Virology 460-461 (2014) 165–172

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Genomic organization and molecular characterization of porcine cytomegalovirus

Weiwei Gu, Nanfang Zeng, Lei Zhou, Xinna Ge, Xin Guo, Hanchun Yang *,[1]

Key Laboratory of Animal Epidemiology and Zoonosis of Ministry of Agriculture, College of Veterinary Medicine and State Key Laboratory of Agrobiotechnology, China Agricultural University, Beijing, People's Republic of China

ARTICLE INFO

Article history:
Received 7 October 2013
Returned to author for revisions
17 October 2013
Accepted 7 May 2014
Available online 4 June 2014

Keywords:
Porcine cytomegalovirus (PCMV)
Genome
Sequencing
Organization and characterization
Comparative analysis

ABSTRACT

A strain of the porcine cytomegalovirus (PCMV), which causes inclusion body rhinitis in newborn piglets, has been characterized with respect to its complete genome sequence. The virus genome is 128,367 bp, containing 79 predicted open reading frames (ORFs). Of these ORFs, 69 have counterparts in human herpesvirus 6A (HHV-6A), 6B (HHV-6B) and 7 (HHV-7), and two ORFs are homologous to other members in the subfamily *Betaherpesvirinae*. Eight ORFs have no homologs in herpesvirus. Homologs had higher identity and possessed similar orientation and location as roseoloviruses. The PCMV genome is a DR-U-DR type, similar to HHV-6A, HHV-6B and HHV-7, but the PCMV DR is shorter and lacks predicted genes and telomere-like sequences. Phylogenetic analyses of several core genes indicate that PCMV could be clustered in a branch with roseoloviruses. We suggest that PCMV could be classified as a member of the genus *Roseolovirus* of the subfamily *Betaherpesvirinae*.

Introduction

Porcine cytomegalovirus (PCMV), first described by Done (1955), is an important pathogen that causes inclusion body rhinitis and generalized infection in newborn piglets, and more severely, fetal death due to *in utero* infection in sows. The virus also can cause subclinical infection in adult pigs. PCMV infections are distributed worldwide, with a high prevalence in swine herds. Similar to other herpesviruses which are difficult to eliminate completely from the host, PCMV can be reactivated *in vivo* by stress (Mueller et al., 2004; Narita et al., 1987). In human xenotransplantation, significant attention has been paid to cross-species transmission of PCMV and its potential risk to public health (Gollackner et al., 2003).

PCMV was classified into the subfamily *Betaherpesvirinae* of the family *Herpesviridae* based on its biological properties *in vivo* and *in vitro* (Roizman et al., 1992). The virus remains as an unassigned species in *Betaherpesvirinae* according to the Ninth Report of the International Committee on Taxonomy of Viruses (ICTV) (Pellett et al., 2012). To date, only three genes encoding for DNA polymerase, major capsid protein, and glycoprotein B have been identified in the PCMV genome. These genes were more identical to human herpesvirus 6A (HHV-6A), 6B (HHV-6B) and 7 (HHV-7) in the genus *Roseolovirus* than with members of the *Cytomegalovirus* or *Muromegalovirus* genera, such as human cytomegalovirus (HCMV), mouse cytomegalovirus (MCMV) and rat cytomegalovirus (RCMV) (Goltz et al., 2000; Rupasinghe et al., 1999, 2001; Widen et al., 2001). These limited findings suggested that PCMV might be included within the genus *Roseolovirus*, along with HHV-6A, HHV-6B and HHV-7 (Rupasinghe et al., 1999, 2001; Widen et al., 2001). Therefore, analyzing and molecularly characterizing the PCMV genome is essential for defining the viral genus and exploring its genetic function and molecular pathogenesis.

