UC Davis
UC Davis Previously Published Works

Title

Idiopathic Chronic Diarrhea in Rhesus Macaques Is Not Associated with Enteric Viral Infections

Permalink

https://escholarship.org/uc/item/0804p756

Journal

Viruses, 13(12)

ISSN

1999-4915

Authors

Delwart, Eric
Tisza, Michael J
Altan, Eda
et al.

Publication Date

2021

DOI

10.3390/v13122503

Peer reviewed

Article

Idiopathic Chronic Diarrhea in Rhesus Macaques Is Not Associated with Enteric Viral Infections

Eric Delwart [1,2,*], **Michael J. Tisza** [3], **Eda Altan** [1,2], **Yanpeng Li** [1,2], **Xutao Deng** [1,2],
Dennis J. Hartigan-O'Connor [4,5] **and Amir Ardeshir** [4]

1   Vitalant Research Institute, 270 Masonic Ave, San Francisco, CA 94118, USA; edaltan@hotmail.com (E.A.);
    alphaleeyp@hotmail.com (Y.L.); XDeng@vitalant.org (X.D.)
2   Department of Laboratory Medicine, UCSF, San Francisco, CA 94118, USA
3   Laboratory of Cellular Oncology, National Cancer Institute, NIH, Bethesda, MD 20892, USA;
    mike.tisza@nih.gov
4   California National Primate Research Center, University of California, Davis, CA 95616, USA;
    dhartigan@ucdavis.edu (D.J.H.-O.); aardeshir@ucdavis.edu (A.A.)
5   Department of Medical Microbiology and Immunology, University of California, Davis, CA 95616, USA
*   Correspondence: delwarte@medicine.ucsf.edu

Abstract: While recent changes in treatment have reduced the lethality of idiopathic chronic diarrhea (ICD), this condition remains one of the most common causes of rhesus macaque deaths in non-human primate research centers. We compared the viromes in fecal swabs from 52 animals with late stage ICD and 41 healthy animals. Viral metagenomics targeting virus-like particles was used to identify viruses fecally shed by each animal. Five viruses belonging to the *Picornaviridae*, one to the *Caliciviridae*, one to the *Parvoviridae*, and one to the *Adenoviridae* families were identified. The fraction of reads matching each viral species was then used to estimate and compare viral loads in ICD cases versus healthy controls. None of the viruses detected in fecal swabs were strongly associated with ICD.

Keywords: idiopathic chronic diarrhea; viral metagenomics; virome; disease association; Cenote-Taker 2

1. Introduction

Idiopathic chronic diarrhea (ICD) is the leading cause of non-medical research mortality and morbidity in captive colonies of rhesus macaques [1,2] and is defined by chronic or recurring non-bloody diarrhea and microscopic colonic ulcers [2–5]. ICD cases are typically diagnosed clinically before one year of age. The histopathological evaluations of the ICD cases show surface epithelium attenuation, goblet cell depletion and crypt branching and microscopic ulcers [6]. The lamina propria is also filled with infiltrates including lymphocytes and plasma cells [5,6]. Rhesus macaques' enterocolitis is a T_H1 biased immune response with concurrent gut microbial dysbiosis [7–10].

Bacterial infections with *Shigella, Campylobacter, Yersinia, Salmonella, and/or Clostridium difficile* as well as parasitic infections, including *Cryptosporidium*, are known causes of diarrhea in macaque colonies [2,11]. Cases of persistent and recurrent diarrhea in which no pathogenic bacteria, parasites or other etiological agents are found are defined as ICD [8].

Rhesus macaques with ICD respond poorly to treatment, including corticosteroids or antibiotics, resulting in frequent hospitalization from dehydration and weight loss and non-responsive animals are euthanized for ethical considerations [6].

A study comparing viral etiology of diarrheal diseases in rhesus macaques found that rhesus and pigtail macaques housed in the monkey farm of the Institute of Labor-atory Animal Sciences, Chinese Academy of Medical Sciences shed enteroviruses, adenoviruses, coronaviruses, and rotaviruses, while others at the Yerkes National Primate Research Center (Yerkes, GA, USA) shed mainly enteroviruses and adenoviruses [12]. Adenoviruses and rotaviruses were also associated with diarrhea in captive rhesus macaques [12–14].

Inoculation with primate caliciviruses induced colitis in juvenile rhesus macaques [15]. SIV (simian immunodeficiency virus)-induced enteropathy was associated with a greater diversity of enteric viruses and parvovirus viremia when compared to SIV-negative rhesus macaques [16].

We previously described the virome in feces of healthy macaques as well as in those with acute diarrhea or ICD and reported that shedding of several picornavirus genotypes were weakly associated with ICD while parvoviruses were weakly associated with healthy animals [17]. During the interval between this and the current studies the treatment for ICD changed and the frequency of ICD dropped significantly. Specifically, antibiotic therapy was no longer performed for animals testing negative for enteropathogens; these animals instead received fluid therapy, and synbiotic or probiotic treatments. Here we extend these earlier studies by analyzing fecal samples from 52 animals with late stage ICD and 41 healthy controls.

2. Materials and Methods

Study design: This study is a case-control to search for the possible differences in the stool virome of ICD and healthy animals. The study was not blinded.

