Identification of a novel bufavirus in domestic pigs by a viral metagenomic approach

Lihong Liu,1 Lukas Schwarz,2 Karin Ullman,1 Harri Ahola,1 Yafeng Qiu,3 Zhiyong Ma3 and Isabel Hennig-Pauka2

1Department of Microbiology (MIK), National Veterinary Institute (SVA), 751 89 Uppsala, Sweden
2University Clinic for Swine, Department for Farm Animals and Veterinary Public Health, University of Veterinary Medicine, Vienna, Austria
3Shanghai Veterinary Research Institute, Shanghai, China

Bufavirus is a single-stranded DNA virus belonging to the genus Protoparvovirus. This study reports the identification and characterization of a porcine bufavirus by a metagenomic approach, and a limited epidemiology investigation of bufavirus in six swine farms. A comparative genome analysis showed a similarity of 93 % to a Hungarian porcine bufavirus. Bayesian and maximum-likelihood analyses of genome sequences showed a close relationship of porcine bufaviruses to human and monkey bufaviruses. Molecular dating of the most recent common ancestors supported a recent introduction of bufaviruses into human and pig populations, respectively. A real-time PCR method was developed to screen 60 faecal samples for the porcine bufavirus DNA, and eight positive samples were found in two neighbouring farms, suggesting a relatively low prevalence (13.3 %). No direct transmission of porcine bufaviruses between two neighbouring farms was found, suggesting that bufaviruses may have spread widely in different geographical regions.

Bufavirus belongs to the genus Protoparvovirus in the subfamily Parvovirinae (Cotmore et al., 2014). The virus has a single-stranded DNA genome, encoding non-structural protein 1 (NS1), a putative structural protein 1 (VP1), a small hypothetical protein of 130 aa residues and a structural protein 2 (VP2), as demonstrated in a reference genome (GenBank accession number: NC_024888). Genome comparison and phylogenetic analysis have grouped human bufaviruses into three genotypes (Yahiro et al., 2014). Bufavirus has first been identified in faeces of children with acute diarrhoea in Burkina Faso, a landlocked country in West Africa, and found in one Tunisian child (Phan et al., 2012). Subsequently, the virus has been discovered sporadically in cases of diarrhoea in Bhutan (Yahiro et al., 2014), Finland (Väisänen et al., 2014), the Netherlands (Smits et al., 2014), Turkey (Altay et al., 2015), Thailand (Chieochansin et al., 2015) and China (Huang et al., 2015). The role of bufaviruses as an etiologic agent of human gastroenteritis has not been firmly established yet. Further metagenomic investigations show distinct lineages of bufavirus in wild shrews and non-human primates from Zambia (Sasaki et al., 2015), in Hungarian bats (Kemenesi et al., 2015) and pigs (Hargitai et al., 2016) and in Chinese rats (Yang et al., 2016). This report describes the identification and characterization of bufavirus in faeces of domestic pigs by a metagenomic approach.