Materials and methods

Virus culture and DNA preparation

PCMV strain BJ09 was isolated from pulmonary alveolar macrophages (PAMs) of a 4-week-old pig without obvious clinical symptoms. Available PAMs which were positive for PCMV according PCR were used to amplify a 314-bp fragment of the DNA polymerase (*DPOL*) gene (Hamel et al., 1999). PT-K75 cells (ATCC) were infected with PCMV by co-culturing with PAMs and then passaged six times with seven days' culture each time to obtain a higher virus titer. For viral DNA preparation, PCMV-infected PT-K75 cells were treated with Triton X-100 (Sigma-Aldrich, St. Louis, MO, USA) and micrococcal nuclease (NEB, Ipswich, MA, USA)

* Corresponding author. Tel./fax: +86 10 6273 1296.
E-mail address: yanghanchun1@cau.edu.cn (H. Yang).
[1] College of Veterinary Medicine, China Agricultural University, No. 2 Yuanmingyuan West Road, Haidian District, Beijing 100193, People's Republic of China.

http://dx.doi.org/10.1016/j.virol.2014.05.014
0042-6822/

before extracting DNA with phenol and chloroform as described previously (Sinzger et al., 1999).

Animal use for this study was approved by The Beijing Municipal Committee of Animal Management and The Ethics Committee of China Agricultural University.

DNA sequencing and assembling

First, 4 μg of PCMV DNA was used to prepare a paired-end library with a 500-bp insert size and to generate read lengths of 90 bp. High-throughput Illumina Genome AnalyzerIIX paired-end sequencing was performed in BGI (Shenzhen, China) to sequence the complete PCMV genome. After filtering, the 90-bp reads were assembled with SOAPdenovo software (BGI) to four contigs. To close the gaps (usually less than 20 bp), conventional PCR was applied by designing specific primers. Simultaneously, with the Genome Walking Kit (Takara, Dalian, China), the gaps between the contigs were closed.

After aligning with HHV-6A, HHV-6B and HHV-7, the 5′-RACE and 3′-RACE amplifications were carried out using a Marathon cDNA amplification Kit (Clontech, Mountain View, CA) as previously described (Davison et al., 2003). Briefly, the genomic DNA was flush-ended with T4 DNA Polymerase (Promega, Madison, WI) and ligated to the Marathon adaptor, and then the adaptor primers plus PCMV specific primers were used to identify the termini. Finally, using viral DNA extracted from PCMV-infected cells as templates, the DR-DR junction was amplified by PCR and sequenced with designed specific primers (5′-ACTGCACCGCTTCCCTTTAA-3′ and 5′-ATCACCAGCAGGCACCGTAT-3′) to confirm the Marathon results.

Genomic analyses

For the complete genome, GeneMarkS, a self-training program for predicting gene starts [Georgia Institute of Technology (http://exon.gatech.edu/genemarks.cgi)] (Besemer et al., 2001), was applied to identify ORFs, all of which had ATG as the start codon. At least 60 codons were analyzed for homologs using the BLASTP program (NCBI). Genes were designated according to their counterparts in HHV-6A, HHV-6B and HHV-7 or their positions within the genome. Searches for poly(A) signals, promoters, and splice sites were conducted by submitting the genomic sequence to PolyADQ, a eukaryotic polyadenylation signal search engine [Cold Spring Harbor Laboratory (http://rulai.cshl.org/tools/polyadq/polyadq_form.html)], the Berkeley Drosophila Genome Project's Neural Network Promoter Prediction program (http://www.fruitfly.org/seq_tools/promoter.html), a human promoter search engine and the Berkeley Drosophila Genome Project Splice Site Prediction by Neural Network, a eukaryotic search engine for donor and acceptor splice sites (http://www.fruitfly.org/seq_tools/splice.html) (Tabaska and Zhang, 1999).

Sequence alignments for phylogenetic analysis were carried out with the ClustalX 1.83 program. Phylogenetic trees were generated using the neighbor-joining method, bootstrap1000, of the MEGA 6 program. All herpesviruses used in our genomic analyses and their accession numbers are as follows: HHV-6A [GenBank no: KC465951.1], HHV-6B [GenBank no: AF157706.1], HHV-7 [GenBank no: AF037218.1], HCMV [GenBank no: AY446894.2], MCMV [GenBank no: GU305914.1], RCMV [GenBank no: AF232689.2], cercopithecine herpesvirus 5 (CeHV-5) [GenBank no: FJ483968.2], macacine herpesvirus 3 (RhCMV) [GenBank no: JQ795930.1], panine herpesvirus 2 (PaHV-2) [GenBank no: AF480884.1], elephantid herpesvirus 1A (ElHV-1A) [GenBank no: KC462165.1], elephantid herpesvirus 1B (ElHV-1B) [GenBank no: KC462164.1] and rhinolophus ferrumequinum betaherpesvirus (RfBHV-1) [GenBank no: JQ814845].

