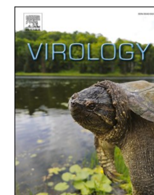
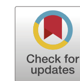
Contents lists available at ScienceDirect

Virology

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





Opportunistic sampling of yellow canary (*Crithagra flaviventris*) has revealed a high genetic diversity of detected parvoviral sequences.

Yuhao Zhang [a], Saranika Talukder [b], Md Safiul Alam Bhuiyan [c], Lei He [a,**], Subir Sarker [d,*]

[a] The Key Lab of Animal Disease and Public Health /Luoyang Key Laboratory of Live Carrier Biomaterial and Animal Disease Prevention and Control, Henan University of Science and Technology, Luoyang, 471023, Henan, China

[b] College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, QLD, 4811, Australia

[c] Faculty of Sustainable Agriculture, Livestock Production, Universiti Malaysia Sabah, Sandakan, Sabah, Malaysia

[d] Biomedical Sciences & Molecular Biology, College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, QLD 4811, Australia

ARTICLE INFO

Handling Editor: Dr. Jasmine Tomar

Keywords:
Parvovirus
Yellow canary
Next-generation sequencing
Genome
Phylogenetics

ABSTRACT

Parvoviruses are known to be significant viral pathogens that infect a wide range of species globally. However, little is known about the parvoviruses circulating in Australian birds, including yellow canaries. Here, we present four parvoviral sequences including three novel parvoviruses detected from 10 yellow canaries (*Crithagra flaviventris*), named canary chaphamaparvovirus 1 and -2 (CaChPV1 and CaChPV2), canary dependoparvovirus 1 and -2 (CaDePV1 and CaDePV2). The whole genome sequences of CaChPV1, CaChPV2, CaDePV1, and CaDePV2 showed the highest identity with other parvoviruses at 76.4%, 75.9%, 84.0%, and 59.1%, respectively. Phylogenetic analysis demonstrated that CaChPV1 and CaChPV2 were clustered within the genus *Chaphamaparvovirus*. Meanwhile, CaDePV1 and CaDePV2 fall within the genus *Dependoparvovirus* and have the closest evolutionary relationship to the bird-associated dependoparvoviruses. Overall, this study enriched our understanding of the genetic diversity among avian parvoviruses within the *Parvoviridae* family.

1. Introduction

Parvoviruses are nonenveloped, icosahedral DNA viruses with diameters ranging from 19 to 24 nm, a relatively small genome consisting of single-stranded DNA arranged in a linear configuration, spanning between 4 and 6 kilobases (kb)(Day and Zsak, 2010; Kisary et al., 1984). The parvovirus genome contains two major expression cassettes, one of which encodes non-structural proteins (NS), the enzymes required for replication, such as the nonstructural protein 1 (NS1), which is a replicase, and also encoding NS2 and NS3 through alternative splicing. The other expression cassette, designated capsid gene (VP1) that encodes structural proteins(Kailasan et al., 2015). Parvoviruses are considered significant viral pathogens capable of infecting a wide range of hosts, including many vertebrate and invertebrate species(Kapgate et al., 2018). The family *Parvoviridae* is divided into three subfamilies, *Parvovirinae*, *Densovirinae*, and *Hamaparvovirinae*, which are classified based on phylogenetic analysis. The subfamily *Parvovirinae* includes ten genera, which comprise viruses infecting mammals, birds, and reptiles. Whereas, the subfamily *Densovirinae* contains eight genera consisting of viruses that infect invertebrate hosts, including insects, crustaceans, and echinoderms; *Hamaparvovirinae* has five genera that contain viruses identified in vertebrate and invertebrate hosts(Cotmore et al., 2014, 2019; Penzes et al., 2020).

Within the subfamily *Hamaparvovirinae*, parvoviruses under the genus *Chaphamaparvovirus*, have been detected in the fecal samples of chickens, turkeys, peafowls, rats, pigs, grus japonensis and eidolon helvums, as well as in a variety of tissue samples, including serums, rectal swabs, nasal swabs, and lung lavages sourced from pigs (Kapgate et al., 2018; M et al., 2019; Palinski et al., 2016; Reuter et al., 2014; Sarker, 2022). In addition, chaphamaparvoviruses were also identified in the liver tissue samples from parrot species, including the rainbow lorikeet (*Trichoglossus moluccanus*) and little corella (*Cacatua sanguinea*) (Chang et al., 2020; Sutherland and Sarker, 2023). Among the ten genera within the subfamily *Parvovirinae*, the genus *Dependoparvovirus* has been found in various species such as snake, muscovy duck, and goose, with additional adeno-associated viruses included in the genus *Dependoparvovirus*(Farkas et al., 2004; Zadori et al., 1995).

The yellow canary (*Crithagra flaviventris*) which belongs to the order

* Corresponding author.
** Corresponding author.
E-mail addresses: helei4280546@163.com (L. He), subir.sarker@jcu.edu.au (S. Sarker).

https://doi.org/10.1016/j.virol.2024.110081
Received 3 November 2023; Received in revised form 18 March 2024; Accepted 3 April 2024
Available online 4 April 2024
0042-6822/© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Passeriformes, family Fringillidae, genus *Crithagra*. Currently, the yellow canary is listed as 'least concern' under the International Union for Conservation of Nature (IUCN) red list of threatened species. Geographically, the existing habitats of yellow canary are mainly located in the southern regions of Africa, such as Angola, Botswana, Lesotho, Namibia, and South Africa(Cardoso et al., 2016). Although the population trend of the yellow canary is currently described as 'stable' by the IUCN, it is essential to be aware of several potential factors affecting the population of this species, such as parvovirus infection, which has been identified as a critical risk factor for the conservation of threatened bird populations(Kapgate et al., 2018; Shan et al., 2022).

Over the past few decades, epidemiological investigations have demonstrated a wide distribution of parvovirus-infected species, including the members of the order Passeriformes(Shan et al., 2022). In detail, there has been only one report of parvovirus identification in *Spinus spinus* of the family Fringillidae, further categorized in the genus *Carduelis*(Shan et al., 2022). To our knowledge, there have been no reports about the parvovirus infection on the species classified under the family Fringillidae, genus *Crithagra*, including the yellow canary. Very little is known about the parvoviruses that may be harbored by yellow canary bird species. Therefore, this study aims to characterize four novel parvoviruses identified during a viral metagenomic study in the faecal materials collected from a group of yellow canary (*Crithagra flaviventris*).

