Virology 560 (2021) 116–123

Contents lists available at ScienceDirect

Virology

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

Four novel picornaviruses detected in Magellanic Penguins (*Spheniscus magellanicus*) in Chile

Juliette Hayer [a,*], Michelle Wille [b,c], Alejandro Font [d], Marcelo González-Aravena [d], Helene Norder [e,f], Maja Malmberg [a,g,**]

[a] Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden
[b] Marie Bashir Institute for Infectious Diseases and Biosecurity, School of Life and Environmental Sciences and School of Medical Sciences, The University of Sydney, Sydney, Australia
[c] Department of Microbiology and Immunology, At the Peter Doherty Institute for Infection and Immunity, The University of Melbourne, Melbourne, Australia
[d] nstituto Antártico Chileno, Plaza Muñoz Gamero, 1055, Punta Arenas, Chile
[e] Department of Infectious Diseases/Virology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Sweden
[f] Region Västra Götaland, Sahlgrenska University Hospital, Department of Clinical Microbiology, Gothenburg, Sweden
[g] Department of Biomedical Sciences and Veterinary Public Health, Swedish University of Agricultural Sciences, Uppsala, Sweden

ARTICLE INFO

Keywords:
Sphenisciformes
Penguins
Picornaviridae
Viral metagenomics
Hepatovirus

ABSTRACT

Members of the *Picornaviridae* family comprise a significant burden on the poultry industry, causing diseases such as gastroenteritis and hepatitis. However, with the advent of metagenomics, a number of picornaviruses have now been revealed in apparently healthy wild birds. In this study, we identified four novel viruses belonging to the family *Picornaviridae* in healthy Magellanic penguins, a near threatened species. All samples were subsequently screened by RT-PCR for these new viruses, and approximately 20% of the penguins were infected with at least one of these viruses. The viruses were distantly related to members of the genera *Hepatovirus*, *Tremovirus*, *Gruhelivirus* and *Crahelvirus*. Further, they had more than 60% amino acid divergence from other picornaviruses, and therefore likely constitute novel genera. Our results demonstrate the vast undersampling of wild birds for viruses, and we expect the discovery of numerous avian viruses that are related to hepatoviruses and tremoviruses in the future.

1. Introduction

Penguins (Order: Sphenisciformes) are unique in the avian world. Penguins are found throughout the temperate regions of the Southern Hemisphere, ranging from the Antarctic continent to as far north as the Galapagos Islands. Despite population declines of many penguin species, largely due to climate change and overfishing of their prey (Trathan et al., 2015), little is known about the pathogens and parasites harboured by these birds. Indeed, infections with various microorganisms are known to play a role in reducing avian populations, such as substantial declines of native Hawaiian birds due to avian malaria, decreased survival of albatross due to avian cholera, and decreased survival of a number of North American passerine species due to infection with *West Nile virus* (Atkinson and LaPointe, 2009; George et al., 2015; Jaeger et al., 2020; Warner, 1968).

To date, studies of viral presence and prevalence among penguins has been limited, opportunistic, and related to sick birds in nature (eg (Kane et al., 2012; Molini et al., 2020)) or in rehabilitation centres (eg (Parsons et al., 2015)). Further, studies of viruses in penguins are highly biased towards the charismatic Antarctic Penguins, with serology studies dating back to the 1970's (Smeele et al., 2018). Due to technological limitations, *i.e.* both serology and PCR based studies allow for the assessment of only known viruses, very little progress has been made revealing the viral communities in these species. This has changed dramatically with the rise of metagenomics and metatranscriptomics (Smeele et al., 2018; Wille et al., 2020), wherein novel and highly divergent viral species may be described. As a result, since 2015, more than 25 different novel viruses have been described in Antarctic and sub-Antarctic penguins (Smeele et al., 2018; Wille et al., 2020). These viruses include members from the *Adenoviridae, Astroviridae,*

* Co-corresponding authors.
** Co-corresponding authors. Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden.
E-mail addresses: juliette.hayer@slu.se (J. Hayer), maja.malmberg@slu.se (M. Malmberg).

https://doi.org/10.1016/j.virol.2021.05.010
Received 11 April 2021; Received in revised form 19 May 2021; Accepted 19 May 2021
Available online 25 May 2021

Caliciviridae, Circoviridae, Coronaviridae, Herpesviridae, Paramyxoviridae, Orthomyxovirdae, Polyomaviridae, Papillomaviridae, Picornaviridae, Picobirnaviridae and *Reoviridae* (de Souza et al., 2019b; Miller et al., 2010; Morandini et al., 2019; Smeele et al., 2018; Wille et al., 2019a, 2020). Beyond Antarctica, evaluation of penguins for viruses is haphazard, although many of the same viral families have been detected through both virology and serology studies of Magellanic Penguins (*Spheniscus magellanicus*), African Penguins (*Spheniscus demersus*), Little Blue Penguins (*Eudyptula minor*), and Galapagos Penguins (*Spheniscus mendiculus*) (Fornells et al., 2012; Kane et al., 2012; Mele et al., 2012; Molini et al., 2020; Morgan et al., 1985; Niemeyer et al., 2013; Offerman et al., 2014; Parsons et al., 2015; Travis et al., 2006; Uhart et al., 2020). It is clear from these studies that the viral diversity of penguins is vastly underappreciated, and thus the role of penguins as potential hosts for an array of viruses are yet to be revealed.

Through metagenomics, a number of novel picornaviruses have recently been described in samples from wild birds. These viruses fall into the genera *Megrivirus, Sapelovirus, Avihepatovirus*, and other highly divergent viruses with unassigned genera (Boros et al., 2015, 2017, 2018; Wang et al., 2019; Wille et al., 2018, 2019b). In penguins, seven picornaviruses have been described: Ross virus, Scott virus, Amundsen virus, Shirase virus, Wedell virus, Pingu virus and penguin megrivirus (*Megrivirus E*), all of which have been detected in apparently healthy Antarctic penguins (de Souza et al., 2019b; Wille et al., 2020). Until recently, avian picornaviruses were exclusively associated with morbidity and mortality in poultry and other birds (pigeons and passerines) (Boros et al., 2014, 2016). Through an increased effort in investigating the virome in healthy wild birds, we are beginning to rewrite the narrative of many viral families. There is mounting evidence to suggest that picornaviruses are not exclusively disease causing, but rather that many picornaviruses from different species detected in wild birds are not associated with any signs of disease.

In this study we aimed to identify and characterise viruses of Magellanic Penguins. We sampled faeces from Magellanic Penguins breeding on the Magdalena island, an important breeding colony for this species, situated in southern Chile, and we used a combination of high throughput metagenomic sequencing, followed by PCR screening to disentangle virus diversity and prevalence. The detection of four novel viruses at high prevalence in penguins without obvious disease, provides further evidence that penguins are reservoir hosts for a multitude of viruses belonging to different virus families and genera, including numerous viruses yet to be revealed. Finally, this finding has important evolutionary implications for the emergence/evolution of the picornavirus supergroup comprising *Hepatovirus* and *Tremovirus*.

