcKLV, only the *RdRp* and *RNA helicase* motifs were identified. For the phylogenetic tree (Fig. 7), only members of the *Dicistroviridae*, *Iflaviridae* and related unclassified viruses were used as reference. Phylogenetically, CcDLV was most closely related to Beihai mantis shrimp 3, and clustered together with other members of the *Dicistroviridae* family. CcKLV, on the other hand was found to reside within the kelpfly-virus related group, while CcPLV1 formed an outgroup together with Beihai sphaeromadae virus 1. CcPLV2 formed an even more distant outgroup, likely to represent a novel viral family.

Hepeviridae

Two shrimp viruses, found in both reared and wild-caught shrimp samples and referred to as Crangon crangon hepe-like virus 1 (CcHLV1) and Crangon crangon hepe-like virus 2 (CcHLV2) are distantly related to members of the *Hepeviridae* family. Their partial genomes were bigger than other traditional hepe-like viruses (6.4–7.2 kb) with 12722 nt and 12222 nt, respectively (Fig. 2f). Both genomes consisted of three ORFs with putative motifs of the methyltransferase gene, RNA helicase and RdRp. Phylogenetically, CcHLV1 and CcHLV2 were most closely related to Beihai hepe-like virus 1 and Wenling hepe-like virus 2/Beihai mantis shrimp virus 2, respectively (both lower than 35% aa similarity, Table 1) as shown in Fig. 8. They fall into a large and diverse group of novel unclassified viruses detected from different crustacean species [25].

Tombus-, *Noda-* and *Narnaviridae* and *Sobemoviruses*

Tombusviridae, is a family of viruses mainly infecting plants [49]. To our knowledge, there are no reports of crustacean viruses, which were assigned to this family, except for the ones described by Shi *et al.* [25]. Most of these viruses are still unclassified. One shrimp virus, found in wild-caught shrimp samples and referred to as Crangon crangon tombus-like virus (CcTLV) is distantly related to members of the *Tombusviridae* family. Only an *RdRp* motif was found in the genome of CcTLV, which consisted of two ORFs (Fig. 2g) [25, 50]. Phylogenetically, CcTLV clusters within the *Tombus*-clade (Fig. 9), which to date is still unclassified.

The genome of *Nodaviridae* is bipartite with RNA 1 being 3.1 kb and RNA 2 1.4 kb. RNA 1 encodes the RdRp, while RNA 2 encodes the capsid protein [51]. In this study, one noda-like virus segment was found (CcNoLV), but the contig was only 1800 bp (Fig. 2h). This might indicate that only the RNA 2

