Evidences for recombination in natural populations of porcine circovirus type 2 in Hong Kong and mainland China

[1]Ching-Man Ma, [1]Chung-Chau Hon, [1]Tsan-Yuk Lam, [1]Vince Yik-Yeung Li, [1]Creany Ka-Wai Wong, [2]Tulio de Oliveira, and [*1]Frederick Chi-Ching Leung

[1]Department of Zoology, The University of Hong Kong, HKSAR, China.
[2]Department of Zoology, The University of Oxford, UK

* To whom correspondence should be addressed: Dr. Frederick C.C. Leung

Postal address: Department of Zoology, Kadoorie Biological Science Building, The University of Hong Kong, China

Telephone: (852) 2299-0825

Fax: (852) 2857-4672

E-mail: fcleung@hkucc.hku.hk

Manuscript Information:

Number of words (summary): 144 words

Number of words (main text, including figures and legends): 2500 words

Number of figures: 3 figures plus 1 supplementary figure

Number of tables: 2 supplementary tables

Number of text pages: 20 pages

Authorship:

(i) All the authors have agreed to its submission and are responsible for its contents.

(ii) All the authors have agreed that the corresponding authors may act on their behalf regarding any subsequent processing of the paper.

Footnotes:

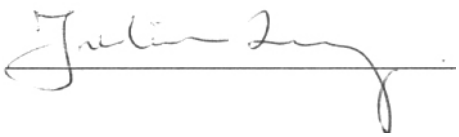
Author contributions: C.M.M. and F.C.C.L. designed research; C.M.M. and C.C.H. wrote the paper; C.M.M., V.Y.Y.L. and C.K.W.W. performed molecular cloning and sequencing; de Oliveira, T. advised the recombination analyses and performed the bootscanning analysis; C.M.M., C.C.H., and T.Y.L. performed the phylogenetic analysis.

Note:

Our paper focuses mainly on recombination in natural PCV2 populations of Hong Kong and mainland China, including extensive statistical analyses on estimations of recombination breakpoints and parental strains. This is a different scope to that of the recent paper in Virology "Molecular evolution of porcine circovirus type 2 genomes: Phylogeny and clonality" (Olvera *et al*., 2007), which contains a preliminarily report on the potential recombination events among GenBank PCV2 sequences.

Conflict of interest statement: No conflicts declared.

Signature of the correspondence author:



Summary:

Porcine circovirus type 2 (PCV2) belongs to the family *Circoviridae*, which is the causative agent of postweaning multisystemic wasting syndrome (PMWS) in pigs. In this study, phylogenetic analyses of three PCV2 complete genomic sequences from Hong Kong suggest that natural recombination happened among different lineages of PCV2. A preliminary investigation of the parental strains of these potential recombinants was carried out using bootscanning Statistical significance of this recombination event was tested and positions of the potential recombination breakpoints were estimated in a maximum likelihood framework. The recombinant breakpoints were estimated to be located within the origin of replication (*ori*) and replicase (*rep*) gene of PCV2. Interestingly, several GenBank sequences of PCV2 in mainland China were found to have a recombination pattern similar to that of the potential PCV2 recombinants from Hong Kong, implying this recombinant genotype might have already been widespread within mainland China.

Main text:

Porcine circovirus (PCV), which belongs to the family *Circoviridae*, is a non-enveloped, single-stranded circular DNA virus with genome of about 1.7 kb. Its genome includes a replicase (*rep*) and a capsid (*cap*) gene for the 2 major open reading frames, ORF1 and ORF2, respectively. Recently, a virus-induced apoptotic gene, ORF3, was newly identified (Liu *et al*., 2005). PCV was classified into PCV type 1 (PCV1) and PCV type 2 (PCV2) based on their genotypes. Serologic tests suggested that PCV1 is widely found in pigs but no evidence has been associated PCV1 to animal diseases (Tischer *et al*., 1987). However, PCV2 is now considered as the causative agent of postweaning multisystemic wasting syndrome (PMWS), which was first identified in Canada during the early 1990s (Harding, 2004). PMWS has become a major cause of wasting disease in swine producing countries including Asia, North America and Europe (Chae, 2004, Liu *et al*., 2002, Mankertz *et al*., 2000, Wen *et al*., 2005).

Point mutation and recombination are major forces of viral evolution. Analysis of the genomes of worldwide PCV2 showed a very high nucleotide sequence identity (> 90%) among them (Fenaux *et al*., 2000, Hamel *et al*., 1998).

Nevertheless, evidences of natural recombination have been reported in members of the family *Circoviridae,* including beak and feather disease virus (Heath *et al*., 2004) and torque teno virus (Manni *et al*., 2002), implying recombination may also contribute to the genetic variations within the family *Circoviridae*. Recently, natural recombination in PCV2 has been proposed in Hungarian wild boars (Csagola *et al*., 2006) and an analysis of worldwide PCV2 GenBank sequences (Olvera *et al*., 2007), though further phylogenetic evidences are needed to validate these hypothesises. Herein, we presented a comprehensive analysis of PCV2 complete genomic sequences from Hong Kong, providing strong evidences of natural recombination among PCV2 of different lineages.

The complete genomic sequences of three PCV2 strains from Hong Kong, namely HKS02-04, HKS03-04 and HKS091-04 (collectively designated HKS04 viruses), with GenBank accession numbers DQ997815, DQ997816 and DQ997817 were reported. These viruses were found in pigs showing PMWS symptoms from three different herds during 2004. Viral DNA was extracted from 500µl of serum using TRI Reagent (Invitrogen) according to manufacturer's instructions. The DNA pellet was then dissolved in 300µl of 8mM NaOH and subjected to PCR

amplification with Platinum High Fidelity Taq DNA Polymerase (Invitrogen) according to manufacturer's instructions. The complete genomes were amplified using primers F-PCVSAC2 and R-PCVSAC2 (Fenaux *et al*., 2002) and the amplicons were cloned into pCR2.1-TOPO (Invitrogen). To avoid misleading results caused by PCR artefacts, five random clones were sequenced for each of the viruses, using M13 universal forward and reserve primers, as well as two internal primers CV1 and CV4 (Fenaux *et al*., 2000). DNA sequencing was performed by using Big Dye terminator cycle sequencing kit (USA Scientific) according to manufacturer's instructions. A limited number of non-recurrent mutations were found among different clones of the same virus, which is thought to be PCR generated sequence mutations. Therefore, the genomic sequences of each virus were represented by the majority consensus of the five clones. The complete genomes of HKS04 viruses were all found to be 1767bp and shared over 99% in nucleotide sequence identity.

