Arch Virol (2010) 155:801–806
DOI 10.1007/s00705-010-0646-8

BRIEF REPORT

Identification and molecular cloning of a novel porcine parvovirus

Andrew K. Cheung · Guang Wu · David Wang · Darrell O. Bayles · Kelly M. Lager · Amy L. Vincent

Received: 8 December 2009 / Accepted: 29 January 2010 / Published online: 26 March 2010

Abstract A novel porcine parvovirus, PPV4, was identified in the lung lavage of a diseased pig coinfected with porcine circovirus type 2. This virus exhibits limited similarity to its closest relative, bovine parvovirus 2, but resembles viruses of the genus *Bocavirus* (bovine parvovirus, canine minute virus and human bocavirus) that encode an additional ORF3. The ORF3 of PPV4 is predicted to encode a protein of 204 amino acid residues, which is similar in size to the ORF3-encoded proteins of the bocaviruses. Whereas the ORF3-encoded proteins of bocaviruses share significant similarity with each other, the PPV4 ORF3 encoded protein does not exhibit homology with any protein in the GenBank non-redundant database.

Keywords *Parvovirinae* · Porcine parvovirus · Phylogeny

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

A. K. Cheung (✉) · K. M. Lager · A. L. Vincent
Virus and Prion Research Unit, National Animal Disease Center, PO Box 70, Ames, IA 50010, USA
e-mail: andrew.cheung@ars.usda.gov

G. Wu · D. Wang
Departments of Molecular Microbiology and Pathology and Immunology, Washington University in St. Louis, St. Louis, MO 63110, USA

D. O. Bayles
Infectious Bacterial Diseases Research Unit, National Animal Disease Center, PO Box 70, Ames, IA 50010, USA

Parvoviruses are small, non-enveloped icosahedral viruses. These viruses are ubiquitous, and they have been isolated from many animal species. Based on their host range, viruses of the family *Parvoviridae* are divided into two subfamilies: *Parvovirinae*, which includes viruses that infect birds and mammals and *Densovirinae*, which includes viruses that infect insects and other arthropods [10]. The *Parvovirinae* and *Densovirinae* viruses do not share any nucleotide sequence homology. The International Committee on Taxonomy of Viruses has further classified the *Parvovirinae* viruses into five genera: *Dependovirus*, *Bocavirus*, *Erythrovirus*, *Parvovirus* and *Amdovirus*. Recently, a new genus, Hokovirus, was proposed to include a group of newly identified porcine isolates (designated PPV3 in this report) and bovine isolates (designated BPV4 in this report) that clustered with the human PARV4 and PARV5 viruses [14].

The genomes of members of the subfamily *Parvovirinae* are linear single-stranded DNA genomes of about 5 kilobases that contain terminal palindromic sequences. In general, the genome contains two major open reading frames (ORFs), coding for the non-structural protein(s), located at the 5′-end, and the capsid protein(s), located at the 3′-end. An additional ORF3 has been observed among viruses of the genus *Bocavirus* [bovine parvovirus (BPV1), canine minute virus (MVC) and human bocavirus (HBoV)], and ORF3 is located in the middle of the viral genome between ORF1 and ORF2 [2, 10]. Recently, it was reported that ORF3 of MCV is essential for viral DNA replication [20]. Parvoviruses have been proposed to replicate their genomes via a modified rolling hairpin model that involves a hairpin transfer mechanism [3, 5, 8]. In this model, two obligatory dimeric genome replicative intermediates (RF) generated via the terminal palindromic sequences are synthesized during viral DNA replication

[6]. These dimeric RFs contain a sense-strand and a minus-strand, i.e., a double-stranded genome, that are covalently linked (via terminal palindromic self-priming) either at the left end or right end of the linear genome and give rise to head-to-head or tail-to-tail concatameric molecules.

In this work, we cloned a novel porcine parvovirus from a circular or a head-to-tail concatameric template. Since it is the fourth (nearly) complete DNA sequence of a porcine parvovirus described, this virus is designated PPV4. Phylogenetically, PPV4 does not cluster with the three previously identified porcine parvoviruses but exhibits limited homology to bovine parvovirus 2 (BPV2) [1]. Incidentally, a recent publication based on partial nucleotide sequence reported identification of a novel porcine boca-like virus [4] among pigs co-infected with porcine circovirus type 2 (PCV2). The ORF3 of this porcine boca-like virus and PPV4 share low nucleotide identity (45%) and amino acid identity (17.5%), which indicate that this boca-like virus and PPV4 are two distinct parvoviruses.