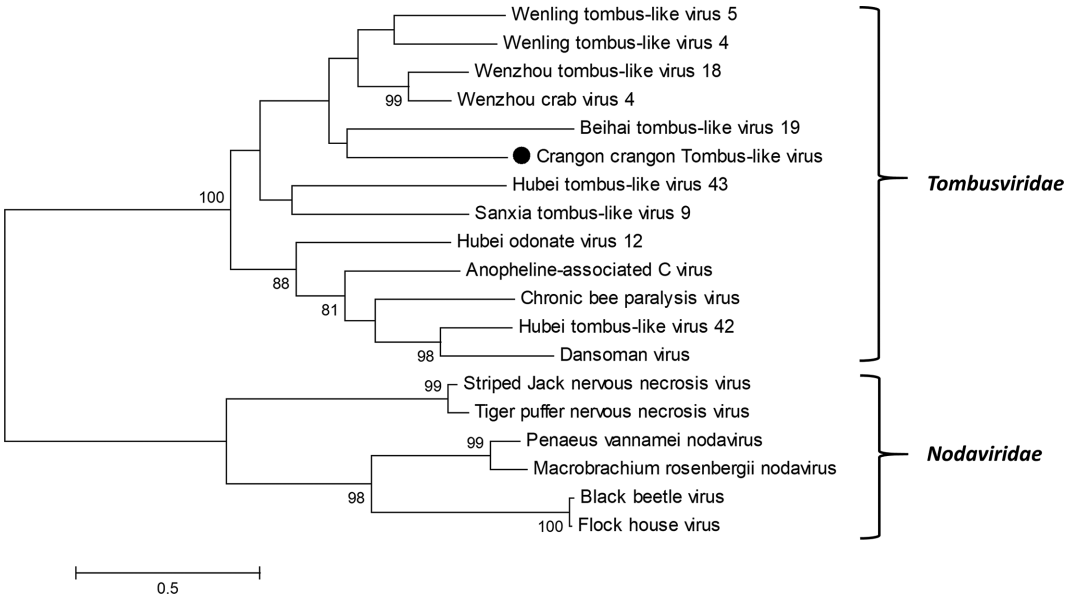

Fig. 9. *Tombus- and Nodaviridae* phylogeny obtained from maximum-likelihood analysis with MEGA7 software, based on *RdRp* aa sequences of selected viruses belonging to *Tombus- or Nodaviridae*. These viruses clustered into two groups: unclassified and *Nodaviridae*. Crangon crangon tombus-like virus clustered within the unclassified group. The robustness of the tree was tested using bootstrap (1000) analysis and the percent values are indicated at the nodes. GenBank accession numbers from these viral aa sequences are shown in Table S7.

fragment was found. Because the lack of sufficient genomic information, no phylogenetic analysis was performed.

Until now, the genus *sobemovirus*, remained unassigned to any family. Several of the *sobemoviruses* are important pathogens and mainly infect plant species [52]. One shrimp virus, referred to as Crangon crangon sobemo-like virus (CcSLV), is related to the *Sobemoviruses*. Two ORFs were found containing only one putative gene motif, e.g. *RdRp* (Fig. 2i). Phylogenetically, CcSLV clusters within a large cluster of sobemo-like viruses detected in different crustacean species (Fig. 10).

Narnaviridae are fungi-infecting viruses [53]. The two Crangon crangon narna-like virus contigs (CcNaLV1 and CcNaLV2) found in this study had a length of 2938 and 2185 nt, respectively. Furthermore, only one putative gene motif was found in the ORF of CcNaLV1, namely *RNA helicase* (Fig. 2j). In the case of CcNaLV2, the RdRp gene motif was found within the only ORF. In Fig. 11, a phylogenetic tree is shown of the *Narnaviridae*. Because most of the viruses residing within the *Narnaviridae* are still unclassified, more research is needed to further subdivide these viruses into different genera.

DISCUSSION

NGS has played an important role in virus discovery. It enabled researchers to explore whole viromes in, on and around a multitude of species. Since the discovery of a bacilliform virus in *C. crangon* named CcBV, this study is the first to identify other viruses in this species [54]. To our knowledge, this is also the first study exploring the viral community of *C. crangon* with the aid of high-throughput sequencing. Our results present genomic evidence on 16 new viruses, including 13 positive single-strand RNA viruses (+ssRNA), one negative single strand RNA virus (-ssRNA), one single-stranded DNA viruses (ssDNA) and one double-stranded DNA virus (dsDNA).

Tissue tropism

During our study, we first explored the virome of shrimp kept in captivity for 5 months. Here, it was decided to determine the virome of both the hepatopancreatic tissue and the muscle tissue. A multitude of viruses was discovered, but based on the heat map, there was a clear difference in viral composition between the two tissues. The difference might be explained by the fact that the hepatopancreas is exposed directly to the environment through feeding and thus can get infected more rapidly than muscle tissue [55]. Viruses also have different attachment proteins, thereby defining their tissue tropism. Our data suggest that some viruses have a tissue tropism for muscle tissue. These viruses belonged to the *Bunyavirales* and *Flaviridae*, namely Crangon crangon bunya-like virus and Crangon crangon flavi-like virus 1 and 2. It seems that these three viruses mainly infect tissues of mesodermal and ectodermal origin. This is similar to what is known about related viruses. In the case of viruses residing within the *Flaviviridae*, mostly epidermal cells and neural cells are infected, while members of *Bunyavirales* mainly infect cells of mesodermal origin, e.g. hemocytes [2, 50, 56, 57]. Crangon crangon picorna-like virus 1 was present in both tissue types, while