To investigate the phylogenetic origin of HKS04 viruses, all available PCV2 complete genomic sequences were downloaded from GenBank. Sequences were screened to exclude patent, artificial, defective and potential PCV2 recombinants

as proposed in Hungary (Csagola *et al*., 2006). The complete genomic sequences (n = 164), as shown in Table S1 of supplementary materials, were aligned using ClustalX with one gap column removed (Thompson *et al*., 1994). A ML phylogeny of 500 bootstrap replicates was constructed based on the complete genomic sequences of worldwide PCV2 strains using PHYML (http://atgc.lirmm.fr/phyml/) under the General Time Reversible nucleotide substitution model (GTR model) (Yang, 1994) with estimated gamma and invariable-site values, and this optimal model was selected using MODELTEST (Posada & Crandall, 1998). Seven well-supported major lineages (bootstrap values > 80%) were identified and designated lineages A, B, C, D, E, F and G respectively (Fig. 1). The mean Kimura's two-parameter (K2P) distance of all taxa is 0.03, while the intra-lineage K2P distance has a mean of 0.01, ranging from 0.008 to 0.015. Preliminary analyses of different genomic regions of HKS04 viruses showed discordant phylogenetic relationships with other PCV2 strains. For instance, HKS04 viruses were clustered within lineage A based on complete genome and ORF2 sequences but they were clustered out of lineage A when ORF1 sequences were considered (data not shown). These conflicting phylogenies implied the possible recombinant origin of HKS04 viruses.

In order to identify the possible parental sequences of HKS04 viruses, an integrated software package, Recombination Detection Program (RDP) was used (Martin & Rybicki, 2000). Involvement of large number of sequences in the RDP analysis results in noisy recombination signals. As the intra-lineage genetic variations are relatively low, i.e. average pairwise K2P distance ranging from 0.008 to 0.015, the RDP analysis was done using phylogenetically distant strains (n=33) from selected from different lineages. Their information is shown in Table S2 of supplementary materials. With the supports of several recombination detection methods implemented in RDP, including Geneconv, RDP and MaxChi methods (cut-off *P*-value of 0.001 with Bonferroni correction), parental strains of HKS04 viruses possibly come from lineage A and lineage E or F. It is noted that the mean genetic distances between lineage A/ lineage E and lineage A/ lineage F, are 0.042 and 0.045 respectively. The specific settings used in each method in RDP are available upon requests. To further investigate this potential recombination event, a recombination detection method called bootscanning analysis implemented in SIMPLOT (Lole *et al*., 1999) was performed on the dataset. Briefly, a set of NJ phylogenies were generated to show the clustering of the query sequences (potential recombinants) and reference sequences (parents) in

a moving window along the alignment with user-defined number of bootstrap replications. In particular, we performed a bootscanning analysis (Fig. 2a) on four groups of PCV2 complete genomic sequences, including a group of potential recombinants, i.e., HKS04 viruses (n = 3), two groups of reference sequences from lineage A (P1, n = 3, accession no. AY682990, AY691679 and AY732494) and lineage F (P2, n = 3, accession no. AY424401, AY424402 and AY424403) and an outgroup from lineage D (O2, n = 2, accession no. AF166528 and AY146991). The bootscanning analysis was performed under the K2P model for 500 replications, with a window size of 400bp, step size of 10bp and transitions / transversion ratio of 2. The result of bootscanning analysis suggested discordances in the phylogenetic origins of the HKS04 viruses, supporting the hypothesis that HKS04 viruses might be a recombinant from different PCV2 lineages.

To this end, we further confirmed this potential recombination event and identified breakpoints using a maximum likelihood (ML) method with statistical significance as described previously (Holmes *et al*., 1999). In brief, an alignment of sequences of two potential parents and a potential recombinant was divided by a partition (breakpoint) and two separated ML phylogenies were constructed for

each of the two regions. Likelihood score of this "recombination model" with a particular breakpoint was combined from the likelihood scores of the two ML phylogenies. This procedure was repeated for all possible breakpoints, which is implemented in LARD (Holmes *et al*., 1999). The breakpoint with highest combined likelihood score was then considered to be the most possible breakpoint. To assess whether the recombination model gave a significantly better fit to the data than the null hypothesis of no recombination, the combined likelihood score of the best "recombination model" was compared with the likelihood of ML phylogeny of the unbroken alignment, i.e. null hypothesis of no recombination, using likelihood ratio test (LRT) (Holmes *et al*., 1999). Simultaneous estimation of multiple breakpoints, e.g., two breakpoints in this case, is computationally intensive. Therefore, the complete genome alignment was divided into two overlapping alignments and only one breakpoint was estimated for each of the trimmed alignments. The trimmed alignments are available upon requests. Due to the fact that HKS04 viruses share over 99% sequence identify, only the LARD results of HKS02-04 were shown for simplicity. Breakpoints of HKS02-04 were estimated to be at nt 1737 and nt 391, designated breakpoint A and B respectively (Fig. 2b), and the estimated breakpoints in other HKS04 viruses were similar ($P <$

1 0.005 in LRT of all cases, data not shown). All nucleotide positions in this
2 manuscript were assigned according to strain NL_PMWS_4 (accession no.
3 AY484416).
4
5 To further assess the statistical significance of these putative recombination
6 events, the likelihood ratios of the our datasets were evaluated against the null
7 distributions of likelihood ratios of the 1000 simulated datasets assuming no
8 recombination (Holmes *et al*., 1999). The simulated datasets were generated using
9 Seq-Gen (Rambaut & Grassly, 1997), with the maximum likelihood model
10 parameters and sequence lengths from the corresponding real dataset. The
11 likelihood ratios of the real and simulated datasets were generated using the
12 breakpoint analysis in LARD as mentioned. The likelihood ratios for all of
13 putative breakpoints in HKS04 viruses were greater than any of the likelihood
14 ratios of the corresponding simulated datasets (Fig. S1 of supplementary
15 materials), implying the discordant phylogenetic relationships for different
16 genomic regions of HKS04 viruses were unlikely to be the result of chance.
17
18 To investigate the phylogenetic origins of the parents of HKS04 viruses, two

