

Article

Revealing the Viral Community in the Hadal Sediment of the New Britain Trench

Hui Zhou ^{1,2,†}, Ping Chen ^{1,†}, Mengjie Zhang ^{1,2}, Jiawang Chen ^{3,*} , Jiasong Fang ^{4,*} and Xuan Li ^{1,2,*}

¹ Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China; zhouhui@sippe.ac.cn (H.Z.); pchen@ips.ac.cn (P.C.); zhangmengjie@sippe.ac.cn (M.Z.)

² University of Chinese Academy of Sciences, Beijing 100049, China

³ Ocean College, Zhejiang University, Zhoushan 316021, China

⁴ Shanghai Engineering Research Center of Hadal Science and Technology, College of Marine Sciences, Shanghai Ocean University, Shanghai 201306, China

* Correspondence: arwang@zju.edu.cn (J.C.); jsfang@shou.edu.cn (J.F.); LiXuan@sippe.ac.cn (X.L.)

† These authors contributed equally to this work.

Abstract: Marine viruses are widely distributed and influence matter and energy transformation in ecosystems by modulating hosts' metabolism. The hadal trenches represent the deepest marine habitat on Earth, for which the viral communities and related biogeochemical functions are least explored and poorly understood. Here, using the sediment samples (8720 m below sea level) collected from the New Britain Trench (NBT), we investigated the viral community, diversity, and genetic potentials in the hadal sediment habitat for the first time by deep shotgun metagenomic sequencing. We found the NBT sediment viral community was dominated by Siphoviridae, Myoviridae, Podoviridae, Mimiviridae, and Phycodnaviridae, which belong to the dsDNA viruses. However, the large majority of them remained uncharacterized. We found the hadal sediment virome had some common components by comparing the hadal sediment viruses with those of hadal aquatic habitats and those of bathypelagic and terrestrial habitats. It was also distinctive in community structure and had many novel viral clusters not associated with the other habitual virome included in our analyses. Further phylogenetic analysis on its Caudovirales showed novel diversities, including new clades specially evolved in the hadal sediment habitat. Annotation of the NBT sediment viruses indicated the viruses might influence microbial hydrocarbon biodegradation and carbon and sulfur cycling via metabolic augmentation through auxiliary metabolic genes (AMGs). Our study filled in the knowledge gaps on the virome of the hadal sediment habitats and provided insight into the evolution and the potential metabolic functions of the hadal sediment virome.

Keywords: New British Trench; hadal sediment; metagenomics; virus



Citation: Zhou, H.; Chen, P.; Zhang, M.; Chen, J.; Fang, J.; Li, X. Revealing the Viral Community in the Hadal Sediment of the New Britain Trench. *Genes* **2021**, *12*, 990. <https://doi.org/10.3390/genes12070990>

Academic Editor: Yan Guo

Received: 5 May 2021

Accepted: 9 June 2021

Published: 29 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The hadal trenches are the deepest part of the ocean, accounting for ~1.5% of the global deep-sea floor and covering ~45% of the oceanic depth range [1]. With unique tectonic, topographic, bathymetric, and hydrographic features, trenches have recently become a hotspot for deep-sea microbial research. Currently, culture-independent high-throughput sequencing and culture-dependent taxonomic characterization have provided insights into the microbial communities and functions of seawater or sediments in hadal environments. Logan et al. compared the microbial community similarities and differences between the Kermadec Trench and the Mariana Trench by amplicon sequencing and culture [2]. Liu et al. characterized the composition and metabolic potential of particle-attached (PA) and free-living (FL) bacterial communities in the Marianas Trench by high-throughput sequencing [3]. Despite evidence that benthic prokaryotes and viruses in the surface centimeters of deep-sea sediments account for 50–80% of the total microbial life on Earth

and 10–30% of the total biota [4,5], their diversity and metabolic potential in these extreme environments remain largely unknown. Even less information is available on the role of viruses in this ecosystem and their interactions with their prokaryotic hosts.

Globally, viruses are the most diverse group in the marine ecosystem and are extremely abundant, tallying around 4×10^{30} [6], with phages estimated to be 5 to 25 times more abundant than bacteria in the surface layer of the ocean [7]. Viruses influence community turnover and resource availability through a range of interactions with their hosts [8]. Viruses can regulate the number of microbial populations by randomly adsorbing, infecting, and then lysing the dominant host bacteria in the microbial community to proliferate, causing some kind of lesion and death of the host cells. They control the overpopulation of the dominant species and thereby regulate and maintain the structure of the microbial community and the balance of species diversity [9]. During infection, the virus may also affect key steps of cellular processes such as metabolism, division, clustered regularly interspaced short palindromic repeat-CRISPR-associated protein (CRISPR-Cas) immunity, motility, and regulation [10]. In deep-sea sediments, viral infestation reduces prokaryotic microbial heterotrophic products by 80%, releasing approximately 0.37–0.63 gigatons of carbon per year globally, suggesting that viruses play an important role in deep-sea ecosystems [11]. However, little has been reported on the diversity, genetic potential, and biogeochemical impact of viruses in the Hadal Trench sediments at depths greater than 6000 m. Elisabetta et al. examined the distribution, activity, and interactions of viruses and prokaryotes in surface sediments at the bottom of the Japanese, Izu-Ogasawara, and Marianas trenches and nearby abyssal plain sites by physicochemical means [12]. However, limited by methodological means, a veil remains over abyssal virus-host interactions, genetic diversities, and biogeochemical impacts on the hadal environment.

The New Britain Trench (NBT), nearing the land of Papua New Guinea, is located in the northern part of the Solomon Sea [13]. It reaches the deepest depth of ~9140 m at Planet Deep [14]. The special geographical location of NBT makes the trench one of the most frequently seismic places in the world. Compared with the Marianas Trench, the NBT is close to land and can receive more heterogeneous input from both land and ocean. Correspondingly, the waters covered by the New Britain Trench have a higher net primary productivity (about three times higher than the Marianas Trench) [13]. Previous studies have investigated the microbiome within the New Britain Trench waters and sediments [15–18]. However, no research has been focused on the hadal viral community in the sediment of the New Britain Trench. Particularly, questions about viral components and their metabolic functions in the New Britain Trench sediment remain unanswered. Thus, the hadal viral community in the sediment of the New Britain Trench is still poorly understood. Here, we adopted a deep metagenomics approach to investigate the hadal virome in the sediment of the New Britain Trench (8720 m). We have investigated the main viral components and the versatile metabolic functions of the NBT virome. Extensive comparative analysis with virome in hadal aquatic habitats, bathypelagic and terrestrial habitats were also conducted.

2. Methods

2.1. Sampling Site Descriptions and Sample Collection

Sediment samples were collected at a depth of 8720 m from the New Britain Trench (154°32.2912' E, 6°14.4920' S) in January 2017 during a cruise aboard M/V Zhang Jian. Sediment cores were collected using a box corer (with a base area of 400 cm² and a height of 25 cm) attached to the hadal lander-II, developed by Shanghai Engineering Research Center of Hadal Science and Technology, Shanghai Ocean University. The samples were either frozen at −80 °C before stored at −20 °C or stored at 4 °C for culturing and isolation of microbes.