2. Material and methods

2.1. Ethics statement

The project was ethically approved by the Chilean National Forestry Corporation (CONAF) for the region of Magallanes, Chile (RESOLUCIÓN No:517/2015).

2.2. Study site

Sampling was carried out on the Magellanic Penguin colony situated on the Magdalena Island in the Strait of Magellan in Chilean Patagonia (−52°55′10″S 70°34′34″W), from 19–November 21, 2015. The colony comprises approximately 63,000 breeding pairs of penguins as the last count in 2007.

Freshly deposited faecal samples (n = 107) were collected from birds comprising 72 Magellanic Penguins, 1 Southern Rockhopper Penguin and, 3 Kelp Gulls. Some penguin faeces (n = 30) were sampled on more than once. An additional 2 samples were collected from the soil around the colony. All samples were collected using sterile plastic tools and placed in 1 ml RNAlater (ThermoFisher) and stored at room temperature for up to 72 h prior to storage in −80 °C. For eight penguins, a duplicate sample was stored dry in sterile tubes without RNAlater for approximately 5 h at 8 °C prior to storage in −80 °C. Samples remained at −80 °C until processing.

2.3. Sample preparation, library preparation and sequencing

Twelve of the faecal samples from Magellanic Penguins were selected and sequenced using viral metagenomics, among which 4 were not stored in RNAlater. These samples represented a diversity of faecal colour and texture. Briefly, samples were extracted using Trizol and chloroform and amplified using Ovation RNA-Seq v2 (NuGEN). Sequencing libraries were constructed using the AB Library Builder System (Ion Xpress™ Plus and Ion Plus Library Preparation for the AB Library Builder™ System protocol, ThermoFisher). Template preparation was performed on the Ion Chef™ System using the Ion 520 & Ion 530 Kit-Chef (ThermoFisher) and sequenced using the Ion S5™ XL System (ThermoFisher).

2.4. Bioinformatics analysis of the metagenomics datasets

The obtained reads were trimmed by quality in 5′ and 3’ and filtered by mean quality using PRINSEQ (v 0.20.4) with a PHRED quality score of 20 (Schmieder and Edwards, 2011). The good quality reads were used to produce *de-novo* assemblies with Megahit version 1.1.1 (Li et al., 2016). A taxonomic classification of the obtained contigs was performed by running Diamond (v 0.9.6) (Buchfink et al., 2015) against the non-redundant protein database (release April 2019) and using the LCA algorithm from Megan 6 (v6.11.7) (Huson et al., 2007) to visualise the classification of each contig. A taxonomic classification at the read level was also performed using Diamond (blastx + LCA, using the output format 102).

2.5. Comparative genomics and phylogenetics

For contigs longer than 6000bp, gene prediction was performed using ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/) and protein domains prediction using InterProScan (https://www.ebi.ac.uk/interpro/search/sequence/) (Jones et al., 2014). Relevant motifs in the 2C, 3C and 3D peptides were identified using conserved domain blast (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Reads were subsequently mapped back to viral contigs using the Burrows Wheeler Aligner BWA-MEM (v 0.7.12) (Li and Durbin, 2009). Internal ribosomal entry site (IRES) were assessed using IRESPred (http://bioinfo.unipune.ac.in/IRESPred/IRESPred.html) (Kolekar et al., 2016) and blast and secondary structure searches implemented in IRESite (www.IRESite.org)(database release 2019-03-18) (Mokrejs et al., 2010). We further queried both the entire 5′ UTR and the 5’ UTR at 100bp intervals using parameter setting as indicated in (Asnani et al., 2015) and in cases where this was unsuccessful, of the short regions of sequence identified in the output of IRESite.

Amino acid sequences of the P1 region were aligned using MAFFT (v7.2.50) integrated in Geneious Prime (www.geneious.com) with the E–INS–I algorithm (Katoh and Standley, 2013). Gaps and ambiguously aligned regions were stripped using trimAL v1.4 (Capella-Gutierrez et al., 2009). The most appropriate amino acid substitution model was then determined (here, LG + F + I + G4), and maximum likelihood trees were estimated using IQ-TREE (http://iqtree.cibiv.univie.ac.at/) with 1000 ultrafast bootstraps (Nguyen et al., 2015).

2.6. Prevalence of four novel picornaviruses

RNA was extracted from the 107 fresh faecal samples using QIAamp Fast DNA Stool mini kit (Qiagen), followed by cDNA synthesis using the High-Capacity cDNA Reverse Transcription (Applied Biosystems). Custom primers were designed for each of the four picornaviruses

revealed (Table S1). For Sphenifaro and Sphenigellan viruses a nested PCR approach was employed, and for Sphenimaju and Sphenilena viruses a semi-nested approach was employed. Detailed reaction methods are presented in the Supplemental Methods.

Viral prevalence was calculated using the *bioconf()* package and statistically evaluated using a Chi squared test in R v 3.5.3 integrated into RStudio 1.1.463.

2.7. Data availability

The 4 assembled viral genomes and their annotations (Supplemental Methods) have been deposited to European Nucleotide Archive (ENA), accession numbers: LR897978, LR897979, LR897980, LR897982. Reads have been deposited to ENA, accession number PRJEB40660.

3. Results

3.1. Sequence data and assembly

We performed metagenomics sequencing of 12 faecal samples collected from 10 different Magellanic Penguins in a colony on Magdalena Island, Chile.

High-throughput sequencing of the samples produced 25, 357, 939 reads (range 517,072–3,558,413 reads) with a mean length of 229 bp representing a total of 5.86 gigabases of sequence data. After quality control and filtering using PRINSEQ 24, 027, 073 reads (range 463,974–3,438,3911 reads) remained (5.75 giga bases). Of these we assembled an average of 10,228 contigs, of which 4% comprised viral contigs (Table S2).

For all samples, about 25% of the reads could be taxonomically classified. For 7 samples, greater than 50% of the reads could be classified. The majority of these reads were microbial, *i.e.* bacterial or viral; however, a number of penguin (host) reads were also identified. For 2 samples, is_034_015 and is_034_016 viral reads were found in 33.8% and 17.3% of all reads. A vast majority of those were classified as viruses belonging to the *Picornaviridae* family, and three previously not described viral genomes were assembled. From sample is_034_018, 33% (397 out of 1179 contigs) were classified as viral, of which 313 were classified as *Picornaviridae*.

Overall, at the contig level, more than 25% of *de novo* assembled contigs were successfully classified. Eight samples had greater than 50% of the contigs classified as bacterial, viral or chordates.

