Infection, Genetics and Evolution 93 (2021) 104914

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

Identification of novel circovirus and anelloviruses from wolverines using a non-invasive faecal sampling approach

Rohan Antonio Bandoo [a], Joshua Bautista [a], Michael Lund [a], Eric Newkirk [b], John Squires [c], Arvind Varsani [a,d,*], Simona Kraberger [a,*]

[a] The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life sciences, Arizona State University, Tempe, AZ 85287-5001, USA
[b] Speedgoat Wildlife Solutions, Missoula, MT 59801, USA
[c] U.S. Forest Service, Rocky Mountain Research Station, 800 East Beckwith Avenue, Missoula, MT 59801, USA
[d] Structural Biology Research Unit, Department of Integrative Biomedical Sciences, University of Cape Town, Rondebosch, 7700 Cape Town, South Africa

ARTICLE INFO

Keywords:
Anelloviridae
Circoviridae
wolverine
Gulo gulo
Mustelidae

ABSTRACT

Viruses in the families *Circoviridae* and *Anelloviridae* have circular single-stranded DNA genomes and have been identified in various animal species. Some members of the *Circoviridae* family such as beak and feather disease and porcine circovirus have been found to cause disease in their host animals. Anelloviruses on the other hand have not been identified to cause disease in their hosts but are highly prevalent in mammalian species. Using a non-invasive sampling approach, we identified novel circovirus and anelloviruses from faecal samples of wolverines dwelling in Montana, USA. Wolverines are forest carnivores that feed on a wide variety of carrion and other prey species, and they occupy diverse habitats across northern Europe to North America. Little is known about viruses associated with wild wolverines. Our investigation of the faecal samples resulted in the identification of a novel circovirus from three out of four wolverine samples, two collected in 2018 and one in 2019. Comparison with other circoviruses shows it is most closely related to a porcine circovirus 3, sharing ~69% identity. Additionally, three anellovirus genomes were recovered from two wolverine faecal samples which share 68-–69% ORF1 nucleotide similarity with an anellovirus from another mustelid species, pine martens. Here we identify novel single-stranded DNA viruses associated with wolverine and open up new avenues for research.

1. Introduction

Wolverines (*Gulo gulo*) are carnivorous animals whose habitat spans the arctic, boreal, and alpine wilderness territories ranging from North America to northern Europe (Aubry et al., 2007; Krebs et al., 2007; Scrafford et al., 2017). Wolverines are the largest land-dwelling species of the family Mustelidae. They have powerful jaws and sharp claws, which they use to scavenge, feeding primarily on the decaying remains of dead animal carcasses during winter. The diet of wolverines consists of a wide range of rodents and ungulates (van Dijk et al., 2008). Their diverse diet and circumpolar distribution make wolverines an interesting candidate for observing the potential pathogens they may encounter in the wild. Studies describing infectious pathogens that are prevalent among wolverine populations are limited. A serological study described the prevalence of wolverine antibodies to canine distemper virus (CDV), canine parvovirus (CPV), and canine adenovirus (CAV) (Dalerum et al., 2005). This study indicated that of the 64 Alaskan wolverine samples collected between 1998 and 2002, four samples had antibodies to CDV (7%), one contained antibodies to CPV (2%), and no samples contained antibodies to CAV (Dalerum et al., 2005). The scarcity of literature describing infectious viral agents in wolverine populations highlights the need to better understand their virome.

In this study, through the analysis of faecal samples collected from wolverine resident living in the Flathead Valley region in the Northern Rocky Mountains of Montana (USA), we identified novel viruses in the families *Circoviridae* and *Anelloviridae*.

