

Research Article

Diverse viromes in polar regions: A retrospective study of metagenomic data from Antarctic animal feces and Arctic frozen soil in 2012–2014

Jun Wang^{a,1}, Jian Xiao^{a,1}, Zheng Zhu^{a,2}, Siyuan Wang^{a,c}, Lei Zhang^a, Zhaojun Fan^a, Yali Deng^a, Zhihong Hu^a, Fang Peng^{b,*}, Shu Shen^{a,*}, Fei Deng^{a,*}^a State Key Laboratory of Virology and National Virus Resource Center, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China^b China Center for Type Culture Collection (CCTCC), College of Life Sciences, Wuhan University, Wuhan, 430072, China^c Xinjiang Key Laboratory of Biological Resources and Genetic Engineering, College of Life Science and Technology, Xinjiang University, Urumqi, 830046, China

ARTICLE INFO

Keywords:

Arctic soil
Antarctica animal feces
Metagenomics
Virus diversity

ABSTRACT

Antarctica and the Arctic are the coldest places, containing a high diversity of microorganisms, including viruses, which are important components of polar ecosystems. However, owing to the difficulties in obtaining access to animal and environmental samples, the current knowledge of viromes in polar regions is still limited. To better understand polar viromes, this study performed a retrospective analysis using metagenomic sequencing data of animal feces from Antarctica and frozen soil from the Arctic collected during 2012–2014. The results reveal diverse communities of DNA and RNA viruses from at least 23 families from Antarctic animal feces and 16 families from Arctic soils. Although the viral communities from Antarctica and the Arctic show a large diversity, they have genetic similarities with known viruses from different ecosystems and organisms with similar viral proteins. Phylogenetic analysis of *Microviridae*, *Parvoviridae*, and *Larvidaviridae* was further performed, and complete genomic sequences of two novel circular replication-associated protein (rep)-encoding single-stranded (CRESS) DNA viruses closely related to *Circoviridae* were identified. These results reveal the high diversity, complexity, and novelty of viral communities from polar regions, and suggested the genetic similarity and functional correlations of viromes between the Antarctica and Arctic. Variations in viral families in Arctic soils, Arctic freshwater, and Antarctic soils are discussed. These findings improve our understanding of polar viromes and suggest the importance of performing follow-up in-depth investigations of animal and environmental samples from Antarctica and the Arctic, which would reveal the substantial role of these viruses in the global viral community.

1. Introduction

Located at the poles of the earth, Antarctica and the Arctic are the coldest places with different topographies and extreme environmental conditions (Yau and Seth-Pasricha, 2019). Surrounded by ocean, Antarctica's massive glaciers buffer the effects of global warming by contributing to the formation of sea ice to drive the Antarctic thermohaline circulation, and the sea ice and microorganisms there play a critical role in carbon sequestration and the global food web (Cavicholi, 2015; Wilkins et al., 2013). The Arctic is an ocean ringed by

landmasses, where a large area ($\sim 15 \times 10^6$ km² in winter) would be covered by the Arctic sea ice (Eric Collins et al., 2010). The global climate and ecological changes would be threatening the survival of living species relying on the presence of sea ice in the Arctic (Pagano and Williams, 2021). Both polar regions are known for a mean annual surface temperature of -20 °C or less, low nutrients, dim light in winter, and high ultraviolet radiation exposure in summer. Despite seasonal fluctuations in light and freshwater melting from ice and variations influenced by temperature, wind, and nutrients, the polar regions contain a diverse range of bacterial, archaeal, and eukaryotic

* Corresponding authors.

E-mail addresses: fangpeng2@aliyun.com (F. Peng), shenshu@wh.iov.cn (S. Shen), df@wh.iov.cn (F. Deng).¹ Jun Wang and Jian Xiao contributed equally to this work.² Current address: National Emerging Infectious Diseases Laboratories, Boston University, Boston, MA 02118, United States; Department of Biochemistry, Boston University School of Medicine, Boston, MA 02118, United States.<https://doi.org/10.1016/j.virs.2022.08.006>

Received 29 April 2022; Accepted 17 August 2022

Available online 24 August 2022

1995-820X/© 2022 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

microbial communities. These microorganisms are important components of polar ecosystems, and thus play an important role in the food web (Cavicchioli, 2015; Gregory et al., 2019; Wilkins et al., 2013; Yau and Seth-Pasricha, 2019).

Viruses play important roles in ecosystem function, microbial mortality, fluctuations in microbial community structure, and horizontal gene transfer of genetic diversity between host cells (Breitbart et al., 2002; Gong et al., 2018; Suttle, 2007; Wommack and Colwell, 2000). Because the rise in global temperature in recent years may have an increasing impact on Antarctica and the Arctic biome, viral communities under extremely cold conditions may also be vulnerable to ongoing global perturbations. Therefore, it is important to characterize the complexity and novelty of viral communities from Antarctica and the Arctic and their potential connection to other environments on earth. Although it has always been difficult to obtain environmental samples from polar regions, great efforts have been made for more than a decade to investigate viromes in the Antarctic and Arctic regions. An initial study on viral communities identified high viral diversity in Antarctica in 2006–2007, which reported the genetic structure of the viral community from a freshwater lake and revealed unexpected genetic richness of multiple viral families, including *Circoviridae*, *Geminiviridae*, *Nanoviridae*, *Microviridae*, *Phycodnaviridae*, and *Mimiviridae*, as well as those of *Caudovirales*, using 454 Roche sequencing (López-Bueno et al., 2009). Later, the viral communities of Antarctic hyperarid desert soils were characterized. The diversity level of hypolithic viruses detected from the cyanobacterium-dominated assemblages situated on the ventral surfaces of quartz pebbles is much higher than that of viromes found in the surface soil of the Miers Valley in eastern Antarctica (Zablocki et al., 2014). Moreover, the viral communities in the soil of ice-free areas of the Mackay Glacier region, north of the McMurdo Dry Valleys, are composed of a high diversity of tailed bacteriophages, phycodnaviruses, and mimiviruses (Adriaenssens et al., 2017). A large number of new single-stranded DNA (ssDNA) viruses have been found in cryoconite holes in the Taylor Valley in Antarctica using metagenomic methods (Sommers et al., 2019). In the Arctic, a few studies were performed mainly to investigate viromes in aquatic ecosystems. For example, metagenomic analyses showed that the Arctic lakes in Spitsbergen are dominated by unknown ssDNA viruses with no close relatives in the database as of 2015. Although these unique DNA viral communities are mostly related to each other, they present minor genetic overlaps with viromes from other environments, suggesting a bipolar distribution and global dispersal capacity for viruses (De Cárcer et al., 2015). Viral communities strikingly dissimilar and highly divergent from known viruses were found in different layers of a highly stratified meromictic lake near the northern limit of the Canadian High Arctic (Labbé et al., 2020), further demonstrating the novelty and complexity of the viral community there. These studies revealed a high diversity of viral communities and microbial richness and complexity of the polar ecosystem.

Owing to difficulties in obtaining and transporting samples from polar regions, the complexity, abundance, and origins of viral communities in Antarctic and Arctic environments are still not well understood. In this study, metagenomic sequencing data were generated by Roche 454 sequencing of fecal samples from wild animals in Antarctica and frozen soils from the Arctic, which were collected from 2012 to 2014. However, very few virus-related sequences were identified by comparison with reference datasets at that time. Until recently, these data were re-analyzed based on a comparison with the current datasets, from which sequences related to a variety of DNA and RNA viruses were found. The complexity, novelty, and functional and genetic relationships of the viral communities from Antarctic animal feces samples and Arctic soil samples were characterized and discussed by comparison with other known viruses. The results reveal that the viral communities in polar regions were highly genetically divergent from, but still correlated with known viruses worldwide, which expanded our understanding of the complexity of polar viral communities as early as nearly ten years ago.

2. Materials and method

2.1. Sample collection from Antarctica and the Arctic

Animal feces samples were collected in the Fildes Peninsula of Antarctica (latitude from 62.15° S to 62.26° S, longitude 58.87° W–59.01° W) in 2014, including seven from migratory birds, four from penguins, and one from a fur seal (Fig. 1A and B, Table 1). All fecal samples were collected with sterile swabs and quickly stored in liquid nitrogen until transported to laboratory for further experiments. In 2012–2013, frozen soil samples were collected from seven different locations on the front edge of glaciers, Midtre Lovénbreen, in the new Olson area, Svalbard Islands of the Arctic (latitude, from 78.88° N to 78.90° N; longitude, from 12.07° E to 12.17° E) (Fig. 1A and C, Table 2). To avoid bias resulting from sampling, soils from the surface layer were collected from 10 to 20 sites at each location, mixed, and sub-packaged. After mixing, approximately 0.5 g of the sample was used for subsequent experiments. The sampling process followed the principles of aseptic procedures. All samples were stored in sterile containers, transported back to the laboratory at a low temperature, and stored at –80 °C until further experiments.