Ethics: All experimental animal procedures were approved by the UC Davis Institutional Animal Care and Use Committee (IACUC Protocol #: 21763) and were in accordance with the requirements of the USDA Animal Welfare Act and the Guide for the Care and Use of Laboratory Animals and the research adhered to the Principals for the Ethical Treatment of Nonhuman Primates, Assurance Number D16-00272(A3433-01). All sample collection activities were conducted per standard operating procedures at the California National Primate Research Center (CNPRC). The UC Davis animal care program is fully accredited by the Association for the Assessment and Accreditation of Laboratory Animal Care, International (AAALACi), registered with USDA, and maintains a Public Health Services Assurance.

Animals: All animals in this study were born, and raised at the CNPRC and samples collected while they are housed in outdoor colony (healthy controls) or from necropsy (ICD). The healthy control animals (*n* = 41) were selected from outdoor housed animals without any gastrointestinal issues (including diarrheal diseases) or antibiotic treatment within the last 6 months prior to sampling. ICD cases (*n* = 52) were selected from the cases that were hospitalized for the third time for non-pathogenic diarrhea within the past 365 days and then euthanized. ICD cases that were euthanized with other accompanying diseases were excluded from the study to control for confounding factors.

Sample collection: Rectal swabs were collected from rhesus macaques while sedated for the bi-annual physical exam (controls) or during necropsy (ICDs) and immediately frozen at −70 °C.

Generation of mNGS libraries: One mL of PBS buffer was added to fecal swabs, vigorously vortexed, and removed from the collection tube. The suspension (500 μL) was then placed in a microcentrifuge tube with 100 μL of zirconia beads and vigorously vortexed. The fecal suspension was centrifuged at 8000× *g* for 5 min at 4 °C, and the supernatant (500 μL) was filtered through a 0.45 μm spin column filter (Millipore, Burlington, MA, USA) to remove bacteria and other large particulates. The filtrate was digested for 1.5 h at 37 °C with a mixture of nuclease enzymes consisting of 14U of Turbo DNase (Ambion, Life Technologies, Carlsbad, CA, USA), 3U of Baseline-ZERO (Epicentre, Madison, WI, USA), 30U of Benzonase (Novagen, Darmstadt, Germany), and 30U of RNase One (Promega, Madison, WI, USA) in 1× DNase buffer (Ambion, Life Technologies, Carlsbad, CA, USA). Then 150 ul of the above mixture was extracted using the MagMAX^TM Viral RNA Isolation kit (Applied Biosystems, Life Technologies, Carlsbad, CA, USA), and nucleic acids were resuspended in 50 μL water with 1 μL RiboLock RNAse inhibitor. A total of 11 μL of nucleic acids were incubated for 2 min at 72 °C with 100 pmol of primer A (5′GTTTCCCACTGGANNNNNNNN3′) followed by a reverse transcription step using 200 units of Superscript III (Invitrogen, Waltham, MA, USA) at 50 °C for 60 min with a

subsequent Klenow DNA polymerase step using 5 units (New England Biolabs, Ipswich, MA, USA) at 37 °C for 60 min. cDNA was then amplified by a PCR step with 35 cycles using AmpliTaq Gold™ DNA polymerase LD with primer A-short (5′GTTTCCCACTGGATA3′) at an annealing temperature of 59 °C. The random amplified products were quantified by Quant-iT™ DNA HS Assay Kit (Invitrogen, Waltham, MA, USA) using a Qubit fluorometer (Invitrogen, USA) and diluted to 1 ng of DNA for library input. The library was generated using the transposon-based Nextera™ XT Sample Preparation Kit using 15 cycles (Illumina, San Diego, CA, USA) and the concentration of DNA libraries was measured by Quant-iT™ DNA HS Assay Kit. The libraries were pooled at equal concentration and size selected for 300 bp–1000 bp using the Pippin Prep (Sage Science, Beverly, MA, USA). The library was quantified using KAPA library quantification kit for Illumina platform (Sigma-Aldrich, MO, USA) and a 10 pM concentration was loaded on the MiSeq sequencing platform for 2×250 cycles pair-end sequencing with dual barcoding.

Bioinformatics: Human and bacterial reads were identified and removed by comparing the raw reads with human reference genome hg38, and bacterial genomes release 66 (collected from ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/, 20 October 2017) using local search mode. The remaining reads were de-duplicated if base positions 5 to 55 were identical. One random copy of duplicates was kept. The sequences were then trimmed for quality and adaptor and primer sequences by using VecScreen [18]. After that, the reads were de novo assembled by Ensemble Assembler [19]. Assembled contigs and all singlet reads were aligned to an in-house viral protein database (collected from ftp://ftp.ncbi.nih.gov/refseq/release/viral/, 20 October 2017) using BLASTx (version 2.2.7) using E-value cutoff of 0.01. The significant hits to the virus were then aligned to an in-house non-virus-non-redundant (NVNR) universal proteome database using DIAMOND to eliminate the false viral hits. Hits with a more significant (lower) E-value to NVNR than to the viral database were removed. The remaining singlets and contigs were compared to all eukaryotic viral protein sequences in GenBank's non-redundant database using BLASTx. The genome coverage of the target viruses was further analyzed by Geneious R11.1.4 software (Biomatters, San Diego, CA, USA).

