Virology Journal

Virome of respiratory secretion from children with unknown etiological acute respiratory disease revealed recombinant human parechovirus and other significant viruses

Ying Liu[1†], Hao Wang[2†], Jie Yang[3], Jian Zeng[3] and Guang-Ming Sun[1*]

Abstract

Using viral metagenomics, viral nucleic acid in 30 respiratory secretion samples collected from children with unknown etiological acute respiratory disease were investigated. Sequences showing similarity to human parainfluenza virus 1, anellovirus, bocavirus, coxsackievirus A4, human parechovirus (HPeV), and alphaflexivirus were recovered from these samples. Complete genomes of one anellovirus, one coxsackievirus A4, three parechoviruses were determined from these libraries. The anellovirus (MW267851) phylogenetically clustered with an unpublished anellovirus (MK212032) from respiratory sample of a Vietnamese patient, forming a separate branch neighboring to strains within the genus *Betatorquevirus*. The genome of coxsackievirus A4 (MW267852) shares the highest sequence identity of 96.4% to a coxsackievirus A4 (MN964079) which was identified in clinical samples from children with Hand, Foot, and Mouth Disease (HFMD). Two (MW267853 and MW267854) of the three parechoviruses belong to HPeV-1 and the other one (MW267855) belongs to HPeV-6. Recombination analysis indicated that an HPeV-1 (MW267854) identified in this study is a putative recombinant occurred between HPeV-1 and HPeV-3. Whether these viruses have association with specific respiratory disease calls for further investigation.

Keywords: Respiratory secretion, Viral metagenomics, Parechovirus, Anellovirus, Coxsackievirus

Background

Globally, acute respiratory disease is the leading cause of morbidity and mortality among children under five years of age [1, 2] and viruses are the main causes of acute respiratory infections with the potential to cause pandemics. The ongoing coronavirus disease 2019 (COVID-19) pandemic emphasizes the need to actively study the virome of unknown etiological acute respiratory disease. Despite intensive laboratory investigations, the etiology of most acute respiratory infections is unknown, which makes clinical management difficult [3]. Viral metagenomics is a is an unbiased virus-detecting technique that is increasingly applied to non-specifically discover both already known and highly divergent viruses [4–6].

Here, using the viral metagenomic technique (Fig. 1), we investigated the virome of respiratory specimens collected from children presenting with acute respiratory infections and fully characterized the complete genomes of human parechovirus, anellovirus, and coxsackievirus identified from these samples.

Findings

During 2019, 30 nasal-throat swab samples were collected from 30 children with age of < 5 years old with acute respiratory symptoms, whose pathogens (including
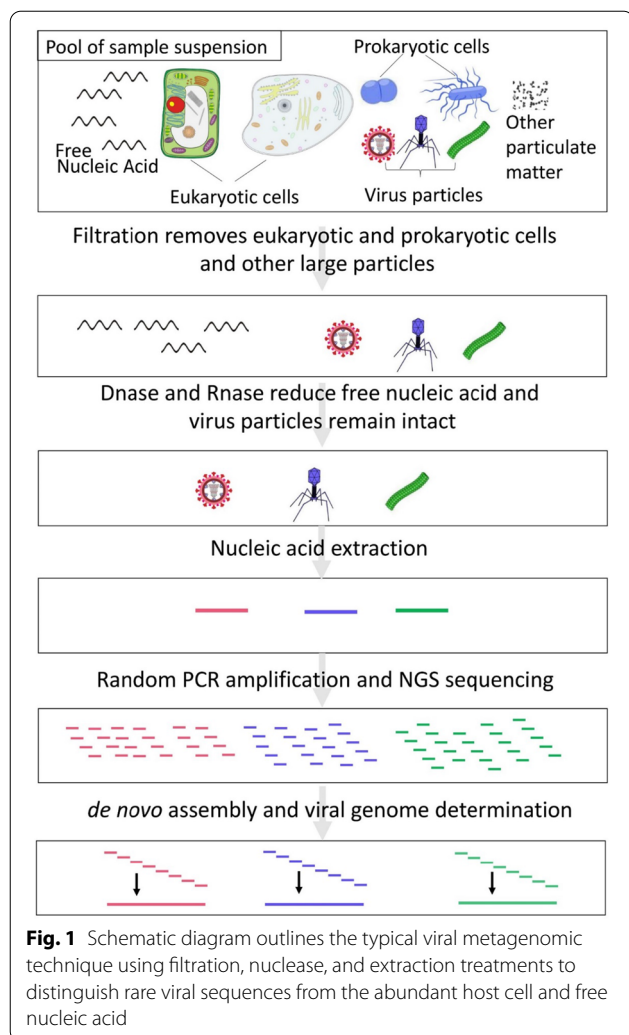
*Correspondence: sgmxhx@163.com
†Ying Liu and Hao Wang have contributed equally to this work
[1] Department of Clinical Laboratory, Xuzhou Central Hospital, 199 Jiefangnan Road, Xuzhou 221009, Jiangsu, China
Full list of author information is available at the end of the article

