

SHORT REPORT

Open Access

# Detection of novel viruses in porcine fecal samples from China

Jie-mei Yu, Jin-song Li, Yuan-yun Ao and Zhao-jun Duan\*

## Abstract

**Background:** Pigs are well known source of human infectious disease. To better understand the spectrum of viruses present in pigs, we utilized the 454 Life Sciences GS-FLX high-throughput sequencing platform to sequence stool samples from healthy pigs.

**Findings:** Total nucleic acid was extracted from stool samples of healthy piglets and randomly amplified. The amplified materials were pooled and processed using a high-throughput pyrosequencing technique. The raw sequences were deconvoluted on the basis of the barcode and then processed through a standardized bioinformatics pipeline. The unique reads (348, 70 and 13) had limited similarity to known astroviruses, bocaviruses and parechoviruses. Specific primers were synthesized to assess the prevalence of the viruses in healthy piglets. Our results indicate extremely high rates of positivity.

**Conclusions:** Several novel astroviruses, bocaviruses and Ljungan-like viruses were identified in stool samples from healthy pigs. The rates of isolation for the new viruses were high. The high detection rate, diverse sequences and categories indicate that pigs are well-established reservoirs for and likely sources of different enteric viruses.

**Keywords:** 454 Pyrosequencing, Astroviruses, Bocavirus, Parechovirus

## The study

Many emerging viruses are of zoonotic origin and can cause epidemics in humans after overcoming the inter-species barrier through mutation or other genetic events, including recombination. The identification of previously unknown viruses in animal hosts is important for understanding their potential for cross-species transmission and emergence. Pigs are known to harbor a diverse array of viruses, including coronaviruses, astroviruses and kobuviruses [1-4], among which coronaviruses and astroviruses are well known viruses that can infect human. To better understand the spectrum of viruses present in pigs, we utilized the 454 Life Sciences GS-FLX high-throughput sequencing platform (454 Life Sciences/Roche, Branford, CT, USA), which has emerged as a non-biased, comprehensive and powerful tool for virus discovery in complex environmental samples [5,6], to sequence stool samples from healthy pigs.

With the cooperation of the Lulong County CDC, stool samples were collected from healthy piglets < 3

months of age from several farms and sporadically distributed families that raised pigs in Lulong County during 2006–2009. The samples were transported on dry ice and stored at  $-80^{\circ}\text{C}$  at the Chinese CDC.

The samples were screened for rotavirus, calicivirus and astrovirus by ELISA or routine PCR. Nine of the samples negative for these viruses were pooled, diluted (1:5 ratio, wt/vol) and then sequentially filtered through 0.45- and 0.22- $\mu\text{m}$  membranes. The pooled sample included two samples that were previously demonstrated to contain a novel porcine bocavirus (6V/7V CHN) that was partially sequenced in our lab [7]. Total nucleic acid was extracted from the pooled sample and cDNA was generated using random hexamers. Random PCR amplification [8] was performed and the amplified cDNA was used as the template for standard library construction and sequencing using Roche Genome Sequencer FLX Titanium pyrosequencing technology.