Quantitation of viral reads: Genetic variation was observed between homologous sequences belonging to the same enterovirus genotypes. For estimating the viral abundance of each viral types using the number of viral reads, closely related contigs (of the same viral genotype) from multiple animals were concatenated (with 100 bp 'N' spacers to prevent reads aligning to the areas where the contigs are joined). The concatenated sequences can be found in supplemental data (Dataset S1). Further information on the % sequence identity of each contig to the most closely related sequences in GenBank are in supplemental data (Table S1). These concatenated contigs where then used to find matching reads using the nucleotide aligner program Bowtie 2 with the seed length parameter set at 30, and the alignment was considered a hit if the read identity to the reference concatemer was greater than 95%. A total of 11 concatenated contigs representing the most commonly detected viruses were generated. For normalization (to account for the variable number of total reads generated from different samples) the number of matching read hits were converted to reads per million (RPM). Heat map was generated using freeware program at heatmapper.ca and the expression function. The RPM were used as input. For the clustering method, the default average linkage was used. For distance measurements, Euclidian distances were used. Both row and column dendrograms are shown. The z-score is the number of standard deviations by which the value of a raw score is above or below the mean value of viral reads per millions (RPM).

Quantitation of de novo assembled virus contigs. The protocols used to predict virus contigs with Cenote-Taker 2 (version 2.0.0) with a minimum contig size of 500 nucleotides and the "standard" database have been described [20,21]. Contigs were dereplicated using the Mummer-based RedRed program [22]. Low-complexity regions of contigs were masked with RepeatMasker [23]. Abundance was calculated with the same methods

as described [21]. Finally, sample-by-sample swarm plots were generated using DaBest python package [24].

3. Results

Fecal swabs were collected during necropsies from 52 macaques suffering from late stage ICD and 41 healthy controls. Fecal material was filtered to enrich for virus-like particles. Host and bacterial nucleic acids were then depleted using nuclease enzymes. Total nucleic acids were extracted from this nuclease digested and filtered materials and amplified using random RT-PCR. DNA was then converted to Illumina-compatible DNA and sequenced (see materials and methods). All raw data are available in GenBank under Bioproject PRJNA608547. The resulting reads were analyzed using BLASTx to identify reads matching all known eukaryotic viruses. Eight different simian viruses were identified. The most prevalent viruses belonged to the *Picornaviridae* family, including members of enterovirus Species A (enterovirus 19, 46, 92), enterovirus Species J (enterovirus 103), and sapelovirus (Species B). Furthermore, also identified but found in fewer animals were sapovirus (*Caliciviridae*), adenovirus (*Adenoviridae*), and bocavirus (*Parvoviridae*). The percent of reads matching these different simian viruses was then plotted in order to compare the eukaryotic virome composition of ICD cases to that of healthy animals (Figure 1). This analysis indicated that for none of the viruses was the median percentage viral read values higher in ICD cases than in healthy controls. Mean percentage viral read values were less consistent with enteroviruses A92 and J103 (and to a lesser extend enterovirus 19) showing larger values in ICD than in healthy controls. Depending on the values compared (median versus mean % viral reads) a consistent correlation was therefore not found for any virus between ICD and higher percentage viral reads.
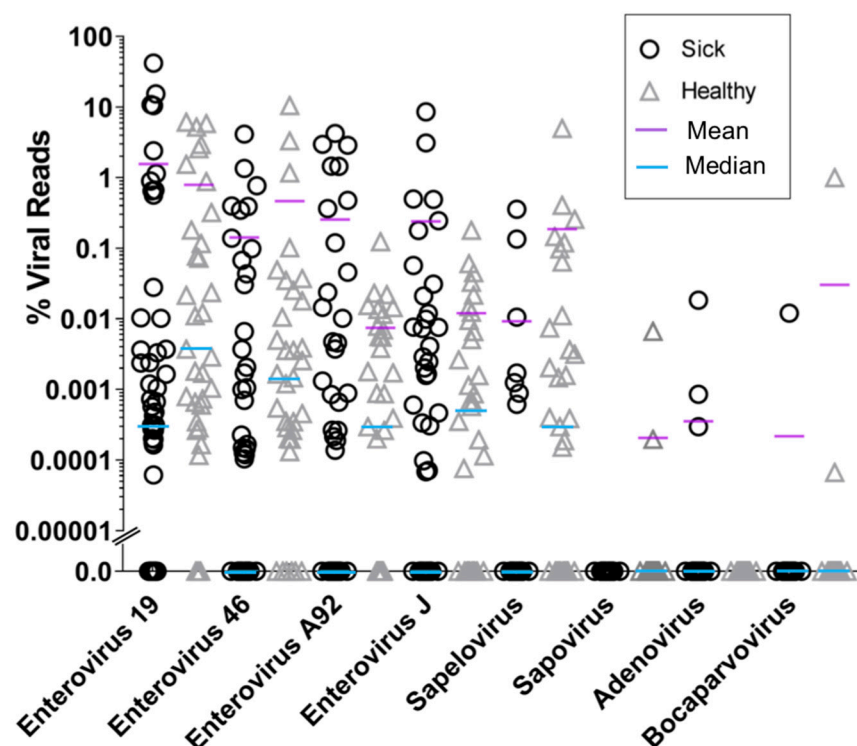

Figure 1. Percentage of reads matching different viruses in fecal swabs from ICD cases (circles) and healthy control (triangles) animals. Viral read numbers were quantified as described in materials and methods.