ML phylogenies were constructed based on the major and minor parental regions of 33 representative PCV2 strains (Fig. 3). The phylogenies were constructed with PHYML (http://atgc.lirmm.fr/phyml/) under the GTR model with 500 bootstrap replications. In the phylogeny based on major parental region, HKS04 viruses were clustered within lineage A with high bootstrap support. On the other hand, in the phylogeny based on minor parental region, HKS04 viruses were clustered with PCV2 from lineage E and lineage F with high bootstrap support, despite the general low bootstrap support of the internal nodes in this phylogeny. In addition, three PCV2 strains (accession no. AY691169, DQ180393 and DQ195679) were showed to have a recombination pattern and phylogenetic origins similar to those of the HKS04 viruses (Fig. 3). Taken together, the above results suggested that HKS04 viruses might be resulted from a recombination event between parental PCV2 that were phylogenetically close to lineage A, and lineage E or lineage F respectively.

Previous studies on sequences of PCV1 showed that it has affinities with plant circoviruses (Meehan *et al*., 1997). Circoviruses were further proposed to be resulted from recombination between a plant nanovirus and a vertebrate infecting

virus, e.g., a single-strand RNA virus calicivirus, during a host-switch event (Gibbs & Weiller, 1999). Moreover, extensive natural recombinations between strains of another member beak and feather disease virus from the family *Circoviridae* have been documented (Heath *et al.*, 2004). These results suggested the possibility of natural recombination in other circoviruses. In the present study, the recombination breakpoints were predicted to be located within the *ori* and *rep* gene (ORF1). Within the *ori* of PCV, there is a stem loop structure flanked by a pair of palindromic sequences. It has been proposed that PCV1 replicated via a rolling-circle melting-pot replication model. This provided insights into the illegitimate recombination of any circular DNAs with an origin-flanking palindrome (Cheung, 2004). Whether the chance of recombination is enhanced with the origin-flanking palindrome remains to be determined.

Until now, there is still no conclusive correlation of PCV2 strains to particular PCV2-assoicated diseases such as PMWS and porcine dermatitis and nephropathy syndrome (Larochelle *et al*., 2002). It has been suggested that all PCV2 strains belong to a single pathogenic genotype (Chae, 2005) and genetic variations of PCV2 might be associated with geographic origins (Fenaux *et al*.,

2000, Meehan *et al*., 2001). Based on our analysis of the worldwide PCV2 complete genomes, correlations of the five major lineages with particular geographic origins were not observed. Interestingly, our results suggested that recombination event(s) happen in PCV2 natural populations. In particular, the common ancestor of HKS04 viruses in this study was resulted from a recombination event between a major parent of lineage A and a minor parent of lineage E or lineage F. Our result is unlikely a laboratory artefact because three HKS04 viruses were cloned and sequenced individually at different time periods. Furthermore, several PCV2 strains from different provinces of China (accession no. AY691169, DQ180393 and DQ195679) were shown to have a similar recombination pattern. As shown in figure 3, the genomic regions of these viruses shared same phylogenetic origins with the corresponding genomic regions of the HKS04 viruses, suggesting HKS04 and its related viruses might be resulted from a single recombination event and this recombinant genotype might have already been widespread within China.

Recombination requires the simultaneous infection of a single cell by two different viral strains. In pigs infected with PCV2, the concurrent coinfections

with other viral and bacterial agents, such as porcine reproductive and respiratory syndrome virus, swine influenza virus, porcine parvovirus, *Haemophilus parasuis, Actinobacillus pleuropneumoniae, Streptococcus suis* and *Mycoplasma hyopneumoniae* have been frequently detected (Kim & Lyoo, 2002). However, the detection of coinfections with divergent PCV2 strains has not been reported. The facts that PCV1 and PCV2 share a highly homologous genomic organization and coinfections are reported (Calsamiglia *et al*., 2002), this raised the possibility of recombination between PCV1 and PCV2. To conclude, the present study provides strong statistical evidences on the presence of natural recombination among different lineages of PCV2. Our result also suggests cautions have to be taken on the interpretation of PCV2 phylogenies based on a single partial genomic region.

Acknowledgements:

This work was supported by a contract research grant from the Agricultural Fisheries and Conservation Department of the The Hong Kong Government,. We are grateful to Dr. M. Mackett for his expert editing the manuscript.

Figure legends:

Fig. 1. Unrooted ML phylogeny of 164 worldwide PCV2 complete genomic sequences. Nodes of different lineages are labelled with their bootstrap values, while the bootstrap values of internal nodes are omitted for simplicity. The scale bar represents the nucleotide substitution per site under the GTR model. Information of all sequences can be found in Table S1 of supplementary materials.

Fig. 2. Recombination analysis of HKS04 viruses. (a) Bootscanning analysis of HKS04 viruses with representative strains of lineage A (P1), lineage F (P2) and lineage D as an outgroup (O2). (b) Schematic diagram showing the recombination pattern of HKS04 viruses. PCV2 strain NL_PMWS_4 (accession no. AY484416) is used as reference for the annotation. Recombination breakpoints A and B of HKS02-04 are estimated using LARD. The minor parent region is indicated by a black color. The ORF1, ORF2 and *ori* are indicated by grey and white colors respectively.

Fig. 3. Phylogenetic origins of the putative major and minor parents of HKS04 viruses. ML phylogeny based on (a) the major parent region, i.e. concatenated alignment of nt 1033-1737 and nt 392-1032; and (b) minor parent region, i.e., nt 1738-391, of HKS04 viruses. Discordant clustering patterns of HKS04 viruses (black dots) with different lineages are observed and other potential PCV2 recombinants are labelled with grey dots (refer to text). Nodes of different lineages with bootstrap values > 80% were represented by asterisks. Representative PCV2 strains are shown with both accession numbers and geographical origins. Aut, Austria; Can, Canada; Chi, China; Fr, France; Hg, Hungary; Jp, Japan; Kor, Korea; Spa, Spain; Tw, Taiwan.