Liu *et al. Virol J*   (2021) 18:122

Page 2 of 6

Fig. 1 Schematic diagram outlines the typical viral metagenomic technique using filtration, nuclease, and extraction treatments to distinguish rare viral sequences from the abundant host cell and free nucleic acid
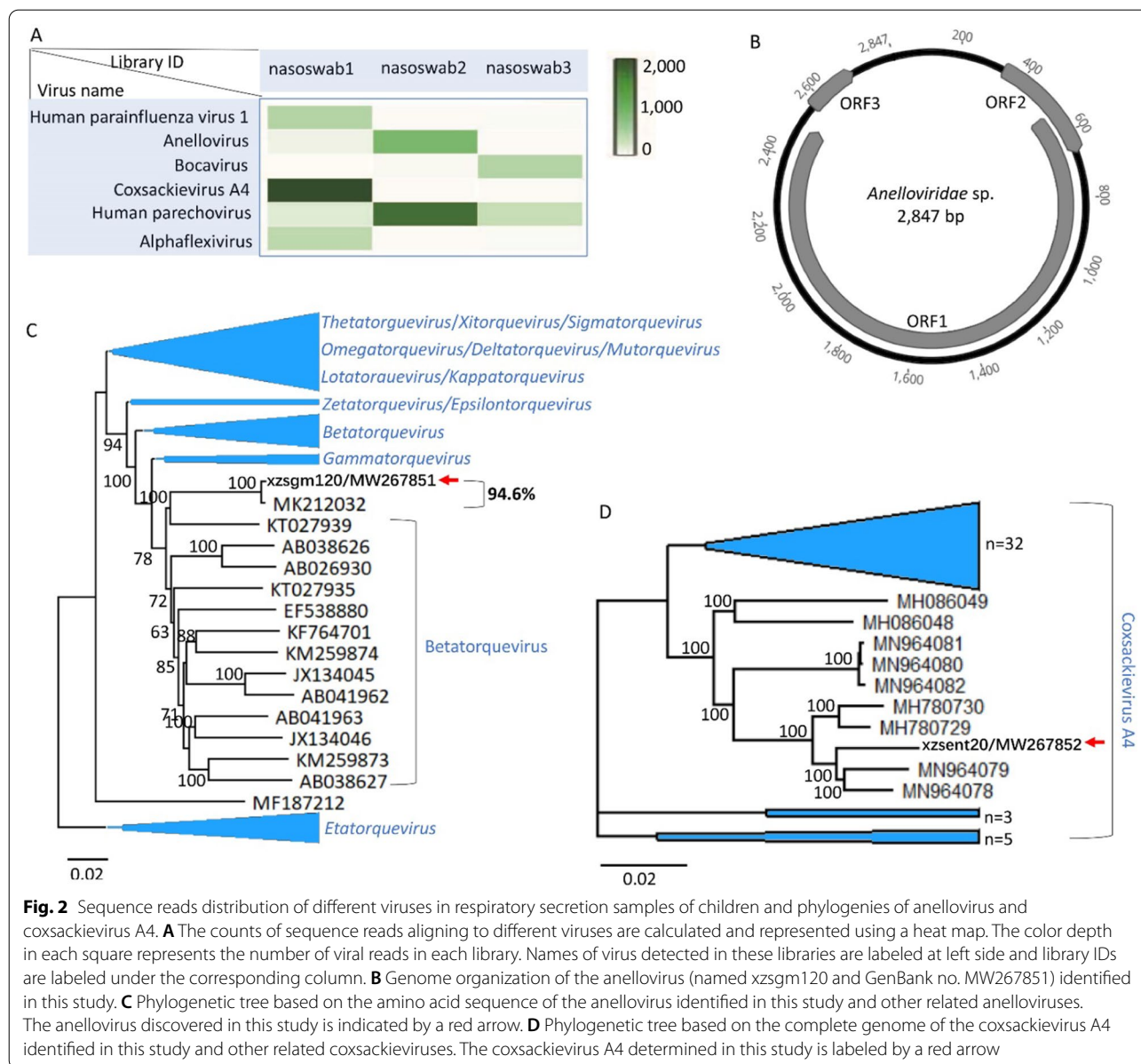
virus and bacteria) were not identified at the Division of Clinical Microbiology of Xuzhou Center Hospital. Ethical Approval was given by the Ethics Committee of Xuzhou Central Hospital with the reference number as xzzx2018090. These samples were then subject to viral metagenomic analysis to investigate the virome. Briefly, tips of the swabs were immersed into 0.5 mL PBS and vigorously vortexed for 5 min and incubated for 30 min at 4 °C. The supernatants were then collected after centrifugation (10 min, 15,000×*g*) and randomly pooled into 3 sample pools each including 10 samples. Viral nucleic acid in the filtered Supernatant pool was then isolated using QIAamp MinElute Virus Spin Kit (Qiagen) according to the manufacturer's protocol. Three libraries were then constructed using Nextera XT DNA Sample Preparation Kit (Illumina) and sequenced using the MiSeqIllumina platform with 250 bases paired ends with dual barcoding for each pool. For bioinformatics analysis,

