



Identification and characterization of a novel bocaparvovirus in tufted deer (*Elaphodus cephalophus*) in China

Ziyuan Dai¹ · Hao Wang² · Shixing Yang¹ · Quan Shen¹ · Xiaochun Wang¹ · Tianji Zhou¹ · Zhanghao Feng¹ · Wen Zhang¹

Received: 7 August 2021 / Accepted: 8 October 2021 / Published online: 14 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2021

Abstract

We used viral metagenomics and next-generation sequencing to identify a novel strain of bocaparvovirus in the intestinal tract of tufted deer (*Elaphodus cephalophus*), tentatively named "Elaphodus cephalophus bocaparvovirus" (ECBOV). A nearly complete genome sequence of 5,354 nucleotides was obtained, which had the typical genome organization and protein motifs of a bocaparvovirus. Sequence comparisons and phylogenetic analysis revealed that ECBOV may be a new ungulate bocaparvovirus. The identification and characterization of viruses in wildlife will facilitate our understanding of genetic evolution and cross-species transmission and thus further reduce the potential threat to human and animal health.

The name "bocaparvovirus" (BOV) is derived from the combination of "bovine parvovirus" and "canine minute virus" [1]. These two parvoviruses were first identified in cattle and dogs in the early 1960s and were both placed in the same genus, "*Bocavirus*", because of their similar genomic organization and high level of sequence similarity. Like RNA viruses, single-stranded DNA viruses have a high rate of evolution because they mutate faster than double-stranded viruses, and genome size appears to correlate negatively with mutation rate [2].

BOVs are commonly present as mixed infections and frequently associated with gastrointestinal and respiratory symptoms in children and neonatal animals, causing serious public-health and economic problems. In China, a two-year molecular epidemiology study showed that human bocaparvovirus (HBOV) and human metapneumovirus (HMPV) are frequently found in children with pneumonia. In recent years, various BOVs have been detected in different hosts

with different geographical distributions. There is a high prevalence of feline bocaviruses in cats with diarrhea in northeastern China [3]. Pneumonia was diagnosed in 96.2% of HBOV-positive and 98.5% HMPV-positive patients [4]. Porcine bocaviruses, first identified in swine with post-weaning multisystemic wasting syndrome, were also detected in fecal samples of piglets with clinical diarrhea in southeastern China [5, 6].

The genus *Bocaparvovirus* belongs to the family *Parvoviridae* [1]. Parvoviruses are non-enveloped viruses with a round shape, T=1 icosahedral symmetry, and diameter of 18–26 nm [7]. Their linear single-stranded DNA (ssDNA) genome is approximately 4–6 kilobases (kb) in length and contains two major open reading frames (ORFs) encoding a non-structural protein (NS1) and a viral protein (VP1) [8]. BOVs have a number of characteristic features in common with other parvoviruses, but unlike most parvoviruses, BOVs have three ORFs in their genome [1, 9]. The third ORF, which encodes a nuclear phosphoprotein, NP1, lies between the ORFs encoding NS and VP.

Tufted deer (*Elaphodus cephalophus*) are small deer belonging to the genus *Elaphodus* in the subfamily Muntiacinae, family Cervidae which are found in the southwestern and southern provinces of China [10]. The tufted deer is the only member of the genus *Elaphodus*, so its genetic resources are an important part of the genetic gene pool of the family Cervidae. Tufted deer are listed as near threatened on the IUCN 2013 Red List of Threatened Species

Handling Editor: Ana Cristina Bratanich.

Ziyuan Dai and Hao Wang contributed equally to this work.

✉ Wen Zhang
z0216wen@yahoo.com

¹ Department of Laboratory Medicine, School of Medicine, Jiangsu University, Zhenjiang 212013, Jiangsu, China

² Department of Clinical Laboratory, The Affiliated Huai'an Hospital of Xuzhou Medical University, Huai'an 223002, Jiangsu, China

VER3.1 – vulnerable due to habitat loss and illegal hunting and trade [11].

In October 2019, we collected fecal samples of four healthy wild tufted deer from the Wannan Wild-Animal First-Aid Center in the south of Anhui province of China. All samples were shipped on dry ice. Prior to viral metagenomics analysis, the collected fecal samples were diluted in 1 mL of Dulbecco's phosphate-buffered saline (DPBS) and vortexed vigorously for 5 min. After centrifugation for 10 min at $15000 \times g$, the supernatants were collected in 1.5-ml centrifuge tubes and stored at -80°C for later use.

