Virus Research 244 (2018) 137–146

Contents lists available at ScienceDirect

Virus Research

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

The gut virome of the protochordate model organism, *Ciona intestinalis* subtype A

Brittany A. Leigh[a,b], Anni Djurhuus[a], Mya Breitbart[a], Larry J. Dishaw[b,*]

[a] University of South Florida, College of Marine Science, St. Petersburg, FL, USA
[b] University of South Florida, Department of Pediatrics, Children's Research Institute, St. Petersburg, FL, USA

ARTICLE INFO

Keywords:
Ciona
Gut
Microbiome
Viral metagenome
Phage

ABSTRACT

The identification of host-specific bacterial and viral communities associated with diverse animals has led to the concept of the metaorganism, which defines the animal and all of its associated microbes as a single unit. Here we sequence the viruses found in the gut (i.e., the gut virome) of the marine invertebrate model system, *Ciona intestinalis* subtype A, in samples collected one year apart. We present evidence for a host-associated virome that is distinct from the surrounding seawater and contains some temporally-stable members. Comparison of gut tissues before and after starvation in virus-free water enabled the differentiation between the *Ciona*-specific virome and transient viral communities associated with dietary sources. The *Ciona* gut viromes were dominated by double-stranded DNA tailed phages (Order *Caudovirales*) and sequence assembly yielded a number of complete circular phage genomes, most of which were highly divergent from known genomes. Unique viral communities were found in distinct gut niches (stomach, midgut and hindgut), paralleling the compartmentalization of bacterial communities. Additionally, integrase and excisionase genes, including many that are similar to prophage sequences within the genomes of bacterial genera belonging to the *Ciona* core microbiome, were prevalent in the viromes, indicating the active induction of prophages within the gut ecosystem. Knowledge of the gut virome of this model organism lays the foundation for studying the interactions between viruses, bacteria, and host immunity.

1. Introduction

In recent years, numerous studies have defined the importance of host-associated microorganisms, including both bacteria and viruses. The bacterial component of this 'microbiome' often outnumbers host cells by an order of magnitude (Turnbaugh et al., 2007) and can influence host nutrient acquisition and metabolism (Nicholson et al., 2012; Tremaroli and Backhed, 2012). The increased recognition of the contributions of bacterial communities to the biology and physiology of animals has motivated our perception of animals as complex metaorganisms (Rohwer et al., 2002; Rosenberg and Zilber-Rosenberg, 2011; Theis et al., 2016). This concept was first introduced into biology to describe the coral animal and all of its associated microbes (Rohwer et al., 2002). Since then, many animal models, including humans, have been documented to contain a stable, core community of bacteria on or within their bodies (Dishaw et al., 2014; Roeselers et al., 2011; Sabree et al., 2012; Schmitt et al., 2012; Turnbaugh and Gordon, 2009), the disruption of which often results in disease (Cho and Blaser, 2012; Sobhani et al., 2011; Tamboli et al., 2004). While bacteria have historically predominated microbiome studies, recent efforts are recognizing a vital role for viruses as well (Abeles and Pride, 2014; Grasis et al., 2014; Minot et al., 2011; Reyes et al., 2010; Thurber et al., 2017). The viruses associated with animal hosts are present as both free viral particles and/or stably integrated proviruses (Feschotte and Gilbert, 2012) and can shape the health of the metaorganism through pathogenesis (Davies et al., 2016) or through influencing the metabolic potential of both the animal host and its associated microbes (Roossinck, 2011). In addition to host immunity (Cullender et al., 2013; Thaiss et al., 2016), nutrient availability (Cohen et al., 2015), and bacterial competition (Flint et al., 2007), viruses, the majority of which are bacteriophages (i.e. phages), likely affect the structure of animal-associated bacterial communities through lytic infection and prophage integration and/or induction. In this report, we describe the gut virome of a marine filter-feeding invertebrate protochordate, *Ciona intestinalis* subtype A, cataloging for the first time the viral communities associated with these early extant chordates and establishing a tractable model system in which to pursue future studies examining the complex dynamics between metazoan hosts and their associated bacterial and viral

* Corresponding author.
E-mail address: ldishaw@health.usf.edu (L.J. Dishaw).

https://doi.org/10.1016/j.virusres.2017.11.015
Received 13 July 2017; Received in revised form 14 November 2017; Accepted 14 November 2017
Available online 15 November 2017
0168-1702/

communities.

Ciona is a well-known developmental model system that has been adapted recently for studies of host-microbe interactions within the gut ecosystem (Dishaw et al., 2014; Dishaw et al., 2011; Dishaw et al., 2016; Liberti et al., 2014). This animal is a marine filter-feeding protochordate that derives nutrients from particulates in the water column, including phytoplankton (Coleman, 1991). *Ciona* is amenable to germ-free mariculture (Leigh et al., 2016) and has been shown previously to maintain a core bacterial community in its gut (Dishaw et al., 2014). Here we report the presence of temporally stable viruses within this metabolically active environment. We also describe evidence for compartmentalization of both viruses and bacteria, with distinct members occupying the stomach, midgut, and hindgut. The observation that *Ciona* maintains a virome with many predicted hosts matching members of the core microbiome and a prevalence of prophages indicates a role for phages as major players in shaping host-associated bacterial communities within the dynamic gut ecosystem of these animals.

2. Materials and methods

2.1. Tissue collections

Adult *Ciona* were wild-harvested in San Diego, California, USA, and shipped overnight to the laboratory in Florida. Upon arrival, ten animals were randomly selected, five of which were immediately harvested for gut tissue while gut contents were cleared from the remaining five in 100 kD filtered virus-free seawater for 24 h, with water changes every 3 h for the first 12 h. In 2014, entire guts were dissected and snap-frozen in liquid nitrogen, while in 2015, guts were trisected (stomach, midgut and hindgut; refer to Fig. 1A for anatomy) using sterile surgical tools, snap-frozen in liquid nitrogen and maintained at −80 °C until processing. Additionally, two water samples of 1 L each were acquired in 2015, one from surrounding seawater in Mission Bay (MB; main collection site) and the other from holding tank water in which they were shipped (CB).

2.2. Viral and bacterial DNA extraction and sequencing

Each tissue pool from five animals was disrupted using the GentleMACS dissociator (Miltenyi Biotec, Bergisch Gladbach, Germany) in 3 mL of sterile 0.02 µm-filtered suspension medium (SM) buffer (100 mM NaCl, 16 mM $MgSO_4$, 4.5 mM Tris Base, pH 7.4). The tissue fragments were pelleted at 6000 RCF for 10 min, and the supernatant containing the microbial fraction was subsequently filtered through a 0.45 µm syringe filter (Merck Millipore, Darmstadt, Germany). The filtrate was then filtered again through a 0.22 µm Sterivex filter (Merck Millipore, Darmstadt, Germany) and the resulting viral effluent collected. To determine specificity of the *Ciona* viromes, each liter of surrounding seawater was 0.22 µm-filtered, and the viruses concentrated using a 100 kD Amicon filter (EMD, Merck Millipore, Darmstadt, Germany). The presence of virus-like particles was verified in 0.22 µm-filtered samples from *Ciona* tissues and water using epifluorescence microscopy (EFM) with SYBR Gold (Invitrogen, Carlsbad, CA, USA) staining as previously described (Patel et al., 2007). Each filtrate, including the Amicon concentrate, was then loaded onto a CsCl gradient to purify viral particles as previously described (Thurber et al., 2009). Briefly, each sample was spun at 61,000 RCF for 3 h, and the 1.2 − 1.5 g/mL fraction was collected using a sterile syringe and needle in a 2 mL sterile tube. To remove bacterial vesicles and free DNA, this fraction was then treated with chloroform (at a final concentration of 20% vol/vol) for 10 min at room temperature, centrifuged for 30 s at max speed (20,000 RCF), and the top aqueous layer was removed and treated with DNase I (2.5 U/µL final concentration) for 3 h at 37 °C, vortexing occasionally. Each sample was then treated with EDTA pH 8.0 (20 mM final concentration) to neutralize the DNase I enzyme. The treated sample was verified to be free of bacterial cells by EFM, and free of bacterial DNA through 16S rRNA gene PCR. Briefly, each sample was tested for bacterial contamination using PCR with 16S rRNA gene primers (27F and 1492R) (Weisburg et al., 1991) using the following PCR conditions: denature at 95 °C for 5 min, cycle 35 times through 94 °C for 30 s, 56 °C for 30 s, 72 °C for 90 s, and end with a final extension at 72 °C for 10 min. Once each sample was confirmed negative for bacteria and positive for viruses, viral DNA was extracted using the Qiagen MinElute Virus Spin Kit (Qiagen, Inc., Valencia, CA, USA). Viral DNA was amplified using a GenomiPhi V2 DNA amplification kit (GE Healthcare Life Sciences, USA) to generate adequate template for sequencing (~1 µg). To minimize bias introduced by Phi29 DNA polymerase, three separate reactions were prepared and pooled. DNA concentrations were then determined using a Qubit and products were visualized using a 1% agarose gel to confirm amplification. Final amplified products were purified using a MinElute PCR Purification Kit (Qiagen, Inc., Valencia, CA, USA). The quality was also assessed using a BioAnalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). Each sample was sequenced with the Illumina MiSeq platform generating mate-pair (2 × 250) libraries (Eurofins MWG Operon LLC, Huntsville, AL, USA). SDC14 was sequenced on an earlier run than all other samples. From the 2015 trisected gut samples, bacterial DNA was extracted from the 0.22 µm Sterivex filters using the MoBio DNeasy PowerSoil Kit

