A novel calicivirus discovered in trumpeter swans (Cygnus buccinator) expands the richness of known avian caliciviruses

Marta Canuti\(^a,1,\)*, Laurie Wilson\(^b\), Victoria Bowes\(^c\), Tony Redford\(^c\), Suzanne C. Dufour\(^a\), Andrew S. Lang\(^a\), Joost T.P. Verhoeven\(^a\)

\(^a\) Department of Biology, Memorial University of Newfoundland, 45 Arctic Ave., St. John’s, NL A1C 5S7, Canada
\(^b\) Environment and Climate Change Canada, Pacific Wildlife Research Centre, 5421 Robertson Rd, RR#1 Delta, BC V4K 3N2, Canada
\(^c\) Animal Health Centre, BC Ministry of Agriculture and Food, 1767 Angus Campbell Road, Abbotsford, BC V3G 2M3, Canada

**Abstract**

Caliciviruses are ssRNA viruses that can infect a wide range of hosts, including birds. While several avian caliciviruses have been discovered, their taxonomy and host distribution are largely unknown. We molecularly characterized a novel calicivirus (trumpeter swan calicivirus: TruSCV) in trumpeter swans over-wintering in south-west British Columbia, Canada. The positivity rate was 20.3% (14/69) and there were no significant differences in infection rates between males (5/34, 14.7%) and females (9/35, 25.7%) or among considered age groups (juveniles: 4/14, 28.6%; sub-adults: 1/9, 11.1%; adults: 9/46, 19.6%). Twelve infected swans died of lead poisoning, one because of starvation, and one from physical injuries. TruSCV complete genome possessed the typical organization and protein motifs of caliciviruses and a type 2 IRES and its closest relative was a virus circulating in Australian ducks. Phylogenetic analyses showed the existence of 34 different but monophyletic avian caliciviruses. These viruses, while having conserved genomic organization and protein motifs, possess different IRES types and group in several divergent clades, with only two of them corresponding to currently defined genera, highlighting the need for epidemiological investigations and systematic analyses to better define their taxonomy. Follow-up studies are needed to elucidate the diversity, distribution, and pathogenic potential of TruSCV.

1. Introduction

Caliciviruses (family Caliciviridae, order Picornavirales) are positive-sense ssRNA viruses measuring approximately 27–40 nm. The non-enveloped caliciviral particles are composed of the major capsid protein VP1 and the minor structural protein VP2 and contain one single-stranded (+) RNA genome (~7-8 kb) including 2–3 open reading frames (ORFs). ORF1 (5’-proximal) encodes a large polyprotein that is post-translationally cleaved by the virus-encoded protease (NS6pro) into six nonstructural proteins (NS1/2-7) and, for most genera, VP1, which is encoded by a separate ORF in a minority of caliciviruses. Finally, all viruses possess an additional smaller ORF coding for VP2 on the 3’ side (Vinjé et al., 2019).

Currently, the family Caliciviridae includes 11 defined genera, two of which encompass exclusively avian viruses (Desselberger, 2019; Vinjé et al., 2019). The two genera including avian viruses, Bavovirus and Nacovirus, were proposed in 2012 by Wolf et al. (2011, 2012) after the identification of novel caliciviruses in chickens and turkeys in Northern Europe. Since then, bavoviruses and nacoviruses have been detected in chickens in America, Asia, and Europe (Kim et al., 2020; Lima et al., 2019, 2017; Wolf et al., 2012). Additionally, several recent virus discovery and metagenomic investigations have identified genomes of novel caliciviruses in various avian hosts with some studies demonstrating high viral diversity and even the likely existence of several novel genera (Canuti et al., 2019; Phan et al., 2013; Shan et al., 2022; Wang et al., 2017; Wille et al., 2019, 2018). However, the taxonomy and host distribution of these viruses remain largely under investigated and undetermined.
During an exploratory virus discovery study performed with the VidION method (Canuti et al., 2021), we discovered, in a trumpeter swan (Cygnus buccinator) sample, a genetic fragment showing homology to avian caliciviruses. The scope of this study was to explore the epidemiology of the discovered virus, to obtain and characterize its complete coding sequence, and to study its phylogenetic relationships with other members of the Caliciviridae. Additionally, to investigate in greater detail the evolution of avian caliciviruses, we explored all publicly available caliciviral sequences identified in birds and assessed the overall caliciviral diversity in avian hosts.