The four samples were combined into one pool for next-generation sequencing. The supernatant enriched in viral particles was purified using a 0.45- μm syringe filter (Millipore), and the filtrate was treated at 37°C for 90 min with a mixture of DNase, RNase, Benzonase, and Baseline-ZERO to digest unprotected nucleic acid [12]. Total nucleic acid was then extracted using a QIAamp MinElute Virus Spin Kit (QIAGEN) as per the manufacturer's instructions. A Nextera XT DNA Sample Preparation Kit (Illumina) was then used to construct a 250-bp paired-end cDNA library, and the sample pool was sequenced using an Illumina MiSeq platform for double barcode sequencing. Paired-end reads generated by MiSeq were debarcoded using Illumina's vendor software and used for further bioinformatics analysis. The data were processed on a 32-node Linux cluster. Low-quality reads, tag sequences, and duplicates were removed, and the cleaned reads were assembled *de novo* within each barcode using the ENSEMBLE assembler [13]. Unassembled reads and contigs were then compared to a custom-built viral proteome database using BLASTx, with an *E*-value cutoff of $<10^{-5}$. To eliminate false-positive viral reads, the candidate viral reads were then compared to a non-virus non-redundant protein database consisting of non-viral protein sequences extracted from an NCBI nr fasta file (based on annotation taxonomy excluding Virus Kingdom). Contigs without significant BLASTx similarity to the viral proteome database were searched against viral protein families in the vFam database [14] using HMMER3 [15–17] to detect remote viral protein similarities.

The reads that were categorized as having been derived from cellular organisms such as archaea, eukaryotes, and bacteria, as well as reads with no obvious resemblance to amino acid sequences in the non-redundant protein database, were removed. A total of 747 Mb of viral nucleotide sequence data was obtained, consisting of 6,134,914 reads. Of these, 119 showed the best match with bocaparvovirus proteins, with a cutoff *E*-value of 10^{-5} . *De novo* assembly using Geneious Prime generated four genome fragment contigs of 1517, 1133, 885, and 251 nt. The approximate locations of the four contigs in the genome were determined by BLASTn. The sequence identity values for the four contigs ranged from 77.78% to 93.52%.

To determine the genome sequence of the novel BOV, nested PCR and primer walking were performed. Primers were designed based on the sequences of the four contigs, and their sequences are shown in Supplementary Table S1. The PCR thermocycler program consisted of 5 min at 95°C , 31 cycles for 30 s at 95°C , 30 s at 50°C (first round) or 55°C (second round), and 60 s or 40 s at 72°C , and a final elongation step at 72°C for 5 min. The premixed enzyme rTaq (TaKaRa) was used in the reaction system. Primers used for PCR were designed based on the virus contigs, using Geneious Prime [18]. In addition, primer walking was performed to obtain the sequences of the 5'UTR and 3' end of the VP gene. The PCR conditions were as follows: 95°C for 5 min, 31 cycles of 95°C for 30 s, 46°C (for the first round) or 48°C (for the second round) for 30 s, and 72°C for 40 s, and a final elongation step at 72°C for 5 min. The PCR results showed that only one of the four fecal samples was positive for the novel BOV, and the viral genome from this positive sample was sequenced. At this point, the genome sequence was still incomplete (4,884 nt in length).

To complete the genome sequence of ECBOV, a next-generation sequencing library based on the ECBOV-positive sample was constructed and sequenced. Using the new metagenomic data, a total of 773 reads were mapped to the partial ECBOV sequence (4,884 nt in length) using Geneious, and a consensus sequence with a length of 5,354 nt was obtained. In addition, we identified sequences of eukaryotic viruses of the families *Parvoviridae* and *Genomoviridae*, as well as bovine serum-associated circular virus and numerous bacteriophages (Supplementary Fig. 1). The nearly complete genome sequence of ECBOV was 5,354 nt long, and the G+C content was 54%. It contained a 315-nt 5'UTR, a complete ORF1, a complete ORF3, a small ORF2, and a 114-nt 3'UTR (Fig. 1). Three putative ORFs were predicted using Geneious Prime. ORF1 encodes a putative nonstructural protein, NS1, with 860 amino acids. ORF3 encodes a putative structural protein, VP1, of 662 amino acids, and ORF2 encodes a putative nonstructural protein, NP1, of 210 amino acids that is predicted to be highly phosphorylated. In NS1, several conserved domains were identified, including a replication initiator domain (xxHxHxxxxx), an SF3 helicase domain with an ATP- or GTP-binding Walker A loop (GxxxxGKT), Walker B loop (xxxxEE), and Walker B' loop (KxxxxGxxxxxxxxK). VP1 was found to contain a calcium-binding loop (YLGPF) and the putative catalytic residues (DxxAxxHDxxY) of a phospholipase A2 (PLA2) domain in the unique region, which, in other parvoviruses, is necessary for infectivity [19] (Fig. 1B).