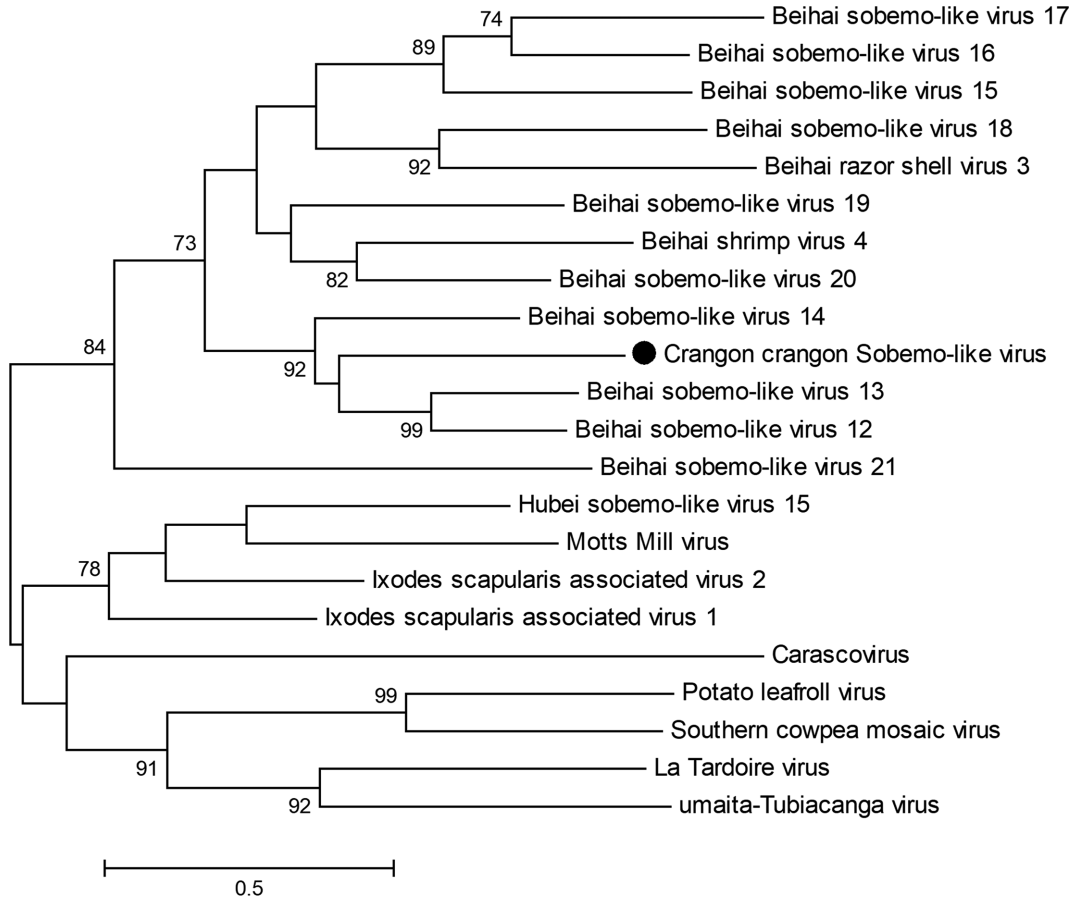

Fig. 10. *Sobemoviruses* phylogeny obtained from maximum-likelihood analysis with MEGA7 software, based on *RdRp* aa sequences of selected *sobemoviruses*. The robustness of the tree was tested using bootstrap (1000) analysis and the percent values are indicated at the nodes. GenBank accession numbers from these viral aa sequences are shown in Table S8.

the other viruses were mainly found in the hepatopancreas of the cultured animals. These viruses belong to *Picornavirales, Parvoviridae, Hepeviridae, Nudiviridae Tombusviridae, Nodaviridae* or *Narnaviridae*. In the case of members of the *Dicistroviridae*, the main target tissue is of mesodermal origin and the viruses mainly reside in the cytoplasm of the connective tissues of the hepatopancreas, gills and intestines [2]. In the case of the *Nodaviridae*, it is surprising to find CcNoLV mainly in the hepatopancreas samples, because the main tissue for infection of related viruses is the tail muscle. Viruses of this family are known to cause white tail disease in, for example, *Macrobrachium rosenbergii* and *P. vannamei* [2, 58, 59]. Members of the *Densovirinae* subfamily (family *Parvoviridae*) are known to mainly infect tissues of mesodermal and ectodermal origin, but in some cases they are also able to infect endodermal tissues [2]. Based on the results in this study, it seems that CcPaLV mainly infects the hepatopancreatic tissue. Members of the *Nudiviridae* mainly infect hepatopancreatic tissue, while tissue tropism of *Hepeviridae* in invertebrates is not described [2]. Further research including more tissue types might refine the tissue tropism of these novel viruses.

Differences in origin

A plethora of novel viruses were found in the first NGS run with *C. crangon* individuals reared for 5 months in captivity. Therefore, we thought it would be interesting to see whether these viruses were also present in wild-caught shrimp and if there would be differences in abundance of some or all viruses. Based on the heat map, it is clear that samples clustered mainly per tissue type (muscle vs hepatopancreas), rather than per origin (cultured vs wild). But besides this observation, some viruses seemed to be more abundant than others.