2.2. Viral metagenomics

Homogenates of each Antarctic fecal sample (0.5 g) and Arctic soil sample (0.5 g) were prepared and diluted in 1 mL precooled 1 × phosphate buffer saline (PBS) and clarified by centrifugation (12,000×g) at 4 °C for 15 min. The supernatants were filtered using 0.45 µm filter membrane filtration and then treated with DNase I (2 U/µL; Ambion, Waltham, MA, USA), RNase I (50 U/µL; Fermentas, Maidstone, United Kingdom), and Benzonase (25 U/µL; NEB, Massachusetts, USA) at 37 °C for 90 min to digest the exogenous nucleic acid. Total RNA was purified from the supernatants using a QIAamp Viral RNA Mini Kit (Qiagen, Dusseldorf, Germany), which may also purify DNA from samples. A cDNA library of each sample was prepared and labeled with different barcodes and subsequently applied to a Roche 454 GS FLX system (GS-FX, Roche Applied Science, Connecticut, USA) for metagenomic sequencing, as previously described (Shen et al., 2018).

The barcoding reads of each sample were obtained by splitting the total reads generated by Roche 454 sequencing according to different barcodes using SeqClean (<http://www.tigr.org/tdb/tgi/software/>), filtered to remove sequences containing more than 10% N, length < 50 bp, or of low complexity to obtain qualified reads. The resulting reads were assembled into contigs using MEGAHIT (version 1.2.9). The data resulting from Roche-454 sequencing were deposited in the China National GeneBank database (CNGDB, accession no. CNA0047375–CNA0047394). The genome sequence of E754 accession no. is CNA0047395 and E226 is CNA0047396.

To find viral-related sequences, the assembled contigs were compared to the NR and NT databases from GenBank available until May 2021 using DIAMOND BLASTX and BLASTN (e value < 10^{–10}), respectively, and were further assigned to the closest viral families according to the least common ancestor of the best result for each contig. Sequences related to phages were identified by DeepVirFinder using contigs of length >1000 bp (Ren et al., 2020) and were further confirmed by manual curation. Open reading frames (ORFs) were predicted using Prodigal (version 2.6.3) (Hyatt et al., 2010). The contig ratio was computed by summarizing the same classified level contig number divided by all contig sequence numbers and are illustrated by ggplot2 package in R.

2.3. Viral sequence homology analysis

The metadata containing the original sources of viruses were downloaded from IMG/VR (v3, containing 2,377,994 sequences) (<https://img.jgi.doe.gov/cgi-bin/vr/main.cgi>), and the virus-related contigs were compared with this metadata to identify viral homologs discovered in different ecosystems or hosts. The functions of putative viral proteins were predicted using Prodigal (V2.6.3) (Hyatt et al., 2010). Functional

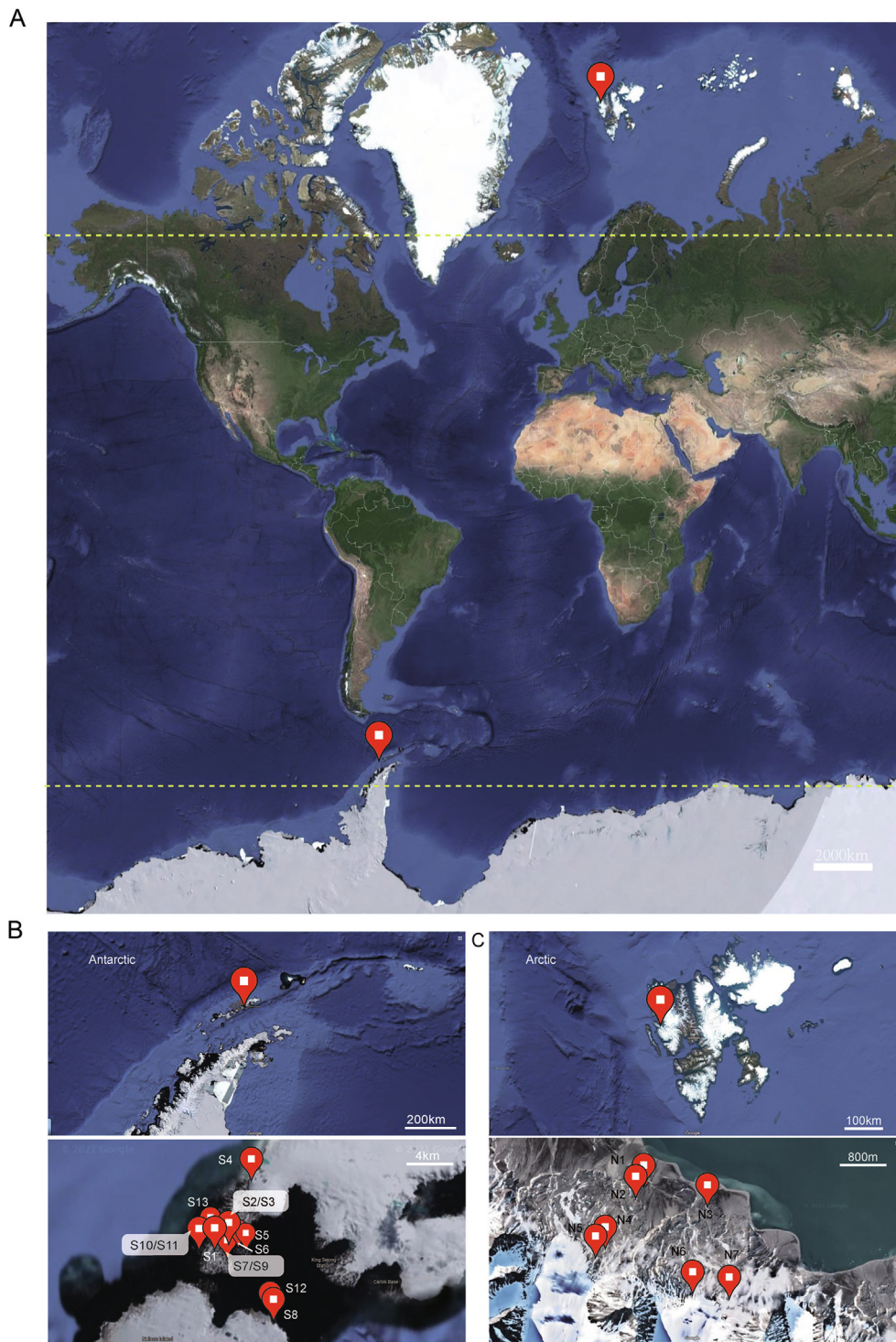


Fig. 1. Maps showing where samples were collected. **A** Global map of sampling sites in Antarctica and the Arctic. The Arctic and Antarctic circles are indicated by yellow lines. **B** and **C** show the sampling sites in the enlarged images.

homology of these proteins was analyzed using Cytoscape 3.7.2 (Otasek et al., 2019) and is presented as clusters and networks.

2.4. Phylogenetic analysis

Sequences from this study and GenBank were aligned using mafft (Katoh and Standley, 2013) and trimmed using trimAI v1.2

(Capella-Gutiérrez et al., 2009). The resulting alignments were used to build trees using IQ-Tree (Nguyen et al., 2015) with automatic determination of the substitution model and 5000 bootstraps. The models used included GTR + F + R5 for *Microviridae*, TIM3 + F + R6 for *Parvoviridae*, Q. pfam + F + R7s for Eukaryotic circular rep-encoding single-stranded (CRESS) DNA virus, and Q. pfam + F + I + G4 for *Lavidaviridae*. Trees were visualized using FigTree and iTOL(Letunic and Bork, 2019).

Table 1
Information of animal feces samples collected from Antarctic.

Pool	Year	Hosts	Reads	Virus-related contigs	Virus reads	Percentage of total reads
S1	2013	Antarctic fur seal	29,810	44	26,176	87.80%
S2	2013	Chinstrap penguin	14,588	33	355	2.40%
S3	2013	Penguin	3972	40	1976	49.70%
S4	2013	Emperor penguin	16,479	59	6807	41.30%
S5	2014	Gentoo penguin	29,095	227	6694	23.00%
S6	2013	Kelp gull	20,195	437	4648	23%
S7	2014	Kelp gull	1824	15	397	21.70%
S8	2014	Stercorarius	50,289	25	373	0.74%
S9	2014	Stercorarius	3921	57	921	23.40%
S10	2014	Tern	27,430	19	5775	21.00%
S11	2014	Tern	4469	26	190	4.20%
S12	2014	Giant petrel	3100	50	389	12.50%
S13	2014	Giant petrel	19,747	21	1616	8.10%

Table 2
Information of frozen soil samples collected from Arctic.

Pool	Year	Reads	Virus-related contigs	Virus reads	Percentage of total reads
N1	2012	72,438	10	126	0.17%
N2	2012	12,001	255	5624	46.80%
N3	2012	10,577	102	1682	15.90%
N4	2012	13,485	111	5772	42.80%
N5	2012	12,128	15	244	2.00%
N6	2012	30,704	8	71	0.23%
N7	2012	71,866	13	129	0.18%

3. Results

3.1. Retrospective analyses of metagenomic sequencing data reveal diverse viral communities in Antarctic animal feces and Arctic soil

Roche 454 sequencing was performed in 2014 after Antarctic animal feces and Arctic soil samples were transported to the laboratory. A total of 448,118 reads were generated from all samples; however, very few virus-related sequences were found in the sequencing data for that year. Retrospective analyses using these data were performed to identify viral sequences by comparison with current datasets, which found viral sequences accounting for varying proportions (0.74%–87.80%) of the total reads from these samples. As a result, 69,965 reads, accounting for 15.6% of the total reads (Tables 1 and 2), were found to be related to viruses belonging to at least 28 viral families.