Phylogenetic analysis was performed using the ECBOV sequence, 21 genomic sequences retrieved from the GenBank database that exhibited sequence similarity in a BLASTx search, bocaparvovirus reference sequences, and outgroup sequences. Amino acid sequence alignments

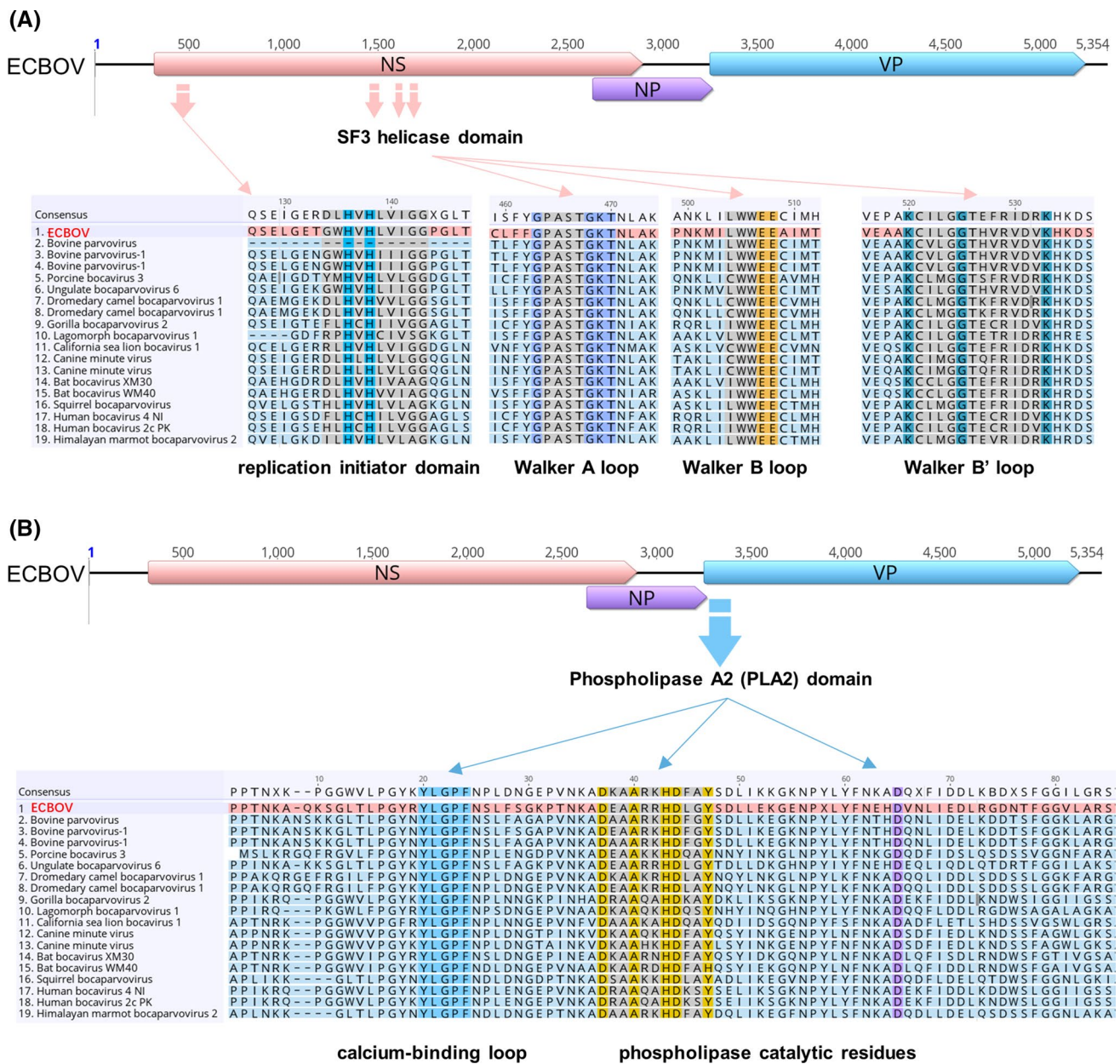


Fig. 1 The genome organization of ECBOV. The NS1 ORF is shown in pink, and the NP ORF is shown in purple. The right end of VP1 is truncated and shown in blue. (A) Identification of the replication initiator domain and SF3 helicase domain in the NS1 protein. (B) Identification of the phospholipase A2 (PLA2) domain in the N-terminal portion of the VP1 protein.

used for all trees were made using MUSCLE, implemented in MEGA 10.2.2. The resulting alignment was trimmed manually, and phylogenetic analysis was performed using MrBayes v3.2 by the mixed-models and Markov chain Monte Carlo (MCMC) methods [20]. All trees were visualized in FigTree v1.4.3, and the midpoints were rooted for clarity. Bayesian consensus trees were