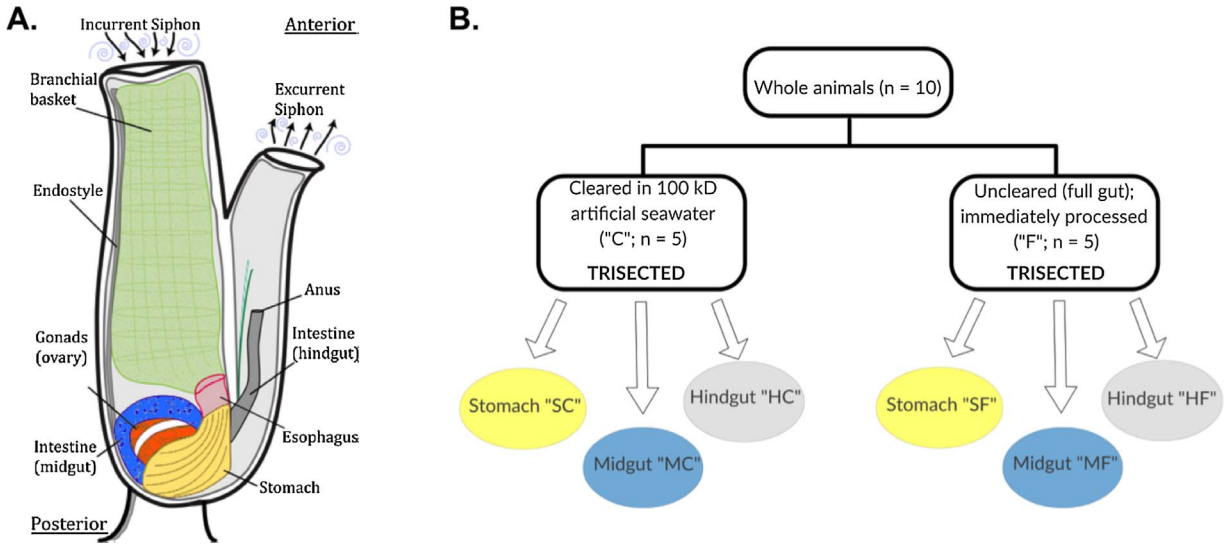

Fig. 1. (A) Schematic of *Ciona intestinalis* subtype A highlighting anatomy of the gut compartments. (B) Summary of tissue sampling approach.

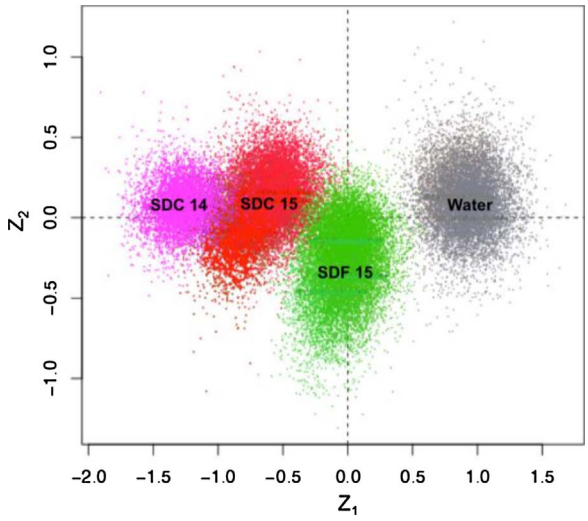

Fig. 2. Bayesian pairwise distance matrix of read similarity among viromes. Each color represents an entire gut or water sample. Compartments in 2015 were combined to form "SDC 15" and "SDF 15". "Water" represents both MB and CB water samples.

(MoBio, Inc., Carlsbad, CA, USA), and the V4-V5 regions of the 16S rRNA gene were amplified using the Fluidigm protocol through the Michigan State University Genomics Core with the primers outlined in Parada et al. (2016). The amplification parameters were as follows: 95 °C for 3 min, followed by 30 cycles of 95 °C for 45 s, 50 °C for 45 s and 68 °C for 90 s, and a final extension of 68 °C for 5 min. The amplicons were verified on an agarose gel and diluted 1:100 for sequencing on the Illumina MiSeq platform.

2.3. Bioinformatics

2.3.1. Viral bioinformatics

Mate-pair reads from the viromes were analyzed using the iVirus pipeline (Bolduc et al., 2017). First the sequences were trimmed using Trimmomatic 0.35.0 and quality checked using FastQC to ensure a PHRED score of 20 and a minimum read length of 100 bp. To assess the amount of shared sequence content between each of the viromes using Bayesian network analyses, the trimmed reads were run through Fizkin implemented through Jellyfish (Marcais and Kingsford, 2011) using default parameters. The resulting ordination plot was used to initially distinguish the viromes based on k-mer similarity among reads (Fig. 2).

De novo assembly of mate-pair reads from each virome individually was completed using SPAdes 3.6.0 with a k-mer value of 57 and default parameters. Assembly quality was determined by QUAST (Gurevich et al., 2013). Because the rolling circle amplification step utilized in library construction biases towards small, circular single-stranded DNA (ssDNA) templates (Kim and Bae, 2011), contigs smaller than 5 kilobases (kb) were removed from the contig-based analyses in this manuscript and will be described elsewhere. Contigs larger than 5 kb were examined with VirSorter, a program used to predict viral sequences based on a number of parameters such as the presence of hallmark viral genes or other known viral-like sequences (Roux et al., 2015). Additionally, all samples were co-assembled with SPAdes 3.6.0 with a k-mer of 57 to generate a single reference file and run through VirSorter. Reads were then mapped back to the VirSorter viral contig outputs greater than 5 kb to estimate the relative abundance of each contig for each sample. BowtieBatch was used to run bowtie2 on all samples of the co-assembled contigs and produced BAM output files read by Read2RefMapper to generate relative abundance and coverage plots for each viral contig within each metagenome (Bolduc et al., 2017). To consider a contig present within an individual sample, reads from that sample had to cover 75% of the viral contig from the co-assembled virome. Venn diagrams were generated using the VennDiagram package (Chen and Boutros, 2011) through R v 3.3.2 software (R Core Team, 2014) to display the overlap of contigs in different gut compartments using the mapping data from Read2RefMapper.

VirSorter also detected linear contigs with matching k-mers at the ends and assumed them to be complete circular genomes. Relative abundance matrices of each of these circular viral genomes within each of the viromes were determined using BowtieBatch and Read2RefMapper as described above. These complete genomes, 29 in total, were then imported into Geneious 8.1.7 (Kearse et al., 2012) and open reading frames (ORFs) were determined using Glimmer3 (Delcher et al., 2007). ORF annotations were improved with the BLASTx algorithm against the non-redundant (NR) protein database in GenBank. For those genomes with the majority of ORFs matching a single known virus, synteny of the full genomes was compared using Mauve progressive genome aligner (Darling et al., 2010) with default parameters.

The VirSorter output for single assemblies was used as input for MetaVir-based taxonomic classification using the RefSeq protein database of complete viral sequences (Roux et al., 2014) and to MG-RAST for functional characterization using Subsystems technology (Keegan et al., 2016). All taxonomic classifications were determined using the top BLASTx hit, with a threshold score of 50 on BLAST bitscore through MetaVir.

2.3.2. Bacterial bioinformatics

To determine the number of shared bacterial operational taxonomic units (OTUs; determined by a 97% identity cutoff) between the different gut compartments, 16S rRNA gene amplicon sequences were trimmed and quality filtered using the DADA2 pipeline (Callahan et al., 2016) through the R v 3.3.2 software (R Core Team, 2014). The SILVA database was utilized to assign taxonomy (Quast et al., 2013) to the resulting OTUs, and data were rarefied to 39,577 reads per sample to produce OTU Venn diagrams as described above for viral contigs.

3. Results

3.1. Ciona gut virome

The seven following *Ciona* gut virome samples were sequenced: SDC 14 (cleared guts from 2014); cleared stomach (SC), midgut (MC), and hindgut (HC) tissues from 2015 (referred to as SDC 15 when pooled); and uncleared ("full") stomach (SF), midgut (MF), and hindgut (HF) tissues from 2015 (referred to as SDF 15 when pooled). The virome sequencing and single sample assembly statistics are shown in Table 1. In 2015, additional viromes were sequenced from the surrounding seawater from the collection site in Mission Bay (MB) and the holding tank water that the animals were shipped in (CB).

Initially, Fizkin k-mer analysis was performed on all trimmed reads generating a Bayesian pairwise distance matrix to provide an overall view of the degree of similarity between the samples. The *Ciona* gut viromes were distinct from the surrounding seawater (Fig. 2) and the uncleared gut viromes (SDF 15) were more similar to the water (MB and CB) than the cleared gut viromes (SDC 14 and SDC 15) were, indicating the contribution of viruses of dietary origin in the full gut virome. Although temporal variability between the cleared gut viromes from 2014 (SDC 14) and 2015 (SDC 15) is evident, these communities have a similar composition.

Due to the higher confidence in taxonomic assignment of longer contigs than short individual reads, the remaining analyses focused on the 3940 contigs (also referred to as nodes) greater than 5 kb that were obtained from separate assembly of each sample (Table 1). Greater than 50% of the reads contributed to contigs > 5 kb, which represented 20–25% of the total generated contigs. Additionally, all *Ciona* viromes were co-assembled to generate a reference list of contigs, with a contig considered present in an individual sample if reads from that sample covered 75% of the viral contig, in which case the coverage (normalized

Table 1
Virome sequencing and single sample assembly details. Only contigs > 5 kb were analyzed in this study. The affiliation ratios refer to the percent of contigs with significant BLASTx similarity to the RefSeq database (BLAST bitscore threshold 50).