paired-end reads of 250 bp generated by MiSeq were debarcoded using vendor software from Illumina. An in-house analysis pipeline running on a 32-nodes Linux cluster was used to process the data. Clonal reads were removed and low sequencing quality tails were trimmed using Phred quality score ten as the threshold. Adaptors were trimmed using the default parameters of VecScreen which is NCBI BLASTn with specialized parameters designed for adapter removal. The cleaned reads were de-novo assembled by SOAPdenovo2 version r240 using Kmer size 63 with default settings. The assembled contigs, along with singlets were aligned to an in-house viral proteome database using BLASTx with an E-value cutoff of $< 10^{-5}$ [4, 7].

For phylogenetic analysis, virus nucleotide sequences (for coxsackievirus and HPeV) or deduced amino acid sequence of ORF1 (for anellovirus) were aligned using CLUSTAL W with the default settings [8]. Phylogenetic trees with 1,000 bootstrap resamples of the alignment data sets was generated using the Neighbour-Joining (N-J) method in MEGA7.0. Bootstrap values for each node were given [9]. For recombination analysis of HPeV, related genomes were retrieved from GenBank and subjected to multiple sequence alignment using MUSCLE in Mega 7.0. The recombination events were first assessed using RDP4.0 [10]. Finding putative recombination events where the genome identified here were involved, the related genomes were selected and realigned and recombination was confirmed by using similarity plot analysis in SimPlot software V. 3.5.1 [11]. Phylogenetic analysis was also performed based on recombinant regions to confirm the recombination event.

Next-generation sequencing (NGS) results indicated that the 3 libraries of 30 respiratory secretion samples generated a total of 1,278,128 sequence reads, where the 3 libraries contained 818,736, 330,598, and 128,794 sequence reads, respectively. BLASTx searching results indicated viral sequences showing similarity to human parainfluenza virus 1, anellovirus, bocavirus, coxsackievirus, human parechovirus (HPeV), and alphaflexivirus were detected in these samples (Fig. 2A), where sequence reads aligning to different viruses were counted and their log10 transformed values are represented using heat map (Fig. 2A). Sequence reads from the same species of virus were assembled within each barcode which generated five different complete virus genomes, including one anellovirus, one coxsackievirus and 3 HPeV genomes.