constructed based on the amino acid sequences of the NS and VP proteins. As shown in Fig. 2, ECBOV branched together with ungulate bocaparvovirus 6 (accession number NC_030402),

bovine parvovirus (NC_001540), and bovine parvovirus 1 (NC_038895 and MW032436). The NS1 protein of ECBOV had 62.66% amino acid sequence identity to NS1 of bovine parvovirus-1 (NC_038895, query coverage: 99%). The VP protein had 63.57% amino acid sequence identity to the corresponding structural protein of ungulate bocaparvovirus 6 (KU172422, query coverage: 100%). The putative NP had 45% to 60% amino acid sequence identity to the NPs of other members of the family *Parvoviridae*. In 2019, new taxonomic guidelines were established for the family

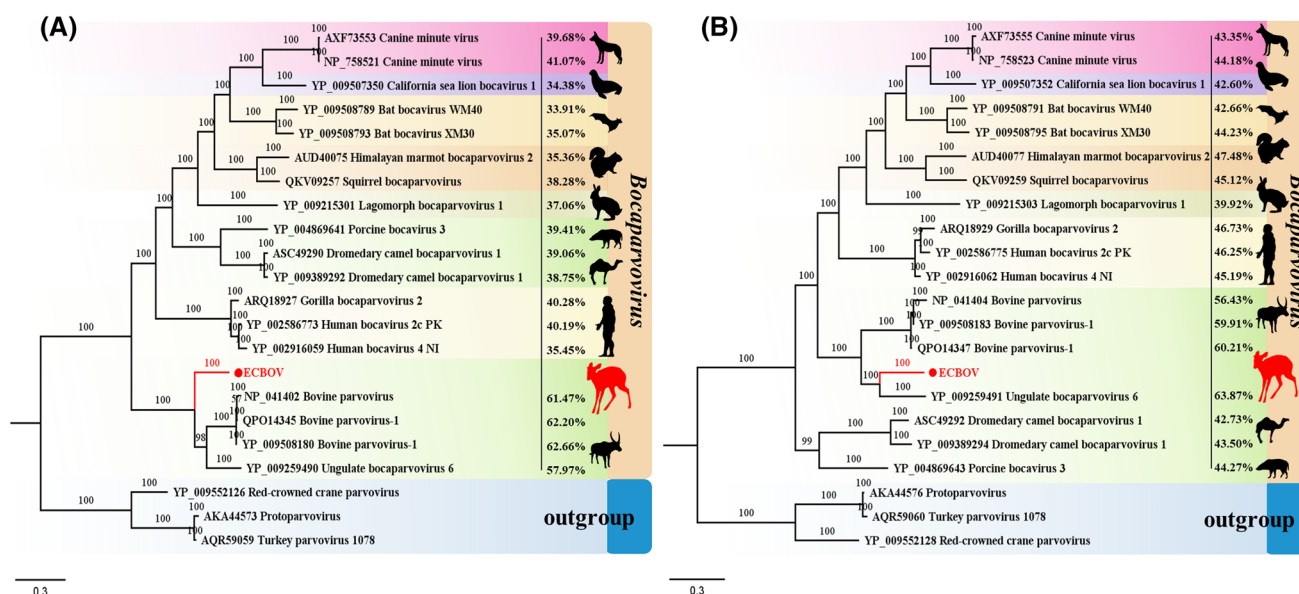


Fig. 2 Bayesian consensus trees based on the NS1 (A) and VP (B) amino acid sequences of parvoviruses. Bootstrap values are shown for each node. The bar indicates 0.3 amino acid substitutions per site. The percentage of sequence identity to ECBOV is shown for each virus.