3.2. Four novel, divergent picornaviruses in Magellanic Penguins

We revealed four novel picornaviruses in faecal samples from Magellanic Penguins which we have referred to as Sphenifaro virus, Sphenigellan virus, Sphenimaju virus and Sphenilena virus for clarity. These viruses were recovered from 3 different genomic libraries corresponding to samples from 3 individuals. These samples were stored without RNAlater and their original sample IDs are NR2/Penguin 52 (library is_034_015), NR4/Penguin 58 (library is_034_016), and NR6/Penguin 62 (library is_034_018). Two viruses (Sphenifaro virus and Sphenilena virus) were identified in a single sample, *i.e.* from the same penguin. Overall, the abundance (proportion of reads based on read mapping back to assemblies) of each virus was low (<1%) with the exception of Sphenimaju virus, which comprised 58.57% of all reads in the sample (is_034_015), with a very high genome coverage of 54,207. The sample "is_034_016" that contained two different viruses (Sphenifaro virus and Sphenilena virus) had low abundance for both viruses, with coverages of 94 and 68 respectively. Each of these viruses comprised 0.29% and 0.21% of the reads in the sample, respectively. Sample "is_034_018" had 0.11% of the reads corresponding to Sphenigellan virus (Table 1), for a coverage of 42.

These new viruses, recovered from the same penguin colony at the same time, are highly divergent from each other, sharing only 32–43% similarity at the amino acid level, suggesting that they represent not only four different species but potentially four different genera. Three of the viruses were most similar to members in the genus *Hepatovirus* in the *Picornaviridae* family when the genomes were analysed by Blastx. However, at the amino acid level the similarity was low, ranging from 35 to 37% (Table 1). The fourth virus, Sphenifaro virus, was most similar to *Crahelivirus A* described in Red-crowned Cranes (*Grus japonensis)* when using Blastx, although as with the Blast results of the other viruses, the amino acid percentage identity was low (37.91%). By phylogenetic analysis these viruses were found in the same supergroup as members of the *Hepatovirus* and *Tremovirus* genera, in addition to newly described and unassigned viruses from Antarctic penguins (Wille et al., 2020) and *Crahelivirus A* and *Gruhelivirus A* from Red-crowned Cranes (Wang et al., 2019) (Fig. 1, Fig. S1). However, as with the blast results, the viruses revealed here shared <30% amino acid similarity in the P1 region to other viruses in this supergroup. Further, all four viruses had long branch lengths in the phylogeny, all this together suggest that each of these new picornaviruses may represent new genera in the *Picornaviridae* family demonstrating the vast undersampling of viruses in this part of the tree.

To investigate the frequency of these four new viruses in other samples sequenced by high throughput sequencing, all libraries were mapped against each of these new viral genomes. In the sample containing Sphenimaju virus (is_034_015), there were over 5000 reads mapping to the Sphenigellan virus. In the sample with both Sphenifaro virus and Sphenilena virus (is_034_016), there were over 1000 reads mapping the Sphenigellan virus and Sphenimaju virus. Overall, there were 4 libraries with no evidence of any of the novel viruses described here, 2 with reads against only one of the novel viruses, and 6 libraries with evidence of more than one of these viruses (Table S3).

Using InterProScan domain prediction we could predict the 3 main regions P1, P2, P3 for three of the viruses (Fig. 2). For Sphenilena virus, it was not possible to predict the exact P2 region location. Mature peptides 1B, 1C and 1D could also be located, as well as 2C, 3C and 3D peptides for all 4 viruses. On the contrary, peptides 2A, 2B, 3A and 3B could not be predicted using InterProScan (Table S4 – S7). We were able to identify all putative cleavage sites of the polyproteins using alignments with other known viruses of the genus *Hepatovirus* as per Wang et al. (2019). Using conserved domain blast we were able to identify both Walker A and Walker B motifs central to NTPase activity in the 2C

Table 1
Metadata for the four novel hepato-like viruses revealed in this study.

Sample Name	Virus Name	Top Blastn hit	Blastn percentage identity	Length (bp)	Total Number of reads in sample	Number of reads (proportion of reads in sample)
is_034_016	Sphenifaro virus	AUW34301 Picornaviridae red crowned crane	37.91%	7201	916,871	2687 (0.29%)
is_034_018	Sphenigellan virus	YP_009164030 Phopivirus	38.76%	7384	1,090,489	1238 (0.11%)
is_034_015	Sphenimaju virus	YP_009179216 Bat hepatovirus	38.37%	7441	2,689,813	1,575,605 (58.57%)
is_034_016	Sphenilena virus	YP_009215780 Tupaia hepatovirus A	35.83%	7048	916,871	1919 (0.21%)