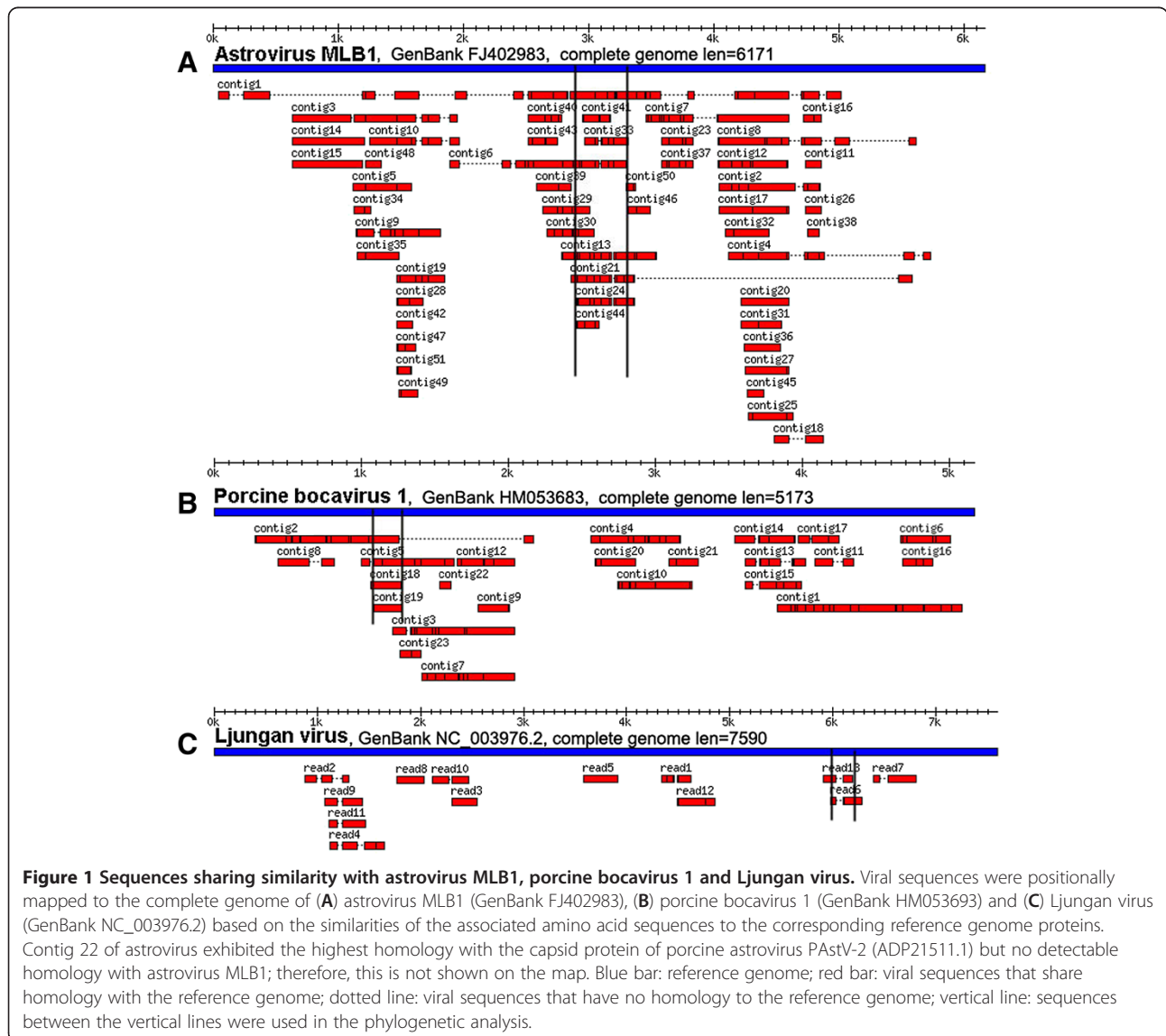
The initial pyrosequencing runs produced in excess of 145,000,000 bases of high-quality nucleotide sequence with an average read length of 360 bp. We used a customized informatics pipeline as described previously [9] with minor modifications to computationally identify

\* Correspondence: zhaojund@126.com  
National Institute for Viral Disease Control and Prevention, China CDC, Beijing 100052, China

viral sequences. Briefly, raw sequence reads were filtered to remove low-quality and repetitive sequences; BLASTn and BLASTx were used to identify sequences with similarity to known viruses in GenBank, then sequences identified as viral were further classified into viral families based on the taxonomy of the best hit. Sequence assembly was performed using Newbler (454 Life Sciences) with parameter be set as 90% nucleotide identity over 50 base pairs. Based on this analysis, 348, 70 and 13 unique reads had limited similarity to known astroviruses, bocaviruses and parechoviruses, respectively. Assembly of the raw reads for astrovirus and bocavirus generated 51 and 23 contigs, respectively. Due to the limited number of reads with similarity to parechoviruses, individual reads rather than contigs were analyzed. Most of the sequences detected were highly divergent compared to the most closely related viruses

(Figure 1 and Additional file 1: Table S1, Additional file 2: Table S2 and Additional file 3: Table S3 in the additional file). The reference genomes were selected to represent the top BLAST match for most of the sequences.