* Corresponding authors at: The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life sciences, Arizona State University, Tempe, AZ 85287-5001, USA.
E-mail addresses: Arvind.varsani@asu.edu (A. Varsani), simona.kraberger@asu.edu (S. Kraberger).

https://doi.org/10.1016/j.meegid.2021.104914
Received 7 April 2021; Received in revised form 21 April 2021; Accepted 11 May 2021
Available online 13 May 2021

2. Materials and methods

2.1. Sample collection, viral DNA extraction and high-throughput sequencing

We opportunistically collected four faecal samples from wolverines in the Flathead Valley of the Northern Rocky Mountains, Montana, USA during the winter of 2018 ($n = 3$) and 2019 ($n = 1$). Faecal samples were stored at −20 °C. 5 ml of SM buffer was added to 5 g of the faecal sample, which was first homogenised and then centrifuged at 10,000 rpm for 10 mins. The supernatant was filtered through a 0.22 μm syringe filter and viral-like particles precipitated with 10% PEG *w/v* and centrifuged at 10,000 rpm for 10 min to form a pellet. The pellet was resuspended in 500 μl of SM buffer, and 200 μl of this was used to extract viral DNA using the High Pure viral nucleic acid kit (Roche Diagnostics, USA). 1 μl of the extracted DNA was then used in a rolling circle amplification (RCA) reaction using TempliPhi 2000 kit (GE Healthcare, USA). RCA products were sequenced on an Illumina NovaSeq 6000 sequencer at Psomagen (USA). The resulting read data were trimmed using Trimmomatic (Bolger et al., 2014) and the trimmed reads were de novo assembled using SPAdes v 3.12.0 (Bankevich et al., 2012). Contigs were then filtered for those with similarities to circoviruses and anelloviruses using BlastX (Altschul et al., 1990) and specific abutting primers were designed based on these contigs.

2.2. Recovery and sequence assembly of circovirus and anellovirus genomes

For recovery of viral genomes from individual samples, polymerase chain reaction (PCR) was performed using the specific abutting primers. For the circoviruses the following primers were used: Wolv_circo1F: 5′TAG CCA TCC TTC TGT TGA CCA GCT GC′3 Wolv_circo1R: 5′-CAG GAT GAC CCA TAT GCT AAC TCC TCC A′-3, and for the anelloviruses the following primers were used; Wolv_anello1F: 5′-CAC CAG ATG TAA ACT ACT TTT CCT TCC-′3 Wolv_anello1R: 5′-GGA TGT GGG TAT AGT ATG AGT TTA GTT CC-′3 and Wolv_anello2F: 5′-GTG GAA AGA TCA AAA TAA GTA CAC CGA TCC-′3 Wolv_anello2R: 5′-CCT CTA TAA CTA TAT AGT CTG GTG GGA AAG-′3. 10 μl PCR reactions were performed using Kapa HiFi DNA polymerase (Kapa Biosystems, USA) water, 1 μl of the specific primer pair, and 1 μl of the RCA template DNA.

The four wolverine samples were screened using each primer pair. PCR amplicons were resolved in a 0.7% agarose gel and amplicons of the correct size were excised from the gel, purified, and cloned into pJET 1.2 vector (ThermoFisher Scientific, USA), and transformed into XL blue *E. coli* cells. The *E. coli* transformants were PCR screened to confirm plasmid with correct insert was present. Colonies harbouring the correct recombinant plasmids were grown overnight in a shaking incubator at 37 °C and the recombinant plasmids were purified and Sanger sequenced at Macrogen Inc. (Korea). The Sanger sequence reads of both circoviruses and anelloviruses were assembled using Geneious Prime 2021.0.3 (Biomatters Ltd. New Zealand).

2.3. Bioinformatic analyses

2.3.1. Circoviruses

Representative genomes of the circoviruses, one genome representing a species, were downloaded from GenBank and linearized at the nonanucleotide sequences. From these replication-associated protein (*rep*) and capsid protein (*cp*) gene sequences were extracted and used to build protein sequence datasets of the Rep and CP.