PPV4 was identified among swine suffering from an acute-onset disease of high mortality in North Carolina during late 2005 [7]. Case reports of this disease indicated that various viral and bacterial pathogens were identified in tissues from swine infected with PCV2. PCR analysis showed that porcine reproductive and respiratory syndrome virus, porcine respiratory coronavirus, and several different serovars of porcine enterovirus were isolated from one or more of the affected herds. PCV2 was identified in essentially all pigs from each of the herds. Tissue homogenates consisting of lung, lymph node, spleen, and heart were prepared from fresh frozen tissues of the affected swine. The tissues were triturated with sterile sea sand via mortar and pestle. Serum-free minimum essential medium was used to reconstitute a 10% solution by weight. The solution was clarified at 1,000×*g* for 10 min and filtered through a 0.45-micron filter. Two colostrum-deprived cesarean-derived (CDCD) pigs, pigs 753 and 754, were inoculated oronasally with 2 ml of the 10% tissue homogenates and monitored for clinical disease. Both pigs developed severe respiratory disease and had to be euthanized. Pig 754 was euthanized on day 16 post-infection and pig 753 was euthanized on day 21 post-infection.

To explore whether there were any unknown viruses present in the diseased pigs, a random shotgun sequencing experiment was conducted. RNA from a lung lavage sample of pig 753 was extracted, randomly amplified and cloned into pCR4.TOPO (Invitrogen, Carlsbad, CA, USA) as described previously [23] and sequenced using standard Sanger chemistry [17] on an ABI 3730xl sequencer. The resulting sequence reads were trimmed to remove vector and primer sequences and then assembled using Phred/Phrap [9]; (Phil Green, http://weeds.mgh.harvard.edu/goodman/doc/). Contigs with sequence similarity to known viral sequences were identified using tBLASTx. Three hundred thirty-four clones were sequenced. Alignment of the sequence reads to the GenBank NR database yielded 19 sequences with limited identity to viruses in the family *Parvoviridae*. Nucleotide sequences belonging to other virus families were not detected. The 19 parvovirus-related sequences were assembled by Phred/Phrap into two contigs that were experimentally verified by generating PCR amplicons the length of the contigs. Deduced amino acid sequence homology was analyzed with the translated BLAST search. Contig 1 was 1,815 base pairs (bp) long and had 34% amino acid identity to its top hit, BPV2 capsid protein. Contig 2 was 2,993 bp long and had 39% identity to the non-structural protein of BPV2. Primers at nt 3,681–3,702 and nt 4,206–4,183 (based on c17 of Fig. 1) were then designed to span the gap between the two contigs. PCR was conducted with total cellular DNA from pig 753 that was prepared using a STAT-60 DNA extraction kit purchased from Tel-Text B, Inc. (Friendswood, TX, USA). The PCR reaction was carried out with 0.1 μg of DNA in the presence of 10 mM Tris–HCl (pH 8.3), 0.2 mM each of the four deoxynucleotide triphosphates, 100 pM each of the primers, and 2.5 U of Taq polymerase. The reaction mixture was heated at 94°C for 1 min, amplified for 34 cycles at 94°C for 15 s, 55°C for 30 s, and 68°C for 5 min 15 s, and then kept at 68°C for 7 min. The resulting contig was 4,851 bp in length.

Based on the DNA sequence obtained from the initial random amplification and the above PCR experiment, diverging primers were designed to obtain additional nucleotide sequences upstream and downstream of the 4,851-nucleotide contig. Nested PCR was conducted with the pig 753 lung lavage DNA. In first-round PCR, primers X (tga acc cat aca ttg tac ca) and Y (gtc tac ttc ttc taa ctg ca) were used (Fig. 1a). In second-round PCR, an aliquot of the first-round products was amplified with the common primer C (ctc cgc ctg ccg aac agc ga), and one of three primers: #1 (ttt cca gat tac atg ggt tc), #2 (tga gac tcc tgg cac aga tg) or #3 (tta cat atg aaa tgg aat gg). In each case, a prominent PCR product of the expected size was obtained. The results demonstrated that a PPV4 molecule in head-to-tail configuration was present in the pig 753 DNA sample. To obtain the unit-length head-to-tail PPV4 sequence, two abutting primers of opposite polarity located at the unique *Kpn*I site of the viral genome were used for PCR. The primers used were PKpnF-gga agt ggt acc ttc tct tac aca tca ac (forward orientation) and PKpnR-gga agt ggt acc tgt ttt cag ttg gtg at (reverse orientation), and the *Kpn*I site is located at nt 3,342–3,348 of the clone 17 (c17) sequence. After amplification, several PCR products were visualized on an agarose gel. One of the products was approximately 6,000 bp long, and the rest were less than 4,000 bp long (Fig. 1b). The 6,000-bp PCR product was excised from the