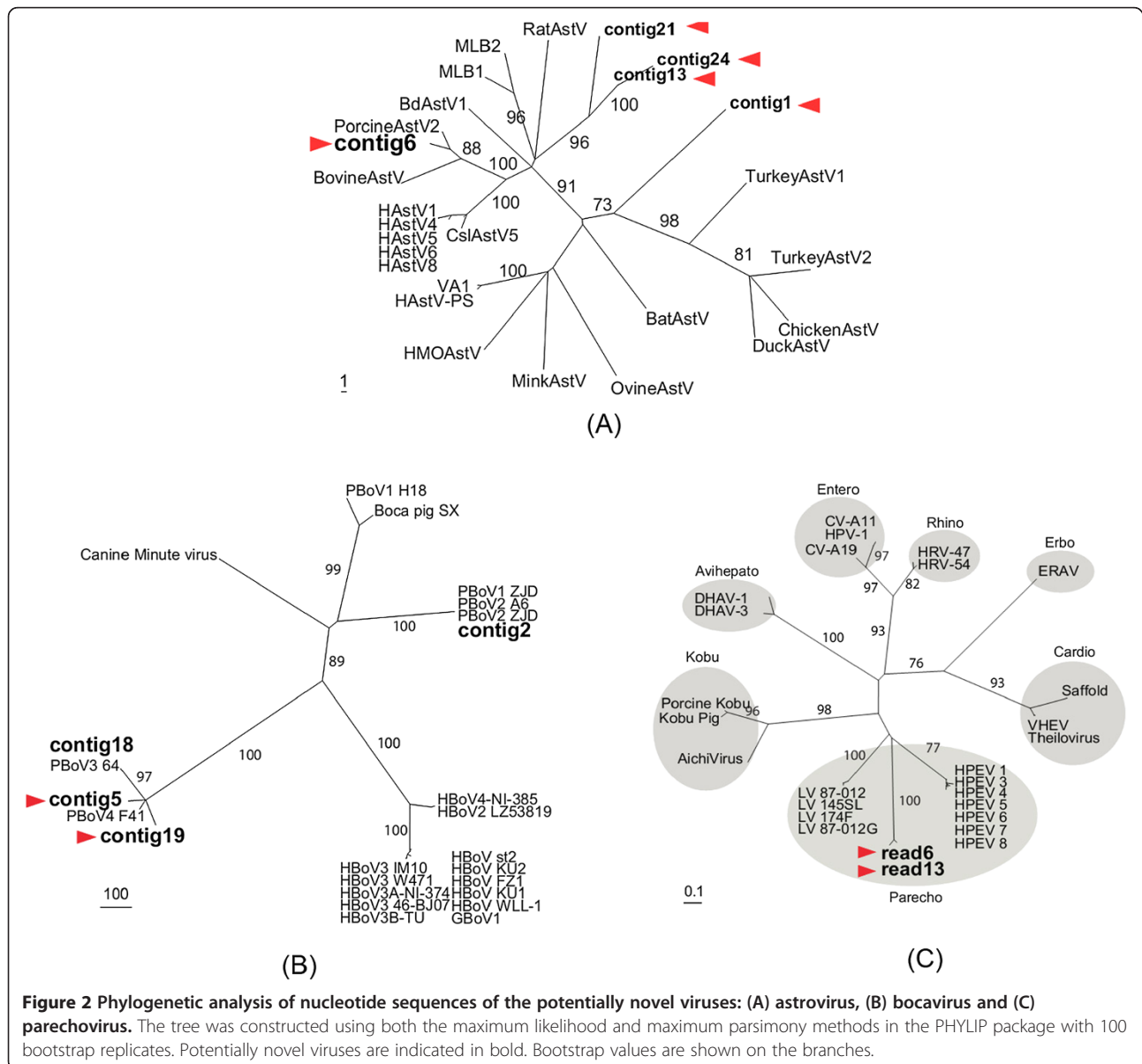
*Astroviruses* have been isolated from a number of host species, including humans, minks, sheep, pigs, chicken, ducks, turkeys [10-12] and more recently from marine mammals, dogs, cheetahs and bats [13-16]. They are generally associated with enteric diseases such as diarrhea and vomiting in a number of mammalian species [17]. The most abundant viral reads in this sequencing library were those from astroviruses. The size of the astrovirus contigs generated ranged from 150 to 5674 bp. These contigs shared 30.1–100% amino acid identity, typically over a small region, with the best hits in the NCBI nr database, which suggests that the sequences are highly diverged from those of known viruses.