References:

Calsamiglia, M., Segalés, J., Quintana, J., Rosell, C. & Domingo, M. (2002). Detection of porcine circovirus type 1 and 2 in semen and tissue samples of pigs with and without postweaning multisystemic wasting syndrome. *J Clin Microbiol* **40**, 1848-1850.

Chae, C. (2004). Postweaning multisystemic wasting syndrome: a review of aetiology, diagnosis and pathology. *Vet J* **168**, 41-9.

Chae, C. (2005). A review of porcine circovirus 2-associated syndromes and diseases. *Vet J* **169**, 326-336.

Cheung, A. K. (2004). Palindrome regeneration by template strand-switching mechanism at the origin of DNA replication of porcine circovirus via the rolling-circle melting-pot replication model. *J Virol* **78**, 9016-29.

Csagola, A., Kecskemeti, S., Kardos, G., Kiss, I. & Tuboly, T. (2006). Genetic characterization of type 2 porcine circoviruses detected in Hungarian wild boars. *Arch Virol* **151**, 495-507.

Fenaux, M., Halbur, P. G., Gill, M., Toth, T. E. & Meng, X. J. (2000). Genetic characterization of type 2 porcine circovirus (PCV-2) from pigs with postweaning multisystemic wasting syndrome in different geographic regions of North America and development of a differential PCR-restriction fragment length polymorphism assay to detect and differentiate between infections with PCV-1 and PCV-2. *J Clin Microbiol* **38**, 2494-503.

Fenaux, M., Halbur, P. G., Haqshenas, G., Royer, R., Thomas, P., Nawagitgul, P., Gill, M., Toth, T. E. & Meng, X. J. (2002). Cloned genomic DNA of type 2 porcine circovirus is infectious when injected directly into the liver and lymph nodes of pigs: characterization of clinical disease, virus distribution, and pathologic lesions. *J Virol* **76**, 541-551.

Gibbs, M. J. & Weiller, G. F. (1999). Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. *Proc Natl Acad Sci U S A* **96**, 8022-7.

Hamel, A. L., Lin, L. L. & Nayar, G. P. (1998). Nucleotide sequence of porcine circovirus associated with postweaning multisystemic wasting syndrome in pigs. *J Virol* **72**, 5262-7.

Harding, J. C. (2004). The clinical expression and emergence of porcine circovirus 2. *Vet Microbiol* **98**, 131-5.

Heath, L., Martin, D. P., Warburton, L., Perrin, M., Horsfield, W., Kingsley, C., Rybicki, E. P. & Williamson, A. L. (2004). Evidence of unique genotypes of beak and feather disease virus in southern Africa. *J Virol* **78**, 9277-84.

Holmes, E. C., Worobey, M. & Rambaut, A. (1999). Phylogenetic evidence for recombination in dengue virus. *Mol Biol Evol* **16**, 405-9.

Kim, J. H. & Lyoo, Y. S. (2002). Genetic characterization of porcine circovirus-2 field isolates from PMWS pigs. *J Vet Sci* **3**, 31-9.

Larochelle, R., Magar, R. & D'Allaire, S. (2002). Genetic characterization and phylogenetic analysis of porcine circovirus type 2 (PCV2) strains from cases presenting various clinical conditions. *Virus Res* **90**, 101-12.

Liu, J., Chen, I. & Kwang, J. (2005). Characterization of a previously unidentified viral protein in porcine circovirus type 2-infected cells and its role in virus-induced apoptosis. *J Virol* **79**, 8262-74.

Liu, Q., Wang, L., Willson, P., O'Connor, B., Keenliside, J., Chirino-Trejo, M., Melendez, R. & Babiuk, L. (2002). Seroprevalence of porcine circovirus type 2 in swine populations in Canada and Costa Rica. *Can J Vet Res* **66**, 225-31.

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W. & Ray, S. C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* **73**, 152-60.

Mankertz, A., Domingo, M., Folch, J. M., LeCann, P., Jestin, A., Segales, J., Chmielewicz, B., Plana-Duran, J. & Soike, D. (2000). Characterisation of PCV-2 isolates from Spain, Germany and France. *Virus Res* **66**, 65-77.

Manni, F., Rotola, A., Caselli, E., Bertorelle, G. & Di Luca, D. (2002). Detecting recombination in TT virus: a phylogenetic approach. *J Mol Evol* **55**, 563-72.

Martin, D. & Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *Bioinformatics* **16**, 562-3.

Meehan, B. M., Creelan, J. L., McNulty, S. & Todd, D. (1997). Sequence of porcine circovirus DNA: affinities with plant circoviruses. *J Gen Virol* **78**, 221-227.

Meehan, B. M., McNeilly, F., McNair, I., Walker, I., Ellis, J. A., Krakowka, S. & Allan, G. M. (2001). Isolation and characterization of porcine circovirus 2

from cases of sow abortion and porcine dermatitis and nephropathy syndrome. *Arch Virol* **146**, 835-42.

Olvera, A., Cortey, M. & Segales, J. (2007). Molecular evolution of porcine circovirus type 2 genomes: phylogeny and clonality. *Virology* **357**, 175-85.

Posada, D. & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817-8.

Rambaut, A. & Grassly, N. C. (1997). Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comput Appl Biosci* **13**, 235-8.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680.

Tischer, I., Peters, D., Rasch, R. & Pociuli, S. (1987). Replication of porcine circovirus: induction by glucosamine and cell cycle dependence. *Arch Virol* **96**, 39-57.

Wen, L., Guo, X. & Yang, H. (2005). Genotyping of porcine circovirus type 2 from a variety of clinical conditions in China. *Vet Microbiol* **110**, 141-6.

Yang, Z. (1994). Estimating the pattern of nucleotide substitution. *J Mol Evol* **39**, 105-11.

Supplementary materials for the manuscripts entitled *"Evidences for recombination in natural populations of porcine circovirus type 2 in Hong Kong and mainland China"*, includes 2 tables and 1 figure.