2. Materials and methods

2.1. Sampling, ethical considerations, and extraction of DNA

Faecal materials from a group of 10 yellow canaries (*Crithagra flaviventris*) were obtained from a pet shop in Victoria during the provision of routine animal care and were stored at −80 °C in a cryovial with RNALater until further processing. The Animal Ethics Committee at La Trobe University was informed that findings from the material (with no bird touching) were to be used in a publication, and a formal waiver of ethics approval was granted. Potential impurities from the faecal samples, such as host cells, bacteria, food particles, and free nucleic acids, were removed and was followed by enrichment of virus particles performed as per stated methods (Sarker, 2021a; Vibin et al., 2018) Briefly, the faecal sample was aseptically resuspended and vigorously homogenized in sterile phosphate-buffered saline (PBS) (1:10) and centrifuged at 2500×*g* for 90 min at 4 °C. The supernatant was filtered using a 0.80 μm syringe filter, and the filtrate was then ultracentrifuged at 178,000×*g* and 30 psi for 1 h at 4 °C using a Hitachi Ultracentrifuge CP100NX (Hitachi Koki Co., Ltd., Tokyo, Japan). The supernatant was discarded, and the pellet was suspended in 130 μL of sterile PBS. The filtrates were then nuclease-treated using 2 μL of benzonase nuclease (25–29 U/μL, purity >90, Millipore; Merck KgaA, Darmstadt, Germany) and 1 μL of micrococcal nuclease (2,000,000 gel units/mL; New England Biolabs, Ipswich, Massachusetts, USA) and incubated at 37 °C for 2 h. The nuclease reaction was stopped by adding 3 μL of 500 mM ethylenediaminetetraacetic acid. The viral nucleic acids were extracted using a QIAamp Viral RNA Mini Kit (Qiagen, Valencia, CA, USA) without carrier RNA, which allowed the simultaneous extraction of viral DNA and RNA. The quantity and quality of the isolated nucleic acids were determined using a Nanodrop and an Agilent Tape Station at the Genomic Platform, La Trobe University.

2.2. Next-generation sequencing

Before library construction, extracted nucleic acids were subjected to cDNA synthesis, and amplification was carried out using the Whole Transcriptome Amplification Kit (WTA2, Sigma-Aldrich, Darmstadt, Germany) as per manufacturer instructions. Amplified polymerase chain reaction (PCR) products were then purified using the Wizard® SV Gel and PCR Clean-Up kit (Promega, Madison, WI, USA). The quantity and quality of the purified products were checked using a Qubit dsDNA high sensitivity assay kit with Qubit Fluorometer v4.0 (Thermo Fisher Scientific, Waltham, MA, USA). The library construction was performed using the Illumina DNA Prep (Illumina, San Diego, CA, USA) as per kit instructions, starting with 250 ng of DNA as measured by a Qubit Fluorometer v4.0 (Thermo Fisher Scientific, USA). The quality and quantity of the prepared library were assessed by the Australian Genome Research Facility, Melbourne, Australia. Cluster generation and sequencing of the library were performed with 150 bp paired-end reads on the Illumina® NovaSeq chemistry, according to the manufacturer's instructions.

2.3. Bioinformatic analyses

The resulting raw sequencing reads were analyzed as per the established pipeline (Athukorala et al., 2021; Sarker et al., 2017, 2019; Sutherland et al., 2019) using Geneious Prime® (version 2023.1.1, Biomatters, New Zealand) and LIMS-HPC system (a High-Performance Computer specialized for genomics research at La Trobe University). Briefly, preliminary quality evaluation for all raw reads was generated and pre-processed to remove ambiguous base calls, and poor-quality reads and trimmed to remove the Illumina adapter sequences. Trimmed sequence reads were mapped against the chicken genome Gallus (GenBank accession no. NC006088.5) to remove likely host DNA contamination. In addition, reads were further mapped to the Escherichia coli bacterial genomic sequence (GenBank accession no. U00096) to remove possible bacterial contamination. Cleaned and unmapped reads were used as input data for de novo assembly using a SPAdes assembler (version 3.10.1) (Bankevich et al., 2012) under the 'careful' parameter in the LIMS-HPC system (a High-Performance Computer specialized for genomics research at La Trobe University). The resulting contigs were compared against the nonredundant nucleotide and protein databases on GenBank using BLASTN and BLASTX (Benson et al., 2013b), respectively, with an E-value threshold of 1×10^{-5} to remove potential false positives. Contigs that had significant BLAST hits with bacteria, eukaryotes, or fungi were filtered out to remove non-viral reads. Virus contigs of interest greater than 500 nucleotides (nt) were imported into Geneious Prime® (version 2023.1.1) for further functional analysis. The detected viruses were annotated using Geneious Prime® (version 2023.1.1), where genus-specific published viruses were used as a reference guideline.

2.4. Comparative genomics and phylogenetic analyses

Genomic comparison of the newly sequenced complete viral genomes of parvoviruses were visualized using Geneious (version 2023.1.1). The sequence identity between the selected viral sequences was identified against representative parvovirus sequences by MAFFT software (Version 11.0.11)(Katoh and Standley, 2013). A comparative analysis of the protein sequences encoded by the predicted ORFs was conducted using BLASTX and BLASTP, leveraging the GenBank database, and the relevant information on parvoviruses used in this study are shown in Table 1 and Table 2.

Phylogenetic analysis of the novel parvoviruses genome sequence identified in this study was performed with other selected parvoviruses genome sequences available in the GenBank database. The amino acid sequences of replication protein (nonstructural protein, NS1) and capsid protein (VP1) were aligned by MAFFT software with the L–INS–I algorithm (gap open penalty 1.53 and offset value 0.123). Thereafter, to determine the best-fit model to construct maximum-likelihood (ML) analysis, the best model testing was performed in MEGA software (Version 11.0.11)(Tamura et al., 2021). Phylogenetic analysis for the amino acid sequences of complete NS1 and VP1 were performed using the WAG model with gamma distribution rate variation and a proportion of invariable sites (WAG + G + I). All positions with less than 95% site coverage were eliminated and the number of bootstrap replicates for the phylogenetic analysis was 1000.

Table 1
Related chaphamaparvovirus genome sequences used in further analysis.