2.2. DNA Extraction, Library Construction, and Sequencing

The column sediments were divided into three layers, the surficial layer N1 (sediment depth 1–5 cm), the mid-layer N2 (5–10 cm), and the deep layer N3 (10–15 cm). According to the standard manufacturer protocol, about 10 g of sediment sample were used to extract DNA for each sediment layer, using the MoBio PowerSoil DNA Isolation kit (MO BIO Laboratories, Carlsbad, CA, USA). Three DNA extractions were performed on every sediment sample and were then mixed for Illumina library construction, which minimizes the variation between different extractions. Two Illumina libraries were built for the extraction mixture of each sediment layer, and two separate runs sequenced each library. The first three Illumina libraries of the three samples were sequenced in one HiSeq run and the second sequence for another run to overcome the variation between different sequencing runs. Notably, PCR amplification was limited to 12 cycles for each Illumina library. The quality and quantity of genomic DNA were examined by an Agilent Bioanalyzer 2100 (Agilent Technologies, Redwood City, CA, USA). Metagenomic shotgun sequencing was performed on the HiSeq Xten instruments (Illumina, San Diego, CA, USA), with 2×150 bp paired-end reads.

2.3. Sequence Processing and Assembly

Raw sequence data were first processed using Trimmomatic (version 0.38) to remove adapter sequences and low-quality bases (Q20) [19,20]. Putative contaminated sequences, like plasmid sequences, human sequences, etc., were removed with bowtie2 (v2.3.4.1) [21]. Duplicated reads generated by PCR amplification were removed using Fastuniq [22]. The final cleaned reads were assembled using MEGAHIT (v1.1.3) [23] with the following parameters: `—k-list 21, 29, 39, 59, 79, 99, 119, 127, 139`.

2.4. Taxonomic Classification of Reads and Determination of Relative Sequence Abundance

The reference database compiled in a GORG-Tropics format, including archaea, bacteria, microbial eukaryotes, and viruses, was downloaded. Taxonomic classification of the sequencing reads was performed using Kaiju (v1.7.0) [24] with greedy-5 mode against the reference database. Reads were assigned with the taxonomic id and functional annotations based on mapped references. The relative sequence abundance of a phylum, class, order, etc., was assessed by summarizing the total number of the corresponding category assigned reads and dividing it with the total number of assigned reads of all the categories.

2.5. Recovering and Annotating Viral Contigs

In order to recover high confidence virus sequences from the assembled contig, we integrated the results of multiple virus software identifications. VirFinder(v1.1) [25], VirSorter (v1.0.5) [26], and VirSorter (v2.1) [27] were employed based on the following criteria: (1) VirFinder score ≥ 0.9 and $p < 0.05$; (2) VirSorter (v1.0.5) categories 1, 2, 4 and 5; (3) VirSorter (v2.1) score > 75 ; (4) both identified by VirSorter (v1.0.5) categories 1–6 and VirFinder score ≥ 0.7 and $p < 0.05$; (5) contig length > 1 k bp. The final recovered viral contigs were analyzed with the taxonomic classification program Kaiju (v1.7.0) with default parameters [24] to get their taxonomic lineage information.

2.6. Comparisons to Viral Sequences from Other Environments and Data Sets

To compare NBT sediment viruses ($n = 40,267$) with other habitats, four ecological environmental virus data included: (1) GOV 2.0 seawater [28] ($n = 310,218$); (2) wetland sediment [29] ($n = 3344$); (3) Stordalen thawing permafrost [30] ($n = 1907$); (4) cold springs [31] ($n = 2885$). For each viral contig, open reading frames (ORFs) were called using Prodigal v2.6.3 [21] and the predicted protein sequences were used as input for vConTACT2 [32]. We followed the protocol published in protocols.io (<https://www.protocols.io/view/applying-vcontact-to-viralsequences-and-visualizi-x5xfq7n> (version 5.0, accessed on 1 April 2021)) for the application of vConTACT2 and visualization of the gene-sharing network in Cytoscape v3.5.0 [33] (edge-weighted spring-embedded model). Viral RefSeq

(v95) was selected as the reference database, and Diamond was used for the protein–protein similarity method. Other parameters were set as default.

2.7. Functional Viromics Analysis

For viral function analysis, viral contigs were compared with the eggNOG database (version 5.0, accessed on 26 March 2021) using the eggNOG-mapper (<http://eggno-mapper.embl.de/>) (version 2.0, accessed on 26 March 2021) to do functional annotation. A threshold of 10^{-5} was used to filter the annotation results. Viral ORFs annotated as the CAZymes family genes or categorized into “metabolic pathways” (determined by KEGG kos) in the eggNOG-mapper annotation results were considered as potential auxiliary metabolic genes (AMG) [20,34].

3. Results and Discussion

3.1. Deep Sequencing of the Hadal Microbiome in the NBT Sediments and Identification of Viral Genome/Metavirome

Sediment samples (depth of 8720 m) were collected from the New Britain Trench in January 2017 (Figure 1, see details in methods) to investigate the community structure, genetic potential, and ecological functions of the New Britain Trench (NBT) sediment virus. Deep shotgun metagenomic sequencing was performed on the NBT sediment samples to study the hadal virome (Figure 1). We divided the sediment samples into three depth layers, i.e., the surficial layer NBT-1 (sediment depth 1–5 cm), the mid-layer NBT-2 (sediment depth 5–10 cm), and the deep layer NBT-3 (sediment depth 10–15 cm). The surface layer of 0–1 cm was re-moved to avoid potential environmental contamination. The meta-genome DNA of each sample layer was then extracted with the proto-cols described in Methods (see section DNA extraction, library construction, and sequencing). Two independent DNA sequencing libraries were built for each sample layer. For each DNA sequencing library, a total of 20–30 Gb sequencing data were obtained. Finally, we obtained ~50 Gb sequencing data for each layer of the NBT sediment (see details in Supplementary Table S1).