A total of 60 samples were collected from 6–10-week-old pigs, on five Austrian farms and one Hungarian farm (Fig. 1). The samples consisted of five diarrhoeic (semi-liquid or liquid form) and five normal faeces per farm. One diarrhoeic and one normal sample from each farm were used for the metagenomic studies and all samples were used for subsequent epidemiological investigation. About 200 mg of faeces were diluted to 1 ml of water, vortexed vigorously for 1 min and centrifuged at 14 000 g for 10 min. Supernatant (400 µl) was treated with 400 U ml−1 DNase I (Roche Applied Science) and 10 µg ml−1 RNase A (Invitrogen) at 37 °C for 2 h. DNA and RNA were extracted from 200 µl of the treated supernatant using a QIAamp DNA mini kit (Qiagen) and TRIzol (Invitrogen)/RNaseasy kit (Qiagen), respectively. Procedures for synthesis and random amplification of tagged cDNA and dsDNA, and removal of the primer tag, were described previously (Blomström et al., 2010). A DNA library was prepared using a Nextera XT DNA sample preparation kit (Illumina), and sequencing was performed on a MiSeq instrument using a MiSeq Reagent...
kit v3 (Illumina), according to the manufacturer’s instructions. Sequencing data were processed with the CLC Genomics Workbench (Qiagen). The RNA/DNA sequencing reads of the same samples were trimmed, based on quality (Table S1, available in the online Supplementary Material), and de novo assembled into contigs, with which nucleotide BLAST (blastsn) against a virus database was performed to identify matching virus species with an E-value <10^-5. The reads from the bufavirus-positive samples were also mapped to porcine bufavirus strain Swine/Zsana/2013/HUN (KT965075) and consensus sequences were extracted. In order to fill in gaps in the consensus sequences, primers (Table S2) were designed to amplify target regions by Phusion high-fidelity DNA polymerase (New England Biolabs) for conventional Sanger sequencing of the amplicons. RDP v4.39 (Martin et al., 2015) was used to detect recombination events prior to maximum-likelihood and Bayesian phylogenetic analyses, as described previously (Liu et al., 2009). In addition, BEAST v1.8.4 (Drummond et al., 2012) was used to infer the evolutionary history of bufaviruses. The resulting trees were sampled, and a maximum clade credibility tree with median node heights was annotated. To confirm the viral metagenomic results and for subsequent epidemiological investigation, a real-time TaqMan assay with forward primer (5'-AGACAAACACGGACTCTCACAG-3'), reverse primer (5'-CTTGGTGTTCTTATTCCCCGAA-3') and probe (5'-FAM-CAACCCTCTGGTCTCGGTACGC-BHQ1-3'), using an Ag Path-ID one-step RT-PCR kit, was developed, which was performed on a Bio-Rad CFX system. The relative positions on the reference sequence (NC_024888) were 4284–4306 (forward primer), 4325–4348 (probe) and 4379–4354 (reverse primer). For real-time PCR testing, nucleic acids were extracted from 200µl of supernatant without nuclease treatment, by the TRizol/RNeasy protocol. Additional extraction of DNA using a QIAamp DNA mini kit (Qiagen) was made for PCR amplification of target regions.

Viral metagenomic analysis revealed several contigs similar to a porcine bufavirus, Swine/Zsana/2013/HUN (KT965075) and human bufaviruses. These contigs were from a pair of diarrhoeic and normal faecal samples from one Austrian farm (no. 5) and one Hungarian farm (no. 6), respectively. Real-time PCR assay confirmed the presence of bufaviruses in the two Austrian samples but not in the Hungarian samples (Table S3), which was likely due to mismatches between primer/probe sequences and the viral sequence. In addition, a total of eight samples (out of 60) from the two neighbouring Austrian farms (nos 4 and 5) were found positive by the assay, suggesting a relatively low prevalence (13.3 %) of bufaviruses in the farms investigated.

After filling the gaps in the consensus sequences obtained from next-generation sequencing reads that were mapped to the porcine bufavirus, a near-complete (4189 nt) genome sequence of the Austrian porcine bufavirus (strain 61) was determined. Based on an alignment with reference bufavirus sequences, the genome (KU867071) was predicted to encode a partial NS1 (579 aa), a putative protein (129 aa) and partial structural proteins VP1 (679 aa) and VP2 (538 aa). Expression of VP1 was likely due to an alternative splicing mechanism, as described previously (Phan et al., 2012). Comparative analysis of the near-complete genome sequence showed that the Austrian porcine bufavirus (KU867071) shared a similarity of 65.9 % to human bufavirus-1 BF7 (JX027295), 65.7 % to human bufavirus-2 BF39 (JX027297), 65.4 % to human bufavirus-3 AHP178 (AB982217), 93 % to porcine bufavirus Swine/Zsana/2013/HUN (KT965075), 58.4 % to Shrew ZM38 (AB937988),

http://jgv.microbiologyresearch.org 1593

Fig. 1. Geographical locations of the six swine farms. The map was drawn using arcinfo (ESRI arccpacT 10.0) and DIVA-GIS (http://www.diva-gis.org/Data) and adapted using Adobe Photoshop CS5.1.
Fig. 2. (a) Bufavirus phylogeny inferred by Bayesian and maximum-likelihood analyses of near-complete genome sequences (4189 nt). The newly determined sequence is denoted by a filled square. An asterisk indicates a maximum posterior probability (1.0) and bootstrapping value of more than 95% (out of 1000 replicates) for the major nodes. Two nodes supported by 80% and 83% bootstrapping values are labelled. The sequence name contains five elements in the order of GenBank accession number, host (except humans), isolate name, country where the virus was discovered and the year when the sample was collected, where an underscore is used to separate each element. This is an unrooted tree. Bar indicates substitutions per site.
bufaviruses (Kemenesi et al., 2012). All three genotypes of human bufaviruses formed a well-supported group whereas the WUHARV parvovirus was in a sister branch. Bufaviruses found in bats, a rat and a shrew were distantly related to the porcine and human bufaviruses. Although the WUHARV parvovirus has been suggested to be a recombinant between bat and human bufaviruses. Although the WUHARV parvovirus was in a sister branch. Bufaviruses found in bats, a rat and a shrew were distantly related to the porcine and human bufaviruses. Although the WUHARV parvovirus has been suggested to be a recombinant between bat and human bufaviruses (Kemenesi et al., 2015), this recombination signal was not detected by our conserved analysis, where a recombination signal should be detected by more than four out of all eight methods implemented in the program and at a highest acceptable P value of 0.001.