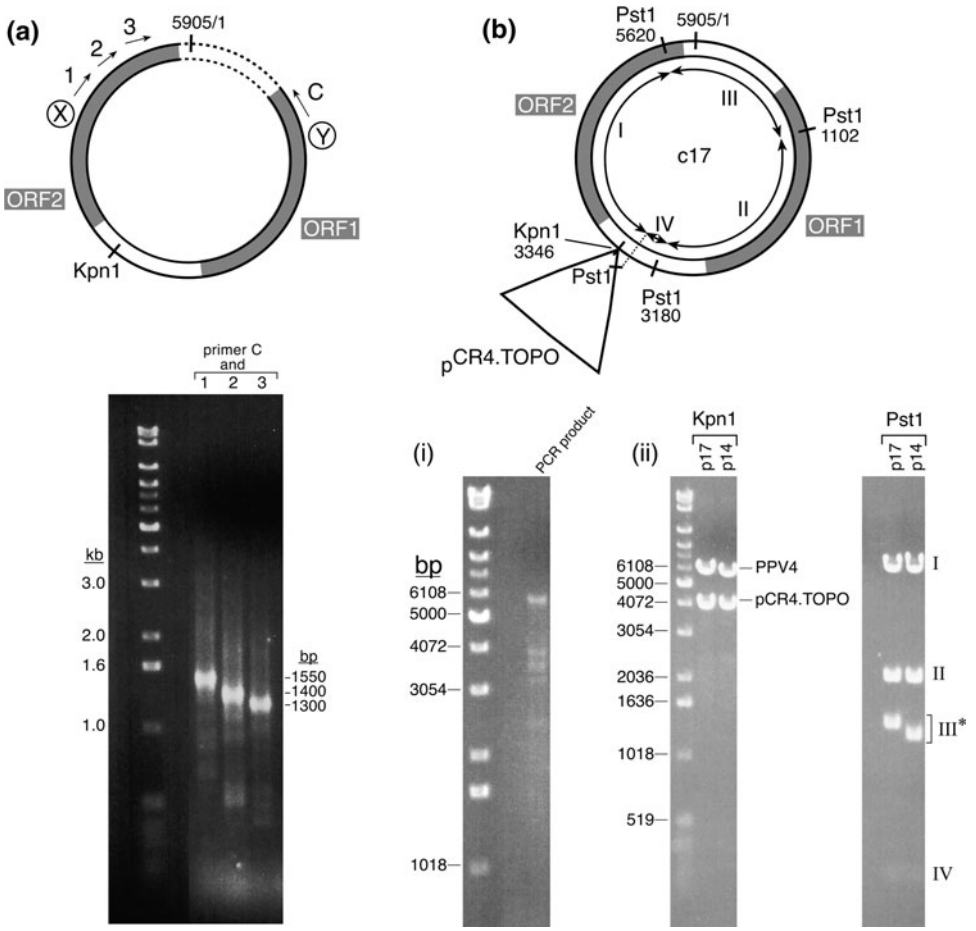

Fig. 1 **a** A schematic diagram denoting the PCR oligonucleotide primers and the strategy to obtain the head-to-tail covalent linkage nucleotide sequence (indicated by dotted lines) from a circular molecule. Nested PCR was performed. The first-round PCR was conducted with primers X and Y (data not shown). *Bottom panel* the second-round nested-PCR was conducted using a common primer C with either primer 1, 2 or 3, as indicated at the top of each lane of the agarose gel. **b** Schematic diagram of plasmid p17, which contains the PPV4 c17 genome inserted into pCR4.TOPO. ORF1, ORF2 and restriction enzyme sites of *Kpn*I and *Pst*I are denoted. *Bottom panel* agarose gel analysis of plasmids p14 and p17. (*i*) PCR products obtained using two divergent primers, PKpnF and PKpnR, at the *Kpn*I site of c17. (*ii*) The restriction enymes (*Kpn*I and *Pst*I) used are indicated at the top of each panel. *Kpn*I digestion releases the entire double-stranded PPV4 genome (5.9 kb) from the cloning vector pCR4.TOPO (4.0 kb). PstI digestion yielded four fragments, labeled *I–IV*, and DNA fragment III, which differs between c14 and c17, is indicated by *asterisk*

gel and cloned into the pCR4.TOPO plasmid. A miniprep plasmid DNA kit (Promega, Madison, WI, USA) was used to isolate nucleic acids from overnight TOP10 (Invitrogen, Carlsbad, CA, USA) bacterial cultures, and DNA electrophoresis was carried out in 1% agarose gel. Two plasmids (p14 and p17) containing the PPV4 DNA clones (c14 and c17), respectively, were obtained. As shown by the *Pst*I restriction enzyme analysis, c14 is slightly smaller than c17. The DNA clones were then sequenced using primer walking. It was confirmed that c14 is 125 nucleotides (nt) shorter than c17 and the deleted sequence is located at nt 127–251. In addition, there are 15 nt differences between these two clones. The GenBank accession numbers for c14 and c17 are GQ387500 and GQ387499, respectively. PCR was also conducted with total cellular DNA purified from the lung lavage of pig 754. The partial PPV4 nucleotide sequence obtained was essentially identical to the parvovirus DNA sequence derived from pig 753. No further experimentation was conducted with the viral sequences from pig 754. To investigate whether PPV4 was present in the field cases of the 2005 outbreak, all of the pigs examined in the original study [7] were tested for PPV4 by PCR. The results showed that only two pigs (n6 and n10) of Farm 1 from North Carolina were positive for PPV4.