Datasets of the full genome, Rep and CP amino acid sequences, together with the circoviruses identified in this study were aligned using MAFFT (Katoh and Standley, 2013). Maximum likelihood phylogenetic trees were constructed from the aligned full genome, Rep and CP datasets using GTR + G + I, LG + G + I, and LG + G + F substitution models (determined in MEGA V7 (Kumar et al., 2016)), respectively with PhyML (Guindon et al., 2010). TreeGraph2 was used to collapse branches with aLRT values below 0.8 support (Stover and Muller, 2010). The phylogenetic trees were rooted (cyclovirus sequences as outgroup) and visualised in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

2.3.2. Anelloviruses

Complete genomes of anelloviruses available in GenBank were downloaded and ORF1 gene sequences were extracted to build an ORF1 dataset. The ORF1 dataset together with the ORF1 of the anelloviruses identified in this study were translated and aligned using MAFFT (Katoh and Standley, 2013). The resulting alignment was used to infer an approximately-maximum-likelihood phylogenetic tree using FastTree (Price et al., 2010) with JTT + CAT model substitution model. Branches with less than 0.6 posterior support were collapsed in Treegraph2 (Stover and Muller, 2010).

2.3.3. Pairwise identities

All pairwise identities (nucleotide and amino acid) were determined using SDT v1.2 (Muhire et al., 2014).

3. Results and discussion

3.1. Identification of novel anelloviruses and circoviruses from wolverine samples

We used a metagenomic approach to identify novel anelloviruses and a circovirus from wolverine faecal samples opportunistically sampled during 2018 and 2019 in the Northern Rocky Mountains of Montana. Three faecal samples from wolverines (ID # 37,38,39) were sampled in the winter of 2018 and one (ID# 40) was sampled in the winter of 2019. Among de novo assembled contigs generated from the high-throughput sequencing data, we identified sequences with similarities to a circovirus and two with similarities to anelloviruses. Based on these contigs we designed abutting primers to recover and screen the four samples for these viruses.

Circoviruses are small single-stranded DNA viruses that have genomes of about 2 kb encoding two known proteins, replication-associated protein (Rep) and a capsid protein (CP) (Breitbart et al., 2017) have been found in various animal species (Rosario et al., 2017). Some circoviruses are known to cause disease in their host, for example beak and feather disease virus (BFDV) which causes malformations of the tissues/matrices that allow proper beak and feather development (Bert et al., 2005).

Anelloviruses are another group of ssDNA viruses commonly found in wildlife. (Bigarre et al., 2005; de Souza et al., 2018; Kaczorowska and van der Hoek, 2020; Ng et al., 2009; Nishiyama et al., 2014). Their genome is also between ~1.7–4 kb in size (Biagini, 2009; Biagini et al., 2011). Anelloviruses are believed to be present in up to 90% of people (de Souza et al., 2018). While these viruses appear rather ubiquitous, no known disease has been directly implicated in an anellovirus infection (Kaczorowska and van der Hoek, 2020). Several studies have documented the prevalence and diversity of anelloviruses in animal populations (Bigarre et al., 2005; Ng et al., 2009; Nishiyama et al., 2014). It is not yet known if the relationship between anelloviruses and animals is pathogenic or more commensal in nature (Kaczorowska and van der Hoek, 2020). Anelloviruses have a large gene (ORF1) whose translation product is involved in the replication and encapsidation of the virus (Biagini et al., 2011).

Three circoviruses were identified in three of the four wolverine faecal samples (ID# 38, 39 sampled in 2018; ID# 40 sampled in 2019). The three anelloviruses were identified in two of the four wolverine faecal samples (ID# 38,39 both sampled in 2018), with two variants present in the same wolverine faecal sample (ID# 38).