	SDC 14	SC	SF	MC	MF	HC	HF
# Reads (Millions)	6.51	4.41	3.08	3.95	3.63	7.32	3.37
Avg seq length (bp)	199	505	696	725	750	453	809
# conitgs (> 500 bp)	17685	18662	29976	34433	36166	20368	32433
# contigs (> 1000 bp)	5826	5887	8413	12101	10577	6656	9946
# contigs (> 5000 bp)	324	363	602	670	901	128	952
Largest contig	37553	95916	65207	79274	114284	18907	114296
GC%	39.63	39.72	39.02	39.41	38.75	38.29	39.55
N50	1340	1517	1275	1448	1500	1113	1692
Affiliated Contigs > 5 kb	305	328	545	615	826	116	883
Unaffiliated Contigs > 5 kb	19	35	57	55	75	12	69
Affilitation Ratio > 5 kb	0.94	0.9	0.91	0.92	0.92	0.91	0.93
Genes predicted	4026	5610	8415	9535	12898	1104	14610
Circular Contigs > 10 kb	1	3	4	2	4	0	15

to the total number of reads) was calculated. This value is referred to as the relative abundance of each contig within a sample.

To examine the temporal variability of the virome, contigs recovered from cleared *Ciona* gut samples collected one year apart were compared. Strikingly, 453 contigs were shared between cleared gut tissues over two separate years (Fig. 3). None of these shared contigs were found to be circular genomes. The top 10 viral contigs in each group shared sequence identity with phages infecting described members of the *Ciona* core microbiome, including *Flavobacteria*, *Shewanella*, *Pseudoalteromonas,* and *Vibrio* (Dishaw et al., 2014).

In an effort to distinguish between *Ciona*-associated viruses and those that originate from dietary contents, the 2015 trisected gut compartments were either sampled immediately (uncleared stomach (SF), midgut (MF), and hindgut (HF)) or after the animal was allowed to clear its gut contents (cleared stomach (SC), midgut (SC), and hindgut (HC)). Although some contigs were shared between the cleared and uncleared samples, each of the states maintained a large number of unique contigs (Fig. 4). While the majority of the identifiable dsDNA viruses in both cleared and full guts belonged to the *Caudovirales* order (Fig. 5A), differences in phage types were seen on finer taxonomic scales. The unclassified dsDNA phage contigs in cleared guts were similar to a number of *Vibrio* phages, mostly to *Vibrio* phage henriette 12B8 which infects *Vibrio splendidus*, and also *Pseudomonas* phages PA11 and vp_PaeP_Tr60_Ab31, both of which infect *Pseudomonas aeruginosa* (Kwan et al., 2006). In contrast, the majority of unclassified dsDNA phage contigs present in uncleared gut samples were most similar to cyanophages and *Persicivirga* phages. The *Persicivirga* phages P12024L and P12024S infect bacteria that degrade polysaccharides from green algae (Barbeyron et al., 2011; Kang et al., 2012) and while contigs related to these phages were present in all gut samples, they were markedly higher in the uncleared samples, perhaps due to the fact that *Ciona* eat fine detritus and phytoplankton (Coleman, 1991). Contigs related to *Cellulophaga* phages were also more abundant in uncleared than in cleared trisected guts from 2015. This recently discovered group of phages was shown to be widespread in low

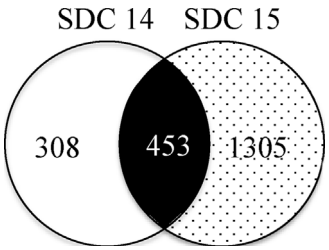

Fig. 3. Virome of the cleared *Ciona* gut between collection years as determined by 75% contig coverage by reads to the co-assembled contigs. Cleared compartments from 2015 were combined to form "SDC 15".

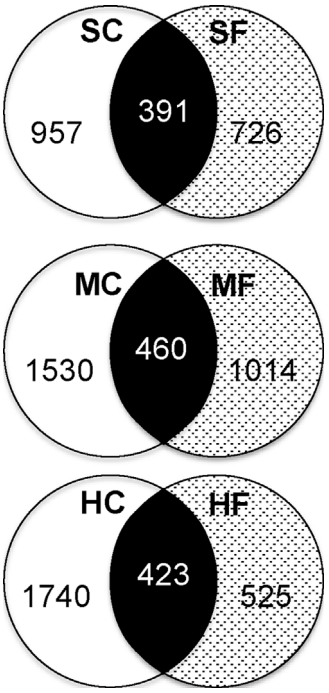

Fig. 4. Shared contigs between cleared and uncleared compartments from 2015 tissues as determined by 75% contig coverage by read mapping.

abundances among marine viral metagenomes (Holmfeldt et al., 2013).

In addition, the relative abundance of contigs similar to *Phycodnaviridae*, viruses that infect eukaryotic algae, was 5–6 x lower in cleared than uncleared guts, consistent with removal of dietary contents through the clearing process. All uncleared samples had large numbers of sequences related to *Chrysochromulina ericina* virus CeV-01B, an algal virus with a 510 kb genome that belongs to the *Phycodnaviridae* (Sandaa et al., 2001). The *Mimiviridae* family of giant viruses, which infect the genus *Acanthamoeba,* were also less abundant in cleared gut samples compared to uncleared.

3.2. Evidence for viral and bacterial compartmentalization in the Ciona gut

Taxonomic comparisons of the contigs highlighted both similarities and differences between gut compartments (Fig. 5A). Although a large number of contigs were present in all gut compartments (601), each possessed a number of unique contigs (SC: 172, MC: 333, HC: 73), with the midgut containing the largest number of unique contigs (Fig. 6A). All compartments were dominated by the *Caudovirales* order of dsDNA viruses, and most of the *Caudovirales* contigs were nearly twice as

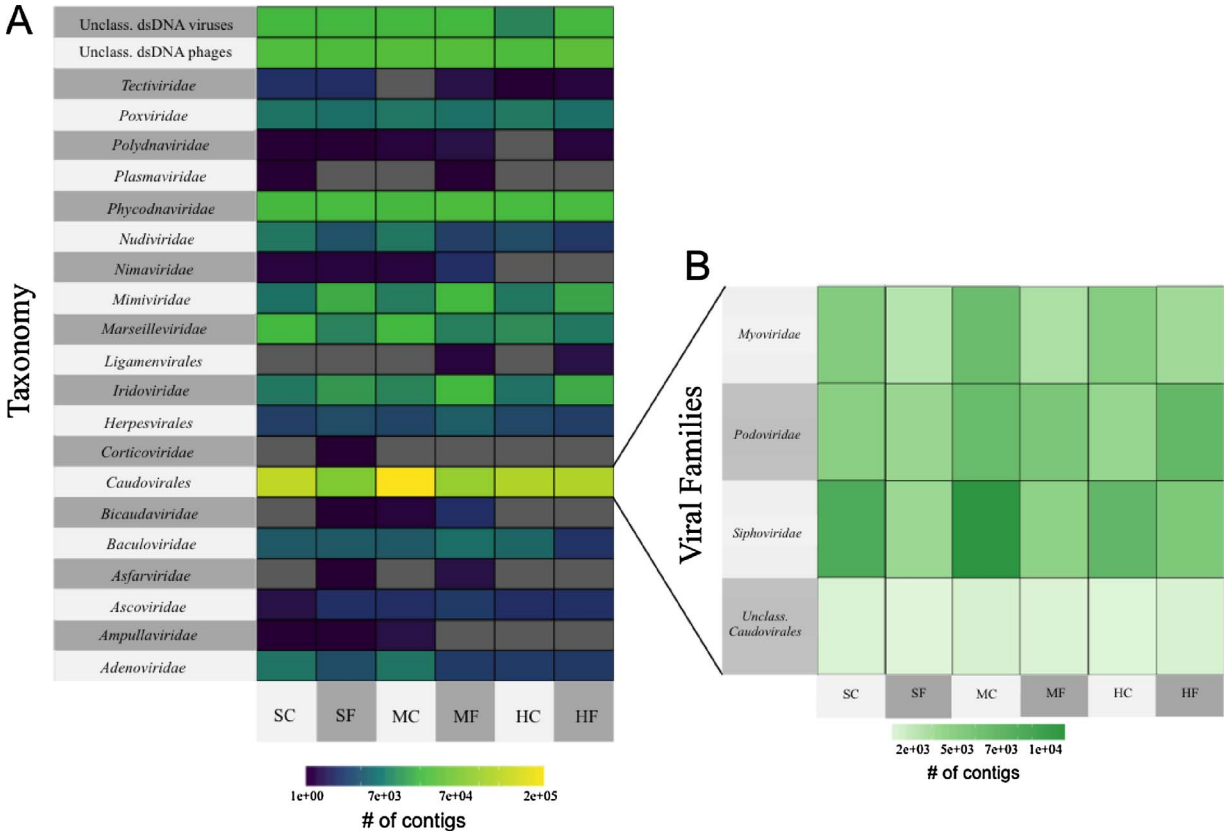

Fig. 5. Taxonomy and abundance of all affiliated contigs from trisected gut samples from (A) all dsDNA viruses and, more specifically, (B) *Caudovirales*. (SC = stomach cleared; SF = stomach full; MC = midgut cleared; MF = midgut full; HC = hindgut cleared;HF = hindgut full).