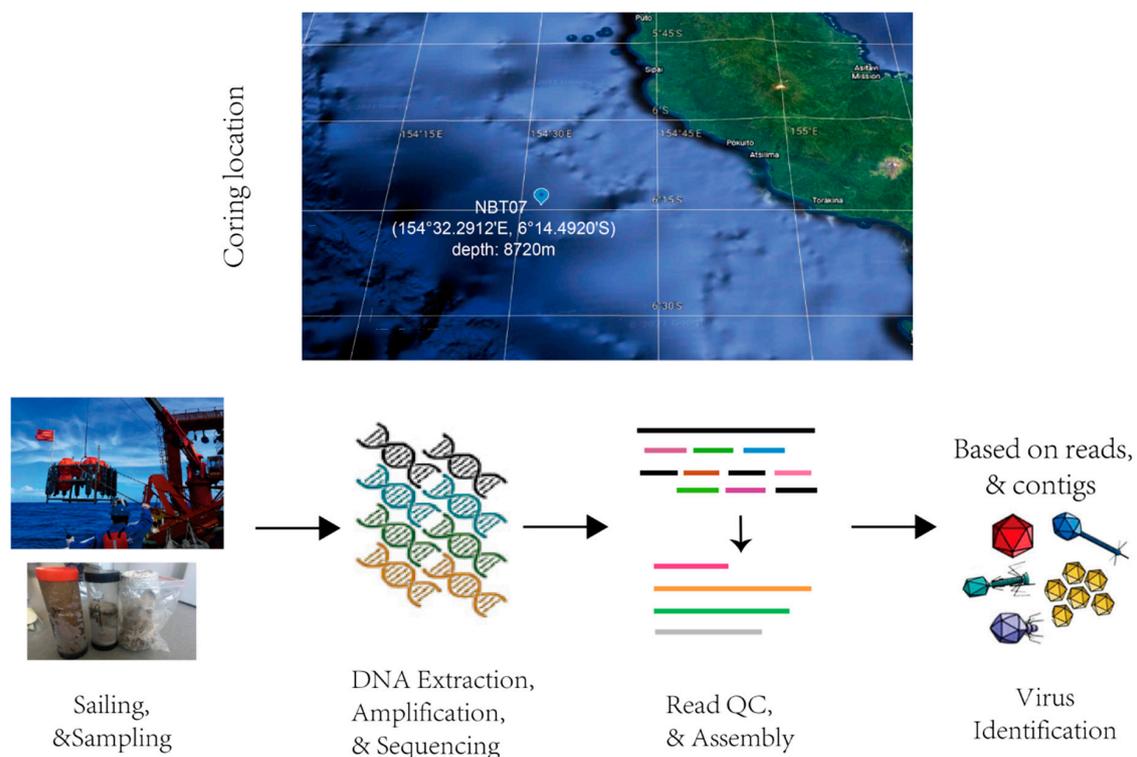


Figure 1. Overview of the study pipeline. Sampling of the hadal sediment in the New Britain Trench (154°32.2912' E, 6°14.4920' S) at a depth of 8720 m. One core was collected in January 2017; the coring location is indicated on the top panel.

Virus identification was based on both the viral reads and the final assembled contigs (Figure 1). After removing adapters and low-quality reads, the sequencing reads were computationally identified and classified by Kaiju [24] based on the least common ancestor (LCA) algorithms. Using the National Center for Biotechnology Information non-redundant (NCBI nr) as the reference database, a total of ~29 M reads were identified and classified as viral sequence fragments. These viral sequences accounted for 0.01% sequence abundance in the NBT sediment microbiome. After co-assembling the total sequencing reads from the samples of the three layers, we integrated three virus contigs classification methods, i.e., VirFinder [25], VirSorter [26], and VirSorter2 [27], to identify the NBT viral contigs (see Methods for details). Then, 714 viral sequences (length >1 kb) were identified by VirSorter, which belonged to categories I, II, IV, and V for “most confident” and “likely” predictions. 13,550 and 29,023 viral sequences (length >1kb) were identified by VirFinder (score ≥ 0.9 and $p < 0.05$) and VirSorter2 (score > 75), respectively. Finally, 40,267 contigs greater than 1 Kb (Supplementary Table S1) were identified as viral contigs in the NBT sediment sample.

3.2. Compositions of the Hadal Viruses in the NBT Sediment Communities

To investigate the viral composition of the NBT sediments, the identified viral reads were further annotated using Kaiju, and the relative abundance was calculated based on the identified viral sequencing reads.

A total of thirty major viral families were identified in the NBT sediment (Figure 2). dsDNA viruses, mainly affiliated with the order Caudovirales (naturally infecting bacteria and archaea), were the most abundant. The most dominant viral family in NBT sediment was Siphoviridae, with an average abundance of 42.9% of the total viral sequences, followed by Myoviridae (~16.4% on average) and Podoviridae (~8.2% on average). The abundance of these three families was not significantly varied across the three layers (Figure 2). Notably, we detected a relatively high abundance (~6.10% on average) of the Nucleocytoplasmic Large DNA Viruses (NCLDV) in the NBT sediment, such as Mimiviridae (~3.37% on average) and Phycodnaviridae (~1.98% on average) (Figure 2, color block framed in black). The NCLDV was reported to comprise an expansive group of viruses that infect diverse eukaryotes [35]. The most dominant NCLDV families found in the NBT sediment, Phycodnaviridae, and Mimiviridae, were reported to be eukaryote-infecting viruses [36]. The six main genera of Phycodnaviridae, Chlorovirus (~0.49%), Coccolithovirus (~0.03%), Phaeovirus (~0.01%), Prasinovirus (~0.71%), Prymnesiovirus (~0.01%), and Raphidovirus (~0.05%), were all found in the NBT sediment. Cafeteriavirus and Mimivirus, the two main genera in Mimiviridae, were also found in the NBT sediment and accounted for a relative abundance of ~3.08% and ~0.06%, respectively. Mimiviridae members, such as *Megavirus chilensis* and Tupanviruses, were reported to infect *Acanthamoeba* in marine [37,38]. Based on our metagenomic analyses, we also detected *Acanthamoeba* (a relative abundance of ~0.005%) in the NBT sediment, which suggested the existence of eukaryote-viruses interaction in the hadal habitats. The relative abundance of NCLDV was found to have a decreasing trend with increasing depth in the three sediment samples (Supplementary Table S2). Giant viruses tend to be overlooked in previous viral genomic studies since samples are typically filtered according to the preconception of typical virion sizes [39]. Our approach is indifferent to the viral components of different particle sizes. Furthermore, we identified viruses that usually rely on co-infected giant dsDNA viruses of the Mimiviridae for their transmission [40], the Lavidaviridae. The relative abundance of Lavidaviridae accounted for ~0.2% on average in NBT sediment. This virus enhances the survival of the viral host by suppressing giant virus replication [41], which provides new perspectives of virus-host interactions in NBT sediment. Viruses with a relative abundance of less than 1% were also annotated, such as the non-tailed virus Autolykiviridae, the filamentous phages Inoviridae, and small fusiform phage Fuselloviridae (Figure 2 and Supplementary Table S2). Autolykiviridae, a kind of virus with double-stranded DNA (dsDNA), represented a novel family within the ancient lineage of double jelly roll (DJR) capsid viruses. They were reported to have a broad host range in contrast to tailed viruses

that kill on average only two hosts in one species [42]. Inoviridae, a kind of single-stranded DNA virus previously reported to be abundant in the oceans [43], were also found in our hadal sediment. Our study comprehensively characterized the compositions of the hadal viruses in the NBT sediment, which opens a new front in the research of microbes-viruses interaction in the hadal habitats. However, the metagenomic strategy can only capture the DNA sequence of the double-stranded DNA (dsDNA) virus and ssRNA (+) virus with dsDNA form during replication. Thus, our study mainly focused on DNA virus research. To further look into the RNA viruses, RNA of the NBT sediment samples should be sequenced and analyzed in the future.