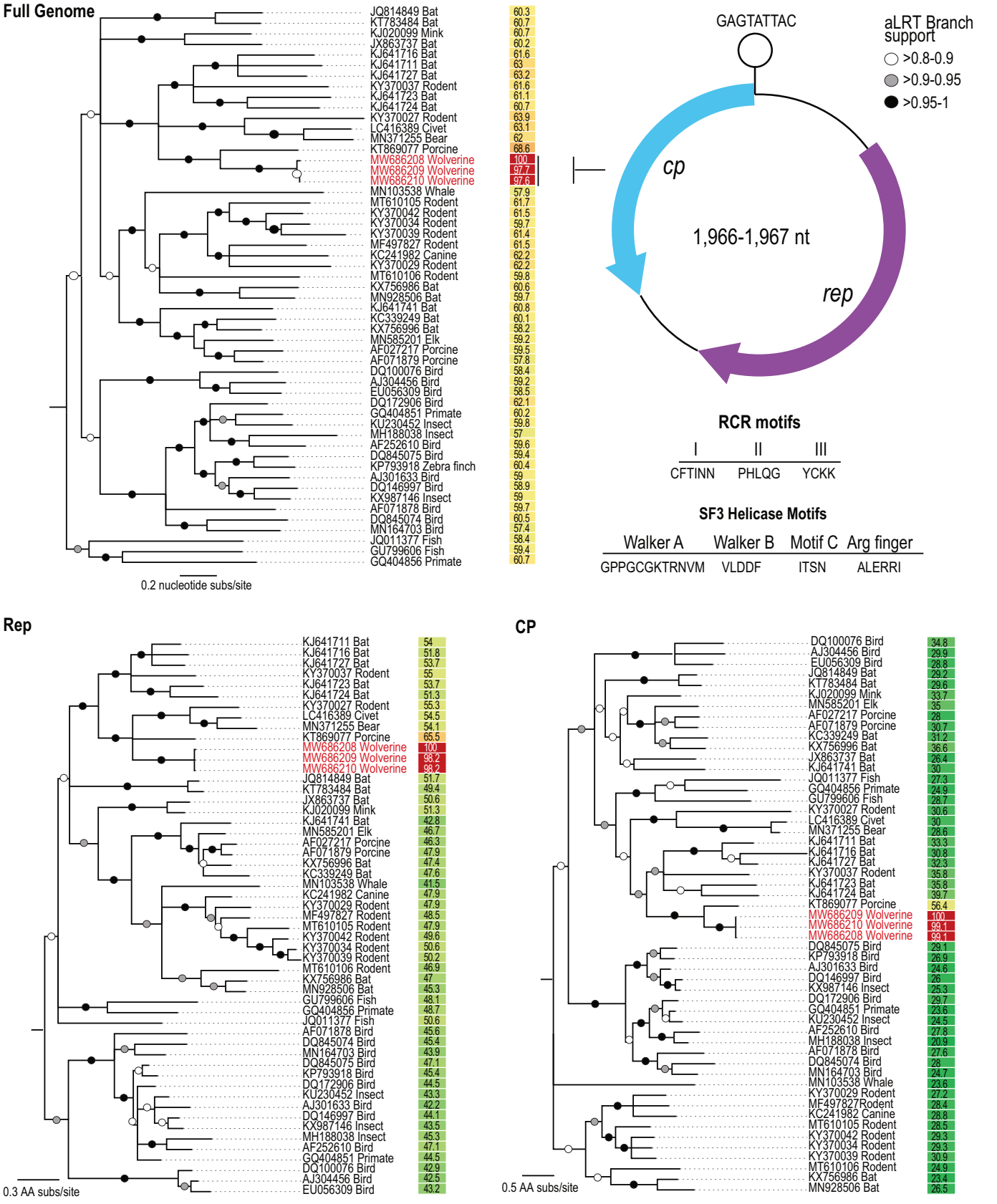

Fig. 1. Maximum likelihood phylogenetic trees of the; full genome nucleotide sequences, Rep amino acid, and CP amino acid sequences showing the phylogenetic relationship between the wolvfaec circovirus sequences and their closest relatives. aLRT values of branch support are shown according to key. Pairwise identity scores are displayed to the right of the phylogenetic tree. The wolvfaec circovirus sequence accession numbers are shown in red. The genome schematic of the wolverine isolates is displayed next to the full genome tree. The genome organization of the circovirus is shown. The Rep amino acid sequences for rolling circle replication and helicase motifs summarized.