Phylogenetic analysis was performed on the PPV4 nucleotide sequences using full-length or nearly full-length genomes from all species of the subfamily *Parvovirinae* as described previously by Lau et al. [14]. Nucleotide and amino acid sequences were aligned with CLUSTAL, and phylogeny was inferred using MEGA4 [21]. The evolutionary history was inferred using the Neighbor-Joining method [16]. The evolutionary distances were computed

using the Maximum Composite Likelihood method [22] and are in units of number of base substitutions per site. The bootstrap consensus tree inferred from 10,000 replicates was taken to represent the evolutionary history of the taxa analyzed [11]. Branches corresponding to partitions reproduced in fewer than 50% of the bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10,000 replicates) is shown next to the branches [11]. Pairwise global comparison of putative amino acid sequences coded for by ORF1, ORF2, and ORF3 were conducted by aligning the sequences for each ORF using the Needleman-Wunsch algorithm [13, 15]. Based on comparative nucleotide sequence, PPV4 did not cluster with any of the known parvoviruses, including the three previously identified porcine isolates (Fig. 2). Instead, PPV4 is most similar to BPV2. Separately, phylogenetic analyses were conducted with the deduced ORF1 or ORF2 amino acid sequences, and essentially identical results were obtained (data not shown).

The c17 PPV4 clone is 5,905 base pairs long. Since the nucleotide sequences of the left and right ends of the linear viral genome have not been determined, the ends of the genome are arbitrarily assigned so that the genome conforms to the general organization of parvoviruses having two major ORFs: ORF1 located at the 5′-end and ORF2 at the 3′-end. PPV4 has an additional ORF3 located between ORF1 and ORF2, which is a feature of the viruses belonging to the genus *Bocavirus*. Analyses were conducted to compare the ORFs of PPV4 with those of other PPVs and BPVs and members of the genus *Bocavirus* at the amino acid level.