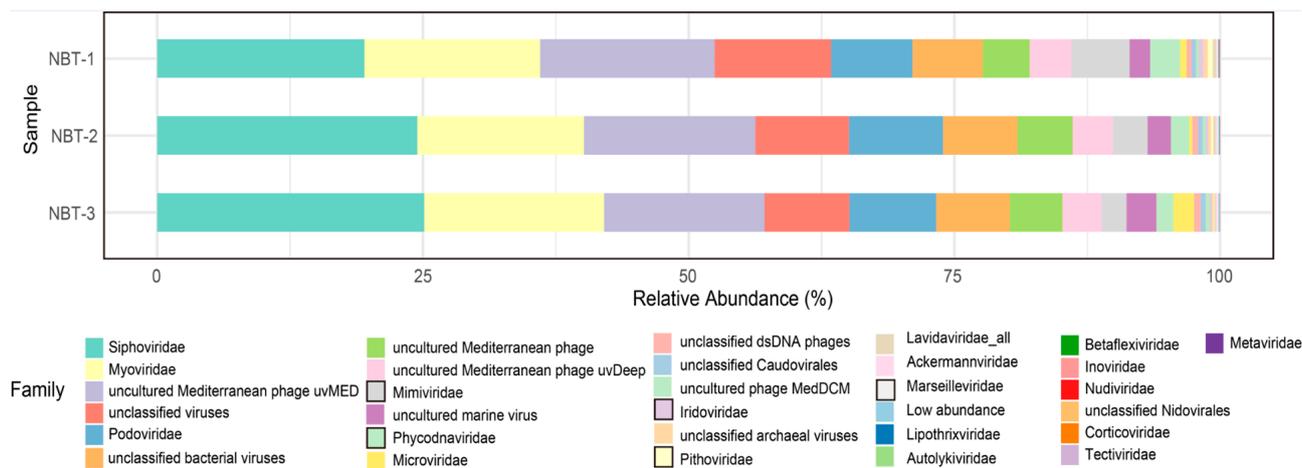


Figure 2. The relative sequence abundance of viral communities of three layers in NBT sediment (top 30). Viral communities shift at the family level. The raw relative abundance data was provided in Supplementary Table S2. Taxonomic assignment of clean reads was performed using Kaiju with NCBI nr database. NBT-1, the surficial layer NBT-1 (sediment depth 1–5 cm); NBT-2, the mid-layer (sediment depth 5–10 cm); NBT-3, the deep layer (sediment depth 10–15 cm). The black frame indicates NCLDV.

3.3. Comparative Analysis of Viromic Compositions between the NBT Sediment and Hadal Aquatic Habitats

To further understand the viromic compositions in hadal sediment, we compared the viromic compositions in the NBT sediment with compositions in hadal aquatic habitats. Due to the lack of metagenomic or viromic data for the NBT aquatic samples, we compared the viromic compositions in the NBT sediment with nine other hadal aquatic samples. These samples included particle-attached (PA), free-living (FL) samples at different water depths (i.e., 4000 m, 9600 m, 10,400 m, and 10,500 m.) in the Mariana Trench [3], and seawater samples at different water depths (i.e., 5000 m, 5700 m, 6000 m) in the Yap Trench [44] (Supplementary Table S3). The relative abundance of virome was calculated with Kaiju using NCBI nr as the reference database (Methods).

Based on the identified viral reads from the twelve hadal metagenomic data sets, variations of viromic community structures were observed among the hadal seawater and the NBT sediments (Figure 3 and Supplementary Table S3). Siphoviridae, which accounted for ~20% on average of the three NBT sediment layers, was the most predominant group in the NBT hadal sediments. However, it only constituted up to ~5% on average and ranked sixth of the major viral groups in the hadal seawater (Figure 3 and Supplementary Table S3). Its relative abundance in the hadal seawater was significantly lower than in the NBT hadal sediments (Mann–Whitney Rank Sum Test, $p < 0.05$, Supplementary Table S3). The second and third major viromic groups in NBT sediment, i.e., Myoviridae and Uncultured_Mediterranean_phage_uvMED, account for about 16% and 15% of the total virome, respectively, while the relative abundances of these two viromic groups increased to higher than 20% in hadal seawater samples (Figure 3 and Supplementary Table S3). Other major viromic groups, such as Podoviridae, was constituted up to ~2% on average of the total

microbial community in the hadal seawater. In comparison, in the NBT hadal sediments, it accounts for 8% in average relative abundance. The relative abundance of Podoviridae in the NBT hadal sediments is significantly higher than in hadal seawater samples (Figure 3 and Supplementary Table S3). Thus, distinct viral community structures were observed between the hadal seawater and the NBT sediments. The non-metric multidimensional scaling (NMDS) plot can show variations in the microbial community structures of different samples. To further explore the variations of viromic community structures among the hadal seawater and sediments, we plot the NMDS plot based on the relative abundance of the virome within each sample (Supplementary Figure S1). The results showed that the three NBT samples are clustering together, and there is a clear separation by NBT hadal sediment samples and hadal seawater samples. Distinct viral community structures were observed between the NBT sediment and hadal aquatic habitats. Previous studies reported that the microbial community structures between the hadal seawater and hadal sediments were distinct [45]. Thus, the hosts of hadal seawater and sediments virome were different, resulting in the distinctive viromic compositions between the NBT sediment and hadal aquatic habitats.