Chapparvoviruses	Abbreviation	Host	Year of detection[a]	Country of origin[a]	GenBank accession number	References
Canary chaphamaparvovirus 1	CaChPV1	*Crithagra flaviventris*	2023	Australia	OR670978	This study
Canary chaphamaparvovirus 2	CaChPV1	*Crithagra flaviventris*	2023	Australia	OR670979	This study
Psittaciform chaphamaparvovirus 1	PsChPV1	*Trichoglossus moluccanus*	2013	Australia	MT457858	Chang et al. (2020)
Psittaciform chaphamaparvovirus 2	PsChPV2	*Neophema elegans/N. splendida*	2020	Australia	MZ364297	Sarker (2021b)
Psittaciform chaphamaparvovirus 3	PsChPV3	*Cacatua sanguinea*	2020	Australia	OK482703	Sutherland and Sarker (2023)
Peafowl parvovirus 1	PfPV1	*Pavo cristatus*	2013	China	MK988619	Liu et al. (2020)
Galliform chaphamaparvovirus 1	GaChV1	*Meleagris gallopavo*	2012	Hungary	KF925531	Reuter et al. (2014)
Chicken chapparvovirus 1	CChV1	*Gallus gallus*	2015	Brazil	MG846441	*[b]
Galliform chaphamaparvovirus 2	GaChV2	*Gallus gallus*	2015	Brazil	MG846443	*
Parvoviridae sp.	RcPV9	*Grus japonensis*	2014	China	KY312548	Wang et al. (2019)
Parvoviridae sp.	RcPV10	*Grus japonensis*	2014	China	KY312549	Wang et al. (2019)
Rodent chaphamaparvovirus 1	RoChV1	*Mus musculus*	2014	USA	MF175078	Williams et al. (2018)
Chiropteran chaphamaparvovirus 1	ChiCPV1	*Desmodus rotundus*	2010	Brazil	NC032097	Souza et al. (2017)
Bat parvovirus	BtPV	*Eidolon helvum*	2013	Cameroon	MG693107	Yinda et al. (2018)
Rodent chaphamaparvovirus 2	RoChV2	*Rodent*	2014	China	KX272741	Yang et al. (2016)
Ungulate chaphamaparvovirus 1	UChPPV1	*Sus*	2015	USA	KU563733	Palinski et al. (2016)
Parvoviridae sp.	EcPV150	*Emberiza chrysophrys*	2018	China	MT138311	Shan et al. (2022)
Parvoviridae sp.	EsPV36	*Emberiza schoeniclus*	2018	China	MT138291	Shan et al. (2022)
Parvoviridae sp.	WbPV1	Wild bird	2018	China	MT138217	Shan et al. (2022)

[a] If the collection date/country was not available, the year/country of submission to GenBank is reported.
[b] unpublished.

Table 2
Related dependoparvovirus genome sequences were used in further analysis.

Dependoparvoviruses	Abbreviation	Host	Year of detection[a]	Country of origin[a]	GenBank accession number	References
Canary dependoparvovirus 1	CaDePV1	*Crithagra flaviventris*	2023	Australia	OR670977	This study
Canary dependoparvovirus 2	CaDePV2	*Crithagra flaviventris*	2023	Australia	OR670980	This study
Adeno-associated virus 2	AdDePV2	*Homo sapiens*	1993	USA	NC001401	Cao et al. (2014)
Goose parvovirus	GoDePV	*Anser anser anser*	2000	USA	NC001701	Zadori et al. (1995)
Muscovy duck parvovirus	MdDePV	*Cairina moschata*	2004	USA	NC006147	Zadori et al. (1995)
Snake adeno-associated virus	SaDePV	*Python regius*	2003	Germany	NC006148	Farkas et al. (2004)
Adeno-associated virus 5	AdDePV5	*Homo sapiens*	1998	USA	NC006152	Chiorini et al. (1999)
Parvoviridae sp.	AiPV1	*Anser indicus*	2016	China	MT138284	Shan et al. (2022)
Parvoviridae sp.	SsPV142	*Spinus spinus*	2018	China	MT138314	Shan et al. (2022)
Parvoviridae sp.	TnPV173	*Turdus naumanni*	2018	China	MT138317	Shan et al. (2022)
Parvoviridae sp.	AiPV4	*Anser indicus*	2016	China	MT138330	Shan et al. (2022)

[a] If the collection date/country was not available, the year/country of submission to GenBank is reported.

3. Results

3.1. Genomes of four parvoviral sequences

A total of four, including three novel parvoviruses genome sequences from faecal materials obtained from 10 yellow canaries (*Crithagra flaviventris*) were fully sequenced using NGS. There were no further eukaryotic viral contigs detected in this study.

Based on the origin, sequence characterization, and evolutionary linkage of these four parvoviruses, named canary chaphamaparvovirus 1 and -2 (CaChPV1 and CaChPV2, respectively), canary dependoparvovirus 1 and -2 (CaDePV1, CaDePV2, respectively). These assembled genomes were all linear single-stranded DNA (ssDNA) molecules with lengths of 4292 bp, 4243 bp, 4813 bp, and 4126 bp for CaChPV1, CaChPV2, CaDePV1 and CaDePV2, respectively. The A + T content of these parvoviral complete genomes detected were found to be 40.7% (CaChPV1), 41.3% (CaChPV2), 54.4% (CaDePV1), and 47.4% (CaDePV2), which corresponds well to other parvoviruses detected in

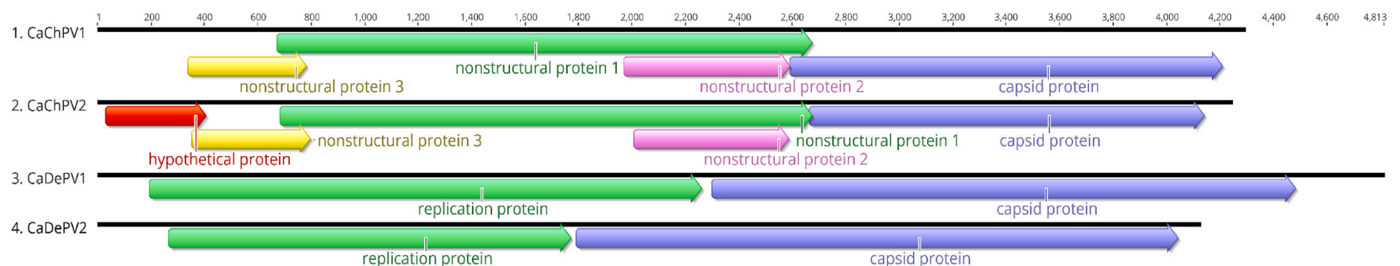


Fig. 1. A schematic illustration of the CaChPV1, CaChPV2, CaDePV1, and CaDePV2 genome was performed using Geneious Prime software (version 2023.1.1). The arrow represents the transcriptional direction of each gene, and the numbers at the top of the illustration represent the corresponding nucleotide positions of these putative novel parvoviruses.

other avian host species (Supplementary Tables S1 and S2). The overall genomic organization of these four novel parvoviral sequenced in this study are similar to other parvoviruses, with two major predicted open-reading frames (ORFs) that typically contain a replication initiator protein or nonstructural protein (NS1) and a viral capsid protein (VP1) (Fig. 1). Additionally, detailed information on the whole genome structures of these parvoviruses is shown in Supplementary Table S3. Like other parvoviruses, the lengths of complete NS1 genes of four detected parvoviral sequences in this study range from 501 to 668 amino acids. As can be seen in Table 3, they encode helicases, including the conserved ATP- or GTP-binding Walker A loop (**GP**xNT**GKT/S**), Walker B (xxx**WEE**), Walker B' (**KQ**xx**E**G**x**xxxxPx**K**), Walker C (**P**xxx**T**x**N**) as motifs, and a conserved replication initiator (endonuclease) motif, xx**H**u**H**xxxx (where, conserved amino acids are indicated in bold letters).