When the data for the five most common infections was plotted using whisker plots, no difference between the ICD and healthy cases could be seen, except again for mean

percentage viral reads for enterovirus A92 representation being lower in healthy animals (Figure 2).
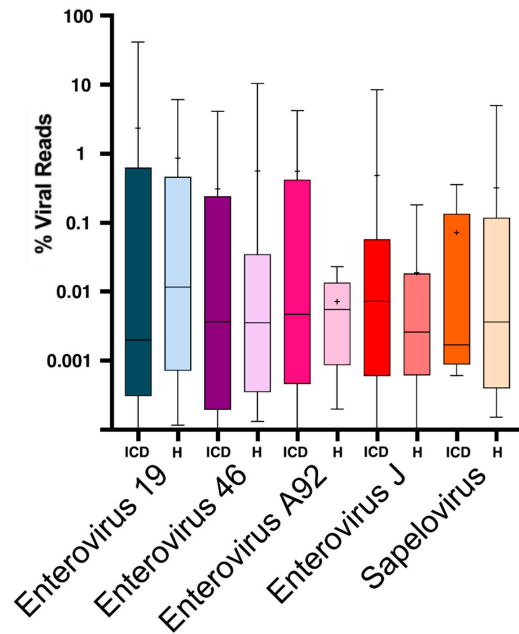

Figure 2. Box and whisker plot of the % viral reads for the five most prevalent viruses. Healthy control animals are labeled H while sick animals are labelled ICD. Boxes indicate the limits of the lower and upper quartile. Horizontal line in each box represents the median % viral reads. The mean % viral reads are indicated by a +. Vertical lines extending from each box represent the minimum and maximum % viral reads.

A hierarchical clustering program (heatmapper.ca) was then used to determine if the viromes of animals with or without ICD clustered separately. No clustering of related viromes based on clinical status (gray and black bar below heat map) was detected further supporting the lack of association between ICD and virome composition (Figure 3).
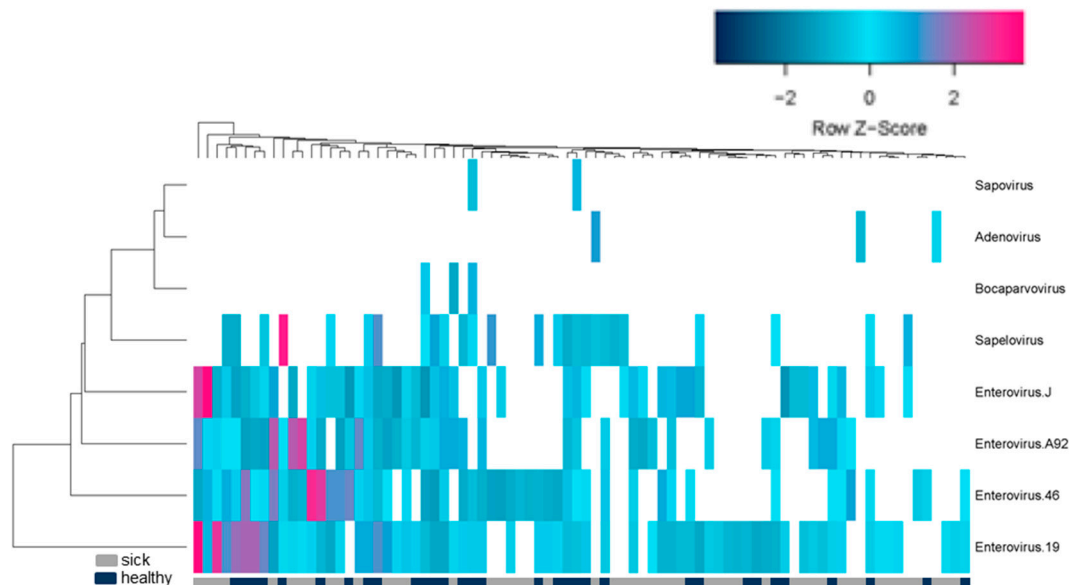
Figure 3. Relative abundance of viruses from rectal swabs. Unsupervised hierarchical clustering analysis of fecal swab viromes suggests that there is no distinction between ICD cases and healthy controls.