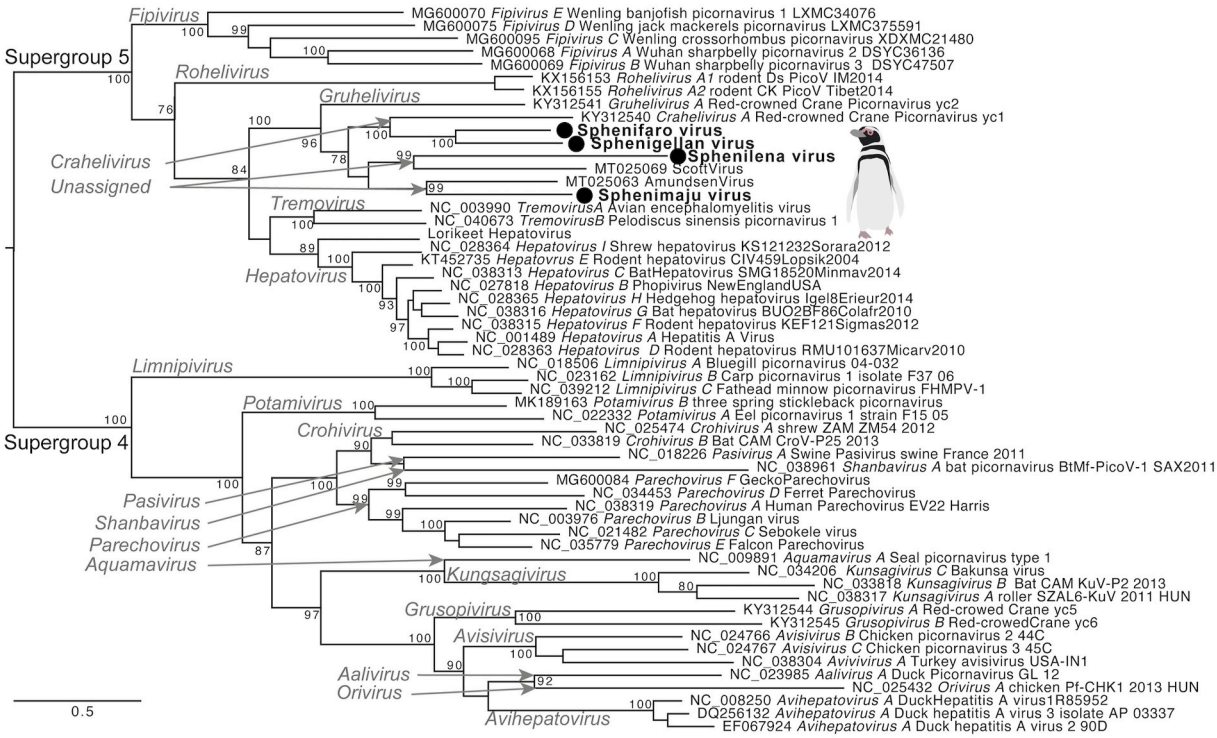

Fig. 1. Maximum likelihood phylogeny of the P1 mature peptide of supergroup 4 and 5 of the *Picornaviridae*. We included representative members of all genera of these groups, as outlined in the 10th update of the ICTV, and the tree was rooted based on these supergroups. Genus names are provided at relevant nodes. Viruses described here are adjacent to a filled circle. Penguin silhouette was developed by M. Wille. Bootstrap values > 70% are shown. The scale bar indicates the number of amino acid substitutions per site.

peptide, the protease active site in the 3C peptide and the RdRp active site in the 3D peptide (Fig. 2).

In addition to the polyprotein, picornaviruses contain a 5′ untranslated region (UTR) containing an internal ribosomal entry site (IRES) and a 3′ UTR. The 5′UTR for Sphenifaro virus, Sphenigellan virus, Sphenimaju virus and Sphenilena virus was 211 bp, 413 bp, 542 bp and 121 bp, respectively. The 5′UTR for the picornaviruses revealed here were generally less than 500bp, and without doing 5′ RACE we are unable to confirm the entire 5′ UTR. Using IRESPred, potential IRES (Type 1) were identified in Sphenifaro virus, Sphenigellan virus and Sphenimaju virus. No IRES was precited for Sphenilena virus, however we recovered only a very short 5′UTR from this virus (121 bp). Using blastn and the blast feature in IRESSite, we found that the 5′UTR of Sphenigellan virus and Sphenimaju virus were similar to hepatoviruses. The 5′ UTR of Spenigellan virus shared 73% nt identity to Phopivirus (*Hepatovirus B*, GenBank accession number KR703607) over 51% of the query sequence (275 nt; positions 199 to 474) and further matched a number of other hepatoviruses. The 5′UTR of Sphenimaju virus shared 73% nt identity to Red-crowned Crane Picornavirus yc-2 (*Gruhelivirus A*, GenBank accession number KY312541) over 64% of the query sequence (403 nt; positions 4 to 407) and further matched a number of other hepatoviruses. IRESsite, but not blastn, found a short match of the 5′ UTR of Sphenilena virus to Avian encephalomyelitis virus (positions 58 to 69 and 39–50). The 3′ UTR of Sphenifaro virus, Sphenigellan virus, Sphenimaju virus and Sphenilena virus was 84 bp, 317 bp, 221 bp and 7 bp, respectively. We identified a polyA tail at the 3′ end of the UTR in Sphenigellan and Sphenimaju virus but we did not identify the entire 3′ UTR for either Sphenifaro or Sphenilena virus. Blastn searches of the 3′ UTR did not show homology to other picornavirus 3′ UTRs.

3.3. The four new picornaviruses are common among magellanic penguins

To reveal the prevalence of the new viruses identified in this study, the 107 faecal samples were screened by RT-PCR for these 4 viruses. The samples were from 72 individual Magellanic Penguins, one Rockhopper penguin (*Eudyptes chrysocome*), and three Kelp gulls (*Larus dominicanus*). Overall, at least one of these 4 viruses were detected in 28 samples (25%; 95% confidence interval [CI] 18–34%). All positive samples were from Magallanic Penguins, of which 22 were positive for at least one of the 4 viruses (21%, 95 CI 14–30%). The prevalence was highest for Sphenilena virus with 11 PCR positive samples. Sphenifaro virus, Sphenigellan virus and Sphenimaju virus were found in 7, 7, and 9 Magellanic Penguin individuals, respectively. There was no statistically significant difference in prevalence between the viruses ($X^2 = 3.01$, df = 3, p = 0.3886) (Fig. 3 A). Interestingly, samples from 7 individuals contained more than one of these viruses. Two individuals were positive for Sphenifaro virus and Sphenigellan virus, one individual for Sphenifaro virus and Sphenimaju virus, three individuals were positive for Sphenimaju virus and Sphenilena virus, finally, one sample contained 3 viruses: Sphenifaro, Spehnigellan, and Sphenilena virus.

Sequencing the ~460bp PCR product demonstrated sequence variability for each of the new viruses. There was an average pairwise identity of 91.6%, 96.5%, 98% and 91% for Sphenifaro, Sphenigellan, Sphenimaju, and Sphenilena virus, respectively. The pairwise similarity for Sphenimaju was higher (*i.e.* less diversity) than that for the other species, despite having a comparable number of PCR products (n = 9) (Fig. 3 B).

4. Discussion

In this study we aimed to reveal the viral diversity of Magellanic Penguins sampled in Chile. By using high throughput sequencing we found four novel viruses belonging to the family *Picornaviridae*. These viruses are highly divergent and share less than 40% of the amino acid sequence in the P1 region with members of *Hepatovirus* and *Tremovirus* that they most closely resemble. The vast majority of the viral reads from most libraries mapped to these novel viruses indicate that these viruses were the dominant species in the faecal virome of the penguins. Further,