abundant in the midgut compared to the two other compartments (Fig. 5A). The *Siphoviridae* was the most abundant virus family in the midgut, followed by *Podoviridae* and then *Myoviridae* (Fig. 5B). The same three phage families dominated the other two compartments, but were present at much lower abundances. The midgut also possessed the highest number of unclassified dsDNA phages, some most closely related to the *Vibrio* phage henriette 12B8, *Persicivirga* phages P12024L and P12024S, *Idiomarinaceae* phage 1N2-2, and *Marinomonas* phage P12026. Consistent with the high number of viral contigs unique to the midgut, the midgut also possessed the largest number of distinct bacterial OTUs compared to the other two compartments (Fig. 6C). The number of reads from each compartment matching the top 50 viral contigs (Fig. 6B) or bacterial OTUs (Fig. 6D) from all three compartments were mapped to assess compartmentalization of viruses and bacteria. While each of the viral contigs shown in Fig. 6B was among the top 50 most abundant viral contigs for one or more of the compartments, their relative abundance in the other compartments often varied dramatically. Some viral contigs that were highly abundant in one compartment were rare or absent in other compartments, revealing a great degree of spatial organization in the *Ciona* gut virome. For example, the *Streptococcus* phage SOCP, which infects a Gram-positive bacteria belonging to *Firmicutes*, a phyla present in over 75% of previously surveyed *Ciona* gut samples (Dishaw et al., 2014), was the most abundant viral contig in the stomach, while it was much less abundant in the midgut and the hindgut. Phages with sequence similarity to the uvMED dataset and *Idiomarinaceae* phage Phi1M2-2 also showed differing abundances, increasing in relative abundance in the midgut and hindgut compared to the stomach. The large changes in the relative abundance of specific viral contigs in the different gut compartments parallels differences in the bacterial communities, with the stomach, midgut and hindgut each maintaining OTUs that are not found in the other two compartments (Fig. 6D). Interestingly, OTUs belonging to the *Flavobacteraceae* family were among the most abundant in the stomach; however, phages that likely use these bacteria as their host were most abundant in the mid- and hindgut. Among the top 50 viral contigs of the hindgut, the two most abundant were only similar to sequences in the metagenomic (env_nr) database from NCBI, indicating the presence of a novel yet abundant group of phages in the *Ciona* gut. Additionally, *Pseudoalteromonas* phage PHS3 and hypothetical viral contigs sharing nucleotide identity with *Pseudoalteromonas* appear to be abundant in the hindgut (ranks 7–9) and the midgut (ranks 3–5) and decreasing in abundance in the stomach (rank 197). An OTU matching *Pseudoalteromonas* spp. was only detectable in the hindgut (36th most abundant OTU in HC). Lastly, a contig with sequence similarity to an integrase gene from *Vibrio vulnificus* (6th most abundant viral contig in HC) was only detectable in the midgut and the hindgut, while no reads mapped back to this contig in the stomach. Correspondingly, OTUs sharing sequence identity with *Vibrio* spp. were also only detectable within the midgut and the hindgut. Large increases in the relative abundance of sequences similar to prophages within the midgut and hindgut suggest these tissues as a site for prophage induction.

3.3. Prophages within the Ciona virome

Prophages are viruses that are stably integrated within bacterial genomes. However, because our methods involved purification of virions, only prophages that were present as free viral particles, presumably as the result of induction, would be identified. Integrases and excisionases, which are hallmark prophage genes, were detected in all of the *Ciona* viromes by comparison to the MG-RAST Subsystems database. All cleared viromes consistently possessed more of these genes than their uncleared counterparts, perhaps due to the clearing process (Fig. **S1**). Many of these prophages shared sequence identity to phage-like genes within the genomes of bacteria belonging to the core *Ciona*

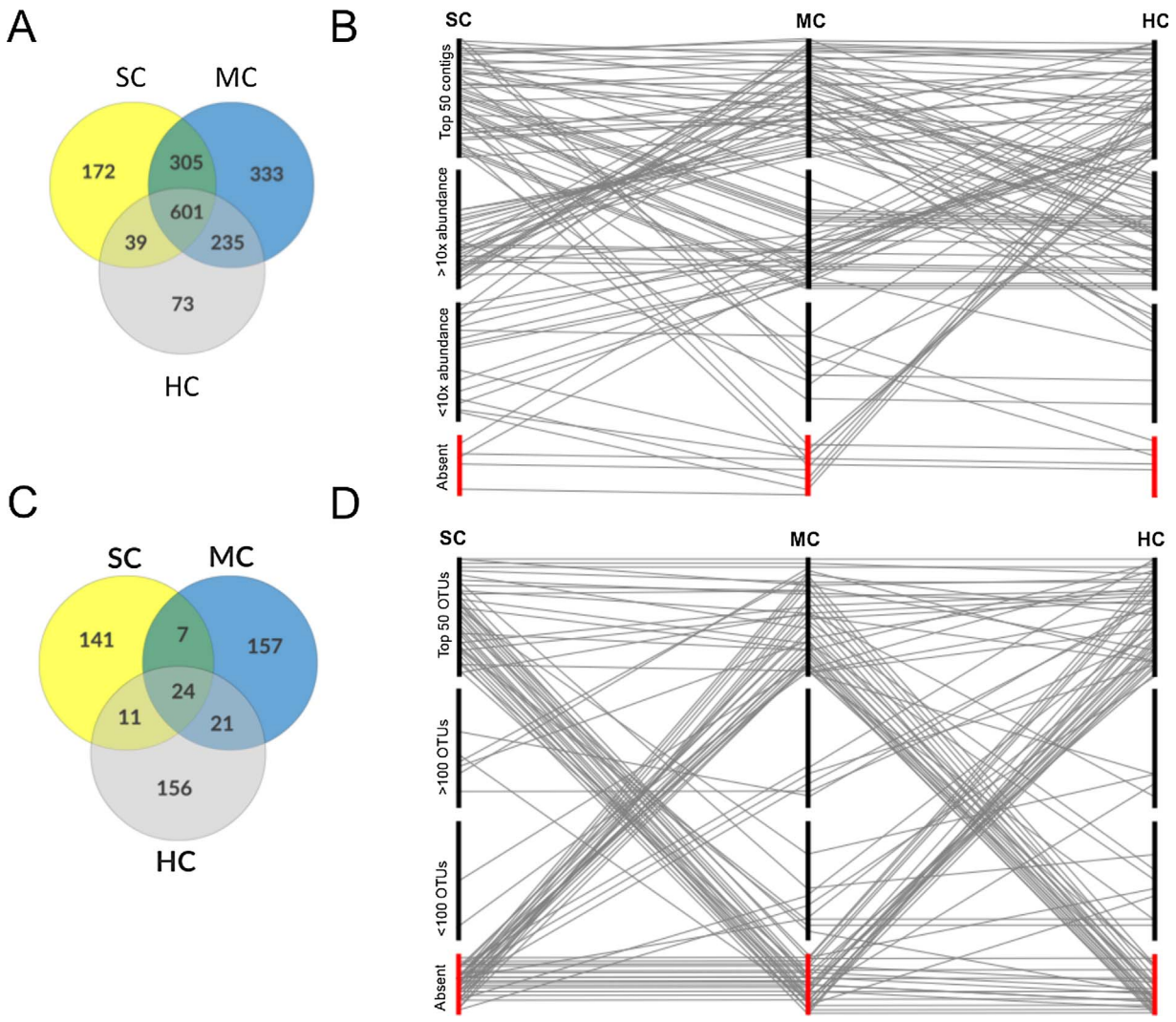

Fig. 6. (A) Compartmentalization of viral contigs from the *Ciona* gut determined by ≥75% contig coverage by reads. (B) Relative abundance of the top 50 viral contigs from each compartment, compared across subsequent compartments as determined by read mapping to contigs from the co-assembly. (C) Compartmentalization of bacterial OTUs from the *Ciona* gut. (D) Relative abundance of the top 50 bacterial OTUs from each compartment, compared across subsequent compartments. (SC = stomach cleared; MC = midgut cleared; HC = hindgut cleared; SF = stomach full; MF = midgut full; HF = hindgut full).

gut microbiome genera and also in this study, including *Vibrio*, *Shewanella*, *Pseudoalteromonas* and *Flavobacteria* (Dishaw et al., 2014).

3.4. Assembly of complete dsDNA viral genomes

A number of recent studies have assembled full-length viral genomes from metagenomic data, largely due to improvements in sequencing and assembly technology (Bolduc et al., 2017; Duhaime and Sullivan, 2012; Nishimura et al., 2017). Here we report 29 circular viral genomes greater than 10 kb (from 18,608–114,157 bp), which we refer to as environmental viral genomes (EVGs) (Table 2) (Nishimura et al., 2017). Uncleared gut compartments yielded the highest number of EVGs (HF: 15; HC: 0; MF: 4; MC: 2; SF: 4; SC: 3); the cleared, entire gut sample from 2014 (SDC 14) yielded only one.

To verify the EVG assemblies, all reads were mapped back to the available circular genomes using Read2RefMapper and BowtieBatch in the iVirus pipeline (Bolduc et al., 2017). Interestingly, the majority of these EVGs encoded ORFs most closely related to "uncultured uvMED phages" sequenced from fosmids constructed from a deep chlorophyll maximum sample from the Mediterranean Sea (Mizuno et al., 2013). Coverage of these EVGs as determined by read mapping varied across samples, with the most abundant (75 x average in cleared and 229 x average in uncleared) being an EVG with shared sequence identity to an uncultured Mediterranean phage uvMED that also included a number of *Endozoicomonas* sp. host genes. Several species of *Endozoicomonas* are among the most abundant bacteria in *Ciona* (Dishaw et al., 2014). Additionally, a circular phage genome (Node 14) that displayed sequence similarity and synteny to *Thalassamonas* phage BA3 (51% aa similarity, BLOSUM62, threshold 0) was present in all compartments of the 2015 samples, but absent in the SDC 14 sample (Fig. 7A). This particular phage, whose host is the known coral pathogen *Thalassamonas loyana*, has been utilized in phage therapy to treat 'white plague' disease in coral (Efrony et al., 2009). The 'Node 13′ contig possessed a number of ORFs similar to *Flavobacterium* phage 11b, which was isolated from Arctic sea ice (Borriss et al., 2007), with an amino acid sequence similarity of 40% (BLOSUM62, threshold 0), mostly matching capsid structural genes (Fig. 7B). The EVGs that did not match the uvMED sample set (total of 12) either had no similar sequences in the database or possessed ORFs related to various *Bacteroidetes* phages, such as *Flavobacteria* and *Cellulophaga* phages. The largest circular contig (Node 1) detected initially within the MF sample (114,157 bp), yielded predicted proteins related to *Cellulophaga* phages, which infect algae-degrading bacteria; and the absence of these sequences in the cleared samples may indicate thorough purging of dietary material.