Firstly, CcNaLV2 and CcSLV were only present in one or a few of the samples in relative low abundances. The main reason could be that these viruses do not infect crustaceans and that they were present in the feed of *C. crangon* or in the environment, especially since the feed was not treated or disinfected. To determine whether these two viruses really infect *C. crangon*, or might be derived from the feed should be further investigated, especially since they are mainly known to infect plant species and fungi. Similarly, Tombus(-like) viruses are only known to infect

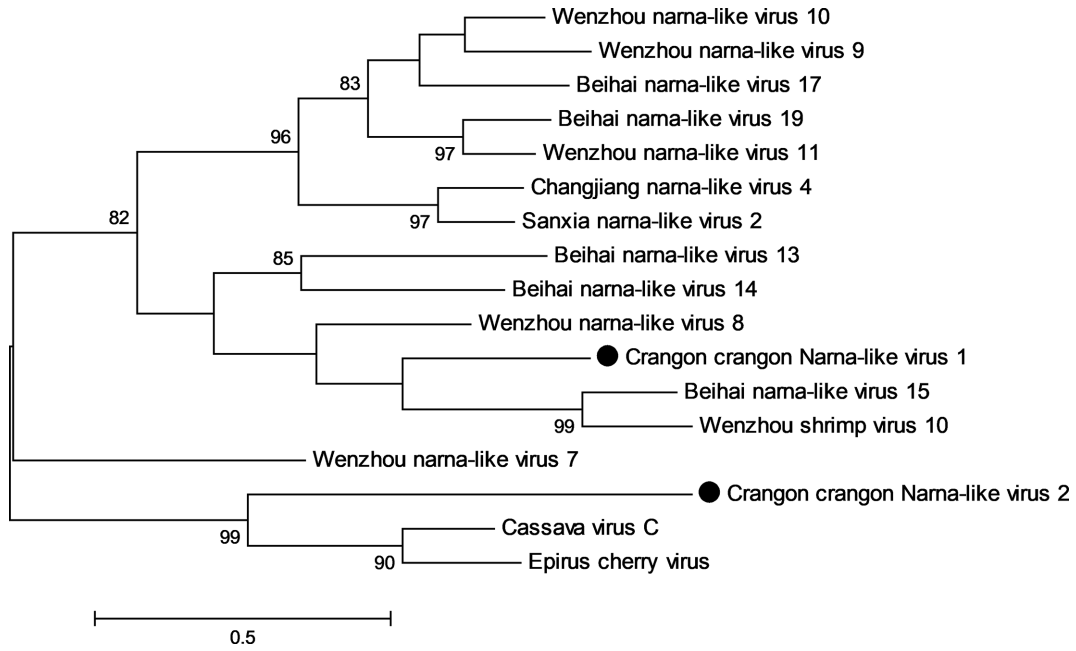

Fig. 11. *Narnaviridae* phylogeny obtained from maximum-likelihood analysis with MEGA7 software, based on *RdRp* aa sequences of selected *Narnaviridae* viruses. The robustness of the tree was tested using bootstrap (1000) analysis and the percent values are indicated at the nodes. GenBank accession numbers from these viral aa sequences are shown in Table S9.

plants, making it unlikely that *C. crangon* was the true host. Furthermore, no descriptions are readily available of crustacean viruses belonging to these three families, questioning their true host species. In the case of CcSLV, it is possible that it infects prey organisms of *C. crangon*, while for CcNaLV2 it is possible that it infects fungi that are present in *C. crangon*. Secondly, most of the viruses were present in high (CcKLV, CcDLV, CcHLV1, CcNV and CcHLV2) or medium (CcNoLV, CcPLV2 and CcNaLV1) abundance in most of the hepatopancreatic tissue pools. This might indicate that these viruses are true viruses of *C. crangon* and form some kind of 'core virome', as was recently observed for mosquitoes [13]. A third observation was the increase in abundance of three viruses, namely CcBLV, CcPLV1 and CcPaLV in the hepatopancreatic tissue of cultured shrimp, compared to wild shrimp. A possible explanation can be that laboratory conditions and anthropogenic stressors can affect the immune system of the shrimp, thereby enhancing or inhibiting viral replication. So it is possible that a virus, which is present in undetectable titers in nature, becomes detectable in culture conditions because of enhanced replication. Stress and high density in aquaculture can result in shedding of high quantities of a particular virus into the environment. This can then lead to a change in viral composition of the environment and subsequently of the reared animal [9]. Thus, it is important to further investigate the changes in viral load between the transfer from nature to tanks, including the possible risks on disease proliferation, growth retardation and/or mortality.