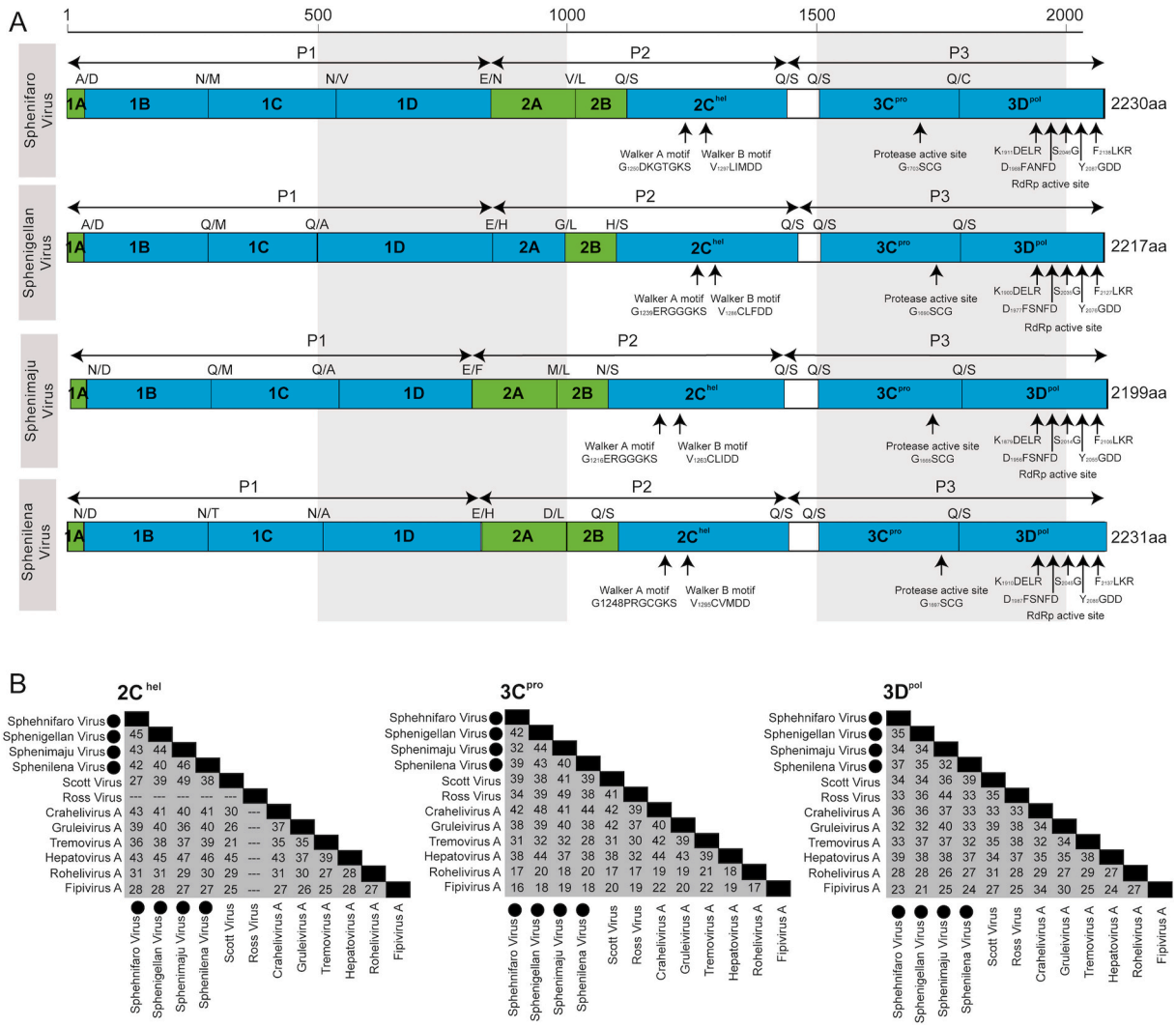

Fig. 2. Genome features for the four novel picornaviruses described in this study. (A) Domains identified using InterProScan are in blue. Detailed results of domain searches using InterProScan are presented in Tables S4–S7. Additional peptides identified through alignments with hepatoviruses and tremoviruses are presented in green. Putative cleavage sites are presented above each schematic. Relevant motifs present in the 2C, 3C and 3D were identified using conserved domain blast and are presented below each schematic (B) Amino acid identity (percentage similarity) of the viruses revealed in this study and select reference sequences of the domains $2C_{hel}$, $3C_{pro}$ and $3D_{pol}$. As Ross virus is incomplete, no percentage identity values were calculated for the 2C. Similarly, Amundsen virus is incomplete and does not have sequence associated with any of the relevant peptides it is not included in panel B. Accession numbers for viruses included in B include: MT025069 Scott virus, MT025070 Ross virus, MT025064 and MT025063 Amundsen virus, KY312540 *Crahelivirus A*, KY312541 *Gruhelivirus A*, NC_003990 *Tremovirus A*, NC_001489 *Hepatovirus A*, KX156153 *Rohelivirus A*, MG600068 *Fipivirus A*.

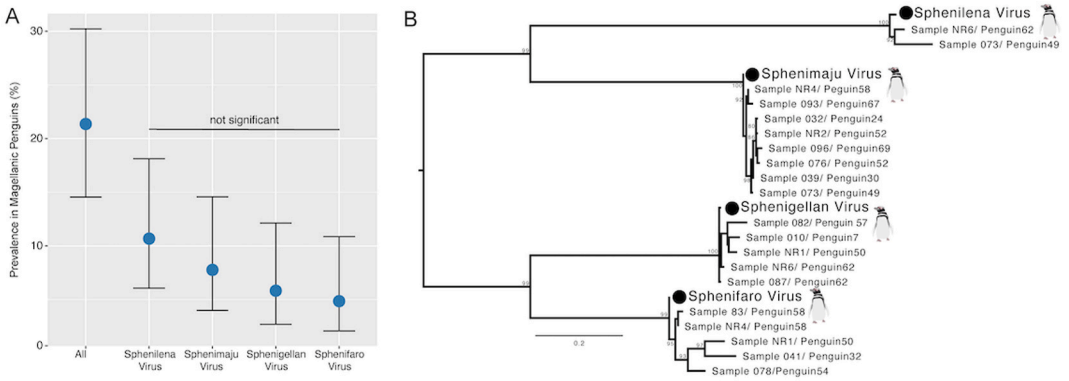

Fig. 3. Diversity of four novel picornaviruses revealed through PCR. (A) Prevalence of each novel virus revealed in this study across samples collected from Magellanic Penguin individuals. The point estimate is presented as a filled circle and error bars correspond to 95% confidence intervals (B) Maximum likelihood tree of ~400bp PCR products sequenced from 20 positive samples. The tree was midpoint rooted for clarity only. Scale bar represents number of nucleotide substitutions per site. Viruses' sequences assembled following metagenomic sequencing are indicated with a filled circle. Samples stored without RNAlater are labelled NR#.

through PCR screening we showed that over 20% of the sampled penguins in this study were shedding at least one of these viruses, and in many individuals we co-detected two or three of these new viruses. The viruses revealed here, in addition to novel viruses from Antarctic penguins likely constitute a number of novel genera. This finding of very divergent avian picornaviruses nested within Supergroup 5 suggests that there are many likely to be more picornaviruses in this group yet to be identified.