1. ORF1: PPV4 ORF1 is 1,794 nt long and capable of coding for a protein of 588 residues. BLAST search showed that it is most closely related to the putative nonstructural protein of BPV2 and the replication proteins of adeno-associated viruses 1 (AAV-1) (45.8% identity) (GenBank accession number FJ872544) of the genus *Dependovirus*. It shares 33.6% amino acid residue identity with BPV2 and less with the other porcine and bovine parvoviruses. PPV4 shares 17.4–20.5% amino acid identity with the bocaviruses, while the bocaviruses share 24.8–33.5% identity among themselves.
2. ORF2: PPV4 ORF2 is 2,184 nt long and capable of coding for a protein of 728 residues. BLAST search showed that it is most closely related to the putative structural protein of BPV2 and the capsid proteins of various parvoviruses. It shares 24.5% amino acid residue identity with BPV2 and less with the other porcine and bovine parvoviruses. PPV4 shares 18.7–20.3% amino acid identity with the bocaviruses, while the bocaviruses share 41.4–45.6% identity among themselves.
3. ORF3: PPV4 ORF3 is 612 nt long and capable of coding for a protein of 204 residues. BLAST search showed that ORF3 did not have significant homology to any protein in the GenBank non-redundant protein database. With respect to coding capacity, it is similar to that of the bocaviruses. The ORF3 of BPV1, HBoV and MCV is capable of coding for a protein of 213, 219 and 188 residues, respectively. The PPV4 ORF3-encoded protein shares only 4.9–11.2% amino acid identity with the bocaviruses, while the bocaviruses share 43.3–47.0% identity among themselves. Whereas the PPV4 ORF3 does not overlap with either major ORF, MVC overlaps with ORF1, HBoV overlaps ORF2 and BPV1 overlaps with both ORF1 and ORF2.

In summary, we report the identification of a novel porcine parvovirus, PPV4. In this study, a 5,905-nt PPV4 sequence was obtained after PCR amplification using two diverging primers from the unique *Kpn*I site of the viral genome. Thus, the templates for the generation of the genomic clones, c14 and c17, were either circular or head-to-tail concatameric molecules in which the left and right ends of the linear genome were covalently linked. Based on the genome organizataion of parvoviruses, it is expected that both the left end and the right end of the linear PPV4 genome are located between nt 1 and 848 (Fig. 1b). However, it is possible that some terminal nucleotides of the linear genome are missing in c17 and c14 during formation of the head-to-tail configuration. The genomes of parvoviruses are single-stranded linear DNA molecules; however, it has been reported that the genomes of AAV-2, a member of the genus *Dependovirus*, can persist in tissues as head-to-tail monomeric and concatameric circles [18, 19]. The presence of head-to-tail concatameric PPV4 genomes in tissues of infected animals suggests that this virus may be able to establish persistence in its host.

Phylogenetic analysis showed that the porcine and bovine parvoviruses exhibit high degrees of diversity. In general, the porcine or bovine isolates do not cluster phylogenetically with their respective viruses from swine or cattle. Among the porcine parvoviruses, PPV1 belongs to the genus *Parvovirus*, PPV3 belongs to the proposed genus Hokovirus, PPV4 and the porcine boca-like virus [4] resemble members of the genus *Bocavirus*, and PPV2 does not cluster with any known isolates. A similar phenomenon has also been reported for the bovine parvoviruses [14]. There are four bovine parvoviruses, but they also do not cluster together phylogenetically (Fig. 2). BPV1 belongs to the genus *Bocavirus*, BPV2 and BPV3 [1, 10] do not cluster with any other known isolates, and BPV4 belongs to the proposed genus Hokovirus. Interestingly, one porcine