Possible causes for high mortality in aquaculture settings

Until now, there is still little information available concerning the viromes of crustacean species. The reason is the reactive approach, meaning that only after a serious viral outbreak one starts searching for the causative pathogen [15]. Furthermore, universal genetic markers for viruses are lacking, hereby impeding the development of nucleic-acid-based detection tools for surveillance of unknown, but potentially dangerous, viral pathogens [7]. The arrival of high-throughput sequencing, however, made it possible to determine the virome of several species and different environments. Furthermore, recent metagenomic studies in invertebrates revealed distinct virome profiles both within and between host phyla [7, 18]. In addition, Cornejo-Granados *et al.* [60] found that the microbiome of *P. vannamei* differed significantly between wild and cultured individuals and between healthy and diseased individuals. Aiming to develop *C. crangon* aquaculture systems with pathogen-free broodstock, we decided to explore the virome of both cultured and wild *C. crangon* with NGS. This research revealed the presence of multiple viruses, belonging to different viral families, in both wild and cultured *C. crangon* stocks. It was also clear that the novel viruses clustered per tissue type (hepatopancreas vs muscle), rather than origin (wild vs cultured). It would be interesting though to determine the virome of other tissues as well, like nervous tissue, gut, gills and ovarian tissue, for example. Furthermore, to study the tissue tropism in more quantitative detail, it would be of interest to develop quantitative real time PCR assays for each novel virus, to screen different tissues,

or use FISH on different tissue sections. Combining different (molecular) techniques, might speed up the full characterization of each of these novel viruses. In the future, this information will be important for developing disease surveillance strategies.

Furthermore, it will be important to determine which viruses are potentially dangerous, especially under sub-optimal rearing conditions [61]. Based on the phylogenetic trees and relatedness to already known crustacean viruses, we might already have an idea of which novel viruses might pose a threat to *C. crangon* aquaculture. For example, Crangon crangon dicistro-like virus or CcDLV resides within the same family as taura syndrome virus (TSV) and mud crab dicistrovirus. In particular, TSV is known to cause sudden and massive mortalities in all life stages of some crustacean species [62]. Furthermore, affected shrimp can have empty guts, are lethargic and present a pink to reddish colour, especially in the tail fan [62]. So, it should be investigated whether these symptoms can be found in *C. crangon* as well, especially when these picorna-like viruses would be present in high concentrations.

A second potentially dangerous novel virus, could be Crangon crangon bunya-like virus or CcBLV. Recently, a few novel bunya-like viruses were described in *Cherax quadricarinatus* [57], *Carcinus maenas* [63] and in *Austropotamobius pallipes* [64]. The Athtabvirus in *C. quadricarinatus*, for example, is thought to be implicated in mortalities after transportation of the animals [57] and in the case of bunya-like brown spot virus in *A. pallipes*, it is thought to cause massive mortality events in wild *A. pallipes* [64]. Again, it would be interesting in the future to further investigate the possible role of CcBLV in *C. crangon* aquaculture.

Once potentially dangerous viruses have been characterized, the possibility of expanding the duplex PCR detection tool to a multiplex PCR method should be investigated as well [30]. In the past, a multiplex RT-PCR assay was successfully designed for the detection of six important shrimp viruses, YHV, TSV, WSSV, HPV, IHHNV and MBV, in penaeid shrimp [65]. Also in other invertebrates, like bumble bees, multiplex PCR tools were successfully designed to monitor the presence of viruses [66].

In conclusion, this study showed that apparently healthy *C. crangon,* caught from the wild, harbour many different novel viruses. In the future, it will be interesting to further investigate these viral communities both in nature and captivity. For example, how these viruses proliferate when shrimp are caught in the wild and transported to rearing facilities. Also the pathogenesis of each of these viruses is lacking and the effect of these viruses when they become virulent. Furthermore, the sequences of these novel viruses can now serve as a marker that will be useful to investigate their potential involvement in occurrences or outbreaks of diseases. Whether or not any of these viruses impact the biology of *C. crangon* in wild populations and/or reared shrimp remains to be solved.

Funding information
This research was supported by the Special Research Fund of the Ghent University, the Research Foundation - Flanders (FWO-Flanders) and the Flanders research institute for agriculture, fisheries and food (ILVO). Olivier Christiaens is a postdoctoral fellow of the Research Foundation - Flanders (FWO-Flanders).

Acknowledgement
We thank David Vuylsteke and Coenraad Deputter (ILVO) for collecting and managing the shrimps.