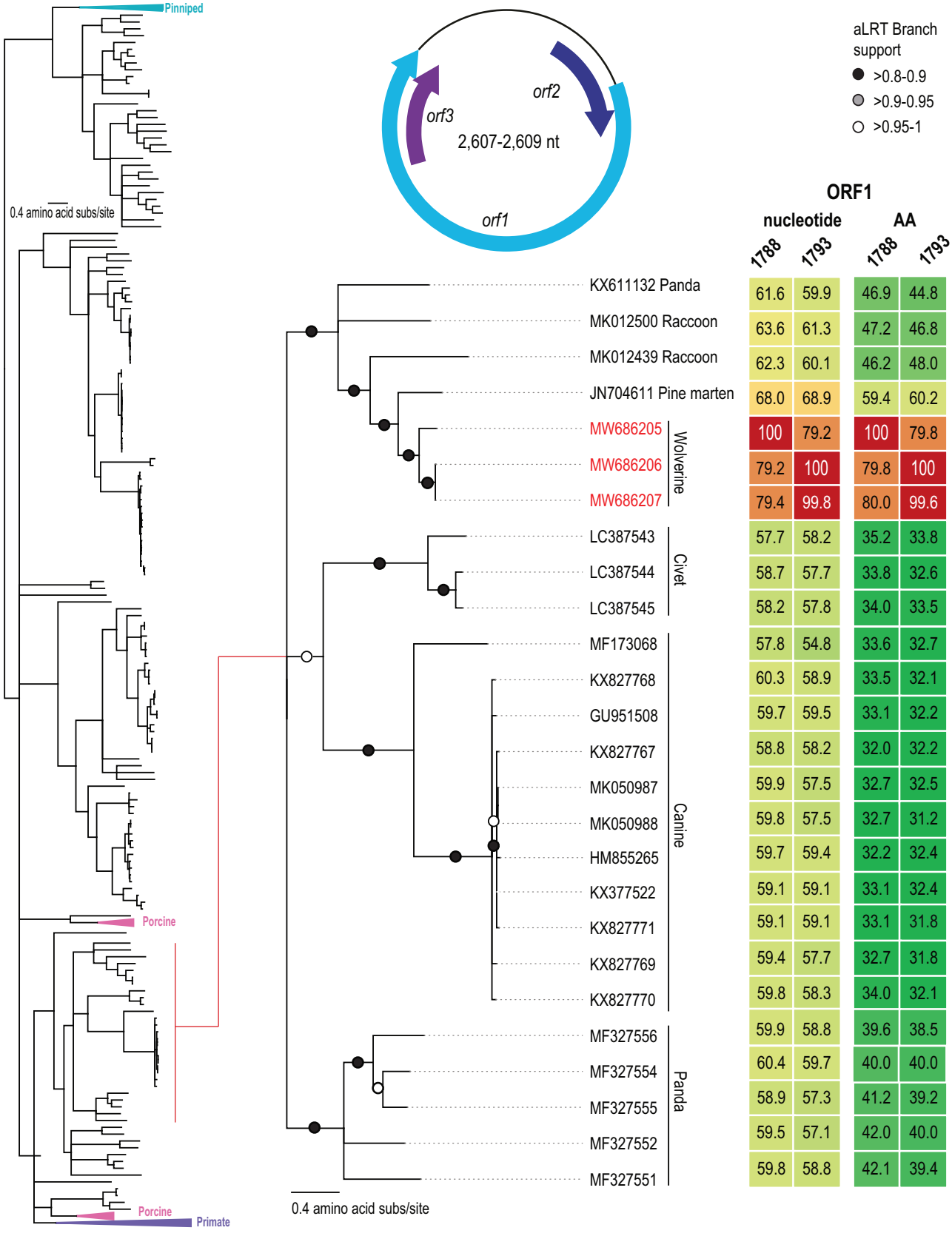

Fig. 2. Phylogenetic tree of ORF1 amino acid sequences of anellovirus sequence together with the three from wolverines. Larger clades have been collapsed, coloured and labelled. The clade containing the wolverine anelloviruses and their relatives has been enlarged. Animals from which each anellovirus sequence has been identified are labelled. Branch support is indicated by key. The genome organization of the anelloviruses from the wolverine is shown.