The charismatic Sphenisciformes have long been a target for virus surveillance, with early studies initiated in Antarctic Penguins in the 1970's (Lang et al., 2016; Smeele et al., 2018). These studies were limited to screening for only previously described viruses, which were viruses of poultry, such as *Influenza A Virus*, Newcastle Disease virus, infectious bursal disease virus (*e.g.* (Lang et al., 2016; Morgan and Westbury, 1981, 1988; Smeele et al., 2018). As a consequence of this focus on poultry-relevant virus surveillance combined with limitations in technology, until very recently, few viruses had been described in penguin species, globally. Our viral catalogues are expanding, without detailed virus ecology studies, it remains unclear the role that penguins may play as hosts to an array of viruses. This particularly occurs for viruses detected in penguins in rehabilitation centres (Parsons et al., 2015). The high prevalence of the four novel viruses identified in this study, and the high frequency of co-detection, indicates that these penguins are likely to be an important reservoir hosts for these picornaviruses. To understand the host range of these viruses, we would strongly encourage for the sampling and surveillance of not only other populations of Magellanic Penguins, but also other penguin species found in South America, and globally. Evidence for connectivity of penguin species and populations as hosts for viruses is sparse, however studies of avian ortho- and metaavulaviruses provide some clues. First, we have now seen repeated detection of *Avian orthoavulavirus 17, 18, 19* in three Antarctic Penguin species and across a number of colonies along the Antarctic peninsula. More interesting was the detection of *Avian metaavulavirus 10*, first in Rockhopper Penguins from the Falkland/-Malvinas Islands (Miller et al., 2010), and more recently in Antarctic Penguins (Wille et al., 2019a). This data suggests that there is capacity of viruses to be shared across multiple penguin species and locations. How these viruses may fit into the larger migratory flyways used by other birds in South America, such as Red Knot (*Calidris canutus*), is very unclear (de Araujo et al., 2014; de Souza et al., 2019a). In addition to ecological questions, of importance is to understand the route of transmission of these viruses. Given the detection in faeces, and the fact that faecal-oral route of transmission is very common in RNA viruses of wild birds, such as *Influenza A virus* and avian ortho- and metaavulaviruses, we may speculate that these viruses are transmitted by the faecal oral route. This, however, would need to be confirmed by dedicated studies. Taken together, we suggest not only more sampling and metagenomic studies in these hosts, but also more consistent and repeated sampling to reveal both virus diversity and virus ecology.

Metagenomic tools have rapidly allowed for the expansion of described avian picornaviruses, but also redefined our understanding of the impact that these viruses have on their hosts. Prior to 2010, almost all described avian picornaviruses caused disease in their hosts. This includes duck hepatitis virus (*Avihepatovirus A*, genus *Avihepatovirus*) (Kim et al., 2006), turkey hepatitis virus (*Megrivirus C,* genus *Megrivirus*) (Honkavuori et al., 2011) and avian sapelovirus (*Anativirus A*, genus *Anativirus,* previously *Sapelovirus*) (Tseng and Tsai, 2007) which cause hepatitis in domestic ducks and turkeys, avian encephalomyelitis virus (*Tremovirus A*, genus *Tremovirus*) causing a neural disorder encephalomyelitis (Marvil et al., 1999) in domestic gallinaceous birds, and a number of viral genera causing gastroenteritis in domestic galliforms (Boros et al., 2016). Indeed, a recent metagenomic study revealed that a chicken with gastroenteritis was shedding six picornaviruses, simultaneously, although it is unclear which, if any, of these viruses was causing the disease (Boros et al., 2016). It is clear that the viruses isolated from poultry cause disease, however metagenomic studies have now revealed more than 20 novel picornavirus species in wild birds, and on all occasions the sampled birds had no signs of disease (Boros et al., 2015, 2017, 2018; de Souza et al., 2019b; Wang et al., 2019; Wille et al., 2018, 2019b, 2020), with the exception of *Poecevirus A* causing avian keratin disorder (Zylberberg et al., 2018). In this study, we demonstrated not only a high viral load (large proportion of reads), but also that a number of penguins were co-detected with at least two of the new picornaviruses described here without any obvious signs of disease. Similarly, in a metagenomic study, Wille et al. (2018) found Red-necked Avocets (*Recurvirostra novaehollandiae*) to be co-detected with nine different viruses, including 3 highly divergent picornaviruses. This suggests that wild birds are able to tolerate high viral loads and diversity in the absence of overt disease (Medzhitov et al., 2012; Råberg, 2014). One hypothesis for the discordance in disease signs in wild birds as compared to poultry is that wild birds have a long history of host-pathogen co-evolution (van Dijk et al., 2015; Wille et al., 2018). Mass produced chickens and ducks, which suffer from disease when infected with an array of virus from different species, are a relatively new host niche in evolutionary time (Gilbert et al., 2017).

The phylogenetic placement of the four viruses revealed here lend evidence to the interpretation of important evolutionary patterns in the supergroup 5 viruses. That the phylogeny of these viruses generally follows the phylogeny of the hosts from which they are sampled, with fish viruses most ancestral, suggests that there has been a process of virus-host co-divergence, that likely extends hundreds of millions of years. It further suggests that we may expect to find numerous avian picornaviruses that fall sister to the *Hepatovirus* and *Tremovirus* genera. It has been estimated that over 99% of viruses are still to be described (Geoghegan and Holmes, 2017), and therefore our data suggests a large number of picornaviruses are yet to be identified particularly from the great diversity of wild birds.

CRediT authorship contribution statement

Juliette Hayer: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Supervision. **Michelle Wille:** Formal analysis, Investigation, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Alejandro Font:** Investigation, Project administration, Resources. **Marcelo González-Aravena:** Project administration, Resources, Writing – review & editing. **Helene Norder:** Methodology, Investigation, Writing – review & editing, Resources, Supervision. **Maja Malmberg:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Supervision.

Acknowledgments

Financial support was obtained from the DEANN project, Marie Curie Action funded by the European Commission – Grant agreement ID: 612583 and The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning, Formas (grant number 221-2012-586).

We acknowledge personnel at the Chilean National Forestry Corporation (CONAF) and assistance by the National Genomic Infrastructure (NGI), which is a part of the Science for Life Laboratory (SciLifeLab), as well as the Swedish National Infrastructure for large Scale Sequencing (SNISS) initiatives. NGI is financially supported by SciLifeLab, Knut and Alice Wallenberg foundation, and the Swedish Research Council.

We acknowledge the SLU Bioinformatics Infrastructure (SLUBI) for providing computing resources and supporting JH.

MW is supported by an Australian Research Council DECRA Fellowship (DE200100977).

We acknowledge the support with molecular methods provided by Alexandra Corduneanu at University of Agricultural Sciences and Veterinary Medicine in Romania and Fernando Pinto at Swedish University of Agricultural Sciences in Sweden.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virol.2021.05.010.

Declaration of interest

The authors declare no conflicts of interest.

References

Asnani, M., Kumar, P., Hellen, C.U., 2015. Widespread distribution and structural diversity of Type IV IRESs in members of Picornaviridae. Virology 478, 61–74.

Atkinson, C.T., LaPointe, D.A., 2009. Introduced avian diseases, climate change, and the future of Hawaiian honeycreepers. J. Avian Med. Surg. 23, 53–63.