**Figure 1 Sequences sharing similarity with astrovirus MLB1, porcine bocavirus 1 and Ljungan virus.** Viral sequences were positionally mapped to the complete genome of (A) astrovirus MLB1 (GenBank FJ402983), (B) porcine bocavirus 1 (GenBank HM053693) and (C) Ljungan virus (GenBank NC\_003976.2) based on the similarities of the associated amino acid sequences to the corresponding reference genome proteins. Contig 22 of astrovirus exhibited the highest homology with the capsid protein of porcine astrovirus PASTV-2 (ADP21511.1) but no detectable homology with astrovirus MLB1; therefore, this is not shown on the map. Blue bar: reference genome; red bar: viral sequences that share homology with the reference genome; dotted line: viral sequences that have no homology to the reference genome; vertical line: sequences between the vertical lines were used in the phylogenetic analysis.

Interestingly, multiple contigs mapped to the same region of the reference genome (GenBank FJ402983), suggesting the presence of multiple variants of astroviruses in the samples. For example, contigs 1, 6, 13, 21 and 24 mapped to the same subregion of ORF1b. A phylogenetic analysis of these five contigs was performed using both maximum likelihood and maximum parsimony methods in the PHYLIP package with 100 bootstrap replicates. The tree demonstrated that contig 1 was most closely related to avian astroviruses, contig 6 was closely related to porcine astrovirus and contigs 13, 21 and 24 formed a separate branch that was distantly related to all the other astroviruses (Figure 2a). The observed diversity suggests the existence of novel astroviruses in pigs.

Human bocavirus (HBoV) was first identified in respiratory samples from patients with respiratory infections in 2005 [18]. Subsequently, various bocaviruses have been found in respiratory and fecal specimens of human or animal origin, including HBoV 2–4, Gorilla bocavirus (GBoV1) and porcine bocavirus (PBoV1–4, 6V, and 7V) [19,20]. The sizes of the 23 bocavirus contigs in this study ranged from 147 to 1256 bp, and they shared 74.6–99.2% nucleotide sequence identities with the best hit in the NCBI nt database. Contigs 2, 5, 18 and 19 mapped to the same portion of the NS1 region of the reference genome but shared less than 80% identity among themselves at the nucleotide level. A phylogenetic analysis of partial NS1 sequences showed that



**Figure 2** Phylogenetic analysis of nucleotide sequences of the potentially novel viruses: (A) astrovirus, (B) bocavirus and (C) parechovirus. The tree was constructed using both the maximum likelihood and maximum parsimony methods in the PHYLIP package with 100 bootstrap replicates. Potentially novel viruses are indicated in bold. Bootstrap values are shown on the branches.

contig 2 was closely related to porcine bocavirus 1–2 while contig 18 was closely related to PBoV3 [20]. Contigs 5 and 19 shared 74.6 and 80.0% nucleotide identities with PBoV4, respectively (Figure 2b). When we compared the contigs to the sequences from porcine bocavirus 6V and 7V, one and two contigs, respectively, exhibited high nucleotide homologies (> 99%) indicating that we were able to recover these viruses once again with our methods. According to the most recent International Committee on Taxonomy of Viruses species demarcation criteria for the genus *Bocavirus* (at least 5% nucleotide sequence divergence in their nonstructural gene), our data indicate the existence of multiple new *Bocavirus* species in pigs.

*Parechovirus* is a genus in the family *Picornaviridae* that comprises two species: human parechovirus (HPEV) and Ljungan virus (LV). HPEV causes mild gastrointestinal or respiratory illness, and has been implicated in cases of myocarditis and encephalitis [21]. LV can cause significant morbidity and mortality in wild rodents as well as in laboratory animals, and it is a suspected human pathogen [22]. Bank voles infected with LV in captivity develop several different pathologic signs and symptoms, including myocarditis, diabetes and encephalitis [23]. Thirteen unique viral reads with similarities to both HPEV and LV were identified in this study. They spanned the LV reference genome (GenBank NC\_003976.2), to which the