We next used Cenote-Taker 2 to identify viral contigs from the Illumina libraries from each animal in the two macaque groups. Cenote-Taker 2 scans open reading frames of contigs for virus hallmark genes, identifying known and divergent viral contigs using a curated dataset and enabling the discovery and annotation of both DNA and RNA viral genomes [20,21]. Such hallmark genes include structural genes such as capsid proteins, genome-packaging genes such as terminases, and genome-replication genes such as RNA-dependent RNA polymerase and replicative helicase [20,21]. The abundance of these viral contigs was compared between ICD and healthy animals by measuring reads per kilobase per millions reads (RPKM) (Figure 4). Only contigs that met minimum thresholds of 0.05 reads per kilobase (RPKM) of transcript in at least 10% all samples were compared in this analysis, as rare contigs are unlikely to originate from viruses with a frequent role in ICD pathogenesis.
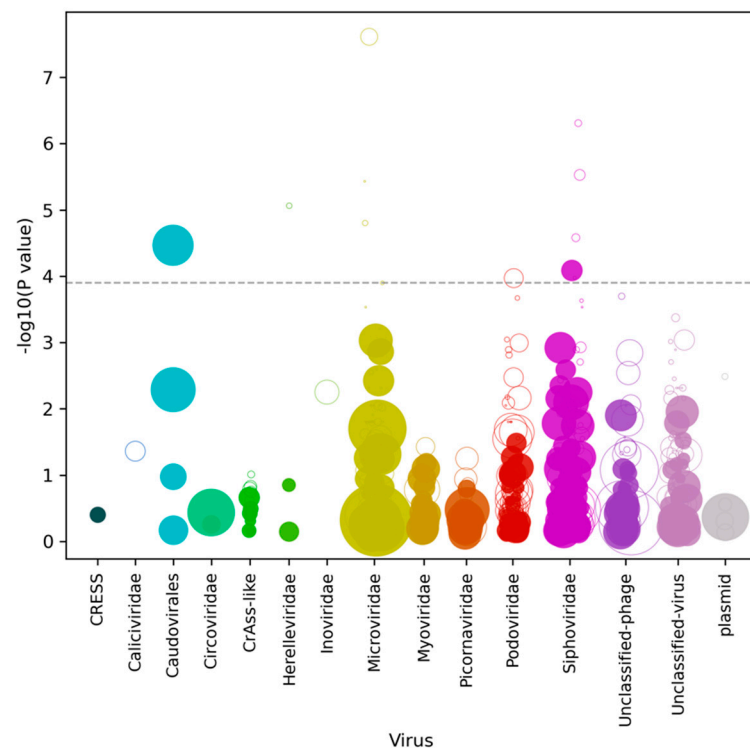

Figure 4. Virome wide associations of different viral families with ICD represented as a Manhattan plot. Each virus contig is represented as a dot along the *x*-axis, with its *y*-axis value being the inverse log10 P value. The size of each dot corresponds to the median relative abundance of the taxon in the disease cohort. Filled dots represent virus contigs found at higher abundance in macaques with ICD while hollow dots represent decreased abundance in macaques with ICD. The dashed gray line represents a false discovery rate threshold of 1%, calculated using the Benjamini–Hochberg procedure.

The 20 virus contigs with the lowest *p* values (i.e., highest $-\log_{10} p$ value) were then compared at a sample-by-sample level (Figure 5A). While twelve virus contigs were significantly different between cohorts after multiple testing correction, it was observed that most relationships represented a minority of samples from one of the cohorts being present at very high RPKM values. Finally, sample-by-sample comparison of eukaryotic viruses (threshold of 10% of samples > 0.05 RPKM) showed no population level differences (Figure 5B). Comparison of abundance of de novo assembled contigs is therefore consistent with the other analyses showing that none of the eukaryotic viruses are associated with ICD.