Tables:

Table S1. Information of the 164 PCV2 strains used for classification analysis

*Strain name	Accession number	Geographic origin	† Lineage	Reference
FRA3	AF201311	France	A	Mankertz et al., 2000
24657 NL	AF201897	Netherlands	A	Wellenberg et al., unpublished
BJ-HB	AF538325	China	A	Wang et al. unpublished
JX/China/04	AM086384	China	A	Yun et al., unpublished
kaozhai	AY122275	China	A	Cui et al., unpublished
PCV2	AY177626	China	A	Wang et al., unpublished
GD-TS	AY181945	China	A	Wang et al., unpublished
HZ0201	AY188355	China	A	Zhou et al., 2005
HZ0202	AY217743	China	A	Zhou et al., unpublished
304	AY256457	Hungary	A	Dan et al., 2003
375	AY256460	Hungary	A	Dan et al., 2003
S2	AY288133	China	A	Cui et al., 2003
P11	AY288134	China	A	Cui et al., 2003
QD	AY291316	China	A	Xin et al., unpublished
SH	AY291318	China	A	Xin et al., unpublished
Fh14	AY321982	France	A	de Boisseson et al., 2004
Fh20	AY321983	France	A	de Boisseson et al., 2004
Fd3	AY321984	France	A	de Boisseson et al., 2004
Fd13	AY321985	France	A	de Boisseson et al., 2004
Fd4	AY321986	France	A	de Boisseson et al., 2004
Fh18	AY321987	France	A	de Boisseson et al., 2004
Fd11	AY321988	France	A	de Boisseson et al., 2004
Fd12	AY321989	France	A	de Boisseson et al., 2004
Fd7	AY321990	France	A	de Boisseson et al., 2004
Fd10	AY321991	France	A	de Boisseson et al., 2004
Fh15	AY321992	France	A	de Boisseson et al., 2004
Fh23	AY321994	France	A	de Boisseson et al., 2004
Fd8	AY321995	France	A	de Boisseson et al., 2004
Fd9	AY321996	France	A	de Boisseson et al., 2004
Fd5	AY321997	France	A	de Boisseson et al., 2004
Fd6	AY321998	France	A	de Boisseson et al., 2004
Fd2	AY321999	France	A	de Boisseson et al., 2004

*Strain name	Accession number	Geographic origin	† Lineage	Reference
Fd1	AY322000	France	A	de Boisseson et al., 2004
Fh21	AY322001	France	A	de Boisseson et al., 2004
Fh22	AY322002	France	A	de Boisseson et al., 2004
Fh19	AY322003	France	A	de Boisseson et al., 2004
NB0301	AY391729	China	A	Zhou et al., unpublished
AUT4	AY424404	Austria	A	Exel et al., unpublished
AUT5	AY424405	Austria	A	Exel et al., unpublished
NL_Control_1	AY484407	Netherlands	A	Grierson et al., 2004
NL_Control_2	AY484408	Netherlands	A	Grierson et al., 2004
NL_Control_3	AY484409	Netherlands	A	Grierson et al., 2004
NL_Control_5	AY484411	Netherlands	A	Grierson et al., 2004
NL_Control_6	AY484412	Netherlands	A	Grierson et al., 2004
NL_PMWS_1	AY484413	Netherlands	A	Grierson et al., 2004
NL_PMWS_2	AY484414	Netherlands	A	Grierson et al., 2004
NL_PMWS_3	AY484415	Netherlands	A	Grierson et al., 2004
NL_PMWS_4	AY484416	Netherlands	A	Grierson et al., 2004
SX0201	AY536755	China	A	Zhou et al., unpublished
HuZhou0301	AY536756	China	A	Zhou et al., unpublished
GX	AY556475	China	A	Zhixin et al., unpublished
JS2003	AY578327	China	A	Yu & Zhang, unpublished
ZhuJi2003	AY579893	China	A	Yu & Zhang, unpublished
ZS	AY596823	China	A	Da et al., unpublished
SX04	AY604430	China	A	Li et al., unpublished
GD	AY613854	China	A	Song et al., unpublished
JH0401	AY641542	China	A	Zhou et al., unpublished
JX0301	AY651850	China	A	Zhou et al., unpublished
ZS0401	AY678532	China	A	Zhou et al., unpublished
BL	AY682990	China	A	Wang et al., unpublished
CHST	AY682992	China	A	Wang et al., unpublished
DG	AY682993	China	A	Wang et al., unpublished
QY	AY682995	China	A	Wang et al., unpublished
ZC	AY682997	China	A	Wang et al., unpublished
JXI	AY686762	China	A	Feng et al., unpublished
ZJ	AY686764	China	A	Feng et al., unpublished
QZ0401	AY691169	China	A	Zhou et al., unpublished
JS	AY691679	China	A	Feng et al., unpublished
JXII	AY732494	China	A	Feng et al., unpublished

*Strain name	Accession number	Geographic origin	† Lineage	Reference
BJW	AY847748	China	A	Liu et al., 2005
Changsha	AY849938	China	A	Huang, unpublished
HD	AY916791	China	A	Jiang et al., unpublished
Henan	AY969004	China	A	Liu et al., unpublished
GD-ZJ	DQ017036	China	A	Song et al., unpublished
QDC	DQ104420	China	A	Lu et al., unpublished
XSC	DQ104422	China	A	Lu et al., unpublished
SD1	DQ141322	China	A	Liu et al., unpublished
fuqing0401	DQ180392	China	A	Zhou et al., unpublished
putian0401	DQ180393	China	A	Zhou et al., unpublished
PCV2	DQ195679	China	A	Shuai & Fang, unpublished
SD2	DQ201639	China	A	Liu et al., unpublished
SD4	DQ201640	China	A	Liu et al., unpublished
DZ	DQ201641	China	A	Liu et al., unpublished
WZ04	DQ201642	China	A	Liu et al., unpublished
JZ	DQ206444	China	A	Liu et al., unpublished
SD3	DQ218419	China	A	Liu et al., unpublished
SD5	DQ218420	China	A	Liu et al., unpublished
SD6	DQ218421	China	A	Liu et al., unpublished
ROM	DQ233257	China	A	Cadar et al., unpublished
TJ	AY181946	China	B	Wang et al., unpublished
SD	AY181947	China	B	Wang et al., unpublished
HB	AY291317	China	B	Xin et al., unpublished
NL_Control_4	AY484410	Netherlands	B	Grierson et al., 2004
HZ0301	AY510375	China	B	Zhou et al., unpublished
SD	AY556473	China	B	Zhixin et al., unpublished
HaiNan	AY556476	China	B	Zhixin et al., unpublished
CHL	AY682991	China	B	Wang et al., unpublished
GZ	AY682994	China	B	Wang et al., unpublished
ST	AY682996	China	B	Wang et al., unpublished
SH	AY686763	China	B	Feng et al., unpublished
JXIII	AY686765	China	B	Feng et al., unpublished
PCV2	AY713470	Germany	B	Knell et al., 2005
YZH	AY943819	China	B	Yu et al., unpublished
No.26	AB072302	Japan	C	Imai et al., unpublished
2-C	AF109398	Canada	C	Hamel et al., 2000
2-D	AF117753	Canada	C	Hamel et al., 2000