3.2. Comparative analyses

Comparative analysis of the protein sequences revealed significant identity (E-value $\leq 10^{-5}$) in the protein sequences of all ORFs among the four parvoviruses. However, there was an ORF in the CaChPV2 genome that encoded for a hypothetical protein (Fig. 1). In the case of CaChPV1, the nonstructural protein 3 (NS3) exhibited the highest amino acid sequence identity (50.36%) to the hypothetical protein of Phoenicopteridae parvo-like hybrid virus (GenBank accession no. QTE03742.1) when analyzed with BLASTX. Furthermore, the NS1, NS2, and VP1 of CaChPV1 displayed the highest amino acid sequence identity with *Parvoviridae* sp. (WbPV1, GenBank accession no. MT138217.1), at 79.3%, 67.3%, and 76.5%, respectively. As for CaChPV2, the NS1 protein exhibited amino acid sequence identities ranging from 33.0% to 75.8% when compared to other related chaphamaparvoviruses, and the highest identity was observed with *Parvoviridae* sp. (EsPV150, GenBank accession no. MT138311) from *Emberiza schoeniclus* (Supplementary Table S1). Similarly, the VP1 protein of CaChPV2 displayed amino acid sequence identities between 30.7% and 72.0%, with the greatest identity seen in *Parvoviridae* sp. (EsPV36, GenBank accession no. MT138291) (Supplementary Table S1). As shown in Supplementary Table S2, at both the nucleotide and amino acid levels, the NS1 and VP1 proteins of CaDePV1 shared the highest sequence identity with *Parvoviridae* sp. (SsPV142, GenBank accession no. MT138314) from *Spinus spinus*. Meanwhile, in the case of CaDePV2, the NS1 protein and VP1 protein exhibited the highest sequence identity at the amino acid level with *Parvoviridae* sp. (TnPV173, GenBank accession no. MT138317) from *Turdus naumanni*, at 74.9% and 63.2%, respectively. Overall, there were significant differences in the NS1 protein between three detected canary parvoviruses (CaChPV1, CaChPV2 and CaDePV2) and other isolated parvoviruses, thus these results indicates that the canary parvoviruses identified in this study should be considered as novel.

Furthermore, an alignment of multiple complete genome sequences unveiled interesting relationships. For the two canary chaphamaparvoviruses identified in this study, the CaChPV1 genome displayed the closest relatedness to WbPV1, exhibiting 76.4% identity (Supplementary Table S1). In contrast, CaChPV2 displayed the highest identity to EsPV36 at 75.4%, followed by EcPV150 at 73.9% and CaChPV1 at 57.2%. Similarly, for the newly discovered canary dependoparvoviruses, the CaDePV1 genome showed the closest relationship to SsPV142, with a striking 84.0% identity (Supplementary Table S2). Meanwhile, CaDePV2 demonstrated the most consistent genome sequence alignment with AiPV1 at 59.1%, followed by TnPV173 at 58.5% and AiPV4 at 56.7%.

3.3. Evolutionary relationships

To explore the evolutionary relationships of the detected parvoviral sequences in this study, we utilized the full-length amino acid sequences of the NS1 protein and VP1 protein to construct phylogenetic trees (Cotmore et al., 2019). The Maximum Likelihood (ML) tree of the NS1 gene unequivocally places the two newly discovered sequences such as CaChPV1, and CaChPV2 in the genus *Chaphamaparvovirus*, subfamily *Hamaparvovirinae* (Fig. 2 and Supplementary Fig. S1). Whereas the sequences belonging to the CaDePV1, and CaDePV2 are placed within the genus *Dependoparvovirus*, subfamily *Parvovirinae* (Fig. 3 and Supplementary Fig. S1), and the detailed genetic relationships observed in these trees align with the comparisons made based on the genome sequences. As shown in Fig. 2, in the ML tree, CaChPV1 and CaChPV2 are situated within a subclade of chaphamaparvoviruses, alongside other chaphamaparvoviruses identified in diverse hosts. Specifically, CaChPV1 exhibits the closest relationships with WbPV1 (GenBank accession no. MT138217), which originates from wild birds. CaChPV2 occupies a central position in a subclade and is most related to EsPV36 and EcPV150 with robust bootstrap support (100%). Both EsPV36 and EcPV150 were isolated from *Emberiza Chrysophrys* in 2018, indicating that the canary chaphamaparvoviruses and *Emberiza Chrysophrys* chaphamaparvoviruses were more closely related in terms of viral evolution (Fig. 2).

The ML tree of the NS1 protein reveals that CaDePV1 and CaDePV2 form a distinct subclade alongside other parvoviruses, including adeno-associated virus, goose parvoviruses, duck parvoviruses, *Parvoviridae* sp, and other dependoparvoviruses (Fig. 3 and Supplementary Fig. S1). Within this subclade of dependoparvoviruses, CaDePV1 demonstrates its closest genetic connection with SsPV142 (GenBank accession no. MT138314) from *Spinus spinus*, boasting robust 100% bootstrap support. Conversely, CaDePV2 shows its closest evolutionary affiliation with TnPV173 (GenBank accession. MT138317) from *Turdus naumanni*, followed by AiPV1 (GenBank accession. MT138284) and AiPV4 (GenBank accession no. MT138330) from *Anser indicus*(Fig. 3).

To better correlate the evolutionary linkage of these four newly identified parvoviruses, the phylogenetic analysis using amino acid sequences of the VP1 protein was performed. As shown in Supplementary Figs. S2 and S3, the results similarly reveal an evolutionary genetic connection between these newly discovered parvoviruses (CaChPV1, CaChPV2, CaDePV1, and CaDePV2) and some other parvoviruses. Among them, it was discovered that the parvoviruses identified from wild birds, *Emberiza schoeniclus*, *Spinus spinus*, and *Turdus naumanni* exhibited the closest resemblance to the newly identified CaChPV1, CaChPV2, CaDePV1, and CaDePV2, respectively, within the corresponding fragment of the genome (Supplementary Figs. S2 and S3). Thereafter, using the same set of VP1 protein sequences, the maximum inter-lineage sequence identity values for the novel parvoviruses were observed to be 76.5% (CaChPV1 vs. WbPV1), 72.1% (CaChPV2 vs.