(**A**)
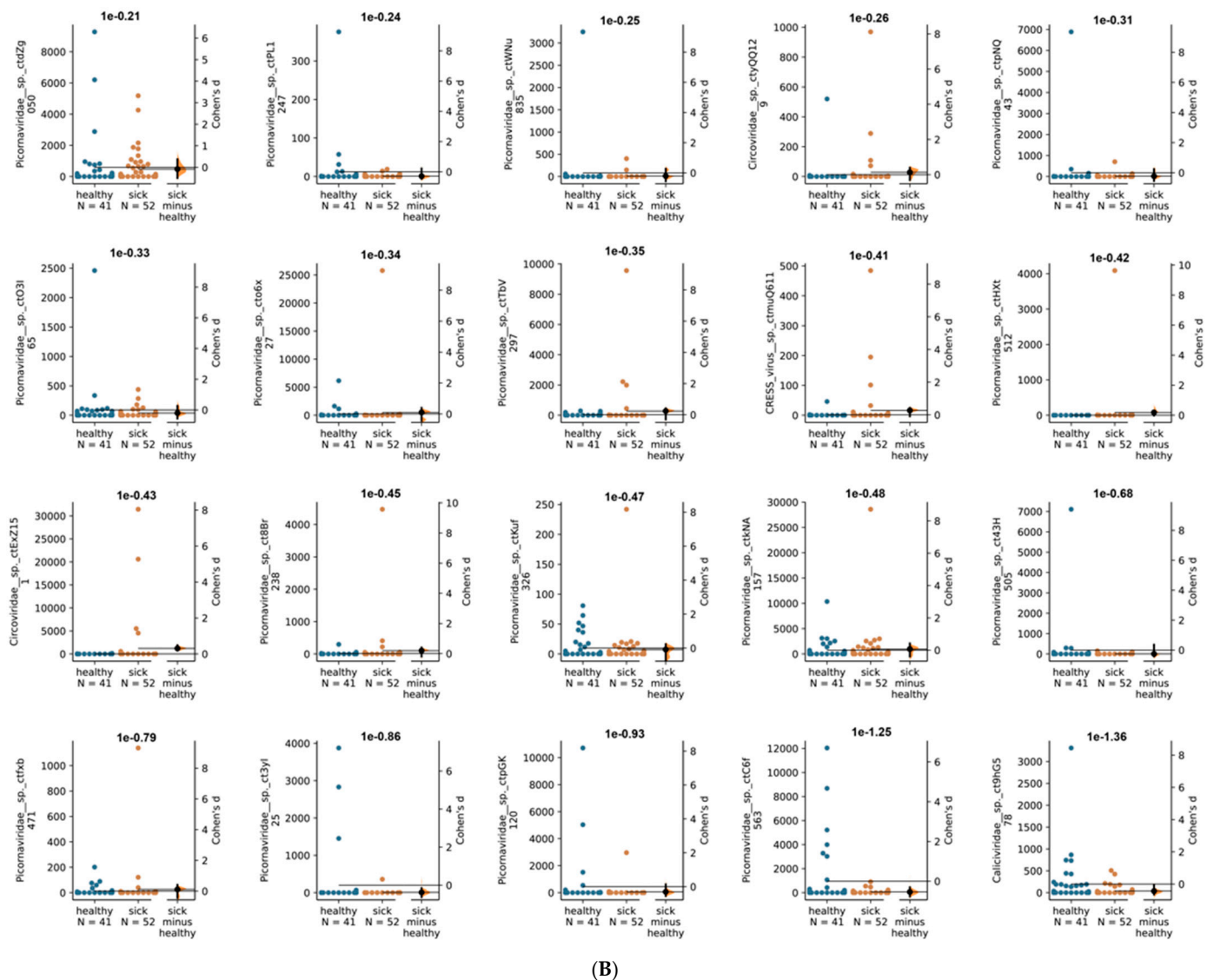
Figure 5. *Cont.*

(**B**)
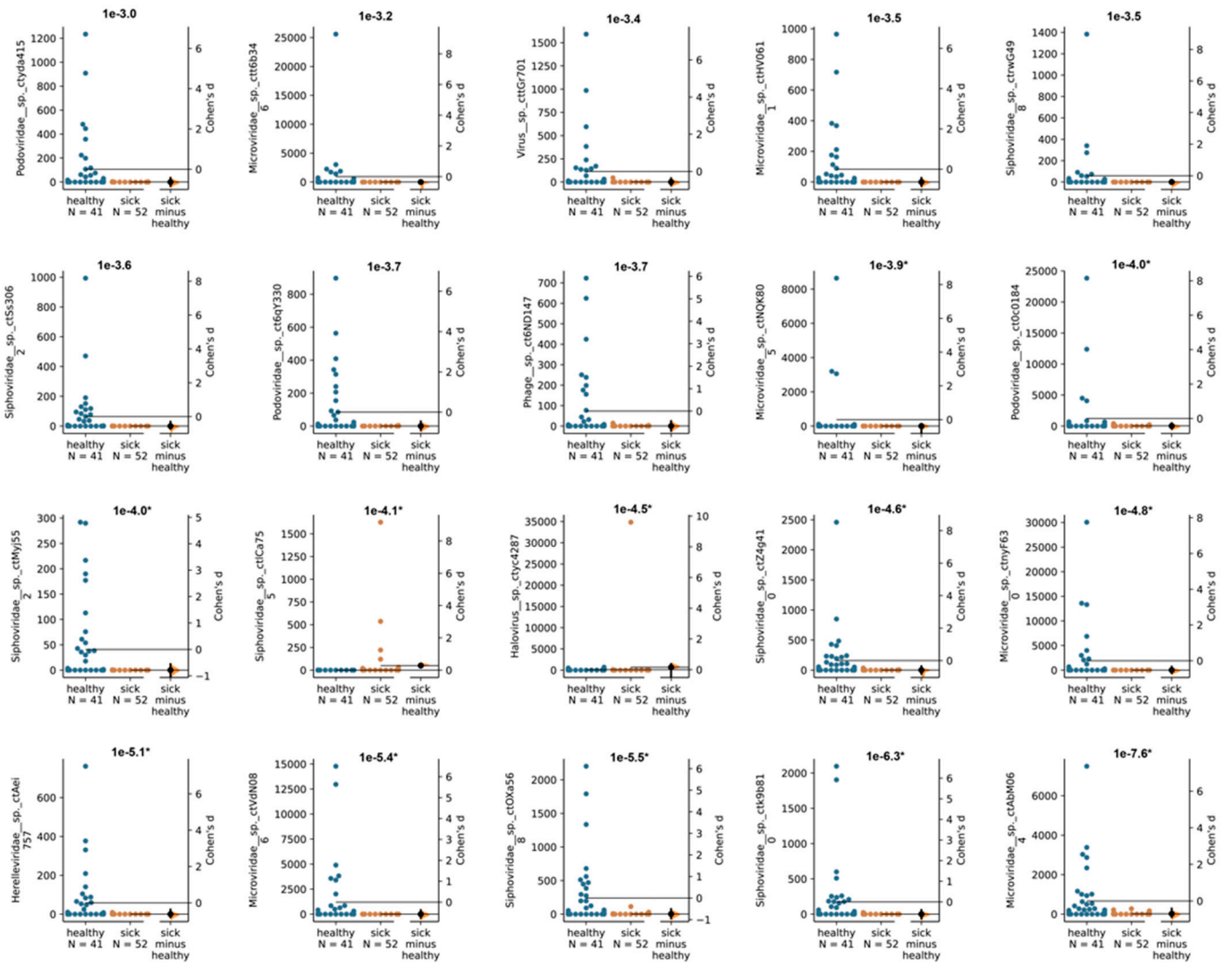
Figure 5. Sample-by-sample comparison of viruses with largest differences between healthy and ICD cohorts. Each plot represents a different virus contig, with RKPM values from healthy and ICD (orange) samples displayed. Cohen's d effect size values are plotted on the right side of each plot, and uncorrected *p* values are reported above each plot. Asterisks next to −log10p value on top of each graph were given to virus contigs with significant *p* values after multiple testing correction. (**A**) 20 virus contigs with the lowest *p* value overall. (**B**) Eukaryotic virus contigs with lowest *p* value. Note that most picornaviruses contigs retrieved from de novo assembly represented sub-genomic fragments with >90% average nucleotide identity to previously discovered macaque viruses.