The circular genome of the anellovirus (named xzsgm120 from library nasoswab2) is 2847 nt long, of which the genome organization is consistent with those of other anelloviruses, including three ORFs (Fig. 2B). The ORF1, the largest ORF in this anellovirus, encodes a 664 amino acid long putative capsid protein. To

Fig. 2 Sequence reads distribution of different viruses in respiratory secretion samples of children and phylogenies of anellovirus and coxsackievirus A4. **A** The counts of sequence reads aligning to different viruses are calculated and represented using a heat map. The color depth in each square represents the number of viral reads in each library. Names of virus detected in these libraries are labeled at left side and library IDs are labeled under the corresponding column. **B** Genome organization of the anellovirus (named xzsgm120 and GenBank no. MW267851) identified in this study. **C** Phylogenetic tree based on the amino acid sequence of the anellovirus identified in this study and other related anelloviruses. The anellovirus discovered in this study is indicated by a red arrow. **D** Phylogenetic tree based on the complete genome of the coxsackievirus A4 identified in this study and other related coxsackieviruses. The coxsackievirus A4 determined in this study is labeled by a red arrow

determine the relationships of xzsgm120 to other anelloviruses, a phylogram was created based on the amino acid sequence encoded by ORF1. The Neighbour-Joining (N-J) tree (with 1,000 bootstrap resamples) based on the amino acid sequence of ORF1 indicated the anellovirus identified here closely clustered with an anellovirus (GenBank MK212032) from a respiratory sample of a Vietnamese patient according to the annotation in GenBank, forming a separate branch neighboring to those strains from the genus *Betatorquevirus* where they shared 94.6% sequence identity based on the complete genome sequence (Fig. 2C).

One complete genome of coxsackievirus (xzsent20) was assembled and well grouped into the cluster of coxsackievirus, sharing the highest nucleotide sequence identity of 96.4% to a coxsackievirus A4 (CV-A4) (MN964079) which was identified from children with Hand, Foot, and Mouth Disease (HFMD) [12].

Three HPeV genomes were generated from the three libraries, respectively. Phylogenetic analysis based on the available complete genomes of 193 HPeVs in GenBank together with the 3 HPeV genomes determined here indicated two (xzsgm20 and xzsgm37) of HPeVs identified in this study were grouped into the cluster of HPeV-1 and the other one (xzsgm13) was closely related

Liu *et al. Virol J*   (2021) 18:122

Page 4 of 6

to HPeV-6, sharing 90.2–95.4% sequence identities based on the complete genome to their best BLASTn matches in GenBank (Fig. 3A). Recombination analysis using RDP4.0 software based on these 196 complete genomes suggested that genomic recombination occurred in one of the HPeV-1 (xzsgm37), where xzsgm37 seems to be a putative recombinant sequence produced by recombination event occurred between two parental lineages represented by an HPeV-1 strain (MH933781) and an HPeV-3 strain (GQ183029), respectively (Fig. 3B). The recombination event was further confirmed by phylogenetic trees based on upstream and downstream sequences of the putative breakpoint, respectively (Fig. 3C, D). To exclude the unnatural recombination due to assembly error during sequence assembly, primers covering the putative breakpoint were used to amplify the sequence fragment of xzsgm37 using cDNA products of library nasoswab01, and the amplified band was subjected to Sanger sequencing, which resulted sequence identical to the original

genome sequence, suggesting this HPeV-1 strain is a natural recombinant.