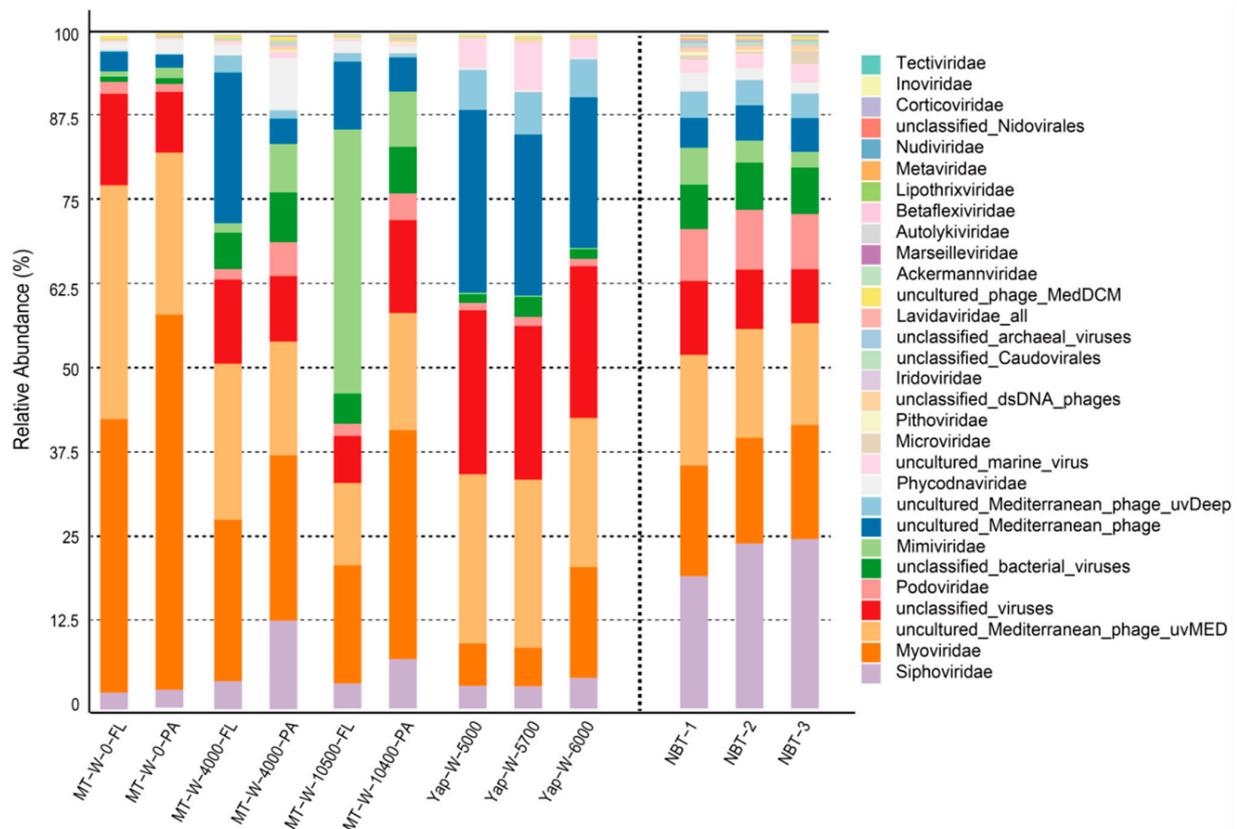


Figure 3. Comparative analysis of viromic compositions between the NBT sediment and hadal aquatic habitats. The relative abundance of the virome was calculated using Kaiju based on the NCBI nr database. Statistical analysis was performed on the relative abundance of the different viruses within the aquatic and sedimental samples (see details in Supplementary Table S3). Sample names are defined by sample location, sample type (seawater is subdivided according to size fraction and sediment is subdivided according to depth if provided), and sampling depth. MT, Mariana Trench; Yap, Yap Trench; W, seawater; PA, particle-attached; FL, free-living. e.g., MT-W-10500-FL is the free-living fraction of Mariana Trench seawater at 10,500 m, NBT-1, the surficial layer NBT-1 (sediment depth 1–5 cm); NBT-2, the mid-layer (sediment depth 5–10 cm); NBT-3, the deep layer (sediment depth 10–15 cm).

3.4. Comparative Analysis of Viromic Compositions between the NBT Sediment and Bathypelagic and Terrestrial Habitats

Gene-sharing network analysis by vContact2 [32] was performed to have a broad perspective of the hadal sediment virome compared to the virome of the bathypelagic and

3.5. Phylogenetic Diversity of the Hadal Virome in the NBT Sediment

Phylogenetic analyses were performed on the major viral groups by using marker genes to assess the diversity and genetic distance among the viruses in the NBT sediment. dsDNA viruses, which are most affiliated with Caudovirales, e.g., Siphoviridae, Myoviridae, and Podoviridae, accounted for the largest fraction in NBT sediment (Figure 2). A previous study reported that the gene coding for terminase large subunits (TerL) could be used to assess the phylogenetic diversity of Caudovirales [47]. Thus, in this study, we used the marker gene TerL to analyze the phylogenetic diversity of Caudovirales in the NBT sediment. In addition, NBT virome contigs homologous to the marker gene TerL (>50% similarity and object coverage ≥ 50) were selected for phylogenetic analysis. Finally, 87 unique NBT sediment viral contigs were used in phylogenetic analyses.

As shown in Figure 5, the NBT viral sequences were distributed widely throughout the phylogenetic tree, with most of the sequences affiliated to the Siphoviridae and Podoviridae. Caudovirales are reported to be the extensively studied virus group to date [47]. In this study, most NBT Caudovirales were phylogenetically distant to the known Caudovirales reference sequences and formed four major NBT clades within the Siphoviridae and Podoviridae. This result indicated that there might be an important uncharacterized diversity for Caudovirales in the hadal habitat, such as the NBT sediment in our study. In addition, some viral reference sequences were found to be included in the NBT clades (Figure 5), which provided clues to determine the hosts of phages from these NBT clades. For example, the largest NBT clade, named NBT clade I, included four viral reference sequences whose host was reported to be Betaproteobacteria or Gammaproteobacteria [48–50], indicating that members of this NBT clade may mainly infect Proteobacteria.

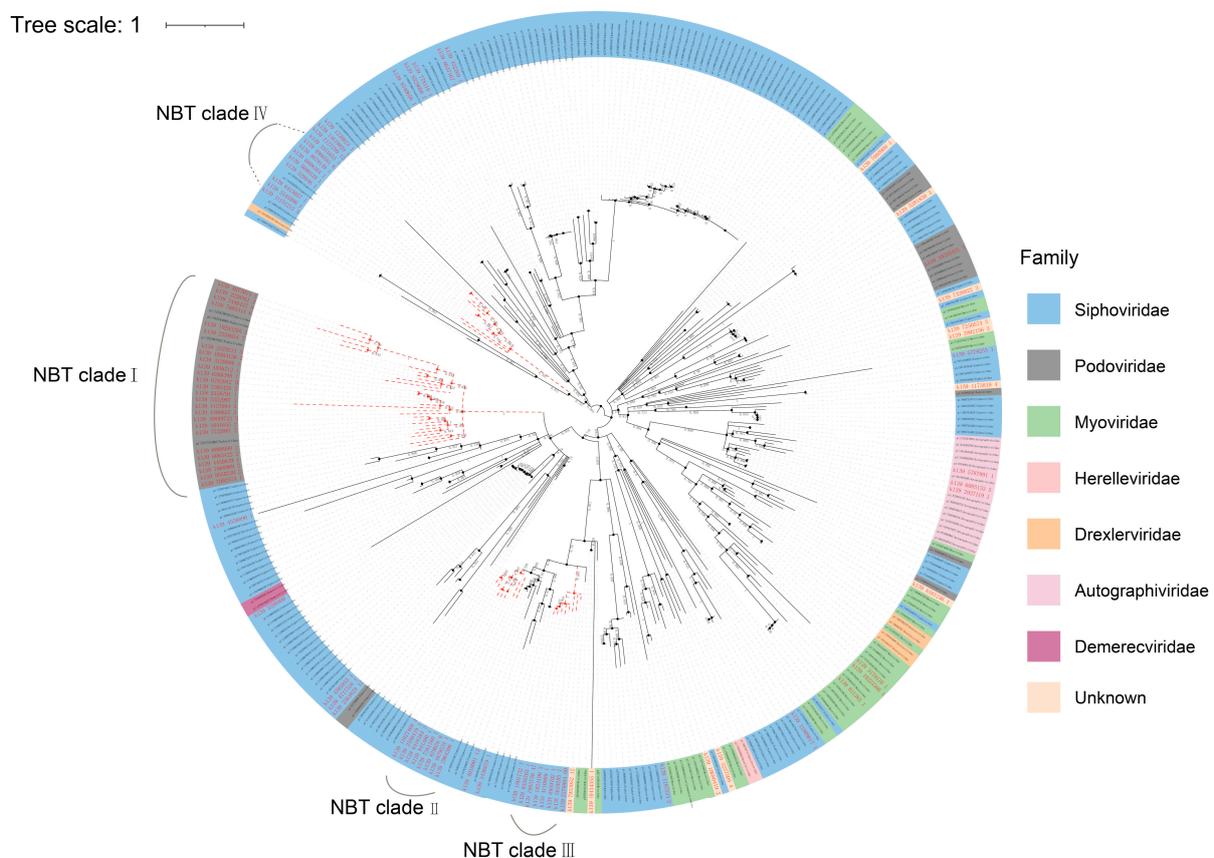


Figure 5. Phylogenetic analysis of Caudovirales based on TerL using the maximum likelihood algorithm. A FastTree approximate maximum-likelihood phylogenetic tree was built using TerL. Reference viral sequences from NCBI are colored in black. Scale bar, one amino acid substitution per site.