4. Discussion

We used sequence similarity of NGS reads derived from viral metagenomics to identify viruses in fecal swabs to test the hypothesis that a specific mammalian virus was associated with ICD. The frequencies of reads matching different viruses were then used as surrogates for viral loads. Comparing the viruses in rhesus macaques with ICD versus those in healthy controls, we were unable to associate any virus with ICD. These results differ from the conclusions of a previous analysis, in which we showed that some enterovirus genotypes were weakly associated with ICD [17]. The ICD fecal swabs samples that were analyzed in this current study were collected during necropsy from severe cases that required medical cull while feces analyzed in the prior study were feces collected from cage pans while the animals were still under treatment. Therefore, it is possible that during the later stage of ICD studied here, shedding of enteric viruses in watery diarrhea is reduced due to more

extensive damage to the lining of the gut and consequent reduction in viral target cells. Earlier sampling of ICD-associated enteric virome may therefore be necessary to identify viruses associated with this condition. In the interval between studies the treatment of ICD cases at the CNPRC also changed with the addition of probiotic and exclusion of antibiotic treatments significantly reducing the pathological outcomes of ICD.

We observed that the median level of viral shedding was generally higher in healthy than in ICD cases (Figures 1 and 2), a result consistent with a reduction in viral target cells in late stage diarrheic animals. In the analysis of virus contigs predicted by Cenote-Taker 2 (Figures 4 and 5), it was also observed that fewer reads to specific phages were generally found in swabs from ICD cases versus healthy animals. Such a reduction in phage shedding may reflect a gut content with fewer bacterial hosts as its content is emptied due to chronic diarrhea.

A single viral cause for ICD is therefore unsupported by our analysis. It remains possible that one or more of the enteric viruses identified here can cause ICD but that the outcome of such infections is highly dependent of hosts factors. A prior study showed that chronic diarrhea disease in rhesus macaques was strongly associated with familial pedigrees indicating an important genetic component for this disease [25]. It is also possible that placental antibodies attenuated enteric infections in infants of previously infected mothers. Wide diversity in the outcome of infections may explain our inability to identify a virus associated with ICD in this rhesus macaque population.