Discussion

Anellovirus is a common virus that can be detected in different tissues and organs of diverse species of animals [13–15]. An important feature of these viruses is the persistent infection in the host [16]. So far, there is no evidence showing a clear link between anellovirus and certain diseases. In the past decade, lots of studies have found that there is a positive correlation between anellovirus and the host immunosuppression. It has been suggested that the titer of anellovirus in plasma could be used as an indicator of immune recovery [14, 17, 18]. In the present study, anellovirus was detected in the respiratory tract of children with acute respiratory symptoms without known pathogens. Phylogenetic analysis revealed that this anellovirus showed a close relationship to another anellovirus which was also
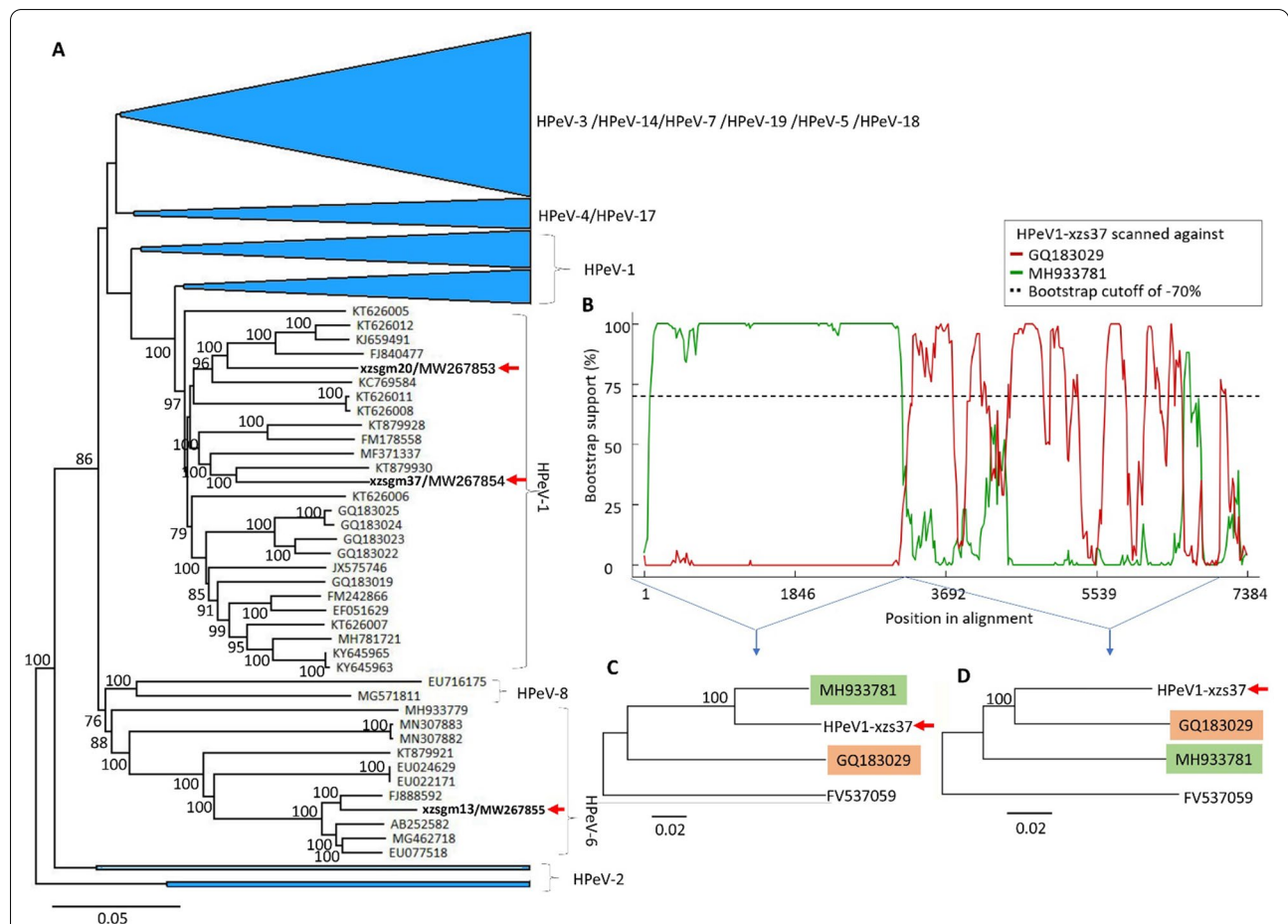

Fig. 3 Phylogeny and recombination of HPeVs identified in respiratory secretion samples of children. **A** Phylogenetic tree based on the complete genome of the HPeVs identified in this study and other representative HPeV strains. **B** Bootscan evidence for the recombination of the HPeV-1 strain xzsgm37 (GenBank no. MW267854). **C** and **D** Phylogenetic trees respectively based on upstream and downstream sequences of the putative breakpoint in the recombination analysis. The HPeVs determined in this study are labeled by a red arrow

Liu *et al. Virol J*   (2021) 18:122

Page 5 of 6

detected in respiratory samples, suggesting this anellovirus may have association with respiratory disease.