Table 2
Relative abundance of all circular contigs from the *Ciona* virome and their closest BLASTx match in the NR database as determined by the majority of ORFs. A single FASTA file with all circular genomes is available as a supplemental file, and annotated genomes are available in MetaVir.

Contig	Length (bp)	SC	MC	HC	SDC 14	SF	HF	MF	Closest BLASTx hit
NODE_1	114157	0	0	0	0	4.92	17.94	25.13	uvMED/Cellulophaga/crAss phage
NODE_4	68049	0	2.37	0	0	5.38	37.1	17.89	uncultured marine virus/many genes to Cellulophaga phages
NODE_6	66215	0	2.96	3.03	0	12.11	31.05	37.28	–
NODE_5	66045	1.99	3.3	3.21	0	10.87	18.57	22.51	uncultured marine virus/uncultured Mediterranean phage uvMED
NODE_7	60707	0	0	0	0	1.69	21.1	10.97	uncultured marine virus
NODE_11	39502	12.77	6.28	14.37	0	0	0	0	–
NODE_158282	37496	0	0	0	79.71	0	0	0	uncultured Mediterranean phage uvMED
NODE_14	36718	10.37	6.81	9.1	0	6.54	18.79	4.6	Thalassamonas phage BA3
NODE_26	36261	1.54	2.01	0	0	7.15	22.53	14.54	Flavobacterium phage
NODE_31	33937	3.83	0	3.29	11.75	3.63	10.83	5.29	–
NODE_34	32853	5.11	3.35	3.67	0	17.32	52.03	35.45	uncultured Mediterranean phage uvMED
NODE_20	32650	6.03	46	72.2	0	82.09	328.01	260.49	–
NODE_25	32410	0	0	0	0	3.11	5.62	4.16	uncultured Mediterranean phage uvMED
NODE_39	31550	2.88	2.38	7.29	85.01	7.19	18.12	12.05	–
NODE_13	31113	3.09	3.16	10.62	5.59	11.74	38.19	46.95	Flavobacterium phage 11b/Cellulophaga phage phi 10:1
NODE_29	30928	0	0	0	0	4.51	7.31	9.87	uncultured Mediterranean phage uvMED
NODE_42	30800	0	0	0	0	4.77	15.12	7.11	uncultured Mediterranean phage uvMED/Burkholderia
NODE_44	30535	78.38	88.92	58.12	0	265.68	224.28	197.3	uncultured Mediterranean phage uvMED/Endozoicomonas
NODE_46	30268	2.44	7.81	7.3	0	55.92	124.66	91.57	uncultured Mediterranean phage uvMED
NODE_33	29858	0	0	0	0	3.56	11.68	10.95	uncultured Mediterranean phage uvMED/Synechococcus
NODE_40	28217	2.22	2.32	0	0	7.04	4.26	10.54	uncultured Mediterranean phage uvMED
NODE_53	28176	0	0	0	0	1.78	10.22	5.57	Flavobacterium phage
NODE_60	27036	1.7	2.9	3.95	0	20.9	146.22	75.22	uncultured Mediterranean phage uvDeep/has fungus gene
NODE_38	24602	4.19	6.5	5.78	3.7	16.23	24.85	41.54	–
NODE_27	24496	5.77	5.28	6.09	3.88	15.86	27.06	46.48	–
NODE_73	24261	0	0	0	0	2.4	18.27	11.94	uncultured Mediterranean phage uvMED
NODE_41	23894	0	8.22	14.26	0	93.19	430.37	292.75	–
NODE_39	19273	13.1	0	0	0	0	0	0	virophage
NODE_54	18608	2.09	0	0	0	10.63	0	0	virophage

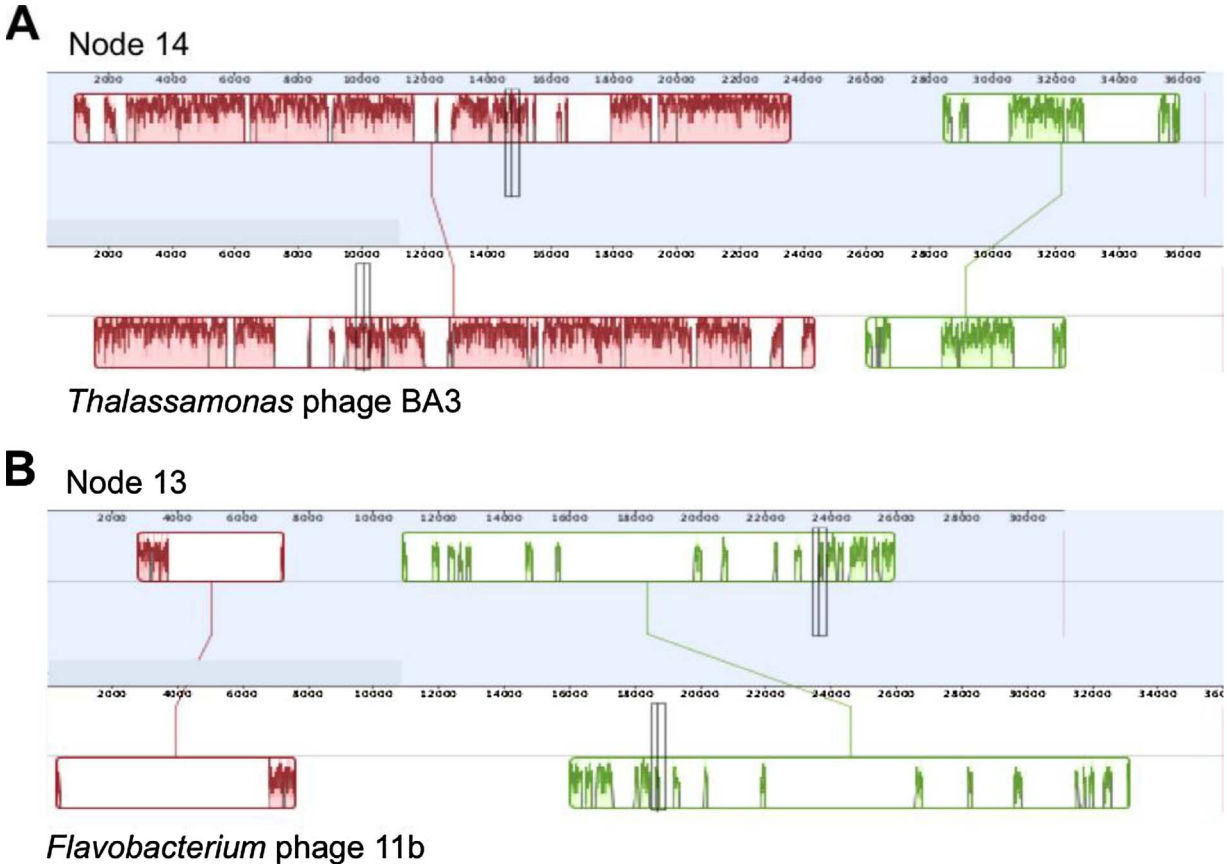

Fig. 7. Whole genome comparison of synteny of two EVGs with their closest BLASTx match in the RefSeq database using Mauve genome alignment. (A) Node 14 to *Thalassamonas* phage BA3. (B) Node 13 to *Flavobacterium* phage 11b. Colored boxes indicate sequence blocks with shared identity, and white regions indicate sequences lacking homology.

This circular contig also possessed two other large ORFs related to an uncultured crAssphage in the NR database, which was originally identified from human fecal material (Dutilh et al., 2014). Reads from all three uncleared compartments matched back to this large circular genome with the most reads recovered from the uncleared midgut compartment (MF) (25 x coverage versus 5 x coverage in SF and 17 x in HF).

Interestingly, two circular contigs (Node 39 at 19,273 bp and Node 54 at 18,608 bp), which both possessed ORFs similar to previously described virophages, were only detected in the stomach (both Nodes present in cleared and only Node 54 in uncleared). Virophages are dsDNA viruses that are dependent on co-infection of another virus, typically a giant virus (Fischer and Suttle, 2011). Node 54 encodes one ORF with similarity to the Sputnik virophage V13 protein (81% of 2517 bp ORF with 28% aa identity) and others similar to the Yellowstone virophage 5 hypothetical protein YSLV5_orf11 (92% of 321 bp ORF with 35% aa identity) and the Qinghai Lake virophage hypothetical protein QLV_03 (71% of 603 bp ORF with 43% aa identity). Node 39 also possessed an ORF similar to the hypothetical protein YSLV_orf11 in the Yellowstone Lake virophage 5 (92% of 321 bp ORF with 34% aa identity) and another similar to the *Phaeocystis globosa* virophage hypothetical protein PGVV_00003 (76% of 816 bp ORF with 32% aa identity). These two virophages possessed 33% amino acid similarity to each other as determined by MUSCLE alignment (BLOSUM62, threshold 0). Although a few genes shared similarity to other known virophages, the majority of these gene sequences shared no sequence identity to other sequences in the databases, suggesting that these are novel virophages.

4. Discussion

Utilizing metagenomics, we report here on the gut virome of the filter-feeding marine protochordate, *Ciona intestinalis* subtype A, a close invertebrate relative of vertebrates. Similar to other studied marine organisms, *Ciona* contains bacterial and viral communities that are distinct from the surrounding environment. In the extensively studied marine environment, phage outnumber bacteria approximately 10:1 (Suttle, 2005). Within the gut mucus environment, phage numbers are estimated to be approximately equal to those of their hosts (Reyes et al., 2012), yet little is known about their ecological and biological roles in this complex and dynamic ecosystem.