*Strain name	Accession number	Geographic origin	† Lineage	Reference
Aust 5	AY754016	Australia	C	Muhling et al., unpublished
Aust 6	AY754017	Australia	C	Muhling et al., unpublished
Aust 7	AY754018	Australia	C	Muhling et al., unpublished
Aust 8	AY754019	Australia	C	Muhling et al., unpublished
Aust 9	AY754020	Australia	C	Muhling et al., unpublished
Aust 10	AY754021	Australia	C	Muhling et al., unpublished
Tainan	AF166528	Taiwan	D	Yang et al., unpublished
V00-52-2	AF305532	Taiwan	D	Weng et al., unpublished
V00-207-2	AF305533	Taiwan	D	Weng et al., unpublished
Chia-Yi	AF364094	Taiwan	D	Wang et al., unpublished
Pingtung-1	AY146991	Taiwan	D	Liao et al., unpublished
Pingtung-3	AY146993	Taiwan	D	Liao et al., unpublished
Pingtung-4	AY180396	Taiwan	D	Liao et al., unpublished
Pingtung-5	AY180397	Taiwan	D	Liao et al., unpublished
SPA1	AF201308	Spain	E	Mankertz et al., 2000
SPA2	AF201309	Spain	E	Mankertz et al., 2000
SPA3	AF201310	Spain	E	Mankertz et al., 2000
212	AY256455	Hungary	E	Dan et al., 2003
336	AY256459	Hungary	E	Dan et al., 2003
2-E	AF109399	Canada	F	Hamel et al., 2000
GER1	AF201305	Germany	F	Mankertz et al., 2000
GER2	AF201306	Germany	F	Mankertz et al., 2000
34464	AF264043	USA	F	Fenaux et al., 2000
HR	AF381176	China	F	Lu & Yang., unpublished
224	AY256456	Hungary	F	Dan et al., 2003
326	AY256458	Hungary	F	Dan et al., 2003
Fh17	AY322004	France	F	de Boisseson et al., 2004
L	AY288135	China	F	Cui et al., 2003
AUT1	AY424401	Austria	F	Exel et al., unpublished
AUT2	AY424402	Austria	F	Exel et al., unpublished
AUT3	AY424403	Austria	F	Exel et al., unpublished
PCV2	NC_005148	Austria	F	Exel et al., unpublished
No.33	AB072301	Japan	G	Imai et al., unpublished
No. 35	AB072303	Japan	G	Imai et al., unpublished
2-B	AF112862	Canada	G	Hamel et al., 2000
IAF-614	AF118095	Canada	G	Ouardani et al., 1999
IAF-4370	AF118097	Canada	G	Ouardani et al., 1999

*Strain name	Accession number	Geographic origin	† Lineage	Reference
ISUVDL 98-15237	AF147751	USA	G	Pogranichnyy et al., unpublished
26606	AF264038	USA	G	Fenaux et al., 2000
26607	AF264039	USA	G	Fenaux et al., 2000
10489	AF264040	USA	G	Fenaux et al., 2000
40856	AF264041	USA	G	Fenaux et al., 2000
40895	AF264042	USA	G	Fenaux et al., 2000
BF	AF381175	China	G	Lu & Yang, unpublished
BX	AF381177	China	G	Lu & Yang, unpublished
IAF2897	AF408635	Canada	G	Ouardani & Dea, unpublished
KSY-1	AF454546	Korea	G	Kim et al., unpublished
SC	AF465211	Taiwan	G	Wang et al., unpublished
JHP	AF520783	Korea	G	Park et al., unpublished
KSY-2	AF544024	Korea	G	Kim et al., unpublished
PCV2	AY094619	USA	G	Cheung, 2003
SZ	AY181948	China	G	Wang et al., unpublished
SA1	AY325495	South Africa	G	Drew et al., unpublished
PCV2	AY699793	USA	G	Fenaux et al., 2004
Aust 11	AY754022	Australia	G	Muhling et al., unpublished
JSC	DQ104419	China	G	Lu et al., unpublished
SHC	DQ104421	China	G	Lu et al., unpublished
DTC	DQ104423	China	G	Lu et al., unpublished

* Strain names not given in GenBank were designated PCV2.

† The lineages were assigned according to figure 1.