amino acid identities ranged from 22.3 to 55.8%. Based on the sequences, we infer the existence of at least two parecho-like viruses. Reads 4, 9 and 11 belong to one virus, and overlap with read 2 over 190–231 bp but with only 81.1–82.7% nucleotide identity. Although each of the read pairs 1 and 12, 3 and 10 and 6 and 13 overlap when aligned with a common reference genome, the two reads in each pair share less than 50% nucleotide identity. Phylogenetic analyses using partial 3C region sequences demonstrated that reads 6 and 13 form a distinct branch in the *Parechovirus* genus (Figure 2c). It is worthy to mention that these divergent sequences represent some of the first parecho-like viruses reported in pigs till now.

The three types of novel viruses were investigated further. According to the target viruses, five sets of primers were designed and synthesized based on the contigs containing the most reads for each viral type (primers A–E in the Table 1). Positive and negative controls were included in every PCR run and all products were sequenced. An additional five sets of primers (primer 2ndA–E in the Table 1), which targeted internal regions amplified in the 1<sup>st</sup> round of PCR, were designed and used in a 2<sup>nd</sup> round of PCR to exclude potential contamination, and the products were sequenced (data not shown). All of the sequences are provided in Additional file 4: S4. The results of the 2<sup>nd</sup> round of PCR validated all of the samples identified as positive in the 1<sup>st</sup> round

**Table 1 List of primers used in both the 1<sup>st</sup> and 2<sup>nd</sup> rounds of PCR in this study**

Virus	Targeted contig/read	Targeted region	Primer name	PCR round	Primer sequence	Size (bp)
Astrovirus	Contig 15	Astroviruses ORF1a region	Set A-F	1 <sup>st</sup>	5'- STTRCCWTGGSTYTGGGAGAT -3'	402
			Set A-R		5'- ATMAGCCTCCAGCAGAAGCA -3'	
	Contig 3		2ndA-F	2 <sup>nd</sup>	5'- ACTGCYTCTACYTWGCAGCAG -3'	199
			2ndA-R		5'- ACATCWGCCARGAACATACC -3'	
	Contig 14	Astroviruses ORF1a region	Set B-F	1 <sup>st</sup>	5'- AGGCTAAACCTCAAGTTAG -3'	360
			Set B-R		5'- CCTACAGGAAAAGTTACTC -3'	
2ndB-F			2 <sup>nd</sup>	5'- CACGGATCTCAATGACATGACC -3'	316	
2ndB-R				5'- GCTATTGCACCTATTTCAGAC -3'		
Bocavirus	Contig 5	Bocaviruses NS1 region	Set C-F	1 <sup>st</sup>	5'- TCTGCCTSAGGTSRGTGAGAA -3'	224
			Set C-R		5'- KATGCATCTCGAGMAYCTCKT -3'	
	Contig 18		2ndC-F	2 <sup>nd</sup>	5'- GAGAAAYTCKCTGGCTAGACAGA -3'	185
			2ndC-R		5'- AYCTCKTCKAGCGTYCGAGA -3'	
	Contig 19	Bocaviruses NS1 region	Set D-F	1 <sup>st</sup>	5'- TCTTAGCTAGCGAGTTCTC -3'	206
			Set D-R		5'- TAGATGAGGTGCTCGAGATGCA -3'	
2ndD-F			2 <sup>nd</sup>	5'- ACCTCATCTAGCGTTCGAGA -3'	159	
2ndD-R				5'- GCCTGCTAAGATCACCAAA -3'		
LV-like virus	Read 12	LV-like virus 2C region	Set E-F	1 <sup>st</sup>	5'- CTCTTCTGCTATGGAAGTCTGCT -3'	452
			Set E-R		5'- TTGCCGAAGTATCAGGCTTC -3'	
			2ndE-R	2 <sup>nd</sup>	5'- CTGCTTGGCAATGATGTGTCGA -3'	404
			2ndE-F		5'- ACAACGTGTGCAGTTTCTGG -3'	