Results and discussion

Genome organization of PCMV BJ09

The raw data with 100 Mb was obtained by the high-throughput sequencing of PCMV BJ09 genome. After filtering and analyzing, the depth of the sequencing approximately reached 500 ×, which is sufficient for genome assembling. To identify the positions of the genomic termini, a 5′-RACE fragment with 510 bp and a 3′-RACE fragment with 523 bp were amplified (Fig. 1A). The locations of the termini were determined and the DR of the PCMV genome was identified by sequencing the amplified fragments. The DR–DR junction fragment with 1375 bp was amplified and sequenced (Fig. 1B).

The complete PCMV BJ09 genome was assembled and deposited into the GenBank database [GenBank no: KF017583]. The virus genome is 128,367 bp, shorter than those of HHV-6A (159,322 bp), HHV-6B (162,114 bp) and HHV-7 (153,080 bp), with a higher G+C content (45.54%), compared with HHV-6A/U1102 (42.44%), HHV-6B/z29 (42.77%), and HHV-7/RK (36.22%). The draft genome of PCMV is shown in Fig. 2. The genome contains a unique 127,641 bp region, flanked by a 363-bp direct repeat (DR) sequence that is shorter than those of HHV-6A, HHV-6B, and HHV-7. Homologous region related to the mammalian telomeric sequence was not identified in the DR of PCMV, whereas this region was found in HHV-6A, HHV-6B and HHV-7 (Megaw et al., 1998). In the unique region, 79 genes/ORFs were predicted by the GeneMarkS program, and the majority of these genes have homologs in HHV-6A, HHV-6B, and HHV-7 in amino acid composition and sequence (Table 1).

The DNA repeats were determined by the sequencing method with 500 bp random fragments. The defined 15 repetitive elements within the PCMV genome are mainly located in two regions —U77–U79 and U86–U100p (Fig. 2). All of them are shorter than 500 bp in length. The details of these repeats are summarized in Table 2. These elements are inferred to affect the transcriptional process of PCMV (Isegawa et al., 1999; Nicholas, 1996). Out of these elements, four (designated R3, R7, R11 and R15) are considered to be perfect consecutive direct repeats; the R15 was found to locate between two poly(A) signals from convergent transcripts, inferring that it might play a role in preventing accidental read-through by RNA polymerase into the oppositely transcribed gene (Klupp et al., 2004). Both the R3 (*KpnI*) of HHV-6A and HHV-6B and R2 (*DraI*) of

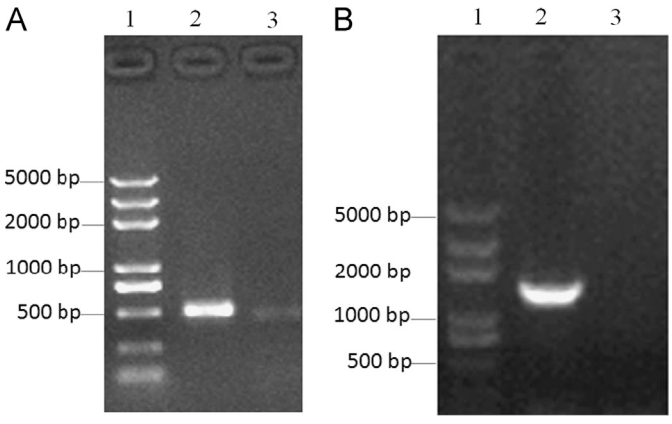

Fig. 1. Identification of the genomic termini positions. (A) The amplifications of 5′-RACE and 3′-RACE by a Marathon cDNA amplification Kit. Lane 1: DNA marker. Lane 2: 3′-RACE product of 523 bp. Lane 3: 5′-RACE product of 510 bp. (B) Amplification of the PCMV DR-DR junction by PCR. Viral DNA was extracted from PCMV-infected cells. PCR was performed using the designed specific primers (5′-ACTGCACCGCTTCCCTTTAA-3′ and 5′-ATCACCAGCAGGCACCGTAT-3′). Lane 1: DNA marker. Lane 2: a 1375 bp product. Lane 3: uninfected cell control.