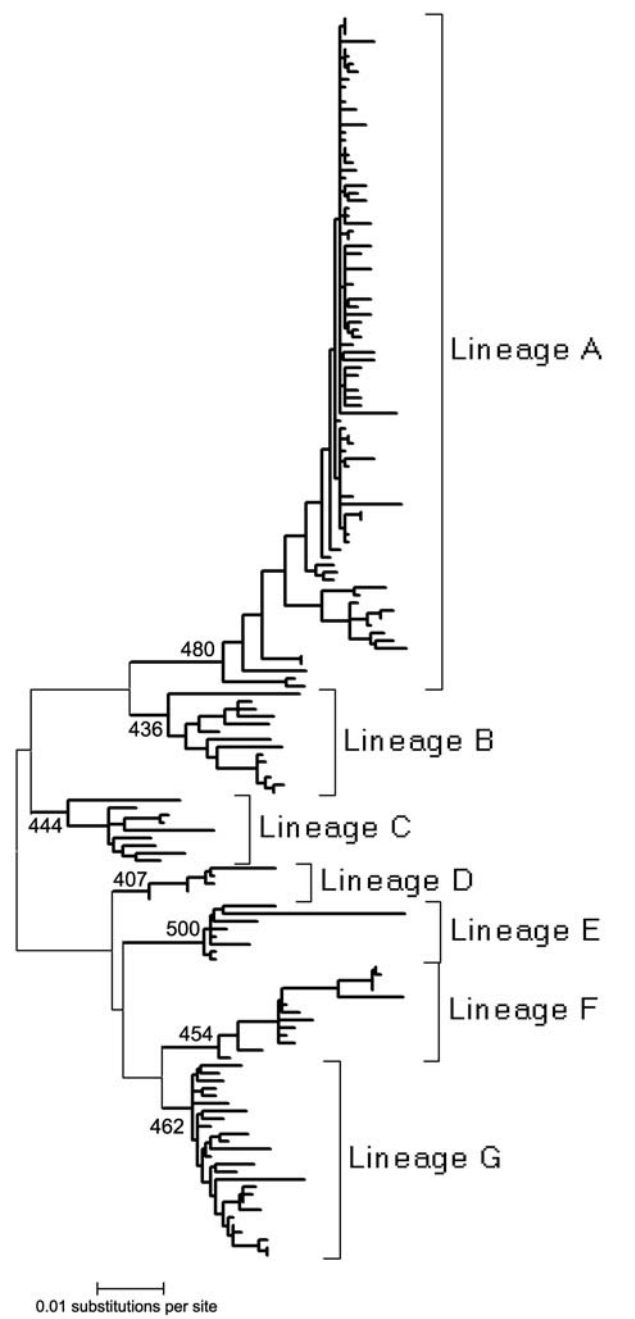
Table S2. Information of the 33 PCV2 strains used for recombination and phylogenetic analyses

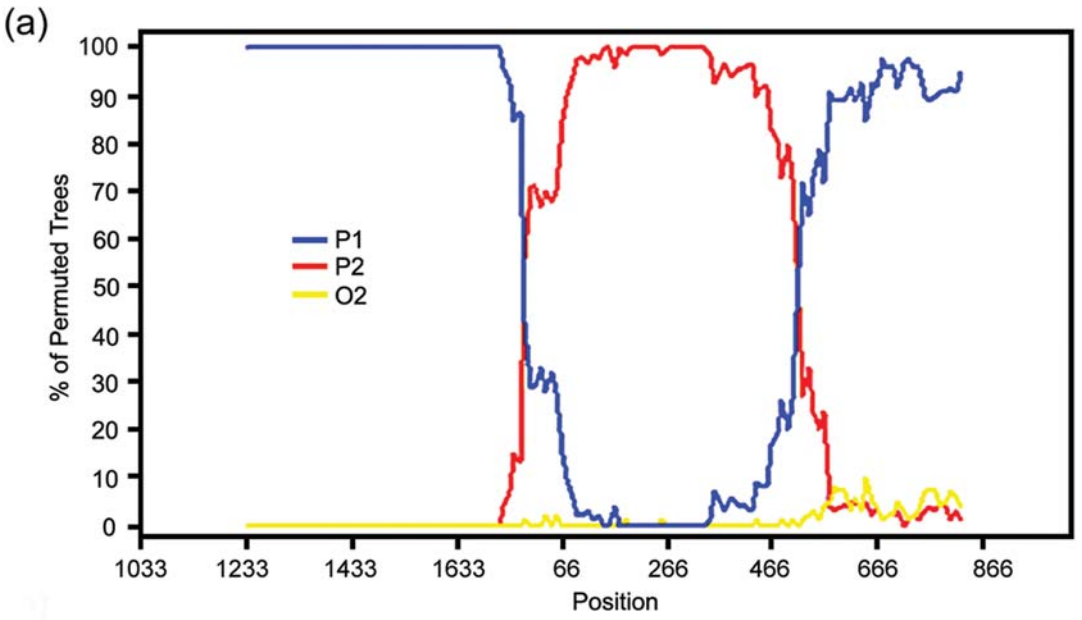
Abbreviation	Accession number	Geographic origin	† Lineage	Reference
AY288134 Chi	AY288134	China	A	Cui et al., 2003
AY321995 Fr	AY321995	France	A	de Boisseson et al., 2004
AY321996 Fr	AY321996	France	A	de Boisseson et al., 2004
AY322000 Fr	AY322000	France	A	de Boisseson et al., 2004
AY322001 Fr	AY322001	France	A	de Boisseson et al., 2004
AY322002 Fr	AY322002	France	A	de Boisseson et al., 2004
AY604430 Chi	AY604430	China	A	Li et al., unpublished
AY682990 Chi	AY682990	China	A	Wang et al., unpublished
AY691169 Chi	AY691169	China	A	Zhou et al., unpublished
AY691679 Chi	AY691679	China	A	Feng et al., unpublished
AY732494 Chi	AY732494	China	A	Feng et al., unpublished
DQ180393 Chi	DQ180393	China	A	Zhou et al., unpublished
DQ195679 Chi	DQ195679	China	A	Shuai & Fang., unpublished
HKS02-04	DQ997815	China	A	this paper
HKS03-04	DQ997816	China	A	this paper
HKS091-04	DQ997817	China	A	this paper
AF166528 Tw	AF166528	Taiwan	D	Yang et al., unpublished
AY146991 Tw	AY146991	Taiwan	D	Liao et al., unpublished
AF201308 Spa	AF201308	Spain	E	Mankertz et al., 2000
AF201309 Spa	AF201309	Spain	E	Mankertz et al., 2000
AF201310 Spa	AF201310	Spain	E	Mankertz et al., 2000
AY256455 Hg	AY256455	Hungary	E	Dan et al., 2003
AY256459 Hg	AY256459	Hungary	E	Dan et al., 2003
AY288135 Chi	AY288135	China	F	Cui et al., 2003
AY424401 Aut	AY424401	Austria	F	Exel et al., unpublished
AY424402 Aut	AY424402	Austria	F	Exel et al., unpublished
AY424403 Aut	AY424403	Austria	F	Exel et al., unpublished
NC 005148 Aut	NC_005148	Austria	F	Exel et al., unpublished
AB072301 Jp	AB072301	Japan	G	Imai et al., unpublished
AB072303 Jp	AB072303	Japan	G	Imai et al., unpublished
AF112862 Can	AF112862	Canada	G	Ouardani et al., 1999
AF118095 Can	AF118095	Canada	G	Ouardani et al., 1999
AF118097 Can	AF118097	Canada	G	Hamel et al., 2000