Parvoviridae [21]. Members of a genus should belong to the same monophyletic lineage based on their NS1 protein sequence, with at least 35–40% amino acid sequence identity between any two members and a coverage rate of at least 80%. Members of the same species should have >85% sequence identity in the NS1 protein. Based on these criteria, ECBOV should be classified as a member of a novel species in the genus *Bocaparvovirus*.

The nearly complete genome sequence of ECBOV was deposited in the GenBank database under accession number MW824450.2. The biological data and information about the tufted deer faeces samples has been deposited in the NCBI database under the BioProject accession number PRJNA705580. The first and second next-generation sequencing data sets are available in the SRA database under the accession numbers SRX10243189 and SRX11527391, respectively.

Knowledge about viral diversity in the family *Parvoviridae* has increased greatly in the past few years, partly due to the application of viral metagenomics technology, which has provided many new genome sequences and revealed entirely new lineages. The traditional taxonomy based on host specificity has not been an appropriate framework for classifying members of the family *Parvoviridae*. A new system was therefore established [21] by the International Committee on Virus Taxonomy that divided the parvoviruses into three subfamilies: *Parvovirinae*, *Densovirinae*, and *Hamaparvovirinae*.

BOVs infect a wide range of hosts, with documented cross-species transmission between human and non-human primates [22]. According to the hosts they infect, BOVs have

been assigned to 28 species, including *Carnivore bocaparvovirus 1–6*, *Chiropteran bocaparvovirus 1–5*, *Lagomorph bocaparvovirus 1*, *Pinniped bocaparvovirus 1–2*, *Primate bocaparvovirus 1–3*, *Rodent bocaparvovirus 1–2*, and *Ungulate bocaparvovirus 1–9*. Phylogenetic analysis showed the closest evolutionary relationship between ECBOV and bovine-associated bocaparvoviruses (BOBOV). BOBOV was first isolated in 1961 from the gastrointestinal tract of normal calves by Abinanti and Warfield [23]. BOBOV was always found in coinfections with other well-characterized virus such as bovine rotavirus or bovine viral diarrhea virus, in cases of gastrointestinal diseases [24], and with bovine coronavirus or bovine rhinitis A and B viruses in the case of respiratory diseases [25]. BOVs have been known in veterinary and clinical medicine since the discovery in 2005 of HBOV, which can cause lower respiratory disease in children [26]. When the virus is released from the respiratory or gastrointestinal tract of an infected host, it can be transmitted horizontally to other individuals. In addition, bocaparvoviruses are able to infect mammalian embryos through vertical placental transmission and may cause fetal death and repeated abortion [27].

In this study, the virome was found to include numerous bacteriophages and eukaryotic viruses, but no additional common mammalian viruses were found. The presence of a large number of phages suggested the presence of numerous bacterial species in the intestines of tufted deer, but whether any of them are associated with disease remains to be investigated. Due to the increasingly frequent interactions between humans and wildlife, further surveillance studies are needed to identify any threats to human and animal

health. While the fecal samples in this study were collected from seemingly healthy tufted deer, previous studies have suggested that HBOV1 is an important pathogen associated with lower respiratory tract illnesses [28]. Persistent infection with BOVs is thought to be responsible for coinfection with other viruses. It is speculated that BOVs may be both passengers and pathogenic agents. In a study of acute wheezing disease in children, Bodewes *et al.* [29] found that rhinoviruses were associated with Th2-type cytokine responses or systemic proinflammatory responses, while coinfection with HBOV1 and rhinoviruses led to non-Th2-type cytokine responses. Another study showed that HBOV NP inhibited the production of IFN- β *in vitro*, which increases susceptibility to viral infection [30]. Whether other viruses act as helper viruses for BOVs is still unknown because multiple viruses are often present. The association of BOVs with specific diseases is inconclusive, and larger prospective studies are needed to investigate the pathogenicity of BOVs.