Table 3
Conserved motifs in NS1 protein.

Conserved motif	CaChPV1	CaChPV2	CaDePV1	CaDePV2
ATP- or GTP-binding Walker A loop (GPxNTGKT/S)	**GP**ATT**GK**T	**GP**SN**TGKS**	**GP**ATT**GK**T	**GP**PTT**GK**T
Walker B (xxx**WEE**)	IIW**WEE**	LAI**WEE**	IIW**WEE**	LIW**WEE**
Walker B′ (KQxxEGxxxxxPxK)	**K**AILG**G**SKVRVDQ**K**	**KQ**IF**EG**MTCSIAI**K**	**K**AILG**G**SKVRVDQ**K**	**K**CILG**G**VSCRVDR**K**
Walker C (PxxxTxN)	**P**VII**T**S**N**	**P**VLM**T**T**N**	**P**VII**T**S**N**	**P**VII**T**S**N**
Replication initiator (endonuclease) motifs: xxHuHxxxx	-V**H**L**H**CCIE	VF**H**T**H**ALLR	-V**H**L**H**CCIE	GA**H**V**H**LMVP
Replication initiator (endonuclease) motifs: YxxK	**Y**LLG	**Y**MM**K**	**Y**LM**K**	**Y**LL**K**

Conserved amino acids are indicated in bold letters.

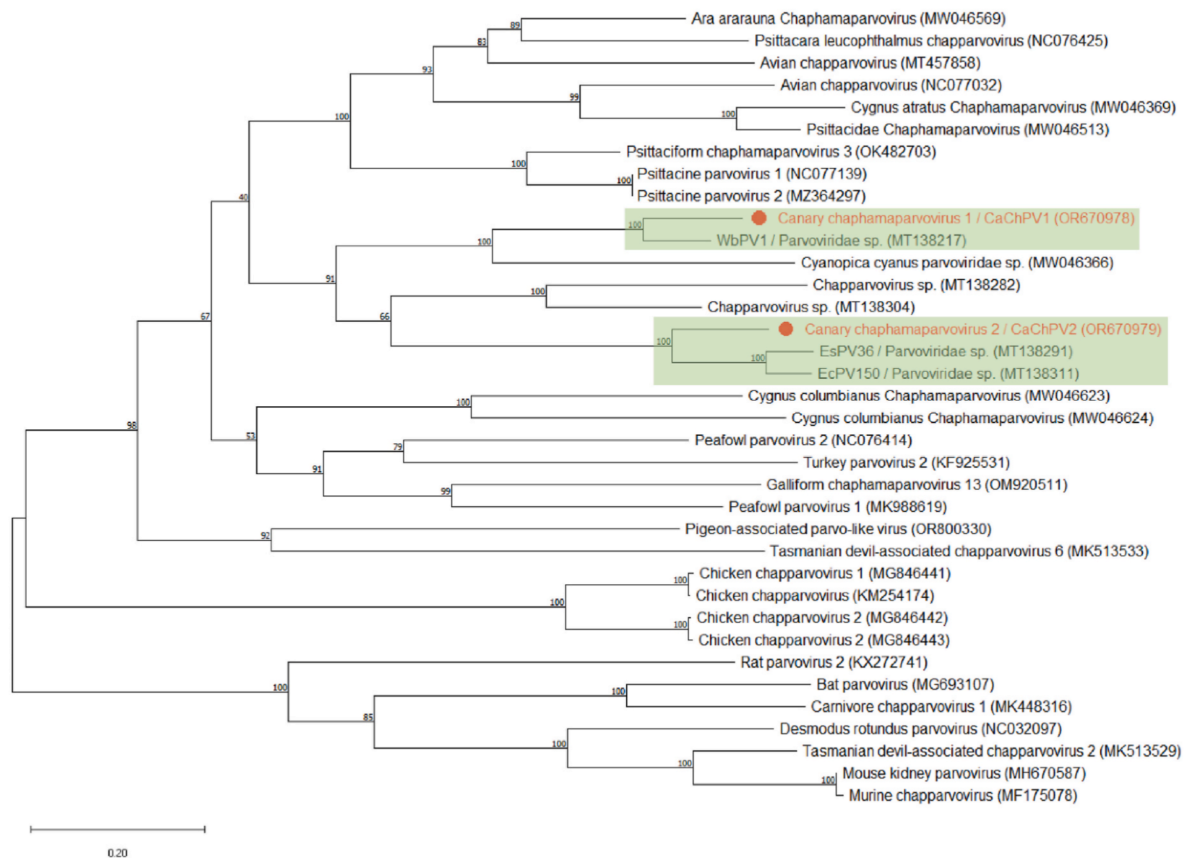


Fig. 2. The maximum likelihood (ML) tree was constructed using the amino acid sequences of complete nonstructural protein (NS1) of chaphamaparvoviruses by MEGA software (Version 11.0.11) with 1000 bootstraps, and the multiple alignments were performed using MAFFT software (Version 7.490). The numbers on the left show bootstrap values as percentages. The labels at the branch tips refer to the parvovirus names, followed by the GenBank accession numbers in parentheses. The position of novel parvoviruses identified in this study are highlighted using red text.

EsPV36), 92.2% (CaDePV1 vs. SsPV142) and 63.7% (CaDePV2 vs. TnPV173), respectively, compared to the most genetically related viruses. These results of sequence identity also mirrored the phylogenetic position of the parvoviruses detected in this study.

4. Discussion

The advancement of metagenomic sequencing technology has opened avenues for genomic-level analysis of numerous viruses, enabling the exploration of the diverse viral landscape within avian species(Sarker, 2021b, 2022; Shan et al., 2022). In this study, we utilized a metagenomic approach to investigate the spectrum of circulating viruses in yellow canaries (*Crithagra flaviventris*) housed in a pet shop in Victoria, Australia. Our finding presents the genomic characterization of four parvoviral sequences (CaChPV1, CaChPV2, CaDePV1, and CaDePV2) in yellow canaries.

The results demonstrated a notable consistency in the genome structures of these newly identified parvoviral sequences, including genome size and number of ORFs, with other characterized parvoviruses available in the GenBank database(Benson et al., 2013a). Nevertheless, CaChPV1, CaChPV2, CaDePV1, and CaDePV2 were also exhibited significant distinctions from other parvoviruses, their whole genome sequences demonstrating the highest sequence identities of 76.4% (CaChPV1 vs. WbPV1), 75.4% (CaChPV2 vs. EsPV36), 84.0% (CaDePV1 vs. SePV142) and 59.1% (CaDePV2 vs. AiPV1), respectively (Supplementary Table S1 and Table S2). In essence, these parvoviral sequences differ sufficiently from known parvoviruses to be considered novel members of the family *Parvoviridae*.