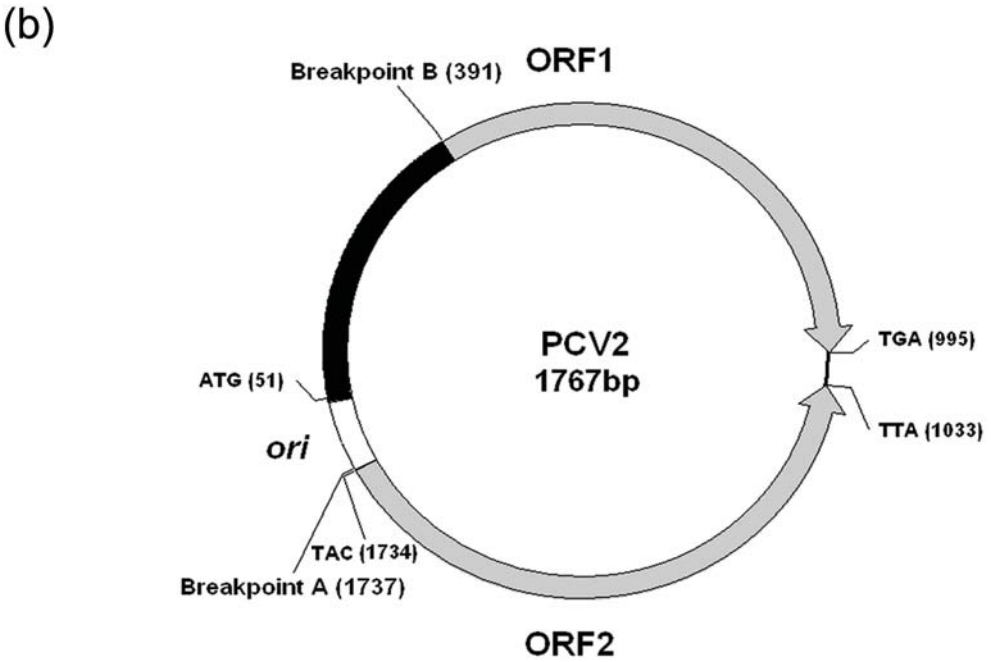
† The lineages were assigned according to figure 1.

Figure legends:

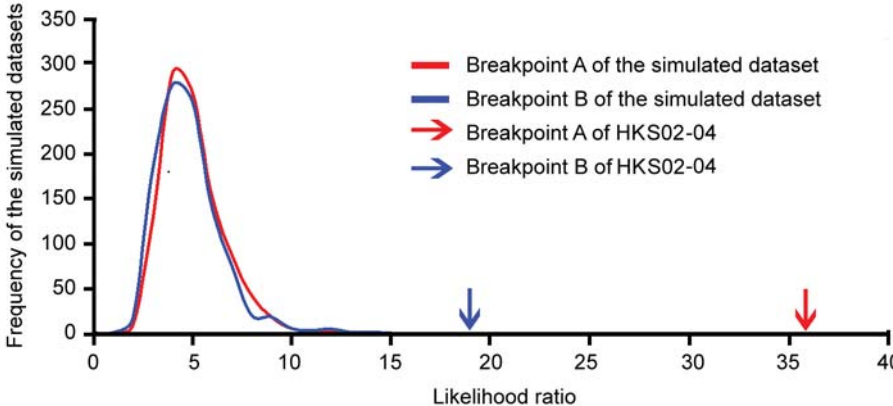
Fig. S1. Null distributions of likelihood ratios expected by chance. The distributions of likelihood ratios for the null hypothesis (i.e. no recombination) are shown. The arrows show the likelihood ratios obtained for breakpoint A and B of HKS02-04. The results in HKS03-04 and HKS091-04 are similar, and therefore only the results of HKS02-04 were shown for simplicity.


Lineage A
480
436
Lineage B
444
Lineage C
407
Lineage D
500
Lineage E
Lineage F
454
462
Lineage G
0.01 substitutions per site


(a)
100
90
80
70
60
50
40
30
20
10
0
% of Permuted Trees
P1
P2
O2
1033
1233
1433
1633
66
266
466
666
866
Position


(b)
Breakpoint B (391)
ORF1
PCV2
1767bp
TGA (995)
TTA (1033)
ATG (51)
ori
TAC (1734)
Breakpoint A (1737)
ORF2


(a)
AF112862 Can
AB072301 Jp
AF118097 Can
AB072303 Jp
AF118095 Can
Lineage G
AY288135 Chi
AY424401 Aut
NC 005148 Aut
AY424403 Aut
AY424402 Aut
Lineage F
AF166528 Tw
AY146991 Tw
Lineage D
AF201309 Spa
AF201308 Spa
AY256455 Hg
AY256459 Hg
AF201310 Spa
Lineage E
HKS091-04
HKS03-04
HKS02-04
AY322001 Fr
AY322002 Fr
AY288134 Chi
AY691169 Chi
AY322000 Fr
AY732494 Chi
AY682990 Chi
AY321996 Fr
AY321995 Fr
AY691679 Chi
AY604430 Chi
DQ195679 Chi
DQ180393 Chi
Lineage A
0.005 substitutions per site
(b)
AY691169 Chi
HKS02-04
HKS091-04
HKS03-04
DQ180393 Chi
DQ195679 Chi
AY288135 Chi
NC 005148 Aut
AY424403 Aut
AY424402 Aut
AY424401 Aut
AF201309 Spa
AF201308 Spa
AY256455 Hg
AF201310 Spa
AY256459 Hg
Lineage E and F
AB072303 Jp
AB072301 Jp
AF118097 Can
AF112862 Can
AF118095 Can
Lineage G
AF166528 Tw
AY146991 Tw
Lineage D
AY604430 Chi
AY288134 Chi
AY321995 Fr
AY322002 Fr
AY322001 Fr
AY321996 Fr
AY322000 Fr
AY691679 Chi
AY732494 Chi
AY682990 Chi
Lineage A
0.005 substitutions per site


Frequency of the simulated datasets
350
300
250
200
150
100
50
0
0
5
10
15
20
25
30
35
40
Likelihood ratio
Breakpoint A of the simulated dataset
Breakpoint B of the simulated dataset
Breakpoint A of HKS02-04
Breakpoint B of HKS02-04