3.2. Circoviruses

The three circoviruses that we identified are 1967 nts in length and were present the typical circovirus genome organization of a CP and a Rep. Pairwise comparison of these three genomes shared ~97.7% pairwise sequence identity and thus we have named all these three as wolvfaec circovirus. Their Reps and CPs share 98.2% and 99.1% pairwise amino acid identity, respectively (Fig. 1). When compared to other circoviruses, the wolvfaec circoviruses share 57–68.6% genome-wide pairwise identity, whereas their Reps share 41.5–65.5%, and CPs from 23.4–56.4% identities. Based on the current classification of circoviruses with an 80% (Rosario et al., 2017) species demarcation threshold for genome-wide comparisons, the wolvfaec circoviruses represent a new species of circovirus. The wolvfaec circoviruses are phylogenetically most closely related to porcine circovirus 3 (GenBank accession KT869077) (Fan et al., 2017). Within the Rep of the wolvfaec circoviruses we identified RCR and SF3 helicase motifs (Fig. 1) which are similar to those of other circoviruses (Rosario et al., 2017). We believe the three circoviruses that we recovered from three wolverine faecal samples in 2018 and 2019, were either infecting wolverines themselves or from a prey animal.

3.3. Anelloviruses

Anelloviruses have been identified in various mammalian species. The Anelloviruses that we recovered from wolverines are 2,607–2,609 nts in length and have three predicted open reading frames; ORF1, ORF2 and ORF3 (Fig. 2). Two of the anelloviruses share 99.8% ORF1 nucleotide identity, and 99.6% amino acid identity, and were identified in sample ID# 38 and 39, both were sampled in 2018. The third anellovirus from sample ID# 38 shares ~79% ORF1 nucleotide and ~ 80% amino acid pairwise identity with the other two. In comparison to the other anellovirus ORF1 nucleotide sequences, the ORF1s of the three anelloviruses identified in this study share 57.1–68.9% nucleotide and 31.2–60.2% amino acid identity (Fig. 2). Therefore, the anelloviruses we identified based on the species demarcation of anelloviruses (Biagini, 2009; Biagini et al., 2011), represents a new anellovirus species, and thus we have named these torque teno mustelid virus 2. Interestingly, they are most closely related to the anellovirus referred to as pine marten torque teno virus 1 (Genbank accession JN704611) (van den Brand et al., 2012).

4. Concluding remarks

Literature pertaining to wolverine infection and pathology is rather limited such as their elusive nature due to the inherent obstacles in investigating infectious agents in wolverine populations. Wolverine habitats are typically remote and widely distributed across a variety of environments across Europe and North America (Scrafford et al., 2017). Utilising a non-invasive faecal sampling approach, we identified a novel circovirus in three wolverine faecal samples, of which two were collected in 2018 and one in 2019, demonstrating that this virus infects wolverines or their prey. Future research is needed to determine if wolverines are the true host and whether this virus is pathogenic. Additionally, we identified three anelloviruses (torque teno mustelid virus 2) belonging to two groups in two wolverine faecal samples from the same year, however, one has two variants that are ~20% dissimilar. These wolverine anelloviruses were most closely related to another mustelid supporting wolverines as the likely host. These findings create new opportunities to understand the relationship between hosts, circoviruses, and anelloviruses. Continued research is needed to explore the wolverine virome.

CRediT authorship contribution statement

Rohan Antonio Bandoo: Investigation, Formal analysis, Conceptualization, Writing - original draft. **Joshua Bautista:** Investigation. **Michael Lund:** Investigation. **Eric Newkirk:** Investigation, Resources. **John Squires:** Investigation, Resources. **Arvind Varsani:** Supervision, Writing - review & editing, Resources. **Simona Kraberger:** Supervision, Writing - review & editing.