Phylogenetic analyses based on amino acid sequences of the parvovirus NS1 and VP1 proteins demonstrate that the four newly identified parvoviruses are part of the genera *Chaphamaparvovirus* and *Dependoparvovirus* (Shan et al., 2022). Although the classification of parvoviruses primarily relies on the protein sequences of the NS1 gene, furthermore, a well-supported ML phylogenetic tree based on the amino acid sequences of the VP1 gene, further confirmed the close evolutionary relationships among these novel parvoviral sequences (Supplementary Figs. S2 and S3)(Penzes et al., 2020). Interestingly, phylogenetic analyses revealed that the four newly identified parvoviruses bear a close evolutionarily relationship to parvoviruses identified in China from 2016 to 2018 (Supplementary Fig. S1). Yet, these viruses were derived from a wide range of hosts, such as *Emberiza chrysophrys*, *Spinus spinus*, *Turdus naumanni*, *Anser indicus*, and wild birds (Tables 1 and 2). These findings underscore the intricate genetic interplay among parvoviruses in various avian species.

Although the infections of many parvoviruses are frequently associated with gastrointestinal disease in a variety of animals(Kapgate et al., 2018; Koo et al., 2015; Sanchez et al., 2023; Sarker, 2021b), pathogenic consequences of parvoviral sequences identified in this study are unknown. In 2019, metagenomic sequencing of 93 fecal samples from wild and captive red-crowned cranes revealed six novel parvoviruses, with four viruses belonging to the genus *Chaphamapparvovirus* and the remaining two to the genus *Aveparvovirus*(Wang et al., 2019). Likewise, virome sequencing of fecal samples from wild birds in Brazil indicated that parvoviruses accounted for roughly 16% of the total dataset generated by the sequencing(M et al., 2019). In 2022, a collection of 3182 fecal samples from wild and breeding birds in China, subjected to metagenomic sequencing, demonstrated that the parvoviruses accounted for approximately 17% of the total number of detected

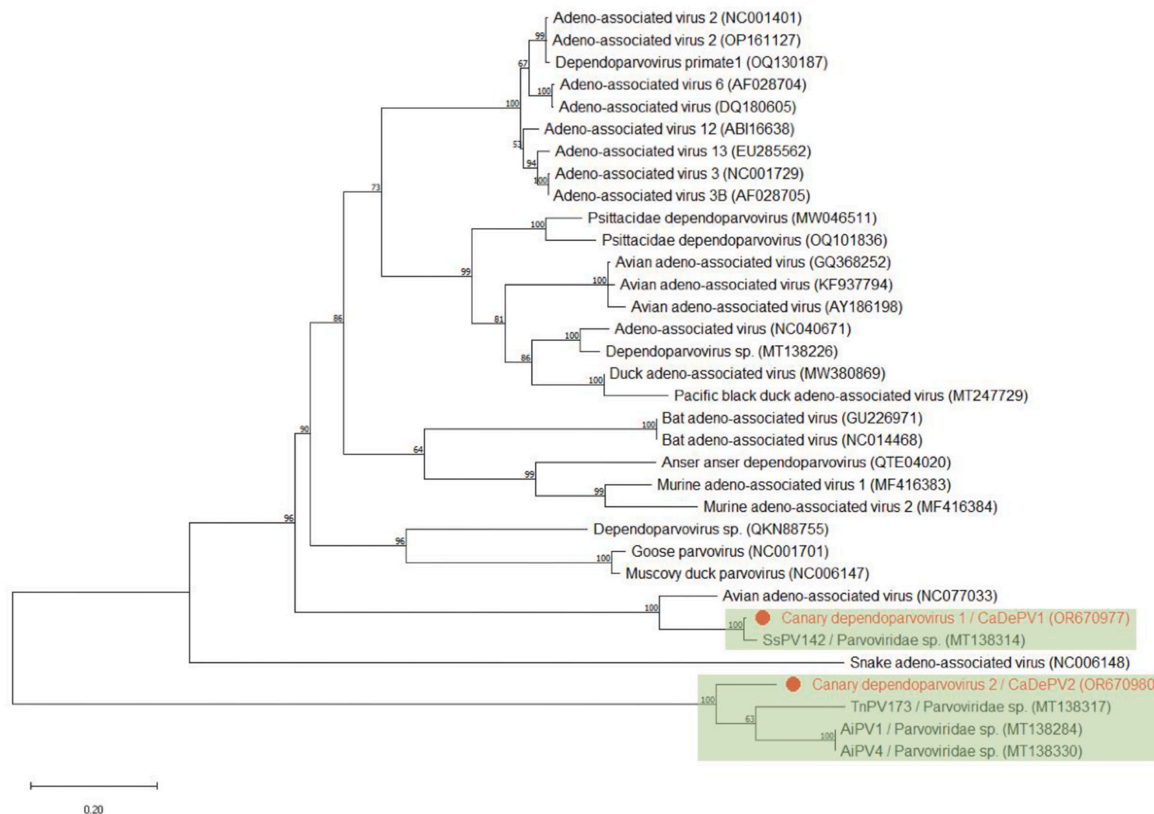


Fig. 3. The maximum likelihood (ML) tree was constructed using the amino acid sequences of complete nonstructural protein (NS1) of dependoparvoviruses by MEGA software (Version 11.0.11) with 1000 bootstraps, and the multiple alignments were performed using MAFFT software (Version 7.490). The numbers on the left show bootstrap values as percentages, and the labels at the branch tips refer to the parvovirus names, followed by the GenBank accession numbers in parentheses. The position of novel parvoviruses identified in this study are highlighted using red text.

viruses (n = 707)(Shan et al., 2022). These findings collectively highlight the presence of parvoviruses in fecal samples from both wild and captive birds, suggesting a diverse range of undetected parvoviruses may be present in these fecal samples. However, the clinical significance of these parvoviruses identified from faecal samples is unclear, as the definitive causal relationship between the viruses and the disease process in the host has not been established. Further virus isolation and animal infection experiments on yellow canaries may provide additional insight into its pathogenicity, which has not been conducted as part of our work yet. Simultaneously, given the limitation of the parvovirus's replication mechanisms in infected cells, studying parvovirus-host interactions remains a major challenge(Kapgate et al., 2018). Thus, further studies on novel avian parvovirus species in both wild and captive birds are essential to determine the evolutionary relationships and pathogenicity of these viruses.