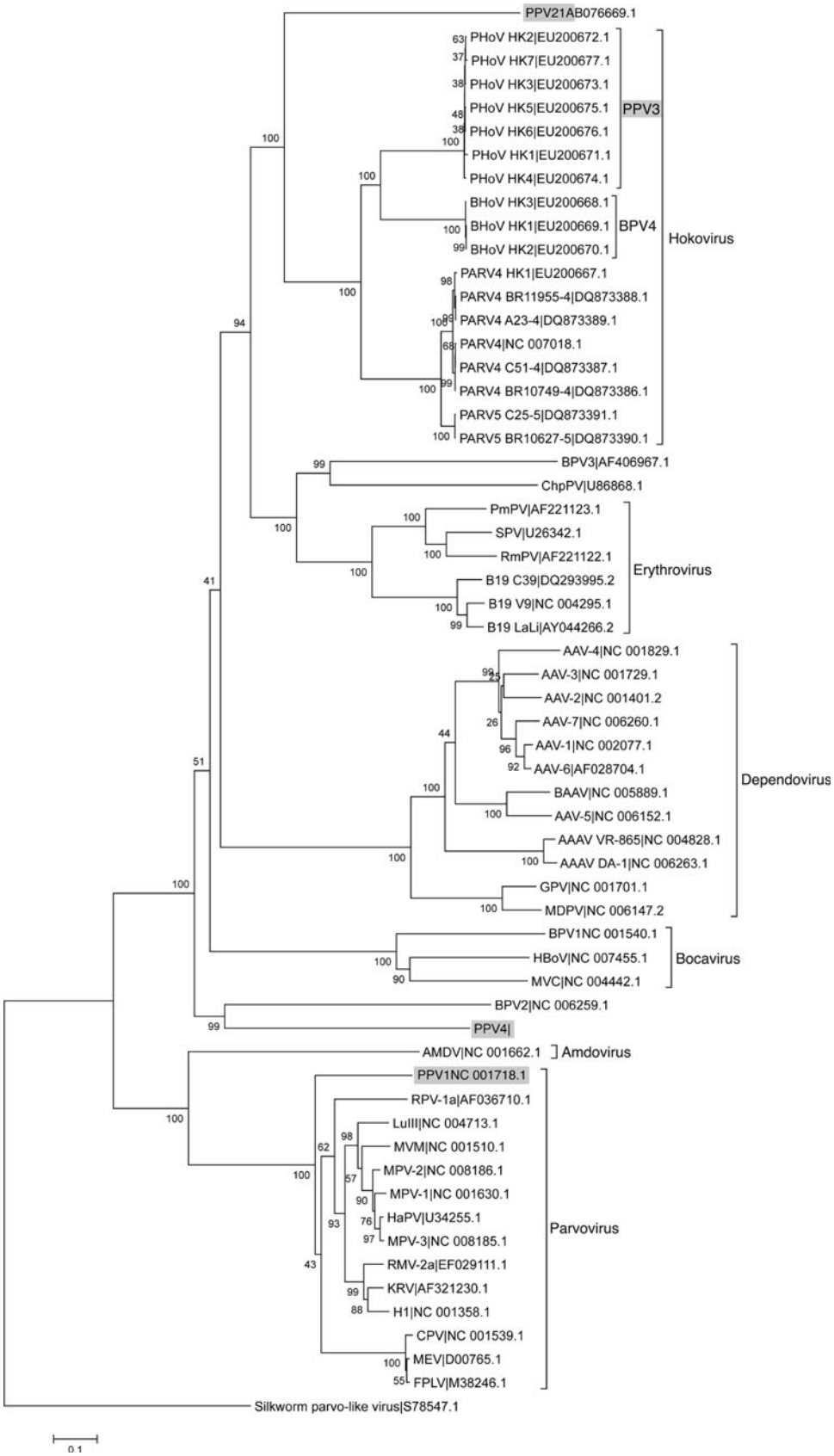

Fig. 2 Evolutionary relationships of PPV4 and other related parvoviruses. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. All positions containing gaps and missing data were eliminated only in pairwise sequence comparisons (pairwise deletion option). There were a total of 6,684 positions in the final dataset. Phylogenetic analysis was conducted using MEGA4 [21]. The porcine parvoviruses are denoted in *shaded boxes*

(PPV3), one bovine (BPV4) and two human parvoviruses (PARV4 and PARV5) are closely related and form the proposed genus Hokovirus [14].

Our analysis revealed several distinct characteristics of PPV4. Whereas PPV4 is most closely related to BPV2 phylogenetically, the coding capacity and genome organization of PPV4 are more similar to those of viruses of the genus *Bocavirus*. PPV4 is most closely related to BPV2 with respect to ORF1 and ORF2 phylogenetically; however, BPV2 does not contain a distinct ORF3 situated between the two major ORFs. Similar to the bocoviruses, PPV4 encodes an additional ORF3 located in the middle of the genome, but the deduced protein of PPV4 ORF3 is quite distinct from the ORF3-encoded proteins of the bocaviruses.