Conflicts of interest
The authors declare that there are no conflicts of interest

References

1. Vago C. A virus disease in crustacea. *Nat* 1966;209:1290.
2. Bateman KS, Stentiford GD. A taxonomic review of viruses infecting crustaceans with an emphasis on wild hosts. *J Invertebr Pathol* 2017;147:86–110.
3. Mokili JL, Rohwer F, Dutilh BE. Metagenomics and future perspectives in virus discovery. *Curr Opin Virol* 2012;2:63–77.
4. Gao SJ, Moore PS, Phil M. Molecular approaches to the identification of unculturable infectious agents. *Emerg Infect Dis* 1996;2:159–167.
5. Martínez-Porchas M, Vargas-Albores F. Microbial metagenomics in aquaculture: a potential tool for a deeper insight into the activity. *Rev Aquacult* 2017;9:42–56.
6. Handelsman J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* 2004;68:669–685.
7. Orosco FL, Lluisma AO. Variation in virome diversity in wild populations of *Penaeus monodon* (Fabricius 1798) with emphasis on pathogenic viruses. *Virus Dis* 2017;28:262–271.
8. Breitbart M, Salamon P, Andresen B, Mahaffy JM, Segall AM *et al*. Genomic analysis of uncultured marine viral communities. *Proc Natl Acad Sci U S A* 2002;99:14250–14255.
9. Munang'andu HM. Environmental viral metagenomics analyses in aquaculture: applications in epidemiology and disease control. *Front Microbiol* 2016;7:1986.
10. Williamson SJ. Viral Metagenomics. In: Bruijn de (editor). *Handbook of Molecular Microbial Ecology II: Metagenomics in Different Habitats*. Wiley-Blackwell; 2011. pp. 4–12.
11. Alavandi SV, Poornima M. Viral metagenomics: a tool for virus discovery and diversity in aquaculture. *Indian J Virol* 2012;23:88–98.
12. Liu S, Chen Y, Bonning BC. RNA virus discovery in insects. *Insect Sci* 2015;8:54–61.
13. Shi C, Beller L, Deboutte W, Yinda KC, Delang L *et al*. Stable distinct core eukaryotic viromes in different mosquito species from Guadeloupe, using single mosquito viral metagenomics. *Microbiome* 2019;7:121.
14. Sritunyalucksana K, Apisawetakan S, Boon-Nat A, Withyachumnarnkul B, Flegel TW. A new RNA virus found in black tiger shrimp *Penaeus monodon* from Thailand. *Virus Res* 2006;118:31–38.
15. Ng TF, Alavandi S, Varsani A, Burghart S, Breitbart M. Metagenomic identification of a nodavirus and a circular ssDNA virus in semi-purified viral nucleic acids from the hepatopancreas of healthy *Farfantepenaeus duorarum* shrimp. *Dis Aquat Organ* 2013;105:237–242.
16. Dunlap DS, Ng TFF, Rosario K, Barbosa JG, Greco AM *et al*. Molecular and microscopic evidence of viruses in marine copepods. *Proc Natl Acad Sci U S A* 2013;110:1375–1380.
17. Rosani U, Gerdol M. A bioinformatics approach reveals seven nearly-complete RNA-virus genomes in bivalve RNA-Seq data. *Virus Res* 2017;239:33–42.
18. Gudenkauf BM, Hewson I. Comparative metagenomics of viral assemblages inhabiting four phyla of marine invertebrates. *Front Mar Sci* 2016;3.
19. Hewson I, Brown JM, Burge CA, Couch CS, LaBarre BA *et al*. Description of viral assemblages associated with the *Gorgonia ventalina* holobiont. *Coral Reefs* 2012;31:487–491.
20. Hewson I, Eaglesham JB, Höök TO, LaBarre BA, Sepúlveda MS *et al*. Investigation of viruses in *Diporeia* spp. from the Laurentian

great lakes and Owasco lake as potential stressors of declining populations. *J Great Lakes Res* 2013;39:499–506.