References

1. Prongay, K.; Park, B.; Murphy, S.J. Risk factor analysis may provide clues to diarrhea prevention in outdoor-housed rhesus macaques (Macaca mulatta). *Am. J. Primatol.* **2013**, *75*, 872–882. [CrossRef]
2. Sestak, K.; Merritt, C.K.; Borda, J.; Saylor, E.; Schwamberger, S.R.; Cogswell, F.; Didier, E.S.; Didier, P.J.; Plauche, G.; Bohm, R.P.; et al. Infectious agent and immune response characteristics of chronic enterocolitis in captive rhesus macaques. *Infect. Immun.* **2003**, *71*, 4079–4086. [CrossRef] [PubMed]
3. Ingle, S.B.; Adgaonkar, B.D.; Ingle, C.R.H. Microscopic colitis: Common cause of unexplained nonbloody diarrhea. *World J. Gastrointest. Pathophysiol.* **2014**, *5*, 48–53. [CrossRef]
4. Elmore, D.B.; Anderson, J.H.; Hird, D.W.; Sanders, K.D.; Lerche, N.W. Diarrhea rates and risk factors for developing chronic diarrhea in infant and juvenile rhesus monkeys. *Lab. Anim. Sci.* **1992**, *42*, 356–359.
5. Blackwood, R.S.; Tarara, R.P.; Christe, K.L.; Spinner, A.; Lerche, N.W. Effects of the macrolide drug tylosin on chronic diarrhea in rhesus macaques (Macaca mulatta). *Comp. Med.* **2008**, *58*, 81–87.
6. Ardeshir, A.; Oslund, K.L.; Ventimiglia, F.; Yee, J.; Lerche, N.W.; Hyde, D.M. Idiopathic microscopic colitis of rhesus macaques: Quantitative assessment of colonic mucosa. *Anat. Rec.* **2013**, *296*, 1169–1179. [CrossRef] [PubMed]
7. Ardeshir, A.; Sankaran, S.; Oslund, K.L.; Hartigan-O'Connor, D.J.; Lerche, N.W.; Hyde, D.M.; Dandekar, S. Inulin treatment leads to changes in intestinal microbiota and resolution of idiopathic chronic diarrhea in rhesus macaques. *Ann. Am. Thorac. Soc.* **2014**, *11* (Suppl. S1), S75. [CrossRef]
8. Broadhurst, M.J.; Ardeshir, A.; Kanwar, B.; Mirpuri, J.; Gundra, U.M.; Leung, J.M.; Wiens, K.E.; Vujkovic-Cvijin, I.; Kim, C.C.; Yarovinsky, F.; et al. Therapeutic helminth infection of macaques with idiopathic chronic diarrhea alters the inflammatory signature and mucosal microbiota of the colon. *PLoS Pathog.* **2012**, *8*, e1003000. [CrossRef]
9. Tamboli, C.P.; Neut, C.; Desreumaux, P.; Colombel, J.F. Dysbiosis in inflammatory bowel disease. *Gut* **2004**, *53*, 1–4. [CrossRef] [PubMed]
10. Round, J.L.; Mazmanian, S.K. The gut microbiota shapes intestinal immune responses during health and disease. *Nat. Rev. Immunol.* **2009**, *9*, 313–323. [CrossRef]
11. Russell, R.G.; Rosenkranz, S.L.; Lee, L.A.; Howard, H.; DiGiacomo, R.F.; Bronsdon, M.A.; Blakley, G.A.; Tsai, C.C.; Morton, W.R. Epidemiology and etiology of diarrhea in colony-born Macaca nemestrina. *Lab. Anim. Sci.* **1987**, *37*, 309–316.
12. Wang, Y.; Tu, X.; Humphrey, C.; McClure, H.; Jiang, X.; Qin, C.; Glass, R.I.; Jiang, B. Detection of viral agents in fecal specimens of monkeys with diarrhea. *J. Med. Primatol.* **2007**, *36*, 101–107. [CrossRef] [PubMed]
13. Stuker, G.; Oshiro, L.S.; Schmidt, N.J.; Holmberg, C.A.; Anderson, J.H.; Glaser, C.A.; Henrickson, R.V. Virus detection in monkeys with diarrhea: The association of adenoviruses with diarrhea and the possible role of rotaviruses. *Lab. Anim. Sci.* **1979**, *29*, 610–616. [PubMed]
14. Wang, K.Y.; Christe, K.L.; Yee, J.; Roberts, J.A.; Ardeshir, A. Rotavirus is associated with decompensated diarrhea among young rhesus macaques (Macaca mulatta). *Am. J. Primatol.* **2019**, *81*, e22948. [CrossRef]
15. Sestak, K.; Feely, S.; Fey, B.; Dufour, J.; Hargitt, E.; Alvarez, X.; Pahar, B.; Gregoricus, N.; Vinje, J.; Farkas, T. Experimental inoculation of juvenile rhesus macaques with primate enteric caliciviruses. *PLoS ONE* **2012**, *7*, e37973. [CrossRef]
16. Handley, S.A.; Thackray, L.B.; Zhao, G.; Presti, R.; Miller, A.D.; Droit, L.; Abbink, P.; Maxfield, L.F.; Kambal, A.; Duan, E.; et al. Pathogenic simian immunodeficiency virus infection is associated with expansion of the enteric virome. *Cell* **2012**, *151*, 253–266. [CrossRef]
17. Kapusinszky, B.; Ardeshir, A.; Mulvaney, U.; Deng, X.T.; Delwart, E. Case-control comparison of enteric viromes in captive rhesus macaques with acute or idiopathic chronic diarrhea. *J. Virol.* **2017**, *91*, e00952-17. [CrossRef]
18. Ye, J.; McGinnis, S.; Madden, T.L. BLAST: Improvements for better sequence analysis. *Nucleic Acids Res.* **2006**, *34*, W6–W9. [CrossRef]
19. Deng, X.; Naccache, S.N.; Ng, T.; Federman, S.; Li, L.; Chiu, C.Y.; Delwart, E.L. An ensemble strategy that significantly improves de novo assembly of microbial genomes from metagenomic next-generation sequencing data. *Nucleic Acids Res.* **2015**, *43*, e46. [CrossRef]
20. Tisza, M.J.; Pastrana, D.V.; Welch, N.L.; Stewart, B.; Peretti, A.; Starrett, G.J.; Pang, Y.S.; Krishnamurthy, S.R.; Pesavento, P.A.; McDermott, D.H.; et al. Discovery of several thousand highly diverse circular DNA viruses. *Elife* **2020**, *9*, e51971. [CrossRef]
21. Tisza, M.J.; Belford, A.K.; Domínguez-Huerta, G.; Bolduc, B.; Buck, C.B. Cenote-Taker 2 democratizes virus discovery and sequence annotation. *Virus Evol.* **2021**, *7*, veaa100. [CrossRef] [PubMed]
22. Delcher, A.L.; Salzberg, S.L.; Phillippy, A.M. Using MUMmer to identify similar regions in large sequence sets. *Curr. Protoc. Bioinform.* **2003**, *10*, 10–13. [CrossRef]
23. Tarailo-Graovac, M.; Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinform.* **2009**, *4*, 4–10. [CrossRef]
24. Ho, J.; Tumkaya, T.; Aryal, S.; Choi, H.; Claridge-Chang, A. Moving beyond *p* values: Data analysis with estimation graphics. *Nat. Methods* **2019**, *16*, 565–566. [CrossRef]
25. Kanthaswamy, S.; Elfenbein, H.A.; Ardeshir, A.; Ng, J.; Hyde, D.; Smith, D.G.; Lerche, N. Familial aggregation of chronic diarrhea disease (CDD) in rhesus macaques (Macaca mulatta). *Am. J. Primatol.* **2014**, *76*, 262–270. [CrossRef]