3.6. Functional Analysis of the Hadal Viromics in NBT Sediments

For the investigation of gene contents and functions of the NBT sediment viruses, we predicted the genes of the NBT virus contigs and annotated them with eggNOG-mapper using the Clusters of Orthologous Groups (COGs) database [51]. While a large proportion of the predicted genes had unknown functions, high-abundance genes were found in the COG categories of ‘replication, recombination and repair’, ‘signal transduction mechanism’, ‘transcription’, ‘cell wall/membrane/envelope biogenesis’, etc. (Figure 6 and Supplementary Table S5).

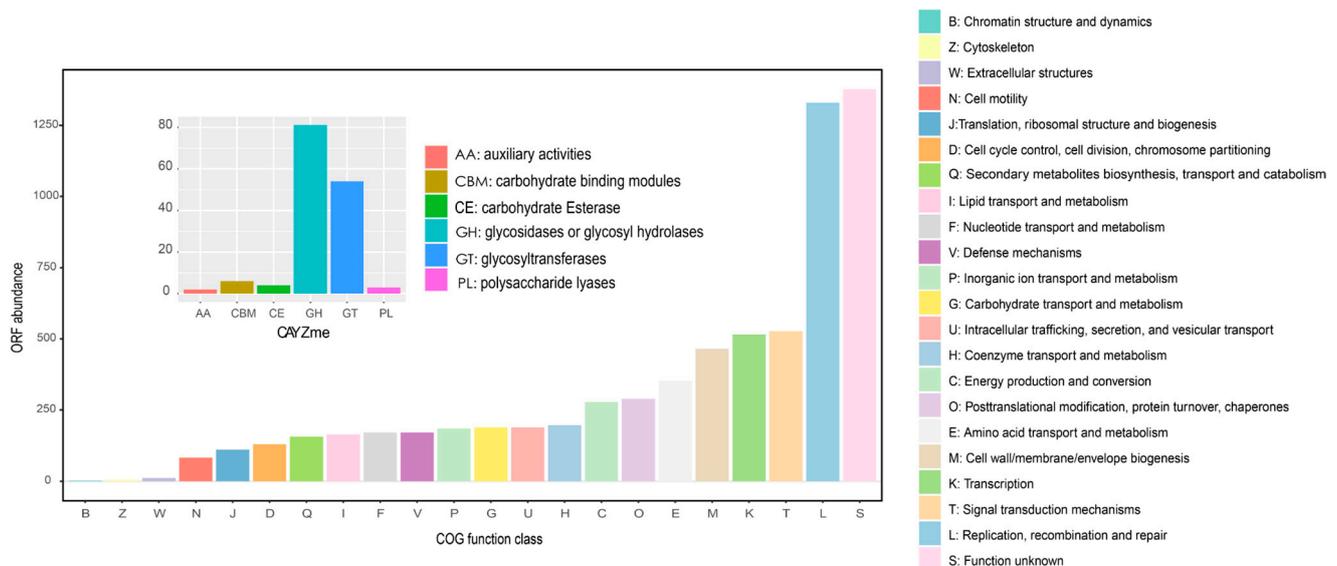


Figure 6. Annotation of viral genomes. COG annotation was performed using eggNOG-mapper. AMGs related to carbohydrate metabolism are illustrated in the inner panel.

Besides the genes for basic viral functions, a special group of virus-encoded genes, the so-called auxiliary metabolic genes (AMGs), can modulate the activities of the hosts upon infection. Overall, the New British Trench sediment viruses tend to encode AMGs for carbohydrate metabolism and cofactor/vitamin, and a significant portion also encoded AMGs for glycan and amino acid metabolism (Supplementary Table S6). The identified carbohydrate metabolism-related genes included 81 GHs (glycosidases or glycosyl hydrolases), 54 GTs (glycosyltransferases), six CBM (carbohydrate-binding modules), four CE (Carbohydrate esterases), three PLs (polysaccharide lyases), and two AA (auxiliary activities). In addition, AMGs related to C1 metabolism (e.g., transketolase), central carbon (e.g., UDP-glucose 4-epimerase), and ABC transporters for carbohydrates were also identified (Supplementary Table S3). Associated hadal sediment microorganisms were reported to be involved in various carbohydrate degradation reactions for processing detrital organic matter supplied from the overlying water column [3,44]. Infection by viruses containing genes related to carbohydrate degradation is consistent, in general, with their supportive role in host metabolism during infection [27,30]. Besides carbohydrate metabolism-related AMGs, we also found AMGs for sulfur metabolism. The *cysC* gene, encoding adenylylsulfate kinase, which reduces APS to produce 3'-phosphoadenosine-5'-phosphosulfate (PAPS), the second step of assimilatory sulfate reduction, was found in the New British Trench sediment viruses, suggesting possible roles to enhance host sulfate reduction. In this study, we explored the metabolic potential of the hadal virus based on the metagenomic data. However, some shortcomings are still inevitable based on such a strategy. For example, the assembling data might include the prophage data, which would introduce false-positive AMG predictions. Since viruses were difficult to isolate from the hadal sediment, analysis based on metagenomic data is currently the most effective method in exploring the virome in hadal sediment. The metagenomic analyses complement the traditional virology and

newer culturomic approaches. However, to further confirm the potential metabolic prediction of the hadal virus based on metagenomic analyses, it would be beneficial to culture these viruses with their corresponding hosts.

4. Conclusions

Due to the challenges of sampling and culturing of hadal sediments viruses, the role of viruses in influencing microbial mortality, ecology, and evolution in hadal sediments remains largely uncharacterized to date. This study adopted a deep shotgun metagenomic sequencing strategy to reveal novel, abundant, and diverse viruses of NBT sediments. Double-stranded DNA viruses, particularly those affiliated to the order Caudovirales, were found to be dominated in the hadal viral community. NCLDV that were often missed or underestimated in early studies due to the technique used to capture viral particles via filtering were detected in the NBT sediment in our study. By comparing the hadal sediment viruses with those of hadal aquatic habitats, we found the hadal sediment virome had some common components but was distinctive in community structure. Comparisons with the virome in bathypelagic and terrestrial habitats revealed that the NBT sediment virus community had many novel viral clusters not associated with the other habitual virome included in our analyses. Based on the phylogenetic analysis on the Caudovirales in the NBT sediment, novel diversities, including new clades specially evolved in the NBT sediment habitat, were found. Virus-encoded AMGs, including genes related to carbon and sulfur metabolism, may enhance metabolism in prokaryotic hosts during infection, potentially altering biogeochemical processes mediated by NBT sediment microbes. Our study filled in knowledge gaps on viral communities' structure, diversity, and metabolic potential in NBT sediments, providing novel insight into viruses' in hadal habitats.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/genes12070990/s1>, Figure S1: An NMDS plot showing variations in the viral community structures of the twelve hadal samples. Figure S2: Venn diagram of shared viral clusters among the five environmental virus data sets (NBT sediment, Tara Ocean seawater, wetland sediment, Stordalen thawing permafrost, and Cold seep). Table S1: Assembly of metagenomic sequencing on the New Britain Trench sediment samples and identification of viral genome/metavirome. Table S2: The relative abundance of virus identified in NBT sediment at a family level. Table S3: Viral community compositions of the NBT sediment and some hadal aquatic habitats. Table S4: Clustered among NBT sediment, cold seep sediments, seawater, wetland, permafrost, and viral RefSeq based on vConTACT2. Table S5: List of eggNOG-mapper annotations for the NBT virome. Table S6: List of putative AMGs of NBT virus.