The gut ecosystem of many animals encompasses a number of unique physical niches for bacteria, which utilize the various nutrients available in each location (Flint et al., 2008; Leach et al., 1973; Looft et al., 2014; Yang et al., 2005), a phenomenon that we also see in the *Ciona* gut. For example, in most animals, the colon or colonic cecum provides an anaerobic environment housing microbes that often form biofilms within the mucus layers and, aside from the mucins, have access to carbon sources not utilized by members of the foregut (Flint et al., 2008). However, this compartmentalization of the microbiome has only been documented for bacteria, and has not yet been shown for viruses, primarily because most viral studies have only examined fecal content (Minot et al., 2011; Norman et al., 2015; Shan et al., 2011). Fecal material often does not represent a complete picture of the bacteria that inhabit the mucosa (Codling et al., 2010; Zoetendal et al., 2002); the same may be true of viruses. In this study, we found that the stomach possessed the lowest number of viral contigs and bacterial OTUs, possibly related to some biophysical restriction in this compartment (e.g., acidity or enzymatic activity) (Dishaw et al., 2016), while the midgut possessed the largest number of unique viral and bacterial sequences. The midgut also is the smallest compartment in the *Ciona* gut by volume, compared to the stomach and hindgut, yet it possesses microvilli and thick mucus that serve as adequate surfaces and substrates for both microbial colonization and viral attachment (Barr et al., 2013; Yonge, 1924). In addition, the midgut is the segment of the digestive tract where the secretory immune molecule, VCBP-C, known to bind bacteria and influence biofilms (Dishaw et al., 2016), has limited expression (Liberti et al., 2014); this feature perhaps may facilitate proliferation and diversification of bacterial and therefore phage communities. In insects, the midgut has been described as the compartment with the most diversity of bacteria (Rani et al., 2009), including those with metabolic genes necessary for digestion (Gaio Ade et al., 2011). The hindgut possesses the lowest number of unique viral contigs in cleared samples and the highest in the uncleared samples. This region is also the longest of the digestive tract and seems to host more transient microbiota.

The large decrease in the abundance of algal viruses within the cleared gut samples supports a dietary origin for many of these viruses. The algal viruses detected here mostly belonged to the *Phycodnaviridae* family, and more specifically were similar to the *Phaeocystis globosa* virus, and to the *Crysochromulina ericina* virus, which infect dense and noxious bloom-forming species (Simonsen, 1997). Furthermore, large numbers of chloroviruses, viruses that have recently been proposed to affect neural activity in humans (Yolken et al., 2014), were detected in uncleared guts, comprising 12–25% of the reads with similarity to the *Phycodnaviridae* family. The Yellowstone lake mimivirus was also detected at a higher abundance in uncleared versus cleared guts, with the highest abundance in the uncleared midgut, comprising 4% of all dsDNA viral reads from this compartment compared to 0.1% in the cleared midgut. A number of virophages infecting the giant mimivirus have been described (Zhou et al., 2015) and quite a few virophages, some with affiliations to the Yellowstone Lake virophages 5, 6 and 7 as well as one similar to the *Phaeocystis globosa* virophage, were also discovered within the *Ciona* virome, including two full circular genomes.

Each of the circular genomes described here provides further evidence that the virosphere is still largely uncharacterized, and these genomes can serve as templates for comparisons of viromes sequenced in the future. Inadequate reference sequence databases make it difficult to characterize these viruses further. The majority of these genomes are most closely related to the uvMED viruses sequenced from the Mediterranean Sea (Mizuno et al., 2013). While the Mediterranean Sea is considerably distant from the Mission Bay site in San Diego, California where the *Ciona* were collected, some of the uvMED phages were found to be globally distributed in low to medium latitude oceanic metagenomes (Mizuno et al., 2013). In addition, a number of complete genomes similar to phages that infect *Flavobacteria* (and their close relatives) were found within the *Ciona* gut. *Flavobacteria* are highly abundant within aquatic habitats and are well-known for their ability to degrade biopolymers such as cellulose and chitin, which are highly abundant in ocean waters (Kirchman, 2002) and within *Ciona* (Dishaw et al., 2016). Members of the core gut community in *Ciona* (Dishaw et al., 2014) such as *Flavobacteria* may serve as hosts for some viruses described here. Additionally, the observation of more prophages in the cleared guts as determined by read mapping to integrase and excisionase genes may be explained by the stressful process of clearing, which inherently starves the animal to remove dietary content and may cause prophage induction. Alternatively, the added sequencing depth of host-associated viruses once the large numbers of algal viruses were removed could also serve as an explanation.

In this report, we describe the presence of recurring viral sequences in the gut of *Ciona* collected two years part, with distinct viral and bacterial distributions in different compartments of the digestive tract. While this study does not provide a complete picture of the *Ciona* gut virome due to the exclusion of viral contigs < 5 kb and the small sample size, data reported here serve as a foundation for understanding viral dynamics within the *Ciona* gut. This study provides an additional perspective on how bacterial and viral communities are distributed within the gut, with an advantage being that *Ciona* is a tractable model system in which future studies can experimentally test hypotheses targeting specific features of gut microbiome interactions. The animal host provides a highly dynamic gut environment through the expression of immune effectors and creation of niches with unique nutrient

profiles. Host-driven influences, combined with interspecific bacterial interactions (e.g., cooperation and/or competition) and viral infection are all important factors in determining the health and function of the metaorganism.

Conceptually, the idea that a gut virome plays important roles within the host microbiome is not new (Breitbart et al., 2003; Turnbaugh et al., 2007; Haynes and Rohwer, 2011; Minot et al., 2011), but the difficulties of studying phage biology *in vivo* has largely hindered these whole-system studies, despite recent successes (Majewska et al., 2015; Maura et al., 2012; Reyes et al., 2013). Phages have been described, *in vitro,* as providing a form of indirect immunity on mucosal surfaces to modulate bacterial settlement and colonization (Barr et al., 2013). Additionally, many *in vitro* experiments have described a role for phages in bacterial biofilm formation (Carrolo et al., 2010; Godeke et al., 2011; Leigh et al., 2017; Nanda et al., 2015; Rice et al., 2009; Tan et al., 2015), a lifestyle often exhibited by resident bacteria within mucus layers. Lysogeny is known to be a prevalent viral state within the gut of healthy humans (Minot et al., 2011; Minot et al., 2013; Reyes et al., 2013). By influencing gene expression or by lysis, prophages have the capacity to influence the outcome of bacterial settlement and biofilm formation during colonization (Carrolo et al., 2010; Godeke et al., 2011; Nanda et al., 2015; Rice et al., 2009; Wang et al., 2010). The presence of prophage hallmark genes, such as integrases and excisionases, which share sequence identity with core gut bacteria in *Ciona* implies that prophages could be serving important roles in the establishment or maintenance of these stable bacterial communities. The utilization of *Ciona* as a model system in future studies will indeed shed light on the role of phages in shaping gut homeostasis.

Data availability

Viral data were deposited in both MetaVir (IDs 7811 (SF), 8143 (SC), 7815 (MF), 7814 (MC), 7812 (HF), 7910 (HC), 7816 (CB), 7819 (MB), and 8255 (SDC14)) and MG-RAST (IDs 4707275.3 (SF), 4707280.3 (SC), 4707282.3 (MF), 4734670.3 (MC), 4707277.3 (HF), 4707278.3 (HC), 4707281.3 (CB), 4707279.3 (MB), 4734673.3 (SDC14)). Bacterial data were deposited in MG-RAST (IDs 4751381.3 (SF), 4751379.3 (SC), 4751382.3 (MF), 4751386.3 (MC), 4751377.3 (HF), and 4751377.3 (HC)).

Funding

These studies were supported by a grant from the National Science Foundation (IOS-1456301) to L.J.D. and M.B. and by a National Science Foundation Graduate Research Fellowship (Award No. 1144244) to B.A.L.

Acknowledgements

The authors would like to thank Matthew Sullivan and his group for training on the iVirus pipeline as well as Ben Bolduc in particular for his constant help in navigating iVirus. The authors would also like to thank Ryan Schenck for his digital illustration of the *Ciona* anatomy.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.virusres.2017.11.015.

References

Abeles, S.R., Pride, D.T., 2014. Molecular bases and role of viruses in the human microbiome. J. Mol. Biol. 426 (23), 3892–3906.

Barbeyron, T., Lerat, Y., Sassi, J.F., Le Panse, S., Helbert, W., Collen, P.N., 2011. Persicivirga ulvanivorans sp. nov., a marine member of the family *Flavobacteriaceae* that degrades ulvan from green algae. Int. J. Syst. Evol. Microbiol. 61 (8), 1899–1905.

Barr, J.J., Auro, R., Furlan, M., Whiteson, K.L., Erb, M.L., Pogliano, J., Stotland, A., Wolkowicz, R., Cutting, A.S., Doran, K.S., Salamon, P., Youle, M., Rohwer, F., 2013. Bacteriophage adhering to mucus provide a non-host-derived immunity. Proc. Natl. Acad. Sci. U. S. A. 110 (26), 10771–10776.

Bolduc, B., Youens-Clark, K., Roux, S., Hurwitz, B.L., Sullivan, M.B., 2017. iVirus: facilitating new insights in viral ecology with software and community data sets imbedded in a cyberinfrastructure. ISME J. 11 (1), 7–14.

Borriss, M., Lombardot, T., Glockner, F.O., Becher, D., Albrecht, D., Schweder, T., 2007. Genome and proteome characterization of the psychrophilic *Flavobacterium* bacteriophage 11b. Extremophiles 11 (1), 95–104.

Breitbart, M., Hewson, I., Felts, B., Mahaffy, J.M., Nulton, J., Salamon, P., Rohwer, F., 2003. Metagenomic analyses of an uncultured viral community from human feces. J. Bacteriol. 185 (20), 6220–6223.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods 13 (7), 581–583.