Although serological evidence and *in vitro* virus replication studies are still needed, considering that ECBOV has been found in an ungulate, the genomic structure is typical of BOVs, and the virus shares 45 to 65% amino acid sequence identity with BOBOV, we propose that this novel virus should be considered a member of a new species. Although this parvovirus may cause only inapparent or subclinical infections, early detection of such viruses in wild animals provides an opportunity to investigate potential threats. Recent human activities, ranging from concentrated animal feeding operations to illegal trade in wild animals, have triggered epidemics caused by zoonotic viruses. Future studies are needed to elucidate the distribution, host range, and cross-species transmission of ECBOV potential to better understand its epidemiological significance.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00705-021-05308-3>.

Acknowledgements This work was supported by National Key Research and Development Programs of China for Virome in Important Wildlife Grant no. 2017YFC1200201, Jiangsu Provincial Key Research and Development Projects Grant no. BE2017693, and Huai'an Natural Science Research Program Grant no. HAB202034.

References

- Manteufel J, Truyen U (2008) Animal bocaviruses: a brief review. *Intervirology* 51:328–334. <https://doi.org/10.1159/000173734>
- Sanjuan R, Domingo-Calap P (2016) Mechanisms of viral mutation. *Cell Mol Life Sci* 73:4433–4448. <https://doi.org/10.1007/s00018-016-2299-6>
- Yi S, Niu J, Wang H, Dong G, Zhao Y, Dong H, Guo Y, Wang K, Hu G (2018) Detection and genetic characterization of feline bocavirus in Northeast China. *Virology* 15:125. <https://doi.org/10.1186/s12985-018-1034-3>
- Zhou J-Y, Peng Y, Peng X-Y, Gao H-C, Sun Y-P, Xie L-Y, Zhong L-L, Duan Z-J, Xie Z-P, Cao Y-D (2018) Human bocavirus and human metapneumovirus in hospitalized children with lower respiratory tract illness in Changsha, China. *Influenza Other Respir Viruses* 12:279–286. <https://doi.org/10.1111/irv.12535>
- Blomström A-L, Belák S, Fossum C, McKillen J, Allan G, Wallgren P, Berg M (2009) Detection of a novel porcine boca-like virus in the background of porcine circovirus type 2 induced postweaning multisystemic wasting syndrome. *Virus Res* 146:125–129. <https://doi.org/10.1016/j.virusres.2009.09.006>
- Li B, Ma J, Xiao S, Fang L, Zeng S, Wen L, Zhang X, Ni Y, Guo R, Yu Z *et al* (2012) Complete genome sequence of a novel species of Porcine Bocavirus, PBoV5. *J Virol* 86:1286–1287. <https://doi.org/10.1128/JVI.06589-11>
- Pérez JJ, de Souza WM, Agbandje-McKenna M, Gifford RJ (2019) An ancient lineage of highly divergent parvoviruses infects both vertebrate and invertebrate hosts. *Viruses*. <https://doi.org/10.1101/571109>
- Mietzsch M, Pérez JJ, Agbandje-McKenna M (2019) Twenty-five years of structural parvovirology. *Viruses*. <https://doi.org/10.3390/v11040362>
- Jartti T, Hedman K, Jartti L, Ruuskanen O, Allander T, Soderlund-Venermo M (2012) Human bocavirus—the first 5 years. *Rev Med Virol* 22:46–64. <https://doi.org/10.1002/rmv.720>
- Zhang W, Cao X, Shen Q, Liu W, Ni S, Hua X, Zhang X (2008) Isolation and characterization of the Cetn1 gene from tufted deer (*Elaphodus cephalophus*). *Biochem Genet* 46:652–662. <https://doi.org/10.1007/s10528-008-9179-1>
- Panyaboriban S, Singh RP, Songsasen N, Padilla L, Brown J, Reed D, Techakumphu M, Pukazhenth B (2016) Reproductive seasonality and sperm cryopreservation in the male tufted deer (*Elaphodus cephalophus*). *Theriogenology* 86:914–923. <https://doi.org/10.1016/j.theriogenology.2016.03.014>
- Zhang W, Li L, Deng X, Kapusinszky B, Pesavento PA, Delwart E (2014) Faecal virome of cats in an animal shelter. *J Gen Virol* 95:2553–2564. <https://doi.org/10.1099/vir.0.069674-0>
- Deng X, Naccache SN, Ng T, Federman S, Li L, Chiu CY, Delwart EL (2015) An ensemble strategy that significantly improves de novo assembly of microbial genomes from metagenomic next-generation sequencing data. *Nucleic Acids Res* 43:e46. <https://doi.org/10.1093/nar/gkv002>
- Skewes-Cox P, Sharpton TJ, Pollard KS, DeRisi JL (2014) Profile hidden Markov models for the detection of viruses within metagenomic sequence data. *PLoS ONE* 9:e105067. <https://doi.org/10.1371/journal.pone.0105067>
- Eddy SR (2009) A new generation of homology search tools based on probabilistic inference. *Genome Inform* 23:205–211
- Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39:W29–W37. <https://doi.org/10.1093/nar/gkr367>
- Johnson LS, Eddy SR, Portugaly E (2010) Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinform* 11:431. <https://doi.org/10.1186/1471-2105-11-431>
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C *et al* (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Zádori Z, Szelei J, Lacoste MC, Li Y, Gariépy S, Raymond P, Allaire M, Nabi IR, Tijssen P (2001) A viral phospholipase A2 is required for parvovirus infectivity. *Dev Cell* 1:291–302
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model