Except for the single strain of human bufavirus 2 (accession number KM580349), the genetic diversities of human bufaviruses within each genotype and of porcine bufavirus are seemingly limited at the moment. This observation indicated that bufaviruses likely entered humans and domestic pigs a relatively short time ago and evolved independently in their respective hosts. To estimate divergence time for most recent common ancestors (tMRCA), a Bayesian evolutionary analysis of the tip-dated (specified as year of collection) sequences was performed using an uncorrelated relaxed clock (exponential distribution), a coalescent tree prior (exponential growth) and under a sequence evolutionary model of GTR+I+G. A total of 336 million states were obtained from six independent runs, and effective sample size (ESS) reached over 200, indicating accurate estimates. A maximum-credibility tree maintaining median node heights is summarized (Fig. 2b), which has the same topology as those from maximum-likelihood and Bayesian analyses (Fig. 2a). The results supported the recent introduction hypothesis: the tMRCA of porcine bufaviruses was dated back to be 29 years while it was 69 years for the human bufaviruses. In other words, based on the dataset, the most recent common ancestor for the porcine bufaviruses might have emerged 29 years before the present time, and for human bufaviruses, 69 years. Separation of the WUHARV parvovirus from human bufaviruses occurred at least 128 years ago, and the tMRCA for Rhesus monkey, human and porcine bufaviruses was estimated to be 167 years before the present time. The estimation of all tMRCA should be interpreted with caution as a large range of the 95% highest posterior density (HPD) was found for most estimations. Based on the position of bat and wildlife bufaviruses in the phylogenetic tree (Fig. 2b), one could speculate that wildlife or bats might be one of the sources from where human and porcine bufaviruses have emerged. However, current limited knowledge of bufaviruses is unable to reveal the emergence, speciation and spread among different hosts, which is in sharp contrast to the well-studied human immunodeficiency virus (HIV) where rich data have enabled scientists to reconstruct the very early spread and epidemic ignition in human populations (Faria et al., 2014). Finally, a high evolutionary rate of bufaviruses (1.6×10^{-3} substitutions per site per year) was determined, which is similar to that of the single-stranded DNA virus porcine circovirus 2 (Firth et al., 2009).

Besides the near-complete genome sequence of porcine bufavirus obtained from the Austrian farm (no. 5), a 376 nt fragment of VP2 sequence (strain 57; KU867072) was determined from the neighbouring farm (no. 4). This short sequence shared a nucleotide similarity of 92.5% with the Austrian strain 61 and of 88% with the Hungarian strain Swine/Zsana/2013/HUN. A direct transmission between the two neighbouring Austrian farms could not be established confidently. It is likely that the bufaviruses might have already spread much more widely in pig populations. In addition, other animals may play a role in the transmission.

In summary, a near-complete genome sequence of porcine bufavirus was determined. Bayesian and maximum-likelihood analyses of genome sequences showed a basal position of porcine bufaviruses relative to human and monkey bufaviruses. Under a relaxed molecular clock, the most recent common ancestor might have emerged around 29 years before the present time for the porcine bufaviruses and 69 years for human bufaviruses, supporting the recent introduction of the virus into swine and human populations, respectively. However, no direct transmission between two neighbouring farms, where a total of eight samples were positive for the porcine bufavirus, could be established confidently, suggesting that porcine bufaviruses may have already spread in different geographical regions.

Acknowledgements

This study was supported by the EU projects ‘RAPIDIA-FIELD’ (grant no. 289364) and ‘LinkTADS’ (grant no. 613804). We are grateful to Dr Georg Duscher from Institute of the Parasitology, University of Veterinary Medicine Vienna, Vienna, for supplying the Austrian map.

References