Carrolo, M., Frias, M.J., Pinto, F.R., Melo-Cristino, J., Ramirez, M., 2010. Prophage spontaneous activation promotes DNA release enhancing biofilm formation in *Streptococcus pneumoniae*. PLoS One 5 (12), e15678.

Chen, H., Boutros, P.C., 2011. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinform. 12, 35.

Cho, I., Blaser, M.J., 2012. The human microbiome: at the interface of health and disease. Nat. Rev. Genet. 13 (4), 260–270.

Codling, C., O'Mahony, L., Shanahan, F., Quigley, E.M., Marchesi, J.R., 2010. A molecular analysis of fecal and mucosal bacterial communities in irritable bowel syndrome. Dig. Dis. Sci. 55 (2), 392–397.

Cohen, L.J., Kang, H.S., Chu, J., Huang, Y.H., Gordon, E.A., Reddy, B.V., Ternei, M.A., Craig, J.W., Brady, S.F., 2015. Functional metagenomic discovery of bacterial effectors in the human microbiome and isolation of commendamide, a GPCR G2A/132 agonist. Proc. Natl. Acad. Sci. U. S. A. 112 (35), E4825–4834.

Coleman, N., 1991. Encyclopedia of Marine Animals. Blandford, London.

Cullender, T.C., Chassaing, B., Janzon, A., Kumar, K., Muller, C.E., Werner, J.J., Angenent, L.T., Bell, M.E., Hay, A.G., Peterson, D.A., Walter, J., Vijay-Kumar, M., Gewirtz, A.T., Ley, R.E., 2013. Innate and adaptive immunity interact to quench microbiome flagellar motility in the gut. Cell Host Microbe 14 (5), 571–581.

Darling, A.E., Mau, B., Perna, N.T., 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5 (6), e11147.

Davies, E.V., James, C.E., Williams, D., O'Brien, S., Fothergill, J.L., Haldenby, S., Paterson, S., Winstanley, C., Brockhurst, M.A., 2016. Temperate phages both mediate and drive adaptive evolution in pathogen biofilms. Proc. Natl. Acad. Sci. U. S. A. 113 (29), 8266–8271.

Delcher, A.L., Bratke, K.A., Powers, E.C., Salzberg, S.L., 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23 (6), 673–679.

Dishaw, L.J., Giacomelli, S., Melillo, D., Zucchetti, I., Haire, R.N., Natale, L., Russo, N.A., De Santis, R., Litman, G.W., Pinto, M.R., 2011. A role for variable region-containing chitin-binding proteins (VCBPs) in host gut-bacteria interactions. Proc. Natl. Acad. Sci. U. S. A. 108 (40), 16747–16752.

Dishaw, L.J., Flores-Torres, J., Lax, S., Gemayel, K., Leigh, B., Melillo, D., Mueller, M.G., Natale, L., Zucchetti, I., De Santis, R., Pinto, M.R., Litman, G.W., Gilbert, J.A., 2014. The gut of geographically disparate *Ciona intestinalis* harbors a core microbiota. PLoS One 9 (4), e93386.

Dishaw, L.J., Leigh, B., Cannon, J.P., Liberti, A., Mueller, M.G., Skapura, D.P., Karrer, C.R., Pinto, M.R., De Santis, R., Litman, G.W., 2016. Gut immunity in a protochordate involves a secreted immunoglobulin-type mediator binding host chitin and bacteria. Nat. Commun. 7, 10617.

Duhaime, M.B., Sullivan, M.B., 2012. Ocean viruses: rigorously evaluating the metagenomic sample-to-sequence pipeline. Virology 434 (2), 181–186.

Dutilh, B.E., Cassman, N., McNair, K., Sanchez, S.E., Silva, G.G., Boling, L., Barr, J.J., Speth, D.R., Seguritan, V., Aziz, R.K., Felts, B., Dinsdale, E.A., Mokili, J.L., Edwards, R.A., 2014. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. Nat. Commun. 5, 4498.

Efrony, R., Atad, I., Rosenberg, E., 2009. Phage therapy of coral white plague disease: properties of phage BA3. Curr. Microbiol. 58 (2), 139–145.

Feschotte, C., Gilbert, C., 2012. Endogenous viruses: insights into viral evolution and impact on host biology. Nat. Rev. Genet. 13 (4), 283–296.

Fischer, M.G., Suttle, C.A., 2011. A virophage at the origin of large DNA transposons. Science 332 (6026), 231–234.

Flint, H.J., Duncan, S.H., Scott, K.P., Louis, P., 2007. Interactions and competition within the microbial community of the human colon: links between diet and health. Environ. Microbiol. 9 (5), 1101–1111.

Flint, H.J., Bayer, E.A., Rincon, M.T., Lamed, R., White, B.A., 2008. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. Nat. Rev. Microbiol. 6 (2), 121–131.

Gaio Ade, O., Gusmao, D.S., Santos, A.V., Berbert-Molina, M.A., Pimenta, P.F., Lemos, F.J., 2011. Contribution of midgut bacteria to blood digestion and egg production in *Aedes aegypti* (Diptera: *Culicidae*) (L.). Parasit. Vectors 4, 105.

Godeke, J., Paul, K., Lassak, J., Thormann, K.M., 2011. Phage-induced lysis enhances biofilm formation in *Shewanella oneidensis* MR-1. ISME J. 5 (4), 613–626.

Grasis, J.A., Lachnit, T., Anton-Erxleben, F., Lim, Y.W., Schmieder, R., Fraune, S., Franzenburg, S., Insua, S., Machado, G., Haynes, M., Little, M., Kimble, R., Rosenstiel, P., Rohwer, F.L., Bosch, T.C., 2014. Species-specific viromes in the ancestral holobiont Hydra. PLoS One 9 (10), e109952.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29 (8), 1072–1075.

Haynes, M., Rohwer, F., 2011. The Human Virome. Metagenomics of the Human Body. pp. 63–77.