3.1.1. Viromes of Antarctic animal feces

In Antarctic animal feces, the identified sequences were related to 14 families of DNA viruses (dsDNA virus: *Autographiviridae*, *Myoviridae*, *Podoviridae*, *Siphoviridae*, *Herelleviridae*, *Lavidaviridae*, *Mimiviridae*, *Phycodnaviridae*, and *Totiviridae*; ssDNA: *Anelloviridae*, *Circoviridae*, *Microviridae*, *Inoviridae* and *Parvoviridae*), nine families of RNA viruses (dsRNA: *Birnaviridae*, *Cystoviridae* and *Partitiviridae*; ssRNA: *Dicistroviridae*, *Marnaviridae*, *Nodaviridae*, *Picornaviridae*, *Tombusviridae* and *Metaviridae*), and unclassified viruses related to *Caudovirales*, *Nucleocytoviricota*, and *Riboviria*. In addition, several sequences were found to be related to unassigned viruses (Fig. 2A).

Of these fecal samples, a large number of sequences related to bacterial phages (contigs among the 13 fecal samples: 6.8%–64%) were found to be related to *Caudovirales*, including *Siphoviridae* (0.9%–37%), *Podoviridae* (0%–24.6%), *Myoviridae* (0%–13.9%), and *Autographiviridae* (0%–3.2%), together with unclassified sequences of this order, and to the families of *Microviridae* (0%–61.6%) and *Inoviridae* (0%–10%) (Fig. 2A, Supplementary Table S1). This suggests that bacterial phages, especially those belonging to *Caudovirales*, might constitute the dominant viral communities in the feces of polar animals. Eight of the 13 fecal samples from migratory seabirds (S6–S13) had more diverse viral communities, from which viral sequences relating to the families of dsDNA virus

(10.4%–81.7%), ssDNA virus (2.2%–65.7%), dsRNA virus (0%–1.3%), and ssRNA virus (0%–8.9%) were found, whereas from the other five samples (S1–S5) from penguins and fur seal, very few sequences of dsRNA virus (0%–0.3%) or ssRNA virus (0%) were identified. In addition, the unclassified viruses of *Riboviria* were found in feces of migratory birds, resulting in 14.8%–54.5% of the total contigs, whereas sequences related to *Riboviria* were 20.9%–43.7% of the fecal samples from penguins and sea furl (Supplementary Table S1).

We subsequently characterized the sequence identities of the identified viral families by comparing them with reference viruses in the current datasets. Viral sequences of the *Myoviridae*, *Siphoviridae*, *Inoviridae*, *Microviridae*, and *Picornaviridae* families showed wide distributions ranging from 29% to 100% to known viruses and were comparable between migratory birds and other land animals. Comparable sequence identities of the *Autographiviridae*, *Mimiviridae*, *Totiviridae*, *Circoviridae*, and *Marnaviridae* families were also found between migratory birds and land animals, which, however, were less similar to the known reference viruses (identities: 29%–71%) (Fig. 2B, Supplementary Table S1). Of these viral families, the sequence identities of *Microviridae* and *Siphoviridae* showed significant differences ($P < 0.5$, Student's *t*-test) between the samples from migratory birds and land animals in Antarctica, whereas this difference was not found for the other viral families. Moreover, sequences of *Totiviridae*, *Dicistroviridae*, *Marnaviridae*, and *Picornaviridae* were only detected in the feces of migratory birds, but were not found in the feces of land animals. Except for the *Picornaviridae* sharing a high degree of identity (33.1%–100%) with known viruses, the *Totiviridae*, *Dicistroviridae*, and *Marnaviridae* relating sequences generally had identities of less than 90% compared with known viruses, suggesting the novelty of these viruses. Sequences of *Mimiviridae* were found in the feces of migratory birds with identities ranging from 33% to 71.6%, whereas only one sequence related to *Mimiviridae* was found in land animal samples (Fig. 2B, Supplementary Table S1).

3.1.2. Arctic soil viromes

Virus-related sequences were found in the Arctic frozen soil, including DNA viruses from 14 families (dsDNA virus: *Alloherpesviridae*, *Demereciviridae*, *Iridoviridae*, *Mimiviridae*, *Myoviridae*, *Phycodnaviridae*, *Podoviridae*, *Siphoviridae*, *Caulimoviridae*; ssDNA: *Circoviridae*, *Parvoviridae*, *Genomoviridae*, *Inoviridae* and *Microviridae*), RNA viruses from two families (dsRNA: *Partitiviridae*; ssRNA: *Metaviridae*), and unclassified viruses related to *Caudovirales*, *Nucleocytoviricota*, and *Riboviria*, as well as a number of sequences related to unassigned viruses (Fig. 3A). Similar to viromes of the fecal samples from the Antarctic, bacteriophages (37.3%–77.1%) constituted the most abundant group of viruses from these soil samples, which included viral sequences relating to *Microviridae* (0%–54.2%), *Siphoviridae* (0%–5.6%), *Myoviridae* (0%–4.0%), and *Podoviridae* (0%–1.4%), as well as a number of viral sequences (22%–43.7%) which were assigned to the order *Caudovirales* but could not be assigned to any of the families of this order. Similarities between the sequences related to *Caudovirales* and known

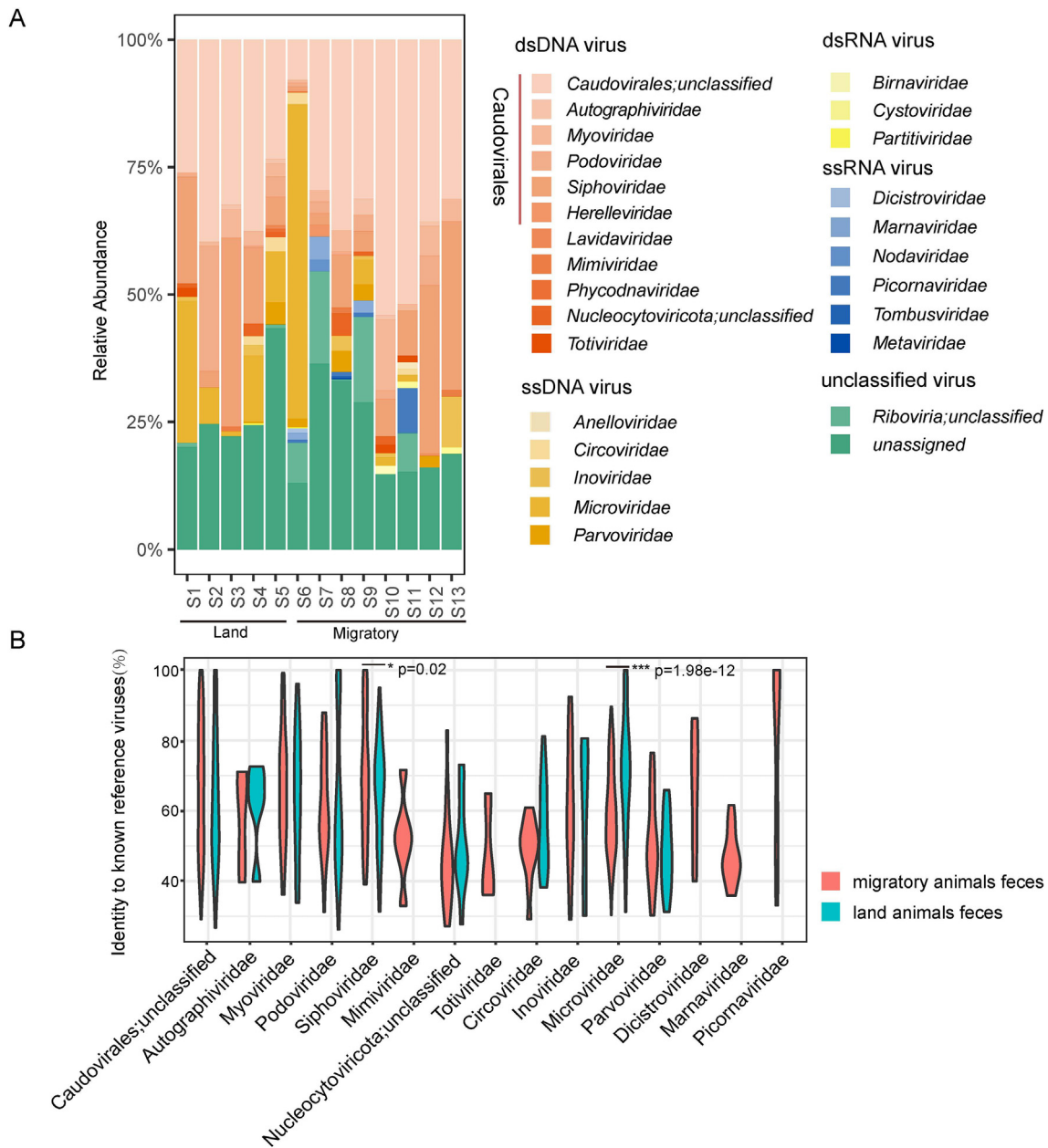


Fig. 2. Viral communities of animal feces from Antarctica and sequence identities of viruses belonging to different families. **A** DNA and RNA viruses of identified families and unassigned viruses are illustrated for each sample. **B** Identities of viral families from migratory birds and land animals compared with known reference viruses.