Boros, A., Pankovics, P., Adonyi, A., Fenyvesi, H., Day, J.M., Phan, T.G., Delwart, E., Reuter, G., 2016. A diarrheic chicken simultaneously co-infected with multiple picornaviruses: complete genome analysis of avian picornaviruses representing up to six genera. Virology 489, 63–74.

Boros, A., Pankovics, P., Matics, R., Adonyi, A., Bolba, N., Phan, T.G., Delwart, E., Reuter, G., 2017. Genome characterization of a novel megrivirus-related avian picornavirus from a carnivorous wild bird, western marsh harrier (*Circus aeruginosus*). Arch. Virol. 162, 2781–2789.

Boros, A., Pankovics, P., Reuter, G., 2014. Avian picornaviruses: molecular evolution, genome diversity and unusual genome features of a rapidly expanding group of viruses in birds. Infect. Genet. Evol. 28, 151–166.

Boros, A., Pankovics, P., Simmonds, P., Kiss, T., Phan, T.G., Delwart, E., Reuter, G., 2018. Genomic analysis of a novel picornavirus from a migratory waterfowl, greater white-fronted goose (Anser albifrons). Arch. Virol. 163, 1087–1090.

Boros, A., Pankovics, P., Simmonds, P., Pollak, E., Matics, R., Phan, T.G., Delwart, E., Reuter, G., 2015. Genome analysis of a novel, highly divergent picornavirus from common kestrel (F*alco tinnunculus*): the first non-enteroviral picornavirus with type-I-like IRES. Infect. Genet. Evol. 32, 425–431.

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60.

Capella-Gutierrez, S., Silla-Martinez, J.M., Gabaldon, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

de Araujo, J., de Azevedo Jr., S.M., Gaidet, N., Hurtado, R.F., Walker, D., Thomazelli, L. M., Ometto, T., Seixas, M.M., Rodrigues, R., Galindo, D.B., da Silva, A.C., Rodrigues, A.M., Bomfim, L.L., Mota, M.A., Larrazabal, M.E., Branco, J.O., Serafini, P., Neto, I.S., Franks, J., Webby, R.J., Webster, R.G., Durigon, E.L., 2014. Avian influenza virus (H11N9) in migratory shorebirds wintering in the Amazon Region, Brazil. PloS One 9 e110141.

de Souza, W.M., Fumagalli, M.J., de Araujo, J., Ometto, T., Modha, S., Thomazelli, L.M., Durigon, E.L., Murcia, P.R., Figueiredo, L.T.M., 2019a. Discovery of novel astrovirus and calicivirus identified in ruddy turnstones in Brazil. Sci. Rep. 9, 5556.

de Souza, W.M., Fumagalli, M.J., Martin, M.C., de Araujo, J., Orsi, M.A., Sanfilippo, L.F., Modha, S., Durigon, E.L., Proenca-Modena, J.L., Arns, C.W., Murcia, P.R., Figueiredo, L.T.M., 2019b. Pingu virus: a new picornavirus in penguins from Antarctica. Virus Evol 5, vez047.

Fornells, L.A., Silva, T.F., Bianchi, I., Travassos, C.E., Liberal, M.H., Andrade, C.M., Petrucci, M.P., Veiga, V.F., Vaslin, M.F., Couceiro, J.N., 2012. Detection of paramyxoviruses in Magellanic penguins (*Spheniscus magellanicus*) on the Brazilian tropical coast. Vet. Microbiol. 156, 429–433.

Geoghegan, J.L., Holmes, E.C., 2017. Predicting virus emergence amid evolutionary noise. Open Biol 7, 170189. https://doi.org/10.1098/rsob.170189.

George, T.L., Harrigan, R.J., LaManna, J.A., DeSante, D.F., Saracco, J.F., Smith, T.B., 2015. Persistent impacts of West Nile virus on North American bird populations. Proc. Natl. Acad. Sci. U. S. A. 112, 14290–14294.

Gilbert, M., Xiao, X.M., Robinson, T.P., 2017. Intensifying poultry production systems and the emergence of avian influenza in China: a 'One Health/Ecohealth' epitome. Arch. Publ. Health 75, 48. https://doi.org/10.1186/s13690-13017-10218-13694.

Honkavuori, K.S., Shivaprasad, H.L., Briese, T., Street, C., Hirschberg, D.L., Hutchison, S. K., Lipkin, W.I., 2011. Novel picornavirus in Turkey poults with hepatitis, California, USA. Emerg. Infect. Dis. 17, 480–487.

Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data. Genome Res. 17, 377–386.

Jaeger, A., Gamble, A., Lagadec, E., Lebarbenchon, C., Bourret, V., Tornos, J., Barbraud, C., Lemberger, K., Delord, K., Weimerskirch, H., Thiebot, J.B., Boulinier, T., Tortosa, P., 2020. Impact of annual bacterial epizootics on albatross population on a remote Island. EcoHealth 17, 194–202.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30, 1236–1240.

Kane, O.J., Uhart, M.M., Rago, V., Pereda, A.J., Smith, J.R., Van Buren, A., Clark, J.A., Boersma, P.D., 2012. Avian pox in magellanic penguins (*Spheniscus magellanicus*). J. Wildl. Dis. 48, 790–794.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Kim, M.C., Kwon, Y.K., Joh, S.J., Lindberg, A.M., Kwon, J.H., Kim, J.H., Kim, S.J., 2006. Molecular analysis of duck hepatitis virus type 1 reveals a novel lineage close to the genus Parechovirus in the family *Picornaviridae*. J. Gen. Virol. 87, 3307–3316.

Kolekar, P., Pataskar, A., Kulkarni-Kale, U., Pal, J., Kulkarni, A., 2016. IRESPred: web Server for prediction of cellular and viral internal ribosome entry site (IRES). Sci Rep-Uk 6.

Lang, A.S., Lebarbenchon, C., Ramey, A.M., Robertson, G.J., Waldenström, J., Wille, M., 2016. Assessing the role of seabirds in the ecology of influenza A viruses. Avian Dis. 60, 378–386.

Li, D., Luo, R., Liu, C.M., Leung, C.M., Ting, H.F., Sadakane, K., Yamashita, H., Lam, T. W., 2016. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102, 3–11.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.

Marvil, P., Knowles, N.J., Mockett, A.P., Britton, P., Brown, T.D.K., Cavanagh, D., 1999. Avian encephalomyelitis virus is a picornavirus and is most closely related to hepatitis A virus. J. Gen. Virol. 80, 653–662.

Medzhitov, R., Schneider, D.S., Soares, M.P., 2012. Disease tolerance as a defense strategy. Science 335, 936–941.

Mele, S., Picardo, K.F., Cunningham, G.B., Hurd, D.D., 2012. The isolation and identification of a causative agent of the feather disorder found in African Penguins (Spheniscus demersus). Paper 15 Biology Faculty/Staff Publications. https://fisherpub.sjfc.edu/biology_facpub/15.

