Virology 491 (2016) 56–63

Contents lists available at ScienceDirect

Virology

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

Genomic analyses reveal that partial sequence of an earlier pseudorabies virus in China is originated from a Bartha-vaccine-like strain

Chao Ye [a,b], Jin-Chao Guo [a], Jia-Cong Gao [a], Tong-Yun Wang [a], Kuan Zhao [a], Xiao-Bo Chang [a], Qian Wang [a], Jin-Mei Peng [a], Zhi-Jun Tian [a], Xue-Hui Cai [a], Guang-Zhi Tong [b,*], Tong-Qing An [a,**]

[a] State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Harbin 150001, China
[b] Shanghai Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Shanghai 200241, China

ARTICLE INFO

Article history:
Received 16 October 2015
Returned to author for revisions
18 January 2016
Accepted 22 January 2016
Available online 11 February 2016

Keywords:
Pseudorabies virus
Diversity
Recombination

ABSTRACT

Pseudorabies virus (PRV), the causative agent of Aujeszky's disease, has gained increased attention in China in recent years as a result of the outbreak of emergent pseudorabies. Several genomic and partial sequences are available for Chinese emergent and European–American strains of PRV, but limited sequence data exist for the earlier Chinese strains. In this study, we determined the complete genomic sequence of one earlier Chinese strain SC and one emergent strain HLJ8. Compared with other known sequences, we demonstrated that PRV strains from distinct geographical regions displayed divergent evolution. Additionally, we report for the first time, a recombination event between PRV strains, and show that strain SC is a recombinant of an endemic Chinese strain and a Bartha-vaccine-like strain. These results contribute to our understanding of PRV evolution.

Introduction

Pseudorabies, caused by pseudorabies virus (PRV), is an economically significant disease in the pig industry worldwide. PRV belongs to the family *Herpesviridae*, subfamily *Alphaherpesvirinae*, genus *Varicellovirus* (Pellett et al., 2011). The family Suidae is the only natural reservoir for the virus, but PRV can also infect many other mammals (Mettenleiter, 2008). Since the early 1980 s, PRV has been circulating globally and has recently become a target of worldwide eradication programs (Muller et al., 2011). As a result, PRV is now considered to be almost eradicated from domestic pigs in most regions of Europe, the United States and New Zealand (Hahn et al., 2010; MacDiarmid, 2000; Muller et al., 2003).

In China, PRV has caused heavy economic losses in the pig industry. The first recorded pseudorabies case was in the 1950s. In the 1970s, the Bartha-K61 (a live attenuated vaccine strain) vaccine (Bartha, 1961) was introduced into China to control the disease, and since the 1990s most pig herds in China have been immunized (An et al., 2013). As a result, PRV has been effectively controlled (Tong and Chen, 1999). However, since October 2011, a number of pseudorabies outbreaks have been reported across China in pig farms vaccinated with Bartha-K61 (An et al., 2013; Luo et al., 2014; Wu et al., 2013; Yu et al., 2014). Previous studies have shown that the Bartha-K61 vaccine could not provide complete protection against emergent PRV strains in China but could offer full protection against infection with PRV strain SC (an earlier virulent strain in China) in immunological protection experiments in sheep (An et al., 2013; Wang et al., 2014). It was therefore be intriguing to analyze the genetic differences and phylogenetic relationships between the emergent strains and the earlier virulent strain (i.e., PRV strain SC).

However, the high GC content and the reiterated repeat sequences on the PRV genome mean that sequencing of PRV is challenging. The first complete sequence of PRV was generated in 2004 by assembling sequence fragments derived from six different strains (Klupp et al., 2004). Recently, several complete PRV genomes have been reported, including three European–American strains (Kaplan, Becker, Bartha) and four Chinese emergent strains

* Correspondence to: Shanghai Veterinary Research Institute, CAAS, No. 518 Ziyue Road, Minhang District, Shanghai 200241 China. Tel.: +86 21 34293436; fax: +86 21 54081818.
** Correspondence to: Harbin Veterinary Research Institute, CAAS, No. 427 Maduan Street, Nangang District, Harbin 150001, China. Tel.: +86 18945686530; fax: +86 451 51997166.
E-mail addresses: gztong@shvri.ac.cn (G.-Z. Tong), antongqing@163.com (T.-Q. An).

http://dx.doi.org/10.1016/j.virol.2016.01.016
0042-6822/

(HeN1, TJ, JS and ZJ01) (Szpara et al., 2011b; Luo et al., 2014; Ye et al., 2015; Gu et al., 2015). There is, however, limited sequence information and no complete genomes available for the earlier Chinese PRV strains. Such information would provide a valuable comparison with recent emergent PRV genome sequences.

Phylogenetic studies to date have demonstrated that there is a wide extent of diversity between the vaccine strain (Bartha) and the European–American wild strains (Kaplan and Becker) (Szpara et al., 2011b). Recently, we provided evidence for the existence of two genotypes, with genotype II consisting of isolates from China and genotype I containing the other isolates from Europe and America (Ye et al., 2015). However, the limited complete PRV genome sequences means that comparative complete genome recombination analysis has not yet been performed.

In this study, we generated two complete PRV genome sequences, an earlier strain (SC) and an emergent strain from Heilongjiang province (HLJ8) by high-throughput sequencing. Comparative genomic analysis suggested that large diversity existed between the isolates, and strain SC was phylogenetically located between the two branches of Chinese emergent strains and European–American strains. A phylogenetic tree based on the *gC* gene content for global isolates reflected the diverse geographic distribution of PRVs. Recombination analysis suggested that PRV evolution involved recombination events and strain SC could be a recombinant of genotype I and II strains.