The pathogenic potential of porcine parvoviruses is not fully understood. PPV1 is ubiquitous in swine, and it causes reproductive failure in pregnant females [24]. PPV2 has been detected in swine serum but has not been incriminated in any disease [12]. The pathogenic nature of PPV3 and PPV4 remains to be determined. PPV3, a member of the proposed genus Hokovirus, has been found in tissues of healthy and sick pigs [14]. In this study, PPV4 was identified in a co-infection with porcine circovirus type 2 (PCV2) from field cases of severe PCV2-associated disease. PCV2 has been associated with multiple disease syndromes in swine and is the primary agent leading to postweaning multisystemic wasting syndrome (PMWS). It has been reported that other pathogens enhance the severity of PMWS during co-infection with PCV2. From this study, it is not clear whether PPV4 can cause disease on its own or if PPV4 contributed to the disease phenomenon in the CDCD pigs. At any event, the detection of PPV4 among diseased pigs that exhibit PCV2-associated PMWS warrants further investigation into the pathogenic nature of PPV4.

Acknowledgments The authors thank W. Boatwright and D. Adolphson for technical assistance and M. Marti and S. Ohlendorf for manuscript preparation.

References

1. Allander T, Emerson SU, Engle RE, Purcell RH, Bukh J (2001) A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. Proc Natl Acad Sci USA 98:11609–11614
2. 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
3. Astell CR, Chow MB, Ward DC (1985) Sequence analysis of the termini of virion and replicative forms of minute virus of mice DNA suggests a modified rolling hairpin model for autonomous parvovirus DNA replication. J Virol 54:171–177
4. Blomstrom A-L, Belak 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
5. Cavalier-Smith T (1974) Palindromic base sequences and replication of eukaryotic chromosome ends. Nature (London) 240:467–470
6. Chen KC, Shull BC, Lederman, Stout ER, Bates RT (1988) Analysis of the termini of the DNA of bovine parvovirus: demonstration of sequence inversion at the left terminus and its implication for the replication model. J Virol 62:3807–3813
7. Cheung AK, Lager KM, Kohutyuk OI, Vincent AL, Henry SC, Baker RB, Rowland RR, Dunham AG (2007) Detection of two porcine circovirus type 2 genotypic groups in United States swine herds. Arch Virol 152:1035–1044
8. Cotmore SF, Tattersall P (1987) The autonomously replicating parvoviruses of vertebrates. Adv Virus Res 33:91–174
9. Ewing B, Green P (1998) Base-calling of automated sequencer traces using Phred, II: error probabilities. Genome Res 8:186–194
10. Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (eds) (2005) Virus taxonomy. 8th report of the International Committee on Taxonomy of Viruses. Academic Press, London
11. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791
12. Hijikata M, Abe K, Win KM, Shimizu YK, Keicho N, Yoshikura H (2001) Identification of new parvovirus DNA sequence in swine sera from Myanmar. Jpn J Infect Dis 54:244–245
13. Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 48:443–453
14. Lau SK, Woo PC, Tse H, Fu CT, Au WK, Chen XC, Tsoi HW, Tsang TH, Chan JS, Tsang DN, Li KS, Tse CW, Ng TK, Tsang OT, Zheng BJ, Tam S, Chan KH, Zhou B, Yuen KY (2008) Identification of novel porcine and bovine parvoviruses closely related to human parvovirus 4. J Gen Virol 89:1840–1848
15. Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. Trends Genet 16:276–277
16. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425
17. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467
18. Schnepp BC, Jensen RL, Chen CL, Johnson PR, Clark KR (2005) Characterization of adeno-associated virus genomes isolated from human tissues. J Virol 79:14793–14803
19. Schnepp BC, Jensen RL, Clark KR, Johnson PR (2009) Infectious molecular clones of adeno-associated virus isolated directly from human tissues. J Virol 83:1456–1464
20. Sun Y, Chen AY, Cheng F, Guan W, Johnson FB, Qiu J (2009) Molecular characterization of infectious clones of the minute virus of canines reveals unique features of Bocaviruses. J Virol 83:3956–3967
21. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599
22. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci USA 101:11030–11035
23. Wang D, Urisman A, Liu YT, Springer M, Ksiazek TG, Erdman DD, Mardis ER, Hickenbotham M, Magrini V, Eldred J, Latreille JP, Wilson RK, Ganem D, DeRisi JL (2003) Viral discovery and sequence recovery using DNA Microarrays. PLoS Biology 1:257–260
24. Wolf VH, Menossi M, Mourão GB, Gatti MS (2008) Molecular basis for porcine parvovirus detection in dead fetuses. Genet Mol Res 7:509–517