21. Gudenkauf BM, Eaglesham JB, Aragundi WM, Hewson I. Discovery of urchin-associated densoviruses (family *Parvoviridae*) in coastal waters of the Big Island, Hawaii. *J Gen Virol* 2014;95:652–658.
22. Hewson I, Button JB, Gudenkauf BM, Miner B, Newton AL *et al*. Densovirus associated with sea-star wasting disease and mass mortality. *Proc Natl Acad Sci U S A* 2014;111:17278–17283.
23. Li C-X, Shi M, Tian J-H, Lin X-D, Kang Y-J *et al*. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *Elife* 2015;4:e05378.
24. Rosario K, Schenck RO, Harbeitner RC, Lawler SN, Breitbart M. Novel circular single-stranded DNA viruses identified in marine invertebrates reveal high sequence diversity and consistent predicted intrinsic disorder patterns within putative structural proteins. *Front Microbiol* 2015;6:696.
25. Shi M, Lin X-D, Tian J-H, Chen L-J, Chen X *et al*. Redefining the invertebrate RNA virosphere. *Nature* 2016;540:539–543.
26. Walker PJ, Winton JR. Emerging viral diseases of fish and shrimp. *Vet Res* 2010;41:51.
27. Walker PJ, Mohan CV. Viral disease emergence in shrimp aquaculture: origins, impact and the effectiveness of health management strategies. *Rev Aquacult* 2009;1:125–154.
28. Van Eynde B, Vuylsteke D, Christiaens O, Cooreman K, Smagghe G *et al*. Improvements in larviculture of *Crangon crangon* as a step towards its commercial aquaculture. *Aquac Res* 2019;50:1658–1667.
29. Stentiford GD, Feist SW. A histopathological survey of shore crab (*Carcinus maenas*) and brown shrimp (*Crangon crangon*) from six estuaries in the United Kingdom. *J Invertebr Pathol* 2005;88:136–146.
30. Van Eynde B, Christiaens O, Delbare D, Cooreman K, Bateman KS *et al*. Development and application of a duplex PCR assay for detection of Crangon crangon bacilliform virus in populations of European brown shrimp (Crangon crangon). *J Invertebr Pathol* 2018;153:195–202.
31. Soetaert M, Chiers K, Duchateau L, Polet H, Verschueren B *et al*. Determining the safety range of electrical pulses for two benthic invertebrates: Brown shrimp (*Crangon crangon* L.) and ragworm (*Alitta virens* S.). *ICES J Mar Sci* 2015;72:973–980.
32. Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L *et al*. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep* 2015;5:16532.
33. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using diamond. *Nat Methods* 2015;12:59–60.
34. Yinda CK, Zeller M, Conceição-Neto N, Maes P, Deboutte W *et al*. Novel highly divergent reassortant bat rotaviruses in Cameroon, without evidence of zoonosis. *Sci Rep* 2016;6:34209.
35. Yinda CK, Rector A, Zeller M, Conceição-Neto N, Heylen E *et al*. A single bat species in Cameroon harbors multiple highly divergent papillomaviruses in stool identified by metagenomics analysis. *Virol Rep* 2016;6:74–80.
36. Yinda CK, Conceição-Neto N, Zeller M, Heylen E, Maes P *et al*. Novel highly divergent sapoviruses detected by metagenomics analysis in straw-colored fruit bats in Cameroon. *Emerg Microbes Infect* 2017;6:e38:1–7.
37. Yinda CK, Zell R, Deboutte W, Zeller M, Conceição-Neto N *et al*. Highly diverse population of *Picornaviridae* and other members of the *Picornavirales*, in Cameroonian fruit bats. *BMC Genomics* 2017;18:249.
38. Ondov BD, Bergman NH, Phillippy AM. Interactive metagenomic visualization in a web browser. *BMC Bioinformatics* 2011;12:385.
39. Maric J. *Long read RNA-Seq mapper (Master thesis)*. Zagreb, Croatia: University of Zagreb; 2017. p. 62.
40. Paulson JN. *MetagenomeSeq: Statistical Analysis for Sparse High-Throughput Sequencing (Vignette)*. United States: University of Maryland; 2019. p. 35.
41. Guu TSY, Zheng W, Tao YJ. Bunyavirus: structure and replication. *Adv Exp Med Biol* 2012;726:245–266.
42. Elliott RM. Molecular biology of the *Bunyaviridae*. *J Gen Virol* 1990;71 (Pt 3:501–522.
43. Schmaljohn CS, Nichol ST. Bunyaviridae. In: Knipe DM, Howley PM (editors). *Field's Virology*. Philadelphia, PA: Lippincott Wiliams & Wilkins Co; 2007. pp. 1741–1789.
44. Elliot RM, Schmaljohn C. Bunyaviridae. In: Knipe DM, Howley PM (editors). *Field's Virology*. Philadelphia, PA: Lippincott Wiliams & Wilkins Co; 2013. pp. 1244–1282.
45. Elliott RM. Orthobunyaviruses: recent genetic and structural insights. *Nat Rev Microbiol* 2014;12:673–685.
46. Shi M, Lin X-D, Vasilakis N, Tian J-H, Li C-X, Chen LJ *et al*. Divergent viruses discovered in arthropods and vertebrates revise the evolutionary history of the *Flaviviridae* and related viruses. *J Virol* 2016;90:659–669.
47. Parry R, Asgari S. Discovery of novel crustacean and cephalopod flaviviruses: insights into the evolution and circulation of flaviviruses between marine invertebrate and vertebrate hosts. *J Virol* 2019;93:e00432–19.
48. Le Gall O, Christian P, Fauquet CM, King AMQ, Knowles NJ *et al*. *Picornavirales*, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T = 3 virion architecture. *Arch Virol* 2008;153:715–727.
49. Scheets K, Jordan R, White KA, Hernández C. Pelarspovirus, a proposed new genus in the family *Tombusviridae*. *Arch Virol* 2015;160:2385–2393.
50. Bonami J-R, Zhang S. Viral diseases in commercially exploited crabs: a review. *J Invertebr Pathol* 2011;106:6–17.
51. Olivier V, Blanchard P, Chaouch S, Lallemand P, Schurr F *et al*. Molecular characterisation and phylogenetic analysis of chronic bee paralysis virus, a honey bee virus. *Virus Res* 2008;132:59–68.
52. Sõmera M, Sarmiento C, Truve E. Overview on Sobemoviruses and a proposal for the creation of the family *Sobemoviridae*. *Viruses* 2015;7:3076–3115.
53. Hillman BI, Cai G. The family narnaviridae: simplest of RNA viruses. *Adv Virus Res* 2013;86:149–176.
54. Stentiford GD, Bateman K, Feist SW. Pathology and ultrastructure of an intranuclear bacilliform virus (IBV) infecting brown shrimp Crangon crangon (Decapoda: Crangonidae). *Dis Aquat Organ* 2004;58:89–97.
55. Fauver JR, Grubaugh ND, Krajacich BJ, Weger-Lucarelli J, Lakin SM *et al*. West African Anopheles gambiae mosquitoes harbor a taxonomically diverse virome including new insect-specific flaviviruses, mononegaviruses and totiviruses. *Virology* 2016;498:288–299.
56. Laureti M, Narayanan D, Rodriguez-Andres J, Fazakerley JK, Kedzierski L. Flavivirus receptors: diversity, identity, and cell entry. *Front Immunol* 2018;9:2180.
57. Sakuna K, Elliman J, Tzamouzaki A, Owens L. A novel virus (order *Bunyavirales*) from stressed redclaw crayfish (*Cherax quadricarinatus*) from farms in northern Australia. *Virus Res* 2018;250:7–12.
58. Bonami J-R, Shi Z, Qian D, Sri Widada J, Widada JS. White tail disease of the giant freshwater prawn, *Macrobrachium rosenbergii*: separation of the associated virions and characterization of MrNV as a new type of nodavirus. *J Fish Dis* 2005;28:23–31.
59. Tang KFJ, Pantoja CR, Redman RM, Navarro SA, Lightner DV. Ultrastructural and sequence characterization of Penaeus vannamei nodavirus (PvNV) from Belize. *Dis Aquat Organ* 2011;94:179–187.
60. Cornejo-Granados F, Lopez-Zavala AA, Gallardo-Becerra L, Mendoza-Vargas A, Sánchez F *et al*. Microbiome of Pacific whiteleg shrimp reveals differential bacterial community composition between wild, aquacultured and AHPND/EMS outbreak conditions. *Sci Rep* 2017;7:11783.