HPeV is a kind of nonenveloped virus with a single-strand positive RNA genome about 7.35 kb in length, including a single ORF [19]. HPeV is wildly prevalent in children population, although most of its infections are asymptomatic, it can be detected in samples from a variety of children's diseases, including skin rash, encephalitis, meningitis, sepsis, and even severe dilated cardiomyopathy [19–21]. HPeVs are now subdivided into > 16 different genotypes, among which some were associated with certain specific diseases [21, 22]. In this study, HPeVs were detected in the respiratory tract of children with acute respiratory disease, which included two different genotypes, HPeV-1, and HPeV-6. HPeV-1 is widely prevalent throughout the world and is often found in children with diarrhea and gastroenteritis [23]. HPeV-6 was first isolated from a cerebrospinal fluid specimen of a 1-year-old girl with Reye syndrome [24]. Detecting HPeV-6 in respiratory tract of children suggested HPeV-6 may transmit through respiratory tract and may cause central nervous system infection.

CV-A4 is classified as human enterovirus A (HEV-A) based on its serotype and is an etiological agent of HFMD. The virus can be detected in throat swabs from herpangina patients and can cause severe central nervous system symptoms [25, 26]. Our data indicated that a complete genome belonging to CV-A4 was present in the respiratory tracts of children with an acute respiratory symptom, suggesting CV-A4 may cause acute respiratory symptom and have potential of causing further infection in central nervous system.

Conclusion

Taken together, we investigated the virome of 30 respiratory secretion samples collected from children with unknown etiological acute respiratory disease and fully characterized five genomes belonging to anellovirus, human parechovirus, and coxsackievirus. Whether these viruses have an association with acute respiratory disease needs further study with a larger sample size and healthy control cohort.

Abbreviations
BLAST: Basic local alignment search tool; NGS: Next-generation sequencing; HPeV: Human parechovirus; NCBI: National Center for Biotechnology Information; ORF: Open reading frame; CV-A4: Coxsackievirus A4; HFMD: Hand, foot, and mouth disease.

Acknowledgements
Not applicable.

Declarations

Ethics approval and consent to participate
Patients signed informed consent forms, and protocols were approved by Medical Ethical Committee at the Xuzhou Central Hospital (Reference code: chxuzhou2017032x).

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Clinical Laboratory, Xuzhou Central Hospital, 199 Jiefangnan Road, Xuzhou 221009, Jiangsu, China. [2]Department of Clinical Laboratory, The Affiliated Huai'an Hospital of Xuzhou Medical University, Huai'an 223002, Jiangsu, China. [3]School of Medicine, Jiangsu University, Zhenjiang 212013, Jiangsu, China.

References
1. Williams BG, Gouws E, Boschi-Pinto C, Bryce J, Dye C. Estimates of worldwide distribution of child deaths from acute respiratory infections. Lancet Infect Dis. 2002;2:25–32.
2. GBD Chronic Respiratory Disease Collaborators JB, Kendrick PJ, Paulson KR, Gupta V, Abrams EM, Adedoyin RA, et al. Prevalence and attributable health burden of chronic respiratory diseases, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Respir Med. 2020;8:585–96. https://doi.org/10.1016/S2213-2600(20)30105-3.
3. Shi T, Arnott A, Semogas I, Falsey AR, Openshaw P, Wedzicha JA, et al. The etiological role of common respiratory viruses in acute respiratory infections in older adults: a systematic review and meta-analysis. J Infect Dis Oxford Acad. 2020;222:S563–9.
4. Zhang W, Yang S, Shan T, Hou R, Liu Z, Li W, et al. Virome comparisons in wild-diseased and healthy captive giant pandas. Microbiome. 2017;5:90.
5. Siqueira JD, Dominguez-Bello MG, Contreras M, Lander O, Caballero-Arias H, Xutao D, et al. Complex virome in feces from Amerindian children in isolated Amazonian villages. Nat Commun. 2018;9:4270.
6. Yang S, Shan T, Xiao Y, Zhang H, Wang X, Shen Q, et al. Digging metagenomic data of pangolins revealed SARS-CoV-2 related viruses and other significant viruses. J Med Virol. 2021;93:1786–91.
7. Wang H, Ling Y, Shan T, Yang S, Xu H, Deng X, et al. Gut virome of mammals and birds reveals high genetic diversity of the family Microviridae. Virus Evol. 2019;5:vez013.
8. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23:2947–8.