- choice across a large model space. *Syst Biol* 61:539–542. <https://doi.org/10.1093/sysbio/sys029>
21. Penzes JJ, Soderlund-Venermo M, Canuti M, Eis-Hubinger AM, Hughes J, Cotmore SF, 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>
 22. Kumakamba C, Ngay Lukusa I, Mbala Kingebeni P, N’Kawa F, Atibu Losoma J, Mulembakani PM, Makuwa M, Muyembe Tamfum JJ, Belais R, Gillis A et al (2018) DNA indicative of human bocaviruses detected in non-human primates in the Democratic Republic of the Congo. *J Gen Virol* 99:676–681. <https://doi.org/10.1099/jgv.0.001048>
 23. Abinanti FR, Warfield MS (1961) Recovery of a hemadsorbing virus (HADEN) from the gastrointestinal tract of calves. *Virology* 14:288–289
 24. Wang M, Yan Y, Wang R, Wang L, Zhou H, Li Y, Tang L, Xu Y, Jiang Y, Cui W et al (2019) Simultaneous detection of bovine rotavirus, bovine parvovirus, and bovine viral diarrhea virus using a gold nanoparticle-assisted PCR assay with a dual-priming oligonucleotide system. *Front Microbiol* 10:2884. <https://doi.org/10.3389/fmicb.2019.02884>
 25. Mitra N, Cernicchiaro N, Torres S, Li F, Hause BM (2016) Metagenomic characterization of the virome associated with bovine respiratory disease in feedlot cattle identified novel viruses and suggests an etiologic role for influenza D virus. *J Gen Virol* 97:1771–1784. <https://doi.org/10.1099/jgv.0.000492>
 26. Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A, Andersson B (2005) Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proc Natl Acad Sci USA* 102:12891–12896
 27. Järplid B, Johansson H, Carmichael LE (1996) A fatal case of pup infection with minute virus of canines (MVC). *J Vet Diagn Investig* 8:484–487
 28. Zhao H, Zhao L, Sun Y, Qian Y, Liu L, Jia L, Zhang Y, Dong H (2012) Detection of a bocavirus circular genome in fecal specimens from children with acute diarrhea in Beijing, China. *PLoS ONE* 7:e48980. <https://doi.org/10.1371/journal.pone.0048980>
 29. Bodewes R, Lapp S, Hahn K, Habierski A, Forster C, König M, Wohlsein P, Osterhaus AD, Baumgartner W (2014) Novel canine bocavirus strain associated with severe enteritis in a dog litter. *Vet Microbiol* 174:1–8. <https://doi.org/10.1016/j.vetmic.2014.08.025>
 30. Zhang Z, Zheng Z, Luo H, Meng J, Li H, Li Q, Zhang X, Ke X, Bai B, Mao P et al (2012) Human bocavirus NP1 inhibits IFN- β production by blocking association of IFN regulatory factor 3 with IFNB promoter. *J Immunol* 189:1144–1153. <https://doi.org/10.4049/jimmunol.1200096>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.