2. Material and methods

2.1. Samples

This study included 69 trumpeter swans that were found dead or sick (and subsequently died while in care or were euthanized because of poor prognosis) in south-west British Columbia between December 2019 and February 2020. The swans were collected from the Canadian side of a mortality event that has been occurring annually since 2001 when swans migrate south from Alaska for their non-breeding season to Whatcom County in Washington State and the Sumas Prairie in British Columbia and access historically spent lead shot present in the local environment (Smith et al., 2009). The swans died mostly because of lead poisoning (58/69 - 56 ascertainment (lethal liver lead levels, -6.05 ppm wet wt) and 2 suspected (elevated liver lead levels, between 2.42 and 6.05 ppm wet wt) – 84.1%), and other causes of death included aspergillosis (5/69, 7.2%), starvation (1/69, 1.5%) and trauma (3/69, 4.4%), while for 2 swans the cause of death remained undetermined. Of the 69 swans, 34 were male (49.3%) and 35 were female (50.7%), 14 were juvenile (20.3%), 9 were sub-adults (13.0%), and 46 were adults (66.7%). Samples included paired oral/clausal swabs preserved in universal transport medium (Starswab Multitrans System, Starplex Scientific, Etobicoke, ON, Canada) and all samples tested negative for avian influenza. Swans were collected from five different locations (62 were from Abbotsford, 3 from Deroche, 2 from Ladner, 1 from Mission, and 1 from Surrey), all situated within 75 km of each other and part of the Pacific Flyway. All samples were collected under Environment Climate Change Canada Scientific Salvage Permits (SC-BC-2019-0012SAL, SC-BC-2020-0012SAL). This study was carried out in accordance with guidelines of the Canadian Council on Animal Care with approved protocol 20-05-AL from the Memorial University Institutional Animal Care Committee.

2.2. Screening and sequencing

The virus was originally discovered with the VidION method in a juvenile female from Surrey that died on November 27, 2019, because of starvation. RNA isolated from pre-treated samples, prepared as described in Canuti et al. (2021), was screened for the virus by a heminested PCR targeting the VP1 region, performed with primers Swalici_F1 (CTCACAAATGTATAAGCGCTTG) and Swalici_R3 (GGAGCAACAGTGCCGTAC) during the first amplification round (amplified fragment: 177 nt) and Swalici_F2 (GCAAATACA- GATCTCAGGTC) and Swalici_R3 during the second step (amplified fragment: 138 nt), using the DreamTaq Green PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA). Reactions were performed in 25 µl final volume using 2.5 µl cDNA and each primer at a final concentration of 0.2 µM and were run for 5 min at 95 °C, followed by 35 or 25 cycles for the first and nested PCRs, respectively, of 30 s at 95 °C, 30 s at 50 °C, and 30 s at 72 °C, followed by a final extension step at 72 °C for 7 min. PCR products were loaded on agarose gel and positives were identified by evaluating the presence of a clear band of the predicted size. Confidence intervals (CI) and statistical significance for differences in positivity rates (number of positive samples over the total number of samples) between groups were calculated using the Mid-p exact two-tailed test with OpenEpi (Dean et al., 2013).

After the original identification, 16 µl of RNA freshly isolated from the first swan identified as positive (SW67) was subjected to DNase I treatment (New England Biolabs, Ipswich, Massachusetts, United States), a reverse transcription with MMLV (Promega, Madison, Wisconsin, United States) and a mixture of random hexamers and poly-(T) primers, a second strand synthesis (Second Strand cDNA Synthesis Kit, New England Biolabs), a final purification with AMPure XP (Beckman Coulter, Brea, CA, USA), and outsourced to the Integrated Microbiome Resource of the Centre for Comparative Genomics and Evolutionary Bioinformatics (Dalhousie University, Halifax, Canada) for Illumina high-throughput sequencing after tagmentation-based library preparation (https://imr.bio/protocols.html). Reads obtained were filtered using Trimmomatic (Bolger et al., 2014) to remove low quality leading and trailing regions, Illumina adapters, and reads with a length of less than 36 nt. Read pairs for which both reads survived were assembled into contigs using SPAdes (Bankevich et al., 2012). Finally, taxonomy was assigned to contigs through discontinuous megablast (Altschul et al., 1990) (NCBI nucleotide collection database downloaded on 24 May 2022) and the results were visualized with MEGAN 6 (Huson et al., 2016).

2.3. Sequence analyses

Motif search and genome annotations were performed with Geneious R11 (Biomatters). The internal ribosomal entry site (IRES) prediction was performed by aligning the 5’UTR sequence of the identified virus to those of other caliciviruses and mFold (Zuker, 2003) was used to build the model using what was predicted by Arhab et al. (2022) as template. A database including all caliciviral sequences obtained from birds available in GenBank as of June 7, 2022 (N = 107) was built (Supplementary Table S1) and predicted protein sequences were aligned to each other and to other reference caliciviruses with MAFFT (E-INS-I algorithm) (Katoh and Standley, 2013). Maximum likelihood phylogenetic trees were built with IQ-TREE 2 (Minh et al., 2020) using the best model for distance estimates identified as the one with the lowest Bayesian information criterion (BIC) with the ModelFinder function (Kalyaanamoorthy et al., 2017). Reference sequences (non-avian caliciviruses) were obtained from the resource section in the ICTV Caliciviridae webpage. Finally, branch robustness was assessed using both ultrafast bootstrap approximation (uBoot) (Hoang et al., 2018) and SH-like approximate likelihood ratio test (SH-aLRT) (Guindon et al., 2010).