Author Contributions: Conceptualization, X.L.; Formal analysis, H.Z., P.C., and M.Z.; Investigation, M.Z.; Project administration, X.L.; Visualization, H.Z., and P.C.; Resources, J.C. and J.F.; Writing—Original draft, H.Z., and P.C.; Writing—Review & editing, J.C., J.F., and X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by grants from the National Key R&D Program of China (2018YFC0310600, 2018YFA0900700), the National Natural Science Foundation of China (31771412, 91951210, 31972881, 41773069), and the Special Fund for strategic bio-resources from Chinese Academy of Sciences (KFJ-BRP-009).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw metagenomic sequencing data analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: We thank crews of M/V Zhangjian for their assistance in sampling.

Conflicts of Interest: The authors declare they have no conflict of interest.

References

1. Jamieson, A.J.; Fujii, T.; Mayor, D.J.; Solan, M.; De Prie, I.G. Hadal trenches: The ecology of the deepest places on Earth. *Trends Ecol. Evol.* **2010**, *25*, 190–197. [[CrossRef](#)]
2. Peoples, L.M.; Sierra, D.; Oladayo, O.; Xia, Q.; Alex, N.; Jessica, B.; Allen, E.E.; Church, M.J.; Bartlett, D.H.; Hauke, S. Vertically distinct microbial communities in the Mariana and Kermadec trenches. *PLoS ONE* **2018**, *13*, e0195102. [[CrossRef](#)]
3. Liu, J.; Zheng, Y.; Lin, H.; Wang, X.; Li, M.; Liu, Y.; Yu, M.; Zhao, M.; Pedentchouk, N.; Lea-Smith, D.J.; et al. Proliferation of hydrocarbon-degrading microbes at the bottom of the Mariana Trench. *Microbiome* **2019**, *7*, 47. [[CrossRef](#)] [[PubMed](#)]
4. Corinaldesi, C. New perspectives in benthic deep-sea microbial ecology. *Front. Mar. Sci.* **2015**, *2*, 17. [[CrossRef](#)]
5. Danovaro, R.; Corinaldesi, C.; Rastelli, E.; Dell'Anno, A. Towards a better quantitative assessment of the relevance of deep-sea viruses, Bacteria and Archaea in the functioning of the ocean seafloor. *Aquat. Microb. Ecol.* **2015**, *75*, 81–90. [[CrossRef](#)]
6. Suttle, C. Viruses in the sea. *Nature* **2005**, *437*, 356–361. [[CrossRef](#)]
7. Fuhrman, J.A. Marine viruses and their biogeochemical and ecological effects. *Nature* **1999**, *399*, 541. [[CrossRef](#)]
8. Suttle, C.A. Marine viruses—Major players in the global ecosystem. *Nat. Rev. Microbiol.* **2007**, *5*, 801–812. [[CrossRef](#)]
9. Winter, C.; Garcia, J.; Weinbauer, M.G.; Dubow, M.S.; Herndl, G.J. Comparison of Deep-Water Viromes from the Atlantic Ocean and the Mediterranean Sea. *PLoS ONE* **2014**, *9*, e100600. [[CrossRef](#)]
10. Sanchez, E.L.; Lagunoff, M. Viral activation of cellular metabolism. *Virology* **2015**, *479–480*, 609–618. [[CrossRef](#)]
11. Danovaro, R.; Dell'Anno, A.; Corinaldesi, C.; Magagnini, M.; Noble, R.; Tamburini, C.; Weinbauer, M. Major viral impact on the functioning of benthic deep-sea ecosystems. *Nature* **2008**, *454*, 1084–1087. [[CrossRef](#)]
12. Manea, E.; Dell'Anno, A.; Rastelli, E.; Tangherlini, M.; Corinaldesi, C. Viral infections boost prokaryotic biomass production and organic C cycling in hadal trench sediments. *Front. Microbiol.* **2019**, *10*, 1952. [[CrossRef](#)]
13. Gallo, N.D.; James, C.; Kevin, H.; Patricia, F.; Douglas, H.B.; Lisa, A.L. Submersible- and lander-observed community patterns in the Mariana and New Britain trenches: Influence of productivity and depth on epibenthic and scavenging communities. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **2015**, *99*, 119–133. [[CrossRef](#)]
14. Davies, H.L.; Keene, J.B.; Hashimoto, K.; Joshima, M.; Stuart, J.E.; Tiffin, D.L. Bathymetry and canyons of the western Solomon Sea. *Geo Mar. Lett.* **1986**, *6*, 181–191. [[CrossRef](#)]
15. Liu, R.; Wang, L.; Liu, Q.; Wang, Z.; Li, Z.; Fang, J.; Zhang, L.; Luo, M. Depth-resolved distribution of Particle-Attached and Free-Living bacterial communities in the water column of the New Britain Trench. *Front. Microbiol.* **2018**, *9*, 625. [[CrossRef](#)]
16. Liu, R.; Wang, Z.; Wang, L.; Li, Z.; Fang, J.; Wei, X.; Wei, W.; Cao, J.; Wei, Y.; Xie, Z. Bulk and active sediment prokaryotic communities in the Mariana and Mussau Trenches. *Front. Microbiol.* **2020**, *11*, 1521. [[CrossRef](#)]
17. Wang, L.; Liu, R.; Wei, X.; Wang, Z.; Shen, Z.; Cao, J.; Wei, Y.; Xie, Z.; Chen, L.; Fang, J. Transitions in microbial communities along two sediment cores collected from the landward walls of the New Britain trench. *Mar. Biol.* **2020**, *167*, 172. [[CrossRef](#)]
18. Liu, Q.; Fang, J.; Li, J.; Zhang, L.; Xie, B.-B.; Chen, X.-L.; Zhang, Y.-Z. Depth-resolved variations of cultivable bacteria and their extracellular enzymes in the water column of the New Britain Trench. *Front. Microbiol.* **2018**, *9*, 135. [[CrossRef](#)]
19. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
20. Chen, P.; Zhou, H.; Huang, Y.; Xie, Z.; Zhang, M.; Wei, Y.; Li, J.; Ma, Y.; Luo, M.; Ding, W.; et al. Revealing the full biosphere structure and versatile metabolic functions in the deepest ocean sediment of the Challenger Deep. *bioRxiv* **2021**. [[CrossRef](#)]
21. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)] [[PubMed](#)]
22. Xu, H.B.; Luo, X.; Qian, J.; Pang, X.H.; Song, J.Y. FastUniq: A Fast De Novo Duplicates Removal Tool for Paired Short Reads. *PLoS ONE* **2012**, *7*, e52249. [[CrossRef](#)]
23. Li, D.; Luo, R.; Liu, C.M.; Leung, C.M.; Ting, H.F.; Sadakane, K.; Yamashita, H.; Lam, T.W. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **2016**, *102*, 3–11. [[CrossRef](#)] [[PubMed](#)]
24. Menzel, P.; Ng, K.L.; Krogh, A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* **2016**, *7*, 11257. [[CrossRef](#)]
25. Ren, J.; Ahlgren, N.A.; Lu, Y.Y.; Fuhrman, J.A.; Sun, F. VirFinder: A novel k-mer based tool for identifying viral sequences from assembled metagenomic data. *Microbiome* **2017**, *5*, 1–20. [[CrossRef](#)] [[PubMed](#)]
26. Roux, S.; Enault, F.; Hurwitz, B.L.; Sullivan, M.B. VirSorter: Mining viral signal from microbial genomic data. *PeerJ* **2015**, *3*, e985. [[CrossRef](#)]
27. Guo, J.; Bolduc, B.; Zayed, A.A.; Varsani, A.; Dominguez-Huerta, G.; Delmont, T.O.; Pratama, A.A.; Gazitúa, M.C.; Vik, D.; Sullivan, M.B.; et al. VirSorter2: A multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* **2021**, *9*, 37. [[CrossRef](#)]
28. Roux, S.; Brum, J.R.; Dutilh, B.E.; Sunagawa, S.; Duhaime, M.B.; Loy, A.; Poulos, B.T.; Solonenko, N.; Lara, E.; Poulain, J.; et al. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* **2016**, *537*, 689–693. [[CrossRef](#)]
29. Paula, D.M.; Danczak, R.E.; Simon, R.; Jeroen, F.; Borton, M.A.; Wolfe, R.A.; Burris, M.N.; Wilkins, M.J. Viral and metabolic controls on high rates of microbial sulfur and carbon cycling in wetland ecosystems. *Microbiome* **2018**, *6*, 138.
30. Emerson, J.B.; Simon, R.; Brum, J.R.; Benjamin, B.; Woodcroft, B.J.; Bin, J.H.; Singleton, C.M.; Solden, L.M.; Naas, A.; Boyd, J.A. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat. Microbiol.* **2018**, *3*, 870–880. [[CrossRef](#)]