Miller, P.J., Afonso, C.L., Spackman, E., Scott, M.A., Pedersen, J.C., Senne, D.A., Brown, J.D., Fuller, C.M., Uhart, M.M., Karesh, W.B., Brown, I.H., Alexander, D.J., Swayne, D.E., 2010. Evidence for a new avian paramyxovirus serotype 10 detected in rockhopper penguins from the Falkland Islands. J. Virol. 84, 11496–11504.

Mokrejs, M., Masek, T., Vopalensky, V., Hlubucek, P., Delbos, P., Pospisek, M., 2010. IRESite-a tool for the examination of viral and cellular internal ribosome entry sites. Nucleic Acids Res. 38, D131–D136.

Molini, U., Aikukutu, G., Roux, J.P., Kemper, J., Ntahonshikira, C., Marruchella, G., Khaiseb, S., Cattoli, G., Dundon, W.G., 2020. Avian influenza H5N8 outbreak in african penguins (*Spheniscus demersus*), Namibia, 2019. J. Wildl. Dis. 56, 214–218.

Morandini, V., Dugger, K.M., Ballard, G., Elrod, M., Schmidt, A., Ruoppolo, V., Lescroël, A., Jongsomjit, D., Massaro, M., Pennycook, J., Kooyman, G.L., Schmidlin, K., Kraberger, S., Ainley, D.G., Varsani, A., 2019. Identification of a novel adélie penguin circovirus at cape crozier (Ross island, Antarctica). Viruses 11, 1088. https://doi.org/10.3390/v11121088.

Morgan, I.R., Westbury, H.A., 1981. Virological studies of adelie penguins (*Pygoscelis adelia*) in Antarctica. Avian Dis. 25, 1019–1026.

Morgan, I.R., Westbury, H.A., 1988. Studies of viruses in penguins in the vestfold hills. Hydrobiologia 165, 263–269.

Morgan, I.R., Westbury, H.A., Campbell, J., 1985. Viral infections of little Blue penguins (*Eudyptula minor*) along the southern coast of Australia. J. Wildl. Dis. 21, 193–198.

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.

Niemeyer, C., Favero, C.M., Kolesnikovas, C.K., Bhering, R.C., Brandao, P., Catao-Dias, J. L., 2013. Two different avipoxviruses associated with pox disease in Magellanic penguins (*Spheniscus magellanicus*) along the Brazilian coast. Avian Pathol. 42, 546–551.

Offerman, K., Carulei, O., van der Walt, A.P., Douglass, N., Williamson, A.L., 2014. The complete genome sequences of poxviruses isolated from a penguin and a pigeon in South Africa and comparison to other sequenced avipoxviruses. BMC Genom. 15, 463.

Parsons, N.J., Gous, T.A., van Wilpe, E., Strauss, V., Vanstreels, R.E., 2015. Herpesvirus-like respiratory infection in African Penguins *Spheniscus demersus* admitted to a rehabilitation centre. Dis. Aquat. Org. 116, 149–155.

Råberg, L., 2014. How to live with the enemy: understanding tolerance to parasites. PLoS Biol. 12 https://doi.org/10.1371/journal.pbio.1001989 e1001989.

Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27, 863–864.

Smeele, 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. Virus Res. 243, 91–105.

Trathan, P.N., García-Borboroglu, P., Boersma, D., Bost, C.-A., Crawford, R.J.M., Crossin, G.T., Cuthbert, R.J., Dann, P., Davis, L.S., Puente, S.D.L., Ellenberg, U., Lynch, H.J., Mattern, T., Pütz, K., Seddon, P.J., Trivelpiece, W., Wienecke, B., 2015. Pollution, habitat loss, fishing, and climate change as critical threats to penguins. Conserv. Biol. 29, 31–41.

Travis, E.K., Vargas, F.H., Merkel, J., Gottdenker, N., Miller, R.E., Parker, P.G., 2006. Hematology, serum chemistry, and serology of Galapagos penguins (*Spheniscus mendiculus*) in the Galapagos Islands, Ecuador. J. Wildl. Dis. 42, 625–632.

Tseng, C.H., Tsai, H.J., 2007. Sequence analysis of a duck picornavirus isolate indicates that it together with porcine enterovirus type 8 and simian picornavirus type 2 should be assigned to a new picornavirus genus. Virus Res. 129, 104–114.

Uhart, M., Thijl Vanstreels, R.E., Gallo, L., Cook, R.A., Karesh, W.B., 2020. Serological survey for select infectious agents in wild Magellanic Penguins (*Spheniscus magellanicus*) in Argentina, 1994-2008. J. Wildl. Dis. 56, 66–81.

van Dijk, J.G.B., Fouchier, R.A.M., Klaassen, M., Matson, K.D., 2015. Minor differences in body condition and immune status between avian influenza virus-infected and noninfected mallards: a sign of coevolution? Ecol Evol 5, 436–449.

Wang, Y., Yang, S., Liu, D., Zhou, C., Li, W., Lin, Y., Wang, X., Shen, Q., Wang, H., Li, C., Zong, M., Ding, Y., Song, Q., Deng, X., Qi, D., Zhang, W., Delwart, E., 2019. The fecal virome of red-crowned cranes. Arch. Virol. 164, 3–16.

Warner, R.E., 1968. The role of introduced diseases in the extinction of the endemic awaiian avifauna. Condor 70, 101–120.
Wille, M., Aban, M., Wang, J., Moore, N., Shan, S., Marshall, J., Gonzalez-Acuna, D., Vijaykrishna, D., Buttler, J., Wang, J., Hall, R.J., Williams, D.T., Hurt, A.C., 2019a. Antarctic penguins as reservoirs of diversity for avian avulaviruses. J. Virol. https://doi.org/10.1128/JVI.00271-00219.
Wille, M., Eden, J.S., Shi, M., Klaassen, M., Hurt, A.C., Holmes, E.C., 2018. Virus-virus interactions and host ecology are associated with RNA virome structure in wild birds. Mol. Ecol. 27, 5263–5278. https://doi.org/10.1111/mec.14918.
Wille, M., Harvey, E., Shi, M., Gonzalez-Acuna, D., Holmes, E.C., Hurt, A.C., 2020. Sustained virome diversity in Antarctic penguins and their ticks. ISME J. 14, 1768–1782. https://doi.org/10.1038/s41396-020-0643-1.
Wille, M., Shi, M., Klaassen, M., Hurt, A.C., Holmes, E.C., 2019b. Virome heterogeneity and connectivity in waterfowl and shorebird communities. ISME J. 13, 2603–2616.
Zylberberg, M., Van Hemert, C., Handel, C.M., DeRisi, J.L., 2018. Avian keratin disorder of Alaska black-capped chickadees is associated with Poecivirus infection. Virol. J. 15, 100.