reference viruses were characterized. The unclassified viruses belonging to *Caudovirales* showed more genetic diversity ranging from ~30% to 100% among the N2 to N6 samples, whereas those from the N1 and N7 samples showed higher similarity to known viruses (65.9%–94.7%) (Fig. 3B). *Demereviridae* was identified only in N6 (identity 60.7%, Supplementary Table S1). *Myoviridae* (27.2%–100%) and *Podoviridae* (43.9%–59.3% for N3 and N4; 47.5%–97.4% for N5 and N6) sequences were not found in any of the seven samples and showed differing identities among them. *Siphoviridae* also had divergent identities to the reference sequences, including 43.2%–62.1% in N2 sample and 30%–100% in samples N3 to N6. The high abundance of viral sequences relating to *Microviridae* in the frozen soil samples indicated they were the dominant viral community in the polar environment; however, their limited identities to the known members of *Microviridae* (58.62%–63.46%) suggested the genetic divergence of this viral community in Arctic soils (Fig. 3B).

3.2. Antarctica and Arctic viral communities share homology with known viruses of different origins and show functional correlations between them

Although Antarctica and the Arctic are geographically isolated from each other, viral sequences may also share evolutionary correlations attributed to genetic homology with known viruses of different origins in the world (Verde et al., 2016). By comparing viral-related sequences to those in the IMG/VR database, we found that the viral sequences from animal fecal samples in Antarctica were mostly related to viruses from aquatic environments (32.25%), followed by humans (11.44%), terrestrial environments (11%), plants (8.38%), invertebrates (2.93%), mammals (2.69%), algae (0.4%), arthropods (0.23%), fungi (0.2%), insects (0.17%), birds (0.13%), air (0.07%), and fish (0.07%) (Fig. 4A). The order of ratios representing the similarity to the viral-related sequences from soil samples in the Arctic to known viruses is mostly similar to that in Antarctica [aquatic environments (26.21%) > terrestrial environments

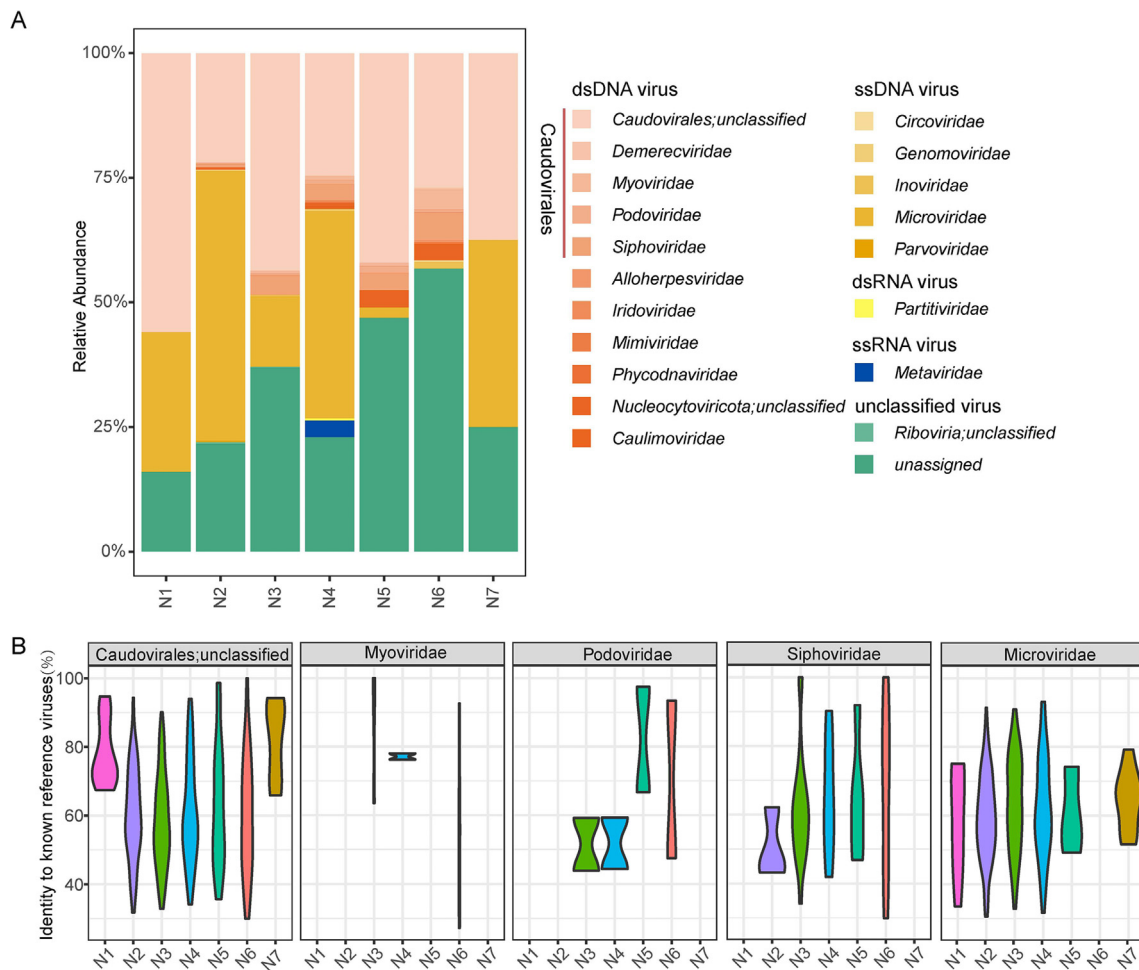


Fig. 3. Viral communities of Arctic soils and sequence identities of phages belonging to different families. **A** DNA and RNA viruses of identified families and unassigned viruses are illustrated for each sample. **B** Identities of phages related to unclassified Caudovirales, Myoviridae, Podoviridae, Siphoviridae, and Microviridae are shown from each of the sampling locations.

(17.58%) > humans (17.51%) > plants (5%) > insects (2.42%) > invertebrates (0.79%) > mammals (0.68%) > air (0.32%) > arthropods (0.32%) > algae (0.11%) > fungi (0.11%)] (Fig. 4A). The results suggest that, despite Antarctica and the Arctic being geographically separated, the polar viromes, regardless of the environment or fauna have genetic homology to the world.

Prediction of viral protein functions further suggested similarities between viruses from Antarctic fecal samples and Arctic soil samples. Viral protein homologs from Antarctica and the Arctic clustered together, resulting in 27 functional clusters with more than 3 nodes (Fig. 4B). Of these clusters, eight represented differential viral structural proteins, including three clusters of DNA pilot proteins (clusters 1, 11, and 15), two capsid proteins (clusters 2 and 10), an internal scaffolding protein (cluster 13), a structural protein (cluster 6), and a tail protein (cluster 20). Nineteen clusters were related to viral non-structural proteins, of which 13 clusters were constituted by viral proteins of unknown functions, and six clusters had putative functions related to virus replication (rep protein) (clusters 4, 5, and 21), helicase (cluster 6), and clarified non-structural proteins (clusters 3 and 12). Moreover, 10 of these clusters (1, 2, 4, 5, 7, 11, 12, 13, 15, and 17) contained virus proteins from both the Antarctic fecal samples and Arctic soil samples, including five clusters of structural proteins (clusters 1, 11, 13, and 15 for DNA pilot protein; cluster 2 for capsid protein) and five clusters of non-structural proteins (clusters 4 and 5 for virus replication; clusters 7 and 17 of unknown function; cluster 12 for clarified non-structural protein), which suggests

that their functional correlations regarding the structure of virus particles, mechanism of DNA translocation, and virus replication. The other 17 clusters were composed of viral proteins solely from the Antarctic fecal samples, of which 11 clusters were proteins of unknown functions, suggesting the virus population in the feces of Antarctic animals bearing high functioning diversities (Fig. 4B).

3.3. Novel viruses identified from Arctic soil and Antarctic animal feces

3.3.1. Microviridae-related viruses

A high abundance of Microviridae-related sequences was found in Antarctic animal feces (up to 60.6%) and Arctic soil samples (up to 54.2%) (Supplementary Table S1), indicating an important role of microviruses in the lives of polar fauna and extreme environments.