31. Li, Z.; Pan, D.; Wei, G.; Pi, W.; Zhang, C.; Wang, J.-H.; Peng, Y.; Zhang, L.; Wang, Y.; Hubert, C.R.J.; et al. Deep sea sediments associated with cold seeps are a subsurface reservoir of viral diversity. *ISME J.* **2021**. [[CrossRef](#)]
32. Jang, H.B.; Bolduc, B.; Zabolocki, O.; Kuhn, J.H.; Roux, S.; Adriaenssens, E.M.; Brister, J.R.; Kropinski, A.M.; Krupovic, M.; Lavigne, R. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat. Biotechnol.* **2019**, *37*, 632–639. [[CrossRef](#)]
33. Shannon, P. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)]
34. Kieft, K.; Zhou, Z.; Anantharaman, K. VIBRANT: Automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome* **2020**, *8*, 1. [[CrossRef](#)] [[PubMed](#)]
35. Koonin, E.V.; Yutin, N. Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. *Intervirology* **2010**, *53*, 284–292. [[CrossRef](#)] [[PubMed](#)]
36. Lang, A.S.; Rise, M.L.; Culley, A.I.; Steward, G.F. RNA viruses in the sea. *FEMS Microbiol. Rev.* **2009**, *33*, 295–323. [[CrossRef](#)]
37. Arslan, D.; Legendre, M.; Seltzer, V.; Abergel, C.; Claverie, J.-M. Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17486–17491. [[CrossRef](#)]
38. Abrahão, J.; Silva, L.; Silva, L.S.; Khalil, J.Y.B.; Rodrigues, R.; Arantes, T.; Assis, F.; Boratto, P.; Andrade, M.; Kroon, E.G.; et al. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* **2018**, *9*, 749. [[CrossRef](#)] [[PubMed](#)]
39. Filee, J.; Pouget, N.; Chandler, M. Phylogenetic evidence for extensive lateral acquisition of cellular genes by Nucleocytoplasmic large DNA viruses. *BMC Evol. Biol.* **2008**, *8*, 320. [[CrossRef](#)]
40. Bekliz, M.; Colson, P.; La Scola, B. The Expanding Family of Virophages. *Viruses* **2016**, *8*, 317. [[CrossRef](#)]
41. Fischer, M.G. The Virophage Family Lavidaviridae. *Curr. Issues Mol. Biol.* **2021**, *40*, 1–24. [[CrossRef](#)] [[PubMed](#)]
42. Kauffman, K.M.; Hussain, F.A.; Yang, J.; Arevalo, P.; Brown, J.M.; Chang, W.K.; VanInsberghe, D.; Elsherbini, J.; Sharma, R.S.; Cutler, M.B.; et al. A major lineage of non-tailed dsDNA viruses as unrecognized killers of marine bacteria. *Nature* **2018**, *554*, 118–122. [[CrossRef](#)] [[PubMed](#)]
43. Labonté, J.; Suttle, C.A. Previously unknown and highly divergent ssDNA viruses populate the oceans. *ISME J.* **2013**, *7*, 2169–2177. [[CrossRef](#)] [[PubMed](#)]
44. Zhang, X.; Xu, W.; Liu, Y.; Cai, M.; Luo, Z.; Li, M. Metagenomics Reveals Microbial Diversity and Metabolic Potentials of Seawater and Surface Sediment from a Hadal Biosphere at the Yap Trench. *Front. Microbiol.* **2018**, *9*, 2402. [[CrossRef](#)]
45. Hiraoka, S.; Hirai, M.; Matsui, Y.; Makabe, A.; Minegishi, H.; Tsuda, M.; Juliarni Rastelli, E.; Danovaro, R.; Corinaldesi, C.; Kitahashi, T.; et al. Microbial community and geochemical analyses of trans-trench sediments for understanding the roles of hadal environments. *ISME J.* **2020**, *14*, 740–756. [[CrossRef](#)]
46. O’Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufu, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **2016**, *44*, D733–D745. [[CrossRef](#)]
47. Jin, M.; Guo, X.; Zhang, R.; Qu, W.; Gao, B.; Zeng, R. Diversities and potential biogeochemical impacts of mangrove soil viruses. *Microbiome* **2019**, *7*, 58. [[CrossRef](#)] [[PubMed](#)]
48. Skennerton, C.T.; Angly, F.E.; Breitbart, M.; Bragg, L.; He, S.; McMahon, K.D.; Hugenholtz, P.; Tyson, G.W. Phage encoded H-NS: A potential achilles heel in the bacterial defence system. *PLoS ONE* **2011**, *6*, e20095. [[CrossRef](#)]
49. Summer, E.J.; Enderle, C.J.; Ahern, S.J.; Gill, J.J.; Torres, C.P.; Appel, D.N.; Black, M.C.; Young, R.; Gonzalez, C.F. Genomic and Biological Analysis of Phage Xfas53 and Related Prophages of *Xylella fastidiosa*. *J. Bacteriol.* **2010**, *192*, 179. [[CrossRef](#)]
50. Clark, A.J.; Pontes, M.; Jones, T.; Dale, C. A possible heterodimeric prophage-like element in the genome of the insect endosymbiont *Sodalis glossinidius*. *J. Bacteriol.* **2007**, *189*, 2949. [[CrossRef](#)]
51. Galperin, M.Y.; Makarova, K.S.; Wolf, Y.I.; Koonin, E.V. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* **2015**, *43*, D261–D269. [[CrossRef](#)] [[PubMed](#)]