3. Results and discussion

3.1. A novel calicivirus in trumpeter swans

A viral genome of 8368 nt was assembled from the 1408,535 obtained Illumina reads (mean coverage 37.5X) and deposited in GenBank under accession number OP271827. The sequence included a 606-nt long 5’UTR untranslated region (UTR), two main ORFs, and a 51-nt long 3’UTR that also included a poly(A) tract. Since a 5’ RACE was not performed, it is possible that the 5’ UTR is not complete. This genome organization is typical of caliciviruses (Desselberger, 2019; Vinje et al., 2019).

A 5’ UTR that is hundreds of nucleotides long can be found in caliciviruses of birds (Arhab et al., 2022) while the 5’UTRs of non-avian viruses are typically much shorter (Alhatlani et al., 2015). Interestingly, as also observed for several other avian caliciviruses (Arhab et al., 2022), the 5’UTR of the identified virus could be folded into a type 2 IRES (Fig. 1), although the apical part of domain J was slightly different from what was predicted for other viruses.

The biggest ORF (7086 nt) encodes the polyprotein (3261 aa), whose predicted sequence contained the typical amino acid motifs associated with enzymatic functions in caliciviral nonstructural proteins. Protein motif localization suggested that the order of the proteins in the
polyprotein is the same as for other caliciviruses (Vinjé et al., 2019). In fact, the helicase/NTPase motif typical of NS3 was identified at aa 585–592 (GPPGIGKT), a 3C-like cysteine protease motif (GDCGLP) typical of the protease NS6 was found at aa 1230–1235, and the three RNA-dependent RNA polymerase (RdRp) motifs typical of NS7 were found at aa 1513–1619 (DYSKWDST, GLPSG, and YGDD). Putative protease cleavage sites were identified based on comparisons with other avian viruses and following what has been identified for other avian caliciviruses (Arhab et al., 2022) at positions Q433 (NS1–2/NTPase), Q774 (NTPase/p29), Q1033 (p29/VPg), Q1110 (VPg/Proterase-RdRp), and Q1791 (Protease-RdRp/VP1). Finally, the putative VP2, encoded by a separate ORF, was 333 aa in size. We named this virus trumpeter swan calicivirus (TruSCV).

Screening swan samples for this virus demonstrated that TruSCV was relatively common in the investigated population as 20.3% (14/69, CI: 12.0–31.0%) of the swans were positive. Nine of the positive swans were female and 5 were male but viral positivity rates in males and females were not significantly different (25.7% (9/35, CI: 13.3–42%) for females and 14.7% (5/34, CI: 5.6–29.6%) for males, p = 0.3). Similarly, no significant differences in positivity were observed between juveniles (4/14, 28.6%, CI: 9.8–55.5), sub-adults (1/9, 11.1%, CI: 0.6–43.9), or adults (9/46 19.6%, CI: 10.0–32.9%). Not surprisingly, as most of the investigated swans died because of lead poisoning, while one animal died because of starvation, and one from physical injuries. Unfortunately, since all swans included in this study were likely weakened by co-occurring conditions, we cannot make conclusions about the pathogenic potential of this virus as viral prevalence might be different among healthy swans. TruSCV was found in samples collected at three locations (11 were from Abbotsford, 1 from Surrey, and 2 from Ladner) indicating that this virus could be common in this part of British Columbia in swans using the Pacific Flyway.