Phylogenetic analysis based on the major capsid protein (MCP) of microviruses shows that a number of Microviridae-related sequences from Antarctic animal feces and Arctic soils clustered with the unique sequence of Antarctic microvirus TYR_006_viral_SP_13, which was identified from cryoconite holes in 2016 (Sommers et al., 2019), and constitutes a new evolutionary clade (clade I) (Fig. 5A). Meanwhile, only one sequence from Antarctic gull fecal samples was closely related to the clade of Antarctic strains, the sequences of which were also identified from cryoconite holes in 2016 (Sommers et al., 2019) (Fig. 5A). Moreover, four sequences from Antarctica and eight from the Arctic were included in the Gokushovirinae derived from chlamydia and human feces,

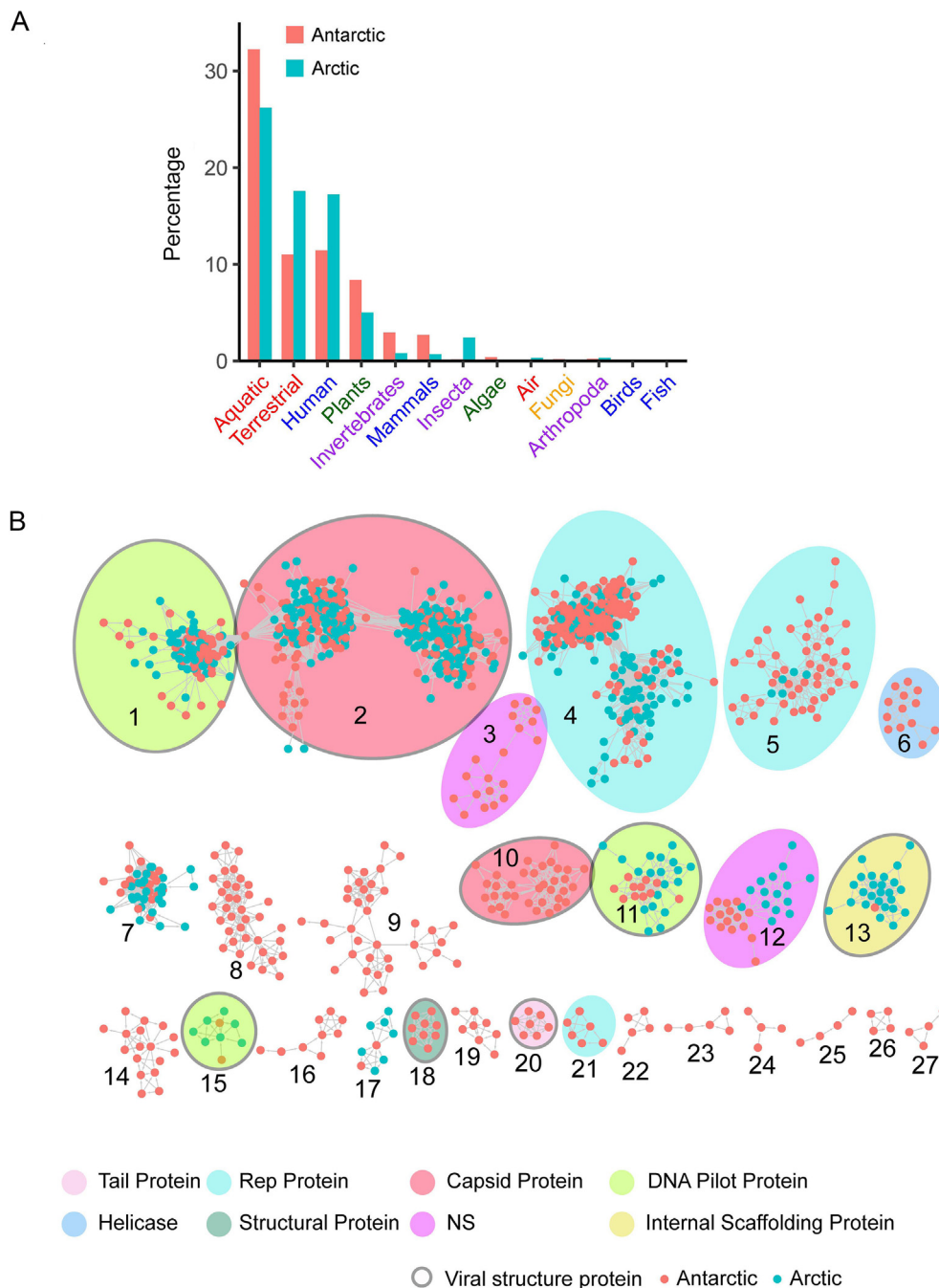


Fig. 4. Genetic correlation and functional homology analysis of viral communities in the polar regions. **A** Genetic correlation of Antarctic and Arctic viruses with known viruses from different ecosystems and organisms using IMG-VR metadata. **B** Prediction of viral protein function of the Antarctic and Arctic viruses. Viral proteins with functional homology are clustered in networks. Each node is a complete protein or a partial protein, and each link represents a hit with an e-value 10^{-10} (node number >3 are shown). Clusters containing proteins of similar putative functions are shaded with the same color so as to be distinguished from other clusters, whereas clusters containing proteins of unknown functions are not shaded. Clusters of structural proteins are circled to distinguish them from non-structural proteins.

whereas ten other sequences from Arctic soils formed a new cluster (clade II) related to but not included in *Gokushovirinea* (Fig. 5A). No sequences were found to be included in *Bullavirinae*, which contains microviruses from *Escherichia coli* (Fig. 5A).

3.3.2. Eukaryotic circular rep-encoding single-stranded (CRESS) DNA virus

CRESS DNA viruses comprise a large group of circular replication-associated protein (rep)-encoding ssDNA viruses. There are bacterial, archaeal, and eukaryotic CRESS DNA viruses. Eukaryotic CRESS DNA

viruses include seven families: *Anelloviridae*, *Bacilladnaviridae*, *Circoviridae*, *Geminiviridae*, *Genomoviridae*, *Nanoviridae*, and *Smacoviridae* (Zhao et al., 2019). Viral sequences related to the families *Anelloviridae* and *Circoviridae* were found in Antarctic feces (Fig. 2A and Supplementary Table S1), and sequences of *Circoviridae* and *Genomoviridae* were found in frozen Arctic soil (Fig. 3A and Supplementary Table S1). The complete genomes of two novel CRESS DNA viruses (E226 and E754) were obtained from Antarctic fecal samples (S6). The genome sequence of E226 (CNGdb accession no. CNA0047396) is 1966 nt in length and

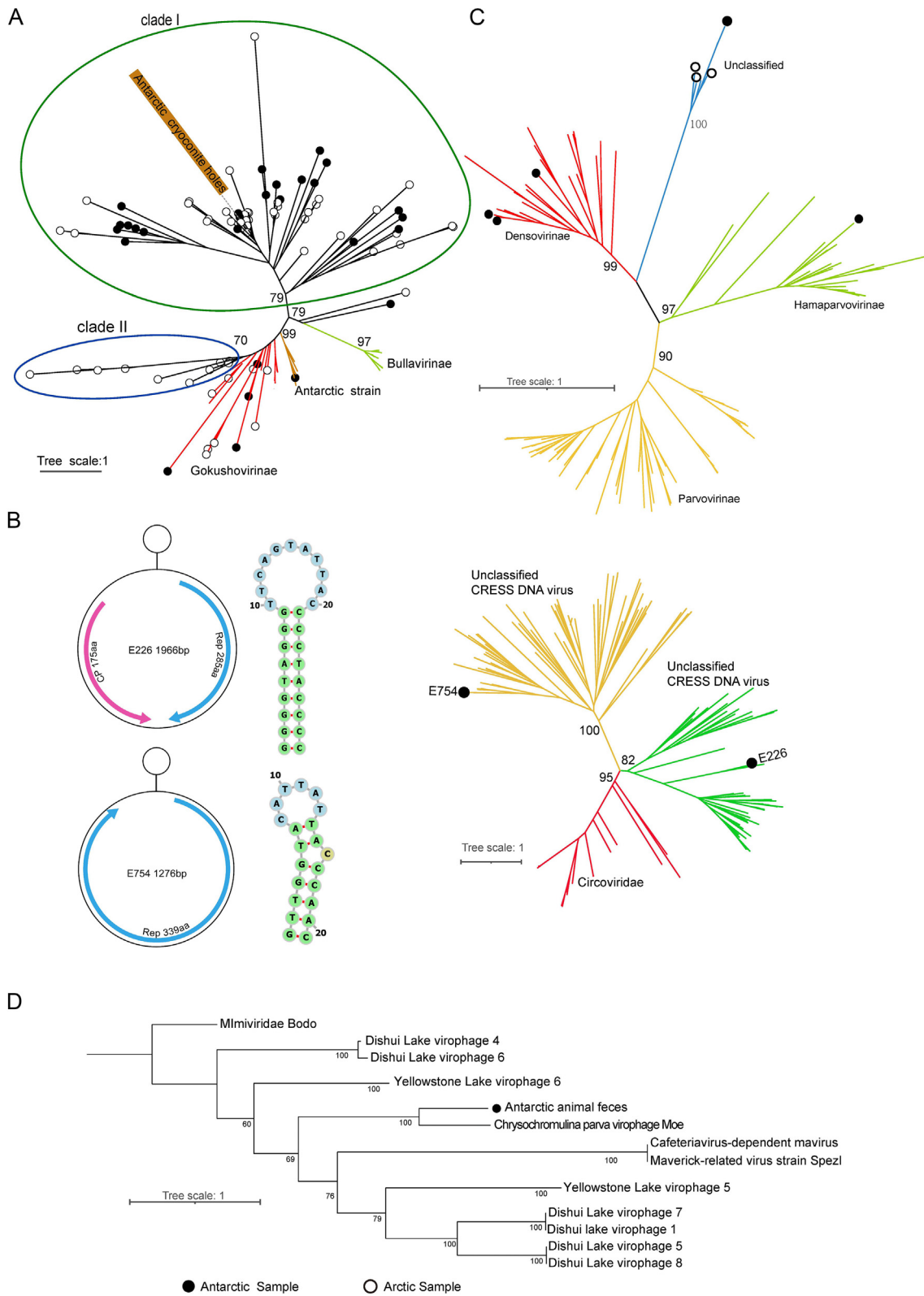


Fig. 5. Phylogeny of viral sequences found in Antarctic animal feces and Arctic soils. Phylogenetic tree of *Microviridae* (A), two novel CRESS DNA viruses related to *Circoviridae* (B), *Parvoviridae* (C), and *Larvidaviridae* (D) were constructed using different models as described in Methods. Sequences from Antarctica are shown in black solid circles and sequences from the Arctic are shown in hollow circles. The genomic organization and structures of the two novel CRESS DNA viruses are illustrated in (B) on the left.