61. Chen Y-H, He J-G. Effects of environmental stress on shrimp innate immunity and white spot syndrome virus infection. *Fish Shellfish Immunol* 2019;84:744–755.
62. Lightner DV. Virus diseases of farmed shrimp in the Western hemisphere (the Americas): a review. *J Invertebr Pathol* 2011;106:110–130.
63. Bojko J, Subramaniam K, Waltzek TB, Stentiford GD, Behringer DC. Genomic and developmental characterisation of a novel bunyavirus infecting the crustacean *Carcinus maenas*. *Sci Rep* 2019;9:12957.
64. Grandjean F, Gilbert C, Razafimafondy F, Vucić M, Delaunay C *et al*. A new bunya-like virus associated with mass mortality of white-clawed crayfish in the wild. *Virology* 2019;533:115–124.
65. Khawsak P, Deesukon W, Chaivisuthangkura P, Sukhumsirichart W. Multiplex RT-PCR assay for simultaneous detection of six viruses of penaeid shrimp. *Mol Cell Probes* 2008;22:177–183.
66. Meeus I, Smagghe G, Siede R, Jans K, de Graaf DC. Multiplex RT-PCR with broad-range primers and an exogenous internal amplification control for the detection of honeybee viruses in bumblebees. *J Invertebr Pathol* 2010;105:200–203.