of PCR. Of the 296 samples, 105 (35.5%), 48 (16.0%), 71 (24.0%), 142 (48.0%) and 40 (13.5%) were positive for primers A, B, C, D and E, respectively. Most of the sequences detected through PCR were highly similar at the nucleotide level to astrovirus (> 80% identity), bocavirus (> 80% identity), and parechovirus (> 90% identity) contigs. There were 41 astrovirus sequences < 80% identical to the original contigs, whereas only two bocaviruses sequences were below this threshold. For parechoviruses, only five sequences showed 85-86% identity to the original contigs. When sequences divergent to the contigs were compared with the NCBI database, some were more similar to known sequences while the best hits of others remained with the contigs. Co-infections with both types of the same virus and different viruses were detected. For example, 35 samples were positive for sets A and B (astroviruses) and 47 samples were positive for sets C and D (bocaviruses), while four samples were positive for all five sets of primers. These results indicate the great diversity of the novel viruses and their high detection rates in the samples. Sequences with the length more than 200 bp were submitted to the Genbank under accession number of JX560993-JX561090.

In this study, we demonstrated the existence of different novel astroviruses, bocaviruses and parechovirus in porcine feces using 454 massively-parallel pyrosequencing. These viruses were highly prevalent in healthy pigs < 30 days old. The great diversity and high detection rate of these viruses in the porcine samples indicate that pigs are well-established reservoirs for many different viruses. Human astroviruses, bocaviruses and parechovirus are well-known pathogens that cause a variety of human diseases. It is unknown if the novel viruses detected in this study are pathogenic or have zoonotic potential. Given the diversity of viruses detected in this preliminary study, more attention should be paid to the viromes of pigs and other agricultural animals due to the potential for transmission of these viruses to humans in close proximity to these animals.

## Additional files

**Additional file 1: Table S1.** Best hits in terms of nr, percent similarities and e values for the contigs. S1: astroviruses.

**Additional file 2: Table S2.** Bocaviruses.

**Additional file 3: Table S3.** Parecho-like virus.

**Additional file 4: S4.** Sequences of the 1<sup>st</sup> and 2<sup>nd</sup> round PCR products of potentially novel viruses from the positive samples. "-" indicates gaps in the sequences.

## Competing interest

The authors declare that they have no competing interests.

## Authors' contributions

JM carried out the data summarizing, participated in the 454 high-throughout sequencing, study designing and drafted the manuscript. JS participate in the 454 high-throughout sequencing. YY carried out the screening of the viruses in the porcine fecal samples. ZJ conceived of the study, and participated in its design. All authors read and approved the final manuscript.

## Acknowledgments

We thank David Wang and Guoyan Zhao from the Department of Pathology and Immunology, Washington University School of Medicine, for performing the bioinformatics analysis. This work was partly supported by the China Mega-Project for Infectious Disease (2011ZX10004).