encodes two ORFs of the replication-associated protein (rep) protein homologous to the capsid protein, whereas the genome of E754 (CNCBdb accession no. CNA0047395) is 1276 nt in length and encodes one ORF of the rep protein (Fig. 5B, left). The virus E226 shares 62.28% amino acid similarity with the replication protein of an uncultured marine virus (AGA18440) identified from Saanich Inlet, British Columbia, Canada in 2013, and 47.46% with the capsid protein of CRESS virus sp. (AXH76836) from abalone in 2018. The rep protein of virus E754 shares 54% identity with the McMurdo Ice Shelf pond-associated circle DNA virus-5 (YP_009047137) found in Antarctica in 2014. The genomes of both viruses had a typical structure featuring the CRESS DNA virus genome topology, each having a stem-loop structure (29 nt and 21 nt in length for E226 and E754, respectively) and a nanonucleotide motif sequence in the non-coding region (Fig. 5B, left). According to the CRESS DNA virus classification system proposed by Rosario (Dayaram et al., 2016; Rosario et al., 2012), the virus E226 belongs to type I CRESS DNA virus (rep protein is encoded in the positive strand and CP in the negative strand) and E754 belongs to type VII (circular molecules encoding only rep). Phylogenetic analysis based on the rep protein shows that E226 and E754 belonged to two different clusters of unclassified CRESS DNA viruses and were most closely related to *Circoviridae* (Fig. 5B, right).

3.3.3. Parvoviridae

Sequences related to the family *Parvoviridae* were found in the Antarctic fecal samples from kelp gulls (S6), giant petrels (S12, S13), terns (S10), and emperor penguins (S4). Meanwhile, parvovirus-relating sequences were also found in the Arctic soil samples, including N2, N4 and N5. A phylogenetic tree was constructed using the sequences of non-structural protein 1 (*NS1*) gene according to the classification standard of parvovirus issued by the International Committee on Taxonomy of Viruses (ICTV) in 2021. It showed that parvoviruses from Antarctica presented divergent evolutionary relationships, as the sequences found in kelp gull feces (S6) and penguin (S3) were clustered with densovirus (*Densovirinae*), which mostly infect arthropod hosts, whereas the sequence from tern feces belonged to the *Hamaparvovirinae*, the members of which were mostly found in animals. In addition, sequences from the Arctic soil samples and Antarctic emperor penguin feces formed a separate clade by clustering with unclassified parvoviruses, which were identified from anal swabs of wild and zoo birds in China in 2018 (Fig. 5C).

3.3.4. Lavidaviridae-related virus

Giant viruses are frequently found in aqueous environments and terrestrial soils (Kerepesi and Grolmusz, 2017). As expected, sequences related to giant viruses belonging to *Mimiviridae* and *Phycodnaviridae* were identified from the Antarctic fecal samples and Arctic soil samples, suggesting their presence in extremely cold environments. dsDNA viruses of the family *Lavidaviridae*, commonly known as virophages, are a fascinating group of eukaryotic viruses that depend on a coinfecting giant virus of *Mimiviridae* for their propagation. The superfamily 3 helicase (S3H) is one of the most conserved virophage genes (Fischer, 2020). Lavidaviridae-related sequences were found in the Antarctic fecal samples (S7), which shared 29%–49% with other known virophages, suggesting the novelty of virophages in the polar region. A phylogenetic tree based on the sequence of the superfamily 3 helicase protein shows that this novel virus is closely related to *Chrysochromulina parva* virophage Moe, which was isolated from a giant virus infecting the freshwater haptophyte *C. parva* in Lake Ontario, Canada in 2011 (Stough et al., 2019) (Fig. 5D). However, it shares limited sequence similarity with the *C. parva* virophage Moe (aa identity: 49%), suggesting distinct genetic diversity of virophages between the Antarctic and other regions. In addition, *Lavidaviridae* fragments were found in the Gentoo penguin fecal sample (S5), with an average identity of 72.6% (Supplementary Table S1).

4. Discussion

Because of the geographical isolation of Antarctica, its organisms might have comparatively low pressure from pathogenic infections (Dobson and Foufopoulos, 2001; Grimaldi et al., 2015). However, early studies initiated in the 1970s using serological methods showed that Antarctic animals, such as seals, penguins, petrels, and skuas, were exposed to various viruses from the viral families *Birnaviridae*, *Flaviviridae*, *Herpesviridae*, *Orthomyxoviridae*, and *Paramyxoviridae* (Smeele et al., 2018), suggesting that these animals are susceptible to some virus species and are likely to become the host of virus infection. Later, researchers discovered a number of viral communities in Antarctic wildlife using probe-based assays based on polymerase chain reaction (PCR), Sanger sequencing, and next-generation sequencing. In a recent study, RNA viromes were characterized in swabs of Antarctic penguins and their parasites, comprised of 107 viral species belonging to at least 11 different families, including *Astroviridae*, *Bunyaviridae*, *Caliciviridae*, *Coronaviridae*, *Herpesviridae*, *Orthomyxoviridae*, *Paramyxoviridae*, *Picobirnaviridae*, *Picornaviridae*, *Reoviridae*, and *Rhabdoviridae* (Wille et al., 2020). Moreover, circoviruses (Levy et al., 2020a), and CRESS DNA viruses (Levy et al., 2020b) were found in cloacal swabs of penguins from the Antarctic. These results suggest that Antarctic wildlife is exposed to higher risks from virus infection than previously reported (Crane et al., 2018; Wille et al., 2020), and indicate that Antarctic penguins might serve as potential reservoirs for a number of viruses more than hosts which were susceptible to viruses.

In Antarctica, penguins and seal furls are the major land animals, although they can also swim in oceans. Migratory birds also remain in the Antarctic region. For example, the south polar skua undergoes trans-equatorial migration and is well adapted to extreme polar weather by breeding all around Antarctica (Devillers, 1977; Ritz et al., 2006). After the breeding season, they migrate northward, cross the Antarctic convergence, and can be seen in boreal oceans as far north as 66° N. In this process, they share the food web and environment with other land animals. Then, it is possible that the migratory birds and land animals as well as other living entities could be affected by the Antarctica-specific microorganisms including viruses. This study investigated the viromes of feces from migratory birds and land animals, revealing the high genetic diversity of the viral communities. The feces of migratory birds contained 13 viral families in common with that of land animals, although the identities differed between them, whereas viruses from 10 other families were found in the feces of migratory birds (Supplementary Table S1, Fig. 2A), but not from land animals. These differences may be attributed to their different diets and environments.

A recent study investigated the viromes of cloacal swabs from Antarctic penguins, as well as their parasite ticks, and suggested the occurrence of potential virus transmission between the vectors and hosts in polar regions (Wille et al., 2020). Ticks could be taken to Antarctica by migratory birds from other continents after a long distance, which may provide another feasible way to enhance gene flow between continents and separate polar regions (Wille et al., 2020). The potential connection of viral communities in Antarctica with other environments and living entities in the world was suggested in this study by comparing viral sequences with other known viruses of various origins. Viruses in the feces of Antarctic animals were linked to aquatic, terrestrial, human, plants, and invertebrates, and similar links were also noted among the viral communities from Arctic soil. Functional correlations of viruses from Antarctica and the Arctic were also suggested.