3.2. Genome organization comparisons and phylogeny of avian caliciviruses

To study the phylogenetic relationships between TruSCV and other caliciviruses, as well as to investigate the overall diversity of avian caliciviruses, a phylogenetic tree was built with the full VP1 protein sequences of 63 avian caliciviruses and 64 reference sequences (Fig. 2). This protein was chosen since it is the one used by the ICTV to define caliciviral genera (Vinjé et al., 2019). The VP1 of avian caliciviruses were monophyletic and these viruses formed several well-supported groups, but only two of those corresponded to caliciviral genera defined by the ICTV (Bavovirus and Nacovirus). Additionally, pairwise sequence identities between sequences within each of these clades were,
in some cases, as low as ~30%. Since the demarcation criteria for caliciviral genera is 40%, these results suggest that several yet unclassified viral genera exist, likely also currently formed by one single member. Overall, 32 different viruses could be identified in this tree and two additional partially sequenced viruses could be identified based on phylogenetic analysis of the RdRp region (Supplementary Fig. S1). In both trees, TruSCV was included in a clade in which the only other members were viruses identified in 2016–2017 in grey teals (Anas gra- cilis) and Pacific black ducks (Anas superciliosa) from Australia (Vibin et al., 2020; Wille et al., 2019) (Figs. 2, Supplementary S1, Supplementary Table S1). Interestingly, our database demonstrated that viruses belonging to the genus Norovirus were also occasionally identified in birds. These were very similar to viruses found in humans (accession numbers MN175617, MN175616, and MF444290- MF444294) and swine (KT326930 and KT326933), but the significance of this finding in terms of viral replication in birds is uncertain, as also highlighted by Summa et al. (2018).

With one exception, possibly due to sequencing error (MT239355), all 39 viruses for which the complete coding sequences were available presented the same genomic organization, with one unique ORF coding for all non-structural proteins and the VP1 and a second smaller ORF coding for VP2 (Fig. 3). The conserved motif found in proteins NS3, NS6, and NS7 were identified in all genomes in the same order. Finally, an analysis performed with 90 polyprotein sequences demonstrated that some of these motifs were more conserved than others with the protease and the first polymerase domains being the most variable. The high variability of sequences at the level of these important motifs within some of the identified clades is also indicative of a great viral diversity and of the potential presence of multiple yet-undefined viral genera and species.

Interestingly, it has been shown that different avian caliciviruses use different types of IRES for translation initiation and it has been hypothesized that they were acquired through horizontal gene transfer (HGT) from other viral families. Both viruses found in clade 4 (TruSCV and grey teal calicivirus) and several viruses within the genus Nacovirus (avocet calicivirus, pink-eared duck calicivirus, Wilkes virus, red-crowned crane calicivirus, and calicivirus strain xftoti59cal1) have all been found to use type 2 IRESs while viruses within clade 5 possess type 2 (duck calicivirus), type 4 (ruddy turnstone calicivirus), and type 5 (calicivirus strain hwf182cal1) IRESs, indicating that HGT could have occurred on multiple occasions. Further efforts should be done to elucidate the secondary structure of the 5’ UTRs of more avian viruses and compare these structures with viral phylogeny to clarify more precisely at which points in the evolution of these viruses HGTs occurred.
4. Conclusions

In this study we genetically characterized a novel avian calicivirus that is present in swans wintering in south-west British Columbia, Canada. The virus is included in a currently undefined caliciviral genus and is a close relative to a virus circulating in Australian ducks. As we only collected preliminary data about virus epidemiology and, since all infected swans died from other conditions, follow-up studies are necessary to elucidate the diversity, distribution, and pathogenic potential of this virus. Our analyses further revealed that the diversity of avian caliciviruses is higher than currently recognized as many yet unclassified viral species and genera exist. As all avian caliciviruses were discovered in the last decade, many more related viruses are likely yet to be discovered. Although genome organization was conserved among clades, different species possessed different IRES types, presumably acquired by multiple HGT events. This also demonstrates how little is currently known about the ecology and evolution of these viruses and highlights the need for more epidemiological investigations and systematic analyses to better define their taxonomy.

Funding

Funding for this research was provided by the Ocean Frontier Institute through a Seed Grant awarded to J.T.P.V. from the Canada First Research Excellence Fund. M.C. was supported by funds from the Joint Mink Research Committee. Sample collection was supported by the Canadian Wildlife Service, Environment and Climate Change Canada, Parks Canada, and the British Columbia Ministry of Agriculture.

CRediT authorship contribution statement

Marta Canuti: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – review & editing. Laurie Wilson: Resources, Writing – review & editing. Victoria Bowes: Resources, Writing – review & editing. Tony Redford: Resources, Writing – review & editing. Suzanne C. Dufour: Writing – review & editing, Supervision, Funding acquisition. Andrew S. Lang: Writing – review & editing, Supervision, Funding acquisition. Joost T.P. Verhoeven: Conceptualization, Methodology, Software, Investigation, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Joost T.P. Verhoeven reports financial support was provided by Ocean Frontier Institute. Marta Canuti reports financial support was provided by Joint Mink Research Committee.

Data availability

Obtained sequences have been submitted to Public Databases and accession numbers were provided.

Acknowledgements

We are grateful to Xiao Jun (Jim) Song and Owain McKibbin for their dedicated efforts recovering sick and dead swans from the study area as well as Wildlife Rescue Association staff for the care they provided to swans admitted to their rehabilitation facility. We thank staff at the BC Ministry of Agriculture Animal Health Lab for assistance with necropsies.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.crmicr.2022.100169.