5. Conclusions

This study reports four previously unidentified parvoviral sequences sourced from faecal materials of 10 yellow canaries (*Crithagra flaviventris*) and designated as CaChPV1, CaChPV2, CaDePV1, and CaDePV2. Comparative examination of their whole genome sequences revealed significant divergence from other known parvoviruses in terms of sequence identity and evolutionary linkage. Except for CaDePV1 which has a 92% NS1 amino acid identity, three other novel parvoviral sequences should be considered as new parvoviruses under the family *Parvoviridae*, under the genera *Chaphamaparvovirus* (CaChPV1 and CaChPV2) and *Dependoparvovirus* (CaDePV2). Our findings not only enrich the genomic information within the family *Parvoviridae* but also shed light on the presence of parvovirus infecting yellow canary, providing valuable insights into the genetic connections of parvovirus infections in this species.

6. Institutional review board statement

The faecal sampling was conducted during routine animal husbandry practice without touching animals. The Animal Ethics Committee at La Trobe University was informed that findings from the material were to be used in a publication, and a formal waiver of ethics approval was granted.

7. Informed consent statement

Not applicable.

Funding

This research did not receive any reportable funding.

9. CRediT authorship contribution statement

Yuhao Zhang: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, preparation, Writing – review & editing, Visualization. **Saranika Talukder:** Methodology, Investigation, Writing – review & editing. **Md Safiul Alam Bhuiyan**: Methodology, Investigation, Writing – review & editing. **Lei He**: Conceptualization, Methodology, Formal analysis, Resources, Data curation, Writing – original draft, preparation, Writing – review & editing, Visualization, Supervision. **Subir Sarker**: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing –

original draft, preparation, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Data availability statement

The novel parvoviruses complete genome sequence and the associated data sets that were generated during this studyaccession number OR670977 (CaDePV1), OR670978 (CaChPV1), OR670979 (CaChPV2) and OR670980 (CaDePV2). T were deposited in GenBank under the he raw sequencing data from this study has been deposited in the NCBI Sequence Read Achieve (SRA) under the accession number SRR26413808 (BioProject accession number: PRJNA1028305, link: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1028305).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Dr. Sarker is the recipient of an Australian Research Council Discovery Early Career Researcher Award (grant number DE200100367) funded by the Australian Government. The Australian Government had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript. The authors also like to acknowledge the LIMS-HPC system (a High-Performance Computer specialized for genomics research in La Trobe University) and La Trobe University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virol.2024.110081.

References

Athukorala, A., Phalen, D.N., Das, A., Helbig, K.J., Forwood, J.K., Sarker, S., 2021. Genomic characterisation of a highly divergent Siadenovirus (psittacine siadenovirus F) from the critically endangered orange-bellied parrot (neophema chrysogaster). Viruses 13. https://doi.org/10.3390/v13091714.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. : a journal of computational molecular cell biology 19, 455–477. https://doi.org/10.1089/cmb.2012.0021.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2013a. GenBank. Nucleic Acids Res. 41, D36–D42. https://doi.org/10.1093/nar/gks1195.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2013b. GenBank. Nucleic Acids Res. 41, D36–D42. https://doi.org/10.1093/nar/gks1195.

Cao, M., You, H., Hermonat, P.L., 2014. The X gene of adeno-associated virus 2 (AAV2) is involved in viral DNA replication. PLoS One 9, e104596. https://doi.org/10.1371/journal.pone.0104596.

Cardoso, P., Stoev, P., Georgiev, T., Senderov, V., Penev, L., 2016. Species conservation profiles compliant with the IUCN red list of threatened species. Biodivers. Data J., e10356 https://doi.org/10.3897/BDJ.4.e10356.

Chang, W.S., Eden, J.S., Hall, J., Shi, M., Rose, K., Holmes, E.C., 2020. Metatranscriptomic analysis of virus diversity in urban wild birds with paretic disease. J. Virol. 94 https://doi.org/10.1128/JVI.00606-20.

Chiorini, J.A., Kim, F., Yang, L., Kotin, R.M., 1999. Cloning and characterization of adeno-associated virus type 5. J. Virol. 73, 1309–1319. https://doi.org/10.1128/JVI.73.2.1309-1319.1999.

Cotmore, S.F., Agbandje-McKenna, M., Canuti, M., Chiorini, J.A., Eis-Hubinger, A.M., Hughes, J., Mietzsch, M., Modha, S., Ogliastro, M., Penzes, J.J., Pintel, D.J., Qiu, J., Soderlund-Venermo, M., Tattersall, P., Tijssen, P., Ictv Report, C, 2019. ICTV virus taxonomy profile: Parvoviridae. J. Gen. Virol. 100, 367–368. https://doi.org/10.1099/jgv.0.001212.

Cotmore, S.F., Agbandje-McKenna, M., Chiorini, J.A., Mukha, D.V., Pintel, D.J., Qiu, J., Soderlund-Venermo, M., Tattersall, P., Tijssen, P., Gatherer, D., Davison, A.J., 2014. The family Parvoviridae. Arch. Virol. 159, 1239–1247. https://doi.org/10.1007/s00705-013-1914-1.

Day, J.M., Zsak, L., 2010. Determination and analysis of the full-length chicken parvovirus genome. Virology 399, 59–64. https://doi.org/10.1016/j.virol.2009.12.027.

Farkas, S.L., Zadori, Z., Benko, M., Essbauer, S., Harrach, B., Tijssen, P., 2004. A parvovirus isolated from royal python (Python regius) is a member of the genus Dependovirus. J. Gen. Virol. 85, 555–561. https://doi.org/10.1099/vir.0.19616-0.

Kailasan, S., Agbandje-McKenna, M., Parrish, C.R., 2015. Parvovirus family conundrum: what makes a killer? Annu Rev Virol 2, 425–450. https://doi.org/10.1146/annurev-virology-100114-055150.

Kapgate, S.S., Kumanan, K., Vijayarani, K., Barbuddhe, S.B., 2018. Avian parvovirus: classification, phylogeny, pathogenesis and diagnosis. Avian Pathol. 47, 536–545. https://doi.org/10.1080/03079457.2018.1517938.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.

Kisary, J., Nagy, B., Bitay, Z., 1984. Presence of parvoviruses in the intestine of chickens showing stunting syndrome. Avian Pathol. : journal of the W.V.P.A 13, 339–343. https://doi.org/10.1080/03079458408418536.

Koo, B.S., Lee, H.R., Jeon, E.O., Han, M.S., Min, K.C., Lee, S.B., Bae, Y.J., Cho, S.H., Mo, J.S., Kwon, H.M., Sung, H.W., Kim, J.N., Mo, I.P., 2015. Genetic characterization of three novel chicken parvovirus strains based on analysis of their coding sequences. Avian Pathol. 44, 28–34. https://doi.org/10.1080/03079457.2014.991693.

Liu, X., Wang, H., Liu, X., Li, Y., Chen, J., Zhang, J., Wang, X., Shen, S., Wang, H., Deng, F., Wang, M., Guan, W., Hu, Z., 2020. Genomic and transcriptional analyses of novel parvoviruses identified from dead peafowl. Virology 539, 80–91. https://doi.org/10.1016/j.virol.2019.10.013.