To the best of our knowledge, little is known about the viral communities in frozen Arctic soils. In our study, we found dsDNA virus, ssDNA virus, small dsRNA virus, and ssRNA virus, which are of great significance in understanding the composition of environmental viruses in the Arctic region and provide a preliminary dataset for exploring the relationship between virome and ecology of the Arctic region in the future. As expected, the viral communities in Arctic soils are mainly

composed of bacteriophages, which may play an important role in driving carbon processing in the ecological environment and in microbiological chemical cycling by interacting with bacteria (Breitbart et al., 2018; Chevallereau et al., 2022). *Microviridae* is known for its wide distribution and can be found in all kinds of ecosystems, such as marine environments (Labonté and Suttle, 2013; Tucker et al., 2011), freshwater habitats (López-Bueno et al., 2009; Roux et al., 2012), the human gut or feces (Roux et al., 2012), and peatlands (Quaiser et al., 2015). The viral sequences related to *Microviridae* identified in Antarctic animal feces and Arctic soils showed huge genetic diversity with other known viruses of this family. In a previous study, Antarctica-derived microviruses were identified from the Antarctic cryoconite hole in 2016 and formed one clade (Antarctica strains) related to *Goukushovirinae* (Sommers et al., 2019). In this study, we found that one viral sequence clustered with these Antarctic strains. In addition, most *Microviridae* sequences showed genetic similarity with only one microvirus sequence from the Antarctic cryoconite hole in 2016 and formed a divergent clade. Because the Antarctica animal feces and Arctic soils were collected in 2012–2014, it revealed the existence of novel *Microviridae*-related viruses in both Antarctica and Arctic ecosystems since 2014, at least two years before the Antarctica strains were identified from cryoconite holes in 2016. The other clade (clade II), composed of viral sequences identified in Arctic soils and closely related to *Goukushovirinae*, further revealed the presence of novel microviruses in polar regions as early as 2013–2014. These studies also demonstrate the huge genetic diversity, novelty, and complexity of microviruses in Antarctica and the Arctic. *Parvoviridae* can be divided into three subfamilies, *Densovirinae*, *Hamaparvovirinae*, and *Parvovirinae*. *Densoviruses* mainly infecting invertebrates were found in sea stars (Hewson et al., 2014), mosquitoes (Boonnak et al., 2015), whereas *Hamaparvovirus* infecting invertebrates were found in penaeid shrimp (Pénzes et al., 2020). This suggests that the *Parvoviridae*-related viruses found in the feces of Antarctic animals may be attributed to their presence in the Antarctic environment or in the diet of Antarctic birds.

This study was a retrospective analysis of the data generated by 454 sequencing in 2014. Owing to the limited amount of data obtained by technical methods and the large difference between the viral sequences and the known virus reference database, only very few viruses were identified at that time. Nevertheless, this study re-analyzed the data and revealed the diversity, complexity, and novelty of viral communities in Antarctica and the Arctic regions present nearly a decade ago. We found that the proportions of viral sequences widely varied among these samples (Tables 1 and 2). The results of sequencing might be affected by whether fresh samples were collected or not. Moreover, there are still unknown sequences from these polar region samples, which cannot be referred to any known viruses. Therefore, the limited proportion of viral sequences may indicate that there could be unknown (novel) viruses that remained to be explored. As little is known about the viromes of Arctic soils, we compared the identified viral families with those from Arctic water samples collected early in 2010 and 2011 (De Cárcer et al., 2015) and Antarctic soils from 2014 to 2015 (Adriaenssens et al., 2017), which were also investigated using metagenomic methods. The RNA viruses found in this study were rarely identified in Arctic freshwater and Antarctic soils, and of the DNA viruses, viruses of high abundance, such as unclassified *Caudovirales* and *Microviridae* in Arctic soil, were of low abundance in Arctic freshwater or not detected in Antarctic soils (Supplementary Table S2). Therefore, the retrospective analysis demonstrated high genetic diversity of viral communities from the polar regions early in 2012–2014, which suggests the importance of retrospective analysis using previous data to identify novel viral sequences.

The limitation of this study is that the samples were collected nearly ten years ago, and each sample was sequenced but might not generate highly informative data of great depth. Nevertheless, the results of this study, together with previous reports (Adriaenssens et al., 2017; De Cárcer et al., 2015; Labbé et al., 2020; López-Bueno et al., 2009; Sommers et al., 2019; Zablocki et al., 2014), show that bacterial phages are always the major viral communities in polar regions, suggesting a critical role for

phages in polar ecosystems. Moreover, a number of unassigned viruses could not be included in any viral taxonomy. Their roles in polar ecosystems and their association with the food web remain to be explored. Nevertheless, based on retrospective analyses of metagenomic sequencing data generated in 2014, this study reveals the genetic diversity, complexity, and novelty of viral communities in Antarctica and Arctic regions as early as 2012–2014, and suggests that these viral communities are genetically and functionally correlated even when their locations are geographically separated. These results improve our understanding of global viral communities and suggest the need of follow-up investigation on polar viromes to reveal their substantial roles in the global viral community.

5. Conclusions

This study performed a retrospective analysis on the metagenomic data from the Antarctic animal feces and Arctic soil samples, which revealed the abundant genetic diversity of DNA and RNA viruses in the polar regions as early as 2012–2014. From these data, novel viruses and sequences of new strains were found. Moreover, the viral communities are genetically and functionally correlated, suggesting the complexity and novelty of viruses in Antarctica and Arctic regions. The results promoted our understanding of polar viromes nearly a decade ago, and suggested the necessity to perform further investigation of viral communities as well as their roles in the global ecosystem.

Data availability

The metagenomic data and the viral genome sequences are available in the China National GeneBank database (<https://db.cngb.org/>) under the accession no. CNA0047375–CNA0047394, CNA0047395, and CNA0047396, respectively. All the data generated during the current study are included in the manuscript.

Ethics statement

This article does not contain any human or animal studies performed by any of the authors.

Author contributions

Jun Wang: data curation, formal analysis, investigation, methodology, software, writing-original draft. Jian Xiao: investigation, methodology, writing-original draft. Zheng Zhu: resource, data curation, investigation. Siyuan Wang: investigation, methodology. Lei Zhang: data curation. Zhaojun Fan: visualization. Yali Deng: software. Zhihong Hu: project administration, resource, supervision. Fang Peng: resource, funding acquisition, project administration, supervision. Shu Shen: conceptualization, funding acquisition, project administration, supervision, validation, writing-review & editing. Fei Deng: conceptualization, funding acquisition, project administration, supervision, validation, writing-review & editing.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the Key Deployment Projects of Chinese Academy of Sciences (KJZD-SW-L11), the International Partnership Program from the Bureau of International Cooperation, Chinese Academy of Sciences (153B42KYSB20200013), the Strategic Priority Science and Technology Program of Chinese Academy of Sciences (Category B) (XDB42000000), the National Natural Science Foundation of China

(grant no. 31900154), and the Chinese polar scientific strategy Research Fund (IC201706).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.virs.2022.08.006>.