Results

Comparison between PRV genomes from geographically distinct isolates

We performed high-throughput sequencing of the purified, randomly fragmented viral DNA with the reads ranging in length from 300 bp to 500 bp. The reads numbers of strains SC and HLJ8 are 10,857 and 10,399 respectively, and the average sequence coverages were 29.6 and 29.1 reads per base for SC and HLJ8, respectively. The whole genome assembly was performed by Shanghai Human Genome Research Center, and finally both genomes were approximately 142 kb in length. To compare geographically distinct PRV genome strains, according to the description of Ye et al. (2015) we assessed the degree of sequence diversity across the genome of our newly sequenced strains (SC and HLJ8) compared with PRV full length sequences from the NCBI nucleotide database (Fig. 1). We noted that these genomes were largely conserved within the U_L and U_S regions, with higher levels of sequence variation observed in the intergenic regions. Regions with the highest levels of variation were localized to the large internal and terminal repeat regions (Fig. 1). This variation could be attributed to both the inherent variation in these regions as well as to difficulties in sequencing and assembling these problematic regions with present sequencing technologies. Compared with the known complete sequences of PRVs, the sequences of strains SC and HLJ8 shared the same genome organization comprising U_L, U_s and internal and terminal repeats, and a similar GC content (73.6% for SC; 73.7% for HLJ8; 73.6% for HeN1; 73.5% for JS; 73.6% for TJ). As an emergent strain, the nucleotide sequence of strain HLJ8 was highly similar to that of other emergent strains. Conversely, strain SC showed a relatively high degree of sequence divergence with the emergent strains, particularly in the internal and terminal repeats and several intergenic regions, whereas other regions such as the coding regions were largely conserved, with the exception of specific genes such as UL36, UL17, UL21, UL44, US1 and IE180 (Fig. 1A).

Evolutionary divergence between PRVs

The generation of nine full-length PRV genome sequences allowed us to study the relatedness of PRV strains circulating in the US, Europe and Asia. Estimates of pair-wise evolutionary divergence between the nine new PRV genomes showed remarkable sequence diversity between strains particularly between genotype I strains and genotype II strains (Table 1), with the maximum nucleotide divergence between strains being 2.93% (Kaplan vs. TJ) across the genome, which was significantly higher than that among herpes simplex virus 1 (HSV-1) (1.31%) and HSV-2 (0.361%) (Kolb et al., 2013; Kolb et al., 2015). Analysis of nucleotide and amino acid diversity within 67 U_L and U_S ORFs among strains SC and HeN1, TJ, ZJ01, HLJ8 and Bartha confirmed relatively high-level sequence variation, with 12 pairs of ORFs exceeding 1.00% diversity at the nucleotide (nt) level and 1.64% diversity at the amino acid (aa) level between strain SC and the four selected emergent strains (Table S1). In contrast, between strains SC and Bartha most of the ORFs showed even higher sequence divergence than that observed above. The level of diversity of these ORFs ranged from 0% to 5.39% at the nt level and 0% to 11.22% at the aa level (Table S1). This included eight ORFs (UL1, UL16, UL34, UL36, UL46, 47, UL49.5 and UL56) that showed over 3.00% diversity at the nt level and 4.68% at the aa level, while seven ORFs (UL20, UL21, UL23, UL24, UL40, UL41 and UL43) showed no diversity and were completely conserved (Table S1). Certain ORFs, such as UL11, UL44 and US1, showed higher diversity ($>1\%$) at both the nt and aa levels than other ORFs among all of the strains analyzed (Table S1).

Diverse inter-strain phylogenies of PRVs

To investigate the phylogenetic relationships between PRV strains, a maximum likelihood tree was constructed that included all nine full-length PRV genomic sequences. It revealed that the emergent strains clustered in an independent branch and diverged from strain SC; but still strain SC appeared more closely related to the Chinese emergent strains (genotype II) than to the European–American strains (genotype I) (Fig. 1B). The Chinese emergent strains also displayed diverse evolution. Strain HeN1 isolated from central China appeared in a relatively distinct branch, the other emergent strains clustered in another distinct branch with HLJ8 and TJ isolated from the northern regions of China clustering in one clade, JS and ZJ01 isolated from the southern areas of China clustering in another clade, which revealed that the PRV strains in China were also geographically divergent (Fig. 1B).

Phylogenetic analysis based on the *gC* gene indicated that generally the PRV strains clustered geographically, which was similar to the results observed for HSV-1 (Szpara et al., 2014). However, the PRV strains appeared to be phylogenetically more diverse within the same geographic area, for example, the genotype I isolates from Italy distributed across six independent clusters, the Belgium isolates also clustered across three separate clades and the American isolates were located in three large branches mixed with the European isolates (Fig. 2). Genotype II was composed mainly of Chinese isolates, which were generally closely related and clustered together. Most of the emergent strains isolated after 2011 including HLJ8 were closely and independently clustered together, of which only several emergent strains together with four earlier strains (Ea, LXB6, LXB88 and HUYD) comprised a small subgroup. Unlike the close relatedness of the emergent strains, the earlier strain Fa (isolated in 1964) and its derived vaccine strain SA215 were phylogenetically diverged from the four earlier strains referred above and also distinct from the emergent strains (Fig. 2).