Holmfeldt, K., Solonenko, N., Shah, M., Corrier, K., Riemann, L., Verberkmoes, N.C., Sullivan, M.B., 2013. Twelve previously unknown phage genera are ubiquitous in global oceans. Proc. Natl. Acad. Sci. U. S. A. 110 (31), 12798–12803.
Kang, I., Jang, H., Cho, J.C., 2012. Complete genome sequences of two *Persicivirga* bacteriophages, P12024S and P12024L. J. Virol. 86 (16), 8907–8908.
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28 (12), 1647–1649.
Keegan, K.P., Glass, E.M., Meyer, F., 2016. MG-RAST: a metagenomics service for analysis of microbial community structure and function. Methods Mol. Biol. 1399, 207–233.
Kim, K.H., Bae, J.W., 2011. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. Appl. Environ. Microbiol. 77 (21), 7663–7668.
Kirchman, D.L., 2002. The ecology of *Cytophaga-Flavobacteria* in aquatic environments. FEMS Microbiol. Ecol. 39 (2), 91–100.
Kwan, T., Liu, J., Dubow, M., Gros, P., Pelletier, J., 2006. Comparative genomic analysis of 18 *Pseudomonas aeruginosa* bacteriophages. J. Bacteriol. 188 (3), 1184–1187.
Leach, W.D., Lee, A., Stubbs, R.P., 1973. Localization of bacteria in the gastrointestinal tract: a possible explanation of intestinal spirochaetosis. Infect. Immun. 7 (6), 961–972.
Leigh, B.A., Liberti, A., Dishaw, L.J., 2016. Generation of germ-free *Ciona intestinalis* for studies of gut-microbe interactions. Front. Microbiol. 7, 2092.
Leigh, B., Karrer, C., Cannon, J.P., Breitbart, M., Dishaw, L.J., 2017. Isolation and characterization of a *Shewanella* phage-host system from the gut of the tunicate, *Ciona intestinalis*. Viruses 9 (3), 60.
Liberti, A., Melillo, D., Zucchetti, I., Natale, L., Dishaw, L.J., Litman, G.W., De Santis, R., Pinto, M.R., 2014. Expression of *Ciona intestinalis* variable region-containing chitin-binding proteins during development of the gastrointestinal tract and their role in host-microbe interactions. PLoS One 9 (5), e94984.
Looft, T., Allen, H.K., Cantarel, B.L., Levine, U.Y., Bayles, D.O., Alt, D.P., Henrissat, B., Stanton, T.B., 2014. Bacteria, phages and pigs: the effects of in-feed antibiotics on the microbiome at different gut locations. ISME J. 8 (8), 1566–1576.
Majewska, J., Beta, W., Lecion, D., Hodyra-Stefaniak, K., Klopot, A., Kazmierczak, Z., Miernikiewicz, P., Piotrowicz, A., Ciekot, J., Owczarek, B., Kopciuch, A., Wojtyna, K., Harhala, M., Makosa, M., Dabrowska, K., 2015. Oral application of T4 phage induces weak antibody production in the gut and in the blood. Viruses 7 (8), 4783–4799.
Marcais, G., Kingsford, C., 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27 (6), 764–770.
Maura, D., Morello, E., du Merle, L., Bomme, P., Le Bouguenec, C., Debarbieux, L., 2012. Intestinal colonization by enteroaggregative *Escherichia coli* supports long-term bacteriophage replication in mice. Environ. Microbiol. 14 (8), 1844–1854.
Minot, S., Sinha, R., Chen, J., Li, H., Keilbaugh, S.A., Wu, G.D., Lewis, J.D., Bushman, F.D., 2011. The human gut virome: inter-individual variation and dynamic response to diet. Genome Res. 21 (10), 1616–1625.
Minot, S., Bryson, A., Chehoud, C., Wu, G.D., Lewis, J.D., Bushman, F.D., 2013. Rapid evolution of the human gut virome. Proc. Natl. Acad. Sci. U. S. A. 110 (30), 12450–12455.
Mizuno, C.M., Rodriguez-Valera, F., Kimes, N.E., Ghai, R., 2013. Expanding the marine virosphere using metagenomics. PLoS Genet. 9 (12), e1003987.
Nanda, A.M., Thormann, K., Frunzke, J., 2015. Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. J. Bacteriol. 197 (3), 410–419.
Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., Pettersson, S., 2012. Host-gut microbiota metabolic interactions. Science 336 (6086), 1262–1267.
Nishimura, Y., Watai, H., Honda, T., Mihara, T., Omae, K., Roux, S., Blanc-Mathieu, R., Yamamoto, K., Hingamp, P., Sako, Y., Sullivan, M.B., Goto, S., Ogata, H., Yoshida, T., 2017. Environmental viral genomes shed new light on virus-host interactions in the ocean. mSphere 2 (2), e00359–16.
Norman, J.M., Handley, S.A., Baldridge, M.T., Droit, L., Liu, C.Y., Keller, B.C., Kambal, A., Monaco, C.L., Zhao, G., Fleshner, P., Stappenbeck, T.S., McGovern, D.P., Keshavarzian, A., Mutlu, E.A., Sauk, J., Gevers, D., Xavier, R.J., Wang, D., Parkes, M., Virgin, H.W., 2015. Disease-specific alterations in the enteric virome in inflammatory bowel disease. Cell 160 (3), 447–460.
Parada, A.E., Needham, D.M., Fuhrman, J.A., 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ. Microbiol. 18 (5), 1403–1414.
Patel, A., Noble, R.T., Steele, J.A., Schwalbach, M.S., Hewson, I., Fuhrman, J.A., 2007. Virus and prokaryote enumeration from planktonic aquatic environments by epifluorescence microscopy with SYBR Green I. Nat. Protoc. 2 (2), 269–276.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glockner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41 (Database issue), D590–596.
R Core Team, 2014. A language and environment for statistical computing. R Foundation for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rani, A., Sharma, A., Rajagopal, R., Adak, T., Bhatnagar, R.K., 2009. Bacterial diversity analysis of larvae and adult midgut microflora using culture-dependent and culture-independent methods in lab-reared and field-collected *Anopheles stephensi*-an Asian malarial vector. BMC Microbiol. 9, 96.
Reyes, A., Haynes, M., Hanson, N., Angly, F.E., Heath, A.C., Rohwer, F., Gordon, J.I., 2010. Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature 466 (7304), 334–338.
Reyes, A., Semenkovich, N.P., Whiteson, K., Rohwer, F., Gordon, J.I., 2012. Going viral: next-generation sequencing applied to phage populations in the human gut. Nat. Rev. Microbiol. 10 (9), 607–617.
Reyes, A., Wu, M., McNulty, N.P., Rohwer, F.L., Gordon, J.I., 2013. Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut. Proc. Natl. Acad. Sci. U. S. A. 110 (50), 20236–20241.
Rice, S.A., Tan, C.H., Mikkelsen, P.J., Kung, V., Woo, J., Tay, M., Hauser, A., McDougald, D., Webb, J.S., Kjelleberg, S., 2009. The biofilm life cycle and virulence of *Pseudomonas aeruginosa* are dependent on a filamentous prophage. ISME J. 3 (3), 271–282.
Roeselers, G., Mittge, E.K., Stephens, W.Z., Parichy, D.M., Cavanaugh, C.M., Guillemin, K., Rawls, J.F., 2011. Evidence for a core gut microbiota in the zebrafish. ISME J. 5 (10), 1595–1608.
Rohwer, F., Seguritan, V., Azam, F., Knowlton, N., 2002. Diversity and distribution of coral-associated bacteria. Mar. Ecol. Prog. Ser. 243, 1–10.
Roossinck, M.J., 2011. The good viruses: viral mutualistic symbioses. Nat. Rev. Microbiol. 9 (2), 99–108.
Rosenberg, E., Zilber-Rosenberg, I., 2011. Symbiosis and development: the hologenome concept. Birth Defects Res. C Embryo Today 93 (1), 56–66.
Roux, S., Tournayre, J., Mahul, A., Debroas, D., Enault, F., 2014. Metavir 2: new tools for viral metagenome comparison and assembled virome analysis. BMC Bioinform. 15, 76.
Roux, S., Enault, F., Hurwitz, B.L., Sullivan, M.B., 2015. VirSorter: mining viral signal from microbial genomic data. PeerJ 3, e985.
Sabree, Z.L., Hansen, A.K., Moran, N.A., 2012. Independent studies using deep sequencing resolve the same set of core bacterial species dominating gut communities of honey bees. PLoS One 7 (7), e41250.
Sandaa, R.A., Heldal, M., Castberg, T., Thyrhaug, R., Bratbak, G., 2001. Isolation and characterization of two viruses with large genome size infecting *Chrysochromulina ericina* (*Prymnesiophyceae*) and *Pyramimonas orientalis* (*Prasinophyceae*). Virology 290 (2), 272–280.
Schmitt, S., Tsai, P., Bell, J., Fromont, J., Ilan, M., Lindquist, N., Perez, T., Rodrigo, A., Schupp, P.J., Vacelet, J., Webster, N., Hentschel, U., Taylor, M.W., 2012. Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. ISME J. 6 (3), 564–576.
Shan, T., Li, L., Simmonds, P., Wang, C., Moeser, A., Delwart, E., 2011. The fecal virome of pigs on a high-density farm. J. Virol. 85 (22), 11697–11708.
Simonsen, S.a.O.M., 1997. Toxicity tests in eight species of *Chrysochromulina* (Haptophyta). Can. J. Bot. 75 (1), 129–136.
Sobhani, I., Tap, J., Roudot-Thoraval, F., Roperch, J.P., Letulle, S., Langella, P., Corthier, G., Tran Van Nhieu, J., Furet, J.P., 2011. Microbial dysbiosis in colorectal cancer (CRC) patients. PLoS One 6 (1), e16393.
Suttle, C.A., 2005. Viruses in the sea. Nature 437 (7057), 356–361.
Tamboli, C.P., Neut, C., Desreumaux, P., Colombel, J.F., 2004. Dysbiosis as a prerequisite for IBD. Gut 53 (7), 1057.
Tan, D., Dahl, A., Middelboe, M., 2015. Vibriophages differentially influence biofilm formation by *Vibrio anguillarum* strains. Appl. Environ. Microbiol. 81 (13), 4489–4497.
Thaiss, C.A., Zmora, N., Levy, M., Elinav, E., 2016. The microbiome and innate immunity. Nature 535 (7610), 65–74.
Theis, K.R., Dheilly, N.M., Klassen, J.L., Brucker, R.M., Baines, J.F., Bosch, T.C., Cryan, J.F., Gilbert, S.F., Goodnight, C.J., Lloyd, E.A., Sapp, J., Vandenkoornhuyse, P., Zilber-Rosenberg, I., Rosenberg, E., Bordenstein, S.R., 2016. Getting the hologenome concept right: an eco-evolutionary framework for hosts and their microbiomes. mSystems 1 (2), e00028–16.
Thurber, R.V., Haynes, M., Breitbart, M., Wegley, L., Rohwer, F., 2009. Laboratory procedures to generate viral metagenomes. Nat. Protoc. 4 (4), 470–483.
Thurber, R.V., Payet, J.P., Thurber, A.R., Correa, A.M., 2017. Virus-host interactions and their roles in coral reef health and disease. Nature Rev. Microbiol. 15 (4), 205–216.
Tremaroli, V., Backhed, F., 2012. Functional interactions between the gut microbiota and host metabolism. Nature 489 (7415), 242–249.
Turnbaugh, P.J., Gordon, J.I., 2009. The core gut microbiome, energy balance and obesity. J. Physiol. 587 (17), 4153–4158.
Turnbaugh, P.J., Ley, R.E., Hamady, M., Fraser-Liggett, C.M., Knight, R., Gordon, J.I., 2007. The human microbiome project. Nature 449 (7164), 804–810.
Wang, X., Kim, Y., Ma, Q., Hong, S.H., Pokusaeva, K., Sturino, J.M., Wood, T.K., 2010. Cryptic prophages help bacteria cope with adverse environments. Nat. Commun. 1, 147.
Weisburg, W.G., Barns, S.M., Pelletier, D.A., Lane, D.J., 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. 173 (2), 697–703.
Yang, H., Schmitt-Wagner, D., Stingl, U., Brune, A., 2005. Niche heterogeneity determines bacterial community structure in the termite gut (*Reticulitermes santonensis*). Environ. Microbiol. 7 (7), 916–932.
Yolken, R.H., Jones-Brando, L., Dunigan, D.D., Kannan, G., Dickerson, F., Severance, E., Sabunciyan, S., Talbot Jr., C.C., Prandovszky, E., Gurnon, J.R., Agarkova, I.V., Leister, F., Gressitt, K.L., Chen, O., Deuber, B., Ma, F., Pletnikov, M.V., Van Etten, J.L., 2014. Chlorovirus ATCV-1 is part of the human oropharyngeal virome and is associated with changes in cognitive functions in humans and mice. Proc. Natl. Acad. Sci. U. S. A. 111 (45), 16106–16111.
Yonge, C.M., 1924. Studies on the comparative physiology of digestion. J. Exp. Biol. 1 (3), 343–389.
Zhou, J., Sun, D., Childers, A., McDermott, T.R., Wang, Y., Liles, M.R., 2015. Three novel virophage genomes discovered from Yellowstone Lake metagenomes. J. Virol. 89 (2), 1278–1285.
Zoetendal, E.G., von Wright, A., Vilpponen-Salmela, T., Ben-Amor, K., Akkermans, A.D., de Vos, W.M., 2002. Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. Appl. Environ. Microbiol. 68 (7), 3401–3407.