References

- Adriaenssens, E.M., Kramer, R., van Goethem, M.W., Makhalyane, T.P., Hogg, I., Cowan, D.A., 2017. Environmental drivers of viral community composition in Antarctic soils identified by viromics. *Microbiome* 5, 83.
- Boonnak, K., Suttiheptumrong, A., Jotekratok, U., Pattanakitsakul, S. nga, 2015. Sphylogenetic analysis reveals genetic variations of densovirus isolated from field mosquitoes in Bangkok and surrounding regions. *Southeast Asian J. Trop. Med. Publ. Health* 46, 207–214.
- Breitbart, M., Bonnaï, C., Malki, K., Sawaya, N.A., 2018. Phage puppet masters of the marine microbial realm. *Nat. Microbiol* 3, 754–766.
- Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D., Azam, F., Rohwer, F., 2002. Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci. U.S.A.* 99, 14250–14255.
- Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973.
- Cavicchioli, R., 2015. Microbial ecology of Antarctic aquatic systems. *Nat. Rev. Microbiol.* 13, 691–706.
- Chevallereau, A., Pons, B.J., van Houte, S., Westra, E.R., 2022. Interactions between bacterial and phage communities in natural environments. *Nat. Rev. Microbiol.* 20, 49–62.
- Crane, A., Goebel, M.E., Krabberger, S., Stone, A.C., Varsani, A., 2018. Novel anelloviruses identified in buccal swabs of Antarctic Fur seals. *Virus Gene.* 54, 719–723.
- Dayaram, A., Galatowitsch, M.L., Argüello-Astorga, G.R., van Bysterveldt, K., Krabberger, S., Stainton, D., Harding, J.S., Roumagnac, P., Martin, D.P., Lefeuvre, P., Varsani, A., 2016. Diverse circular replication-associated protein encoding viruses circulating in invertebrates within a lake ecosystem. *Infect. Genet. Evol.* 39, 304–316.
- De Cárcer, D.A., López-Bueno, A., Pearce, D.A., Alcamí, A., 2015. Biodiversity and distribution of polar freshwater DNA viruses. *Sci. Adv.* 1, e1400127.
- Devillers, P., 1977. The skuas of the northern American Pacific coast. *Auk* 94, 417–429.
- Dobson, A., Foufopoulos, J., 2001. Emerging infectious pathogens of wildlife. *Philos. Trans. R. Soc. B Biol. Sci.* 356, 1001–1012.
- Eric Collins, R., Rocap, G., Deming, J.W., 2010. Persistence of bacterial and archaeal communities in sea ice through an Arctic winter. *Environ. Microbiol.* 12, 1828–1841.
- Fischer, M.G., 2020. The viroplasm family laidaviroidae. *Curr. Issues Mol. Biol.* 40, 1–24.
- Gong, Z., Liang, Y., Wang, M., Jiang, Y., Yang, Q., Xia, J., Zhou, X., You, S., Gao, C., Wang, J., He, J., Shao, H., McMinn, A., 2018. Viral diversity and its relationship with environmental factors at the surface and deep sea of Prydz Bay, Antarctica. *Front. Microbiol.* 9, 2981.
- Gregory, A.C., Zayed, A.A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., Ardyna, M., Arkhipova, K., Carmichael, M., Craud, C., Dimier, C., Dominguez-Huerta, G., Ferland, J., Kandels, S., Liu, Y., Marec, C., Pesant, S., Picheral, M., Pisarev, S., Poulain, J., Tremblay, J.E., Vik, D., Acinas, S.G., Babin, M., Bork, P., Boss, E., Bowler, C., Cochrane, G., de Vargas, C., Follows, M., Gorsky, G., Grimsley, N., Guidi, L., Hingamp, P., Judicone, D., Jaillon, O., Kandels-Lewis, S., Karp-Boss, L., Karsenti, E., Not, F., Ogata, H., Poulton, N., Raes, J., Sardet, C., Speich, S., Stemann, L., Sullivan, M.B., Sunagawa, S., Wincker, P., Culley, A.I., Dutilh, B.E., Roux, S., 2019. Marine DNA viral macro- and microdiversity from Pole to Pole. *Cell* 177, 1109–1123.e14.
- Grimaldi, W.W., Seddon, P.J., Lyver, P.O., Nakagawa, S., Tompkins, D.M., 2015. Infectious diseases of Antarctic penguins: current status and future threats. *Polar Biol.* 38, 591–606.
- Hewson, I., Button, J.B., Gudenkauf, B.M., Miner, B., Newton, A.L., Gaydos, J.K., Wynne, J., Groves, C.L., Hender, G., Murray, M., Fradkin, S., Breitbart, M., Fahsbender, E., Lafferty, K.D., Kilpatrick, A.M., Miner, C.M., Raimondi, P., Lahner, L., Friedman, C.S., Daniels, S., Haulena, M., Marliave, J., Burge, C.A., Eisenlord, M.E., Harvell, C.D., 2014. Densovirus associated with sea-star wasting disease and mass mortality. *Proc. Natl. Acad. Sci. U.S.A.* 111, 17278–17283.
- Hyatt, D., Chen, G.L., LoCasio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinf.* 11, 119.
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Kerepesi, C., Grolmusz, V., 2017. The “Giant Virus Finder” discovers an abundance of giant viruses in the Antarctic dry valleys. *Arch. Virol.* 162, 1671–1676.
- Labbé, M., Girard, C., Vincent, W.F., Culley, A.I., 2020. Extreme viral partitioning in a marine-derived high Arctic lake. *mSphere* 5, e00334, 20.
- Labonté, J.M., Suttle, C.A., 2013. Metagenomic and whole-genome analysis reveals new lineages of gokushoviruses and biogeographic separation in the sea. *Front. Microbiol.* 4, 404.
- Letunic, I., Bork, P., 2019. Interactive Tree of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259.
- Levy, H., Fiddaman, S.R., Djurhuus, A., Black, C.E., Krabberger, S., Smith, A.L., Hart, T., Varsani, A., 2020a. Identification of circovirus genome in a chinstrap penguin (*Pygoscelis antarcticus*) and Adélie Penguin (*Pygoscelis adeliae*) on the Antarctic peninsula. *Viruses* 12, 858.
- Levy, H., Fontenele, R.S., Harding, C., Suazo, C., Krabberger, S., Schmidlin, K., Djurhuus, A., Black, C.E., Hart, T., Smith, A.L., Varsani, A., 2020b. Identification and distribution of novel circoviruses and circular molecules in four penguin species in south Georgia and the antarctic peninsula. *Viruses* 12, 1029.
- López-Bueno, A., Tamames, J., Velázquez, D., Moya, A., Quesada, A., Alcamí, A., 2009. High diversity of the viral community from an Antarctic lake. *Science* 326, 858–861.
- Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274.
- Otasek, D., Morris, J.H., Bouças, J., Pico, A.R., Demchak, B., 2019. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biol.* 20, 185.
- Pagano, A.M., Williams, T.M., 2021. Physiological consequences of Arctic sea ice loss on large marine carnivores: unique responses by polar bears and narwhals. *J. Exp. Biol.* 224 (Pt Suppl. 1), jeb228049.
- Pérez, J.J., Pham, H.T., Chipman, P., Bhattacharya, N., McKenna, R., Agbandje-McKenna, M., Tijssen, P., 2020. Molecular biology and structure of a novel penaeid shrimp densovirus elucidate convergent parvoviral host capsid evolution. *Proc. Natl. Acad. Sci. U.S.A.* 117, 20211–20222.
- Quaiser, A., Dufresne, A., Ballaud, F., Roux, S., Zivanovic, Y., Colombet, J., Sime- Ngando, T., Francez, A.J., 2015. Diversity and comparative genomics of Microviridae in Sphagnum-dominated peatlands. *Front. Microbiol.* 6, 375.
- Ren, J., Song, K., Deng, C., Ahlgren, N.A., Fuhrman, J.A., Li, Y., Xie, X., Poplin, R., Sun, F., 2020. Identifying viruses from metagenomic data using deep learning. *Quant. Biol.* 8, 64–77.
- Ritz, M.S., Hahn, S., Janicke, T., Peter, H.U., 2006. Hybridisation between South polar skua (*Catharacta maccormicki*) and Brown skua (*C. Antarctica lonnbergi*) in the Antarctic Peninsula region. *Polar Biol.* 29, 153–159.
- Rosario, K., Duffy, S., Breitbart, M., 2012. A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Arch. Virol.* 157, 1851–1871.
- Roux, S., Enault, F., Robin, A., Ravet, V., Personnic, S., Theil, S., Colombet, J., Sime- Ngando, T., Debros, D., 2012. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS One* 7, e33641.
- Shen, S., Duan, X., Wang, B., Zhu, L., Zhang, Yanfang, Zhang, J., Wang, J., Luo, T., Kou, C., Liu, D., Lv, C., Zhang, L., Chang, C., Su, Z., Tang, S., Qiao, J., Moming, A., Wang, C., Abudurexiti, A., Wang, H., Hu, Z., Zhang, Y., Sun, S., Deng, F., 2018. A novel tick-borne phlebovirus, closely related to severe fever with thrombocytopenia syndrome virus and Heartland virus, is a potential pathogen article. *Emerg. Microb. Infect.* 7, 95.
- Smele, Z.E., Ainley, D.G., Varsani, A., 2018. Viruses associated with Antarctic wildlife: from serology based detection to identification of genomes using high throughput sequencing. *Viruses Res.* 243, 91–105.
- Sommers, P., Fontenele, R.S., Kringen, T., Krabberger, S., Porazinska, D.L., Darcy, J.L., Schmidt, S.K., Varsani, A., 2019. Single-stranded DNA viruses in antarctic cryoconite holes. *Viruses* 11, 1022.
- Stough, J.M.A., Yutin, N., Chaban, Y.V., Moniruzzaman, M., Gann, E.R., Pound, H.L., Steffen, M.M., Black, J.N., Koonin, E.V., Wilhelm, S.W., Short, S.M., 2019. Genome and environmental activity of a chrysochromulina parvovirus and its viroplasm. *Front. Microbiol.* 10, 703.
- Suttle, C.A., 2007. Marine viruses - major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812.
- Tucker, K.P., Parsons, R., Symonds, E.M., Breitbart, M., 2011. Diversity and distribution of single-stranded DNA phages in the north atlantic ocean. *ISME J.* 5, 822–830.
- Verde, C., Giordano, D., Bellas, C.M., di Prisco, G., Anesio, A.M., 2016. Polar marine microorganisms and climate change. *Adv. Microb. Physiol.* 69, 187–215.
- Wilkins, D., Yau, S., Williams, T.J., Allen, M.A., Brown, M.V., Demaere, M.Z., Lauro, F.M., Cavicchioli, R., 2013. Key microbial drivers in Antarctic aquatic environments. *FEMS Microbiol. Rev.* 37, 303–335.
- Wille, M., Harvey, E., Shi, M., Gonzalez-Acuña, D., Holmes, E.C., Hurt, A.C., 2020. Sustained RNA virome diversity in Antarctic penguins and their ticks. *ISME J.* 14, 1768–1782.
- Wommack, K.E., Colwell, R.R., 2000. Viroplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114.
- Yau, S., Seth-Pasricha, M., 2019. Viruses of polar aquatic environments. *Viruses* 11, 189.
- Zablocki, O., van Zyl, L., Adriaenssens, E.M., Rubagotti, E., Tuffin, M., Cary, S.C., Cowan, D., 2014. High-level diversity of tailed phages, eukaryote-associated viruses, and viroplasm-like elements in the metaviromes of antarctic soils. *Appl. Environ. Microbiol.* 80, 6888–6897.
- Zhao, L., Rosario, K., Breitbart, M., Duffy, S., 2019. Eukaryotic circular rep-encoding single-stranded DNA (CRESS DNA) viruses: ubiquitous viruses with small genomes and a diverse host range. *Adv. Virus Res.* 103, 71–133.