Liu *et al. Virol J*   (2021) 18:122

Page 6 of 6

9. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33:1870–4.
10. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. 2015;1:vev003.
11. Thoi TC, Than VT, Kim W. Whole genomic characterization of a Korean human parechovirus type 1 (HPeV1) identifies recombination events. J Med Virol. 2014;86:2084–91.
12. Guo W-P, Chen G-Q, Xie G-C, Du L-Y, Tang Q. Mosaic genome of human coxsackievirus A4 associated with herpangina and HFMD in Yancheng, China, 2016 and 2018. Int J Infect Dis. 2020;96:538–40.
13. De Vlaminck I, Khush KK, Strehl C, Kohli B, Luikart H, Neff NF, et al. Temporal response of the human virome to immunosuppression and antiviral therapy. Cell. 2013;155:1178–87.
14. Li L, Deng X, Linsuwanon P, Bangsberg D, Bwana MB, Hunt P, et al. AIDS alters the commensal plasma virome. J Virol. 2013;87:10912–5.
15. Young JC, Chehoud C, Bittinger K, Bailey A, Diamond JM, Cantu E, et al. Viral metagenomics reveal blooms of anelloviruses in the respiratory tract of lung transplant recipients. Am J Transplant. 2015;15:200–9.
16. Legoff J, Resche-Rigon M, Bouquet J, Robin M, Naccache SN, Mercier-Delarue S, et al. The eukaryotic gut virome in hematopoietic stem cell transplantation: new clues in enteric graft-versus-host disease. Nat Med. 2017;23:1080–5.
17. Spandole S, Cimponeriu D, Berca LM, Mihăescu G. Human anelloviruses: an update of molecular, epidemiological and clinical aspects. Arch Virol. 2015;160:893–908.
18. Wang X-C, Wang H, Tan S-D, Yang S-X, Shi X-F, Zhang W. Viral metagenomics reveals diverse anelloviruses in bone marrow specimens from hematologic patients. J Clin Virol. 2020;132:104643.
19. Chieochansin T, Vichiwattana P, Korkong S, Theamboonlers A, Poovorawan Y. Molecular epidemiology, genome characterization, and recombination event of human parechovirus. Virology. 2011;421:159–66.
20. Joki-Korpela P, Hyypiä T. Parechoviruses, a novel group of human picornaviruses. Ann Med. 2001;33:466–71.
21. Benschop K, Thomas X, Serpenti C, Molenkamp R, Wolthers K. High prevalence of human Parechovirus (HPeV) genotypes in the Amsterdam region and identification of specific HPeV variants by direct genotyping of stool samples. J Clin Microbiol. 2008;46:3965–70.
22. Zhao X, Shi Y, Xia Y. Genome analysis revealed novel genotypes and recombination of the human parechoviruses prevalent in children in Eastern China. Gut Pathog. 2016;8:52.
23. Chen H, Yao Y, Liu X, Xiao N, Xiao Y, Huang Y, et al. Molecular detection of human parechovirus in children with acute gastroenteritis in Guangzhou, China. Arch Virol. 2014;159:971–7.
24. Watanabe K, Oie M, Higuchi M, Nishikawa M, Fujii M. Isolation and characterization of novel human parechovirus from clinical samples. Emerg Infect Dis. 2007;13:889–95.
25. Li J-S, Dong X-G, Qin M, Xie Z-P, Gao H-C, Yang J-Y, et al. Outbreak of febrile illness caused by coxsackievirus A4 in a nursery school in Beijing, China. Virol J. 2015;12:92.
26. Sarkar JK, Biswas ML, Chatterjee SN, Guha SK, Chakravarty SK. Coxsackie virus from blood of two cases of encephalitis. Indian J Med Res. 1966;54:905–9.

Publisher's Note