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

Molecular work was supported by funds from the Biodesign Institute and School of Life Sciences at Arizona State University awarded to AV. Samples were collected by a lynx research team from the US Forest Service Rocky Mountain Research Station.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Aubry, K.B., McKelvey, K.S., Copeland, J.P., 2007. Distribution and Broadscale habitat relations of the wolverine in the contiguous United States. J. Wildl. Manag. 71, 2147–2158.
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. 19, 455–477.
Bert, E., Tomassone, L., Peccati, C., Navarrete, M.G., Sola, S.C., 2005. Detection of beak and feather disease virus (BFDV) and avian polyomavirus (APV) DNA in psittacine birds in Italy. J. Vet. Med. B Infect. Dis Vet. Public Health 52, 64–68.
Biagini, P., 2009. Classification of TTV and related viruses (anelloviruses). Curr. Top. Microbiol. Immunol. 331, 21–33.
Biagini, P., Bendinelli, M., Hino, S., Kakkola, L., Mankertz, A., Niel, C., Okamoto, H., Raidal, S., Teo, C.G., Todd, D., 2011. Family Anelloviridae, Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Scientific Publ. Co, pp. 331–341.
Bigarre, L., Beven, V., de Boisseson, C., Grasland, B., Rose, N., Biagini, P., Jestin, A., 2005. Pig anelloviruses are highly prevalent in swine herds in France. J. Gen. Virol. 86, 631–635.
van den Brand, J.M., van Leeuwen, M., Schapendonk, C.M., Simon, J.H., Haagmans, B.L., Osterhaus, A.D., Smits, S.L., 2012. Metagenomic analysis of the viral flora of pine marten and European badger feces. J. Virol. 86, 2360–2365.
Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. https://doi.org/10.1093/bioinformatics/btu170.
Breitbart, M., Delwart, E., Rosario, K., Segales, J., Varsani, A., Ictv Report, C., 2017. ICTV virus taxonomy profile: Circoviridae. J. Gen. Virol. 98, 1997–1998.
Dalerum, F., Shults, B., Kunkel, K., 2005. A serologic survey for antibodies to three canine viruses in wolverines (*Gulo gulo*) from the Brooks Range, Alaska. J. Wildl. Dis. 41, 792–795.
van Dijk, J., Gustavsen, L., Mysterud, A., May, R., Flagstad, O., Broseth, H., Andersen, R., Andersen, R., Steen, H., Landa, A., 2008. Diet shift of a facultative scavenger, the wolverine, following recolonization of wolves. J. Anim. Ecol. 77, 1183–1190.
Fan, S., Ku, X., Chen, F., Wang, Y., Yu, X., He, Q., 2017. Complete genome sequence of a novel porcine Circovirus Type 3 strain, PCV3/CN/Hubei-618/2016, isolated from China. Genome Announc. 5.
Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.
Kaczorowska, J., van der Hoek, L., 2020. Human anelloviruses: diverse, omnipresent and commensal members of the virome. FEMS Microbiol. Rev. 44, 305–313.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Krebs, J., Lofroth, E.C., Parfitt, I.A.N., 2007. Multiscale habitat use by wolverines in British Columbia, Canada. J. Wildl. Manag. 71, 2180–2192.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.
Muhire, B.M., Varsani, A., Martin, D.P., 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9, e108277.
Ng, T.F.F., Suedmeyer, W.K., Wheeler, E., Gulland, F., Breitbart, M., 2009. Novel anellovirus discovered from a mortality event of captive California Sea lions. J. Gen. Virol. 90, 1256–1261.
Nishiyama, S., Dutia, B.M., Stewart, J.P., Meredith, A.L., Shaw, D.J., Simmonds, P., Sharp, C.P., 2014. Identification of novel anelloviruses with broad diversity in UK rodents. J. Gen. Virol. 95, 1544–1553.
Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS One 5, e9490.

Rosario, K., Breitbart, M., Harrach, B., Segales, J., Delwart, E., Biagini, P., Varsani, A., 2017. Revisiting the taxonomy of the family Circoviridae: establishment of the genus Cyclovirus and removal of the genus Gyrovirus. Arch. Virol. 162, 1447–1463.

Scrafford, M.A., Avgar, T., Abercrombie, B., Tigner, J., Boyce, M.S., 2017. Wolverine habitat selection in response to anthropogenic disturbance in the western Canadian boreal forest. For. Ecol. Manag. 395, 27–36.

de Souza, W.M., Fumagalli, M.J., de Araujo, J., Sabino-Santos Jr., G., Maia, F.G.M., Romeiro, M.F., Modha, S., Nardi, M.S., Queiroz, L.H., Durigon, E.L., Nunes, M.R.T., Murcia, P.R., Figueiredo, L.T.M., 2018. Discovery of novel anelloviruses in small mammals expands the host range and diversity of the Anelloviridae. Virology 514, 9–17.

Stover, B.C., Muller, K.F., 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinformatics 11, 7.