Received: 27 April 2012 Accepted: 26 November 2012

Published: 30 January 2013

## References

1. Luo Z, Roi S, Dastor M, Gallice E, Laurin MA, L'Homme Y: **Multiple novel and prevalent astroviruses in pigs.** *Vet Microbiol* 2011, **149**:316-323.
2. Morin M, Turgeon D, Jollette J, Robinson Y, Phaneuf JB, Sauvageau R, Beauregard M, Teuscher E, Higgins R, Lariviere S: **Neonatal diarrhea of pigs in Quebec: infectious causes of significant outbreaks.** *Can J Comp Med* 1983, **47**:11-17.
3. Reuter G, Boldizsar A, Kiss I, Pankovics P: **Candidate new species of Kobuvirus in porcine hosts.** *Emerg Infect Dis* 2008, **14**:1968-1970.
4. Yu JM, Jin M, Zhang Q, Li HY, Li DD, Xu ZQ, Li JS, Cui SX, Yang SH, Liu N, Duan ZJ: **Candidate porcine Kobuvirus, China.** *Emerg Infect Dis* 2009, **15**:823-825.
5. Buckwalter MR, Nga PT, Gouilh MA, Fiette L, Bureau JF, Laird ME, Buchrieser J, Ozden S, Cheval J, Eloit M, et al: **Identification of a novel neuropathogenic Theiler's murine encephalomyelitis virus.** *J Virol* 2011, **85**:6893-6905.
6. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, et al: **Genome sequencing in microfabricated high-density picolitre reactors.** *Nature* 2005, **437**:376-380.
7. Cheng WX, Li JS, Huang CP, Yao DP, Liu N, Cui SX, Jin Y, Duan ZJ: **Identification and nearly full-length genome characterization of novel porcine bocaviruses.** *PLoS One* 2010, **5**:e13583.
8. Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, DeRisi JL: **Microarray-based detection and genotyping of viral pathogens.** *Proc Natl Acad Sci USA* 2002, **99**:15687-15692.
9. Finkbeiner SR, Allred AF, Tarr PI, Klein EJ, Kirkwood CD, Wang D: **Metagenomic analysis of human diarrhea: viral detection and discovery.** *PLoS Pathog* 2008, **4**:e1000011.
10. Lukashov VV, Goudsmit J: **Evolutionary relationships among Astroviridae.** *J Gen Virol* 2002, **83**:1397-1405.
11. Reuter G, Pankovics P, Boros A: **Identification of a novel astrovirus in a domestic pig in Hungary.** *Arch Virol* 2011, **156**:125-128.
12. Luo Z, Roi S, Dastor M, Gallice E, Laurin MA, L'Homme Y: **Multiple novel and prevalent astroviruses in pigs.** *Vet Microbiol* 2010, **149**:316-323.
13. Rivera R, Nollens HH, Venn-Watson S, Gulland FM, Wellehan JF Jr: **Characterization of phylogenetically diverse astroviruses of marine mammals.** *J Gen Virol* 2010, **91**:166-173.
14. Zhu HC, Chu DK, Liu W, Dong BQ, Zhang SY, Zhang JX, Li LF, Vijaykrishna D, Smith GJ, Chen HL, et al: **Detection of diverse astroviruses from bats in China.** *J Gen Virol* 2009, **90**:883-887.
15. Atkins A, Wellehan JF Jr, Childress AL, Archer LL, Fraser WA, Citino SB: **Characterization of an outbreak of astroviral diarrhea in a group of cheetahs (*Acinonyx jubatus*).** *Vet Microbiol* 2009, **136**:160-165.
16. Toffan A, Jonassen CM, De Battisti C, Schiavon E, Kofstad T, Capua I, Cattoli G: **Genetic characterization of a new astrovirus detected in dogs suffering from diarrhoea.** *Vet Microbiol* 2009, **139**:147-152.
17. Moser LA, Schultz-Cherry S: **Pathogenesis of astrovirus infection.** *Viral Immunol* 2005, **18**:4-10.
18. Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A, Andersson B: **Cloning of a human parvovirus by molecular screening of respiratory tract samples.** *Proc Natl Acad Sci USA* 2005, **102**:12891-12896.
19. Lau SK, Woo PC, Yip CC, Li KS, Fu CT, Huang Y, Chan KH, Yuen KY: **Co-existence of multiple strains of two novel porcine bocaviruses in the same pig, a previously undescribed phenomenon in members of the**

- family Parvoviridae, and evidence for inter- and intra-host genetic diversity and recombination. *J Gen Virol* 2011, **92**:2047–2059.
20. McKillen J, McNeilly F, Duffy C, McMenamy M, McNair I, Hjertner B, Millar A, McKay K, Lagan P, Adair B, Allan G: **Isolation in cell cultures and initial characterisation of two novel bocavirus species from swine in Northern Ireland.** *Vet Microbiol* 2011, **152**:39–45.
  21. Joki-Korpela P, Hyypia T: **Parechoviruses, a novel group of human picornaviruses.** *Ann Med* 2001, **33**:466–471.
  22. Tolf C, Gullberg M, Johansson ES, Tesh RB, Andersson B, Lindberg AM: **Molecular characterization of a novel Ljungan virus (Parechovirus; Picornaviridae) reveals a fourth genotype and indicates ancestral recombination.** *J Gen Virol* 2009, **90**:843–853.
  23. Niklasson B, Nyholm E, Feinstein RE, Samsioe A, Hornfeldt B: **Diabetes and myocarditis in voles and lemmings at cyclic peak densities—induced by Ljungan virus?** *Oecologia* 2006, **150**:1–7.

doi:10.1186/1743-422X-10-39

**Cite this article as:** Yu et al.: Detection of novel viruses in porcine fecal samples from China. *Virology Journal* 2013 **10**:39.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