M, A.D., Jm, F.S., C, R.B., D, S.T., F, L.M., B, M.R., Nagata, T., F, S.C., 2019. Faecal virome analysis of wild animals from Brazil. Viruses 11. https://doi.org/10.3390/v11090803.

Palinski, R.M., Mitra, N., Hause, B.M., 2016. Discovery of a novel Parvovirinae virus, porcine parvovirus 7, by metagenomic sequencing of porcine rectal swabs. Virus Gene. 52, 564–567. https://doi.org/10.1007/s11262-016-1322-1.

Penzes, J.J., Soderlund-Venermo, M., Canuti, M., Eis-Hubinger, A.M., Hughes, J., Cotmore, S.F., Harrach, B., 2020. Reorganizing the family Parvoviridae: a revised taxonomy independent of the canonical approach based on host association. Arch. Virol. 165, 2133–2146. https://doi.org/10.1007/s00705-020-04632-4.

Reuter, G., Boros, A., Delwart, E., Pankovics, P., 2014. Novel circular single-stranded DNA virus from Turkey faeces. Arch. Virol. 159, 2161–2164. https://doi.org/10.1007/s00705-014-2025-3.

Sanchez, C., Domenech, A., Gomez-Lucia, E., Mendez, J.L., Ortiz, J.C., Benitez, L., 2023. A novel dependoparvovirus identified in cloacal swabs of monk parakeet (myiopsitta monachus) from urban areas of Spain. Viruses 15. https://doi.org/10.3390/v15040850.

Sarker, S., 2021a. Metagenomic detection and characterisation of multiple viruses in apparently healthy Australian Neophema birds. Sci. Rep. 11, 20915 https://doi.org/10.1038/s41598-021-00440-1.

Sarker, S., 2021b. Molecular and phylogenetic characterisation of a highly divergent novel parvovirus (psittaciform chaphamaparvovirus 2) in Australian neophema parrots. Pathogens 10. https://doi.org/10.3390/pathogens10121559.

Sarker, S., 2022. Characterization of a novel complete-genome sequence of a galliform chaphamaparvovirus from a free-range laying chicken clinically diagnosed with spotty liver disease. Microbiol Resour Announc 11, e0101722. https://doi.org/10.1128/mra.01017-22.

Sarker, S., Das, S., Lavers, J.L., Hutton, I., Helbig, K., Imbery, J., Upton, C., Raidal, S.R., 2017. Genomic characterization of two novel pathogenic avipoxviruses isolated from pacific shearwaters (*Ardenna* spp.). BMC Genom. 18, 298. https://doi.org/10.1186/s12864-017-3680-z.

Sarker, S., Isberg, R.S., Moran, L.J., Araujo, D.R., Elliott, N., Melville, L., Beddoe, T., Helbig, J.K., 2019. Crocodilepox virus evolutionary genomics supports observed poxvirus infection dynamics on saltwater crocodile (Crocodylus porosus). Viruses 11. https://doi.org/10.3390/v11121116.

Shan, T., Yang, S., Wang, H., Wang, H., Zhang, J., Gong, G., Xiao, Y., Yang, J., Wang, X., Lu, J., Zhao, M., Yang, Z., Lu, X., Dai, Z., He, Y., Chen, X., Zhou, R., Yao, Y., Kong, N., Zeng, J., Ullah, K., Wang, X., Shen, Q., Deng, X., Zhang, J., Delwart, E., Tong, G., Zhang, W., 2022. Virome in the cloaca of wild and breeding birds revealed a diversity of significant viruses. Microbiome 10, 60. https://doi.org/10.1186/s40168-022-01246-7.

Souza, W.M., Romeiro, M.F., Fumagalli, M.J., Modha, S., de Araujo, J., Queiroz, L.H., Durigon, E.L., Figueiredo, L.T.M., Murcia, P.R., Gifford, R.J., 2017. Chapparvoviruses occur in at least three vertebrate classes and have a broad biogeographic distribution. J. Gen. Virol. 98, 225–229. https://doi.org/10.1099/jgv.0.000671.

Sutherland, M., Sarker, S., 2023. Liver virome of a Little Corella (Cacatua sanguinea) reveals coinfection with a novel parvovirus and two beak and feather disease viruses. Aust. Vet. J. 101, 366–372. https://doi.org/10.1111/avj.13271.

Sutherland, M., Sarker, S., Vaz, P.K., Legione, A.R., Devlin, J.M., Macwhirter, P.L., Whiteley, P.L., Raidal, S.R., 2019. Disease surveillance in wild Victorian cacatuids reveals co-infection with multiple agents and detection of novel avian viruses. Vet. Microbiol. 235, 257–264. https://doi.org/10.1016/j.vetmic.2019.07.012.

Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: molecular evolutionary genetics analysis version 11. Mol. Biol. Evol. 38, 3022–3027. https://doi.org/10.1093/molbev/msab120.

Vibin, J., Chamings, A., Collier, F., Klaassen, M., Nelson, T.M., Alexandersen, S., 2018. Metagenomics detection and characterisation of viruses in faecal samples from Australian wild birds. Sci. Rep. 8, 8686.

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. https://doi.org/10.1007/s00705-018-4037-x.

Williams, S.H., Che, X., Garcia, J.A., Klena, J.D., Lee, B., Muller, D., Ulrich, W., Corrigan, R.M., Nichol, S., Jain, K., Lipkin, W.I., 2018. Viral diversity of house mice in New York city. mBio 9. https://doi.org/10.1128/mBio.01354-17.

Yang, S., Liu, Z., Wang, Y., Li, W., Fu, X., Lin, Y., Shen, Q., Wang, X., Wang, H., Zhang, W., 2016. A novel rodent Chapparvovirus in feces of wild rats. Virol. J. 13, 133. https://doi.org/10.1186/s12985-016-0589-0.

Yinda, C.K., Ghogomu, S.M., Conceicao-Neto, N., Beller, L., Deboutte, W., Vanhulle, E., Maes, P., Van Ranst, M., Matthijnssens, J., 2018. Cameroonian fruit bats harbor divergent viruses, including rotavirus H, bastroviruses, and picobirnaviruses using an alternative genetic code. Virus Evol 4, vey008. https://doi.org/10.1093/ve/vey008.

Zadori, Z., Stefancsik, R., Rauch, T., Kisary, J., 1995. Analysis of the complete nucleotide sequences of goose and muscovy duck parvoviruses indicates common ancestral origin with adeno-associated virus 2. Virology 212, 562–573. https://doi.org/10.1006/viro.1995.1514.