Virus Genes 17:3, 213–218, 1998

Molecular Cloning and Nucleotide Sequence of 3′-Terminal Region of Classical Swine Fever Virus LPC Vaccine Strain

MIN-LIANG WONG,[1] JAU-JIN LIU,[1] CHIENJIN HUANG,[2] JUEI-WEN CHEN,[1] & TIEN-JYE CHANG[1]*

[1] *Department of Veterinary Medicine*

[2] *Graduate Institute of Veterinary Microbiology National Chung Hsing University Taichung 402, Taiwan*

Received March 19, 1998; Accepted June 15, 1998

Abstract. A cDNA of the 3′-terminus of classical swine fever virus (LPC vaccine strain) was cloned and sequenced. The 3431 nucleotides and deduced amino acid sequences were compared with those of other pestiviruses, and the similarity of nucleotide sequences and deduced amino acid sequences were found to be 84–95% and 95–98%, respectively. Similar to other isolates of classical swine fever virus, the sequenced region included the non-structural gene p58 (NS5A) and part of p76 (NS5B) gene. The p76 gene of LPC vaccine strain also contained a highly conserved motif G-D-D (Gly-Asp-Asp) that is present in the RNA replicase of positive-stranded RNA viruses. With the sequence data currently available, we carried out a phylogenetic analysis and obtained a genealogical relationship among members of the classical swine fever virus. The sequence has been submitted to GeneBank with an accession number AF001986.

Key words: pestivirus, classical swine fever virus; RT-PCR, sequencing

Classical swine fever virus (CSFV) is the causative agent of classical swine fever (also named hog cholera), an economically important disease of swine. And it is a member of the genus Pestivirus which includes bovine viral diarrhea virus (BVDV) and border disease virus (BDV) of sheep. Together with the genus Flavivirus and the hepatitis C virus group, the genus Pestivirus is included in the family of Flaviviridae (7,10). To date, the complete nucleotide sequences of CSFV Alfort and Brescia strains and full-length cDNA clone of Alfort/187 strain have been reported (7,10,12). The viral genome consists of a ssRNA of about 12.5 kb, which comprises one large open reading frame coding for one putative polyprotein. The genome of the CSFV is capable of producing at least 4 viral structural proteins and 7 non-structural proteins in the order from 5′-end to 3′-end (13,14,16). For CSFV, the p76 protein derived from the putative polymerase region locates in the 3′ region of the genome has been determined (7,8).

* Corresponding author.

In consideration of the epizootic importance and diagnostic difficulties of CSFV, it is needed to develop novel methods of rapid detection, identification and classification of the virus. In order to characterize and group the pestiviruses, limited regions of the CSFV genome from many other isolates have been amplified by reverse transcription-polymerase chain reaction (RT-PCR). The sequences of RNA polymerase are well conserved between different genotypes of CSFV (5,11), suggesting that the polymerase coding regions seem to be appropriate for phylogenetic analysis. In this study, The amplified products of a 3431 bp fragment from the non-structural gene p58 (NS5A) and part of the polymerase gene (p76, NS5B) of the viral genome were sequenced and a molecular phylogenetic relationship was also analyzed.

The LPC vaccine strain, obtained from Taiwan Provincial Research Institute for Animal Health, Tamsui, Taiwan, was used in this study. Viruses were harvested from spleens of experimentally infected rabbits. Total RNA together with genomic

viral RNA was extracted by the acid guanidinium thiocyanate-phenol-chloroform method (1). The extracted RNA was used as a template for the synthesis of cDNA of LPC strain and the primers (Table 1) derived from the Alfort and Brescia strains sequences (7,10) or random primers were used. Six sets of primers, P1/C6, P2/C4, P3/C5, P4/C3, P4/C2 and P4/C1, were designed to amplify the 3′ non-structural region (Fig. 1a). As shown in Fig. 1b, RT-PCR amplifications of RNA from the LPC strain produced the predicted DNA fragments that correspond to the 3′-terminal regions of genome. The result showed that DNA fragments of the expected size were produced, i.e., 528 bp for P1/C6, 1887 bp for P2/C4, 134 bp for P3/C5, 1455 bp for P4/C1, 803 bp for P4/C2 and 490 bp for P4/C3. These DNA fragments were digested with a variety of restriction enzymes and subcloned into the pGEM7zf(+) vector (Promega) (Fig. 1B). Clones containing DNA fragments described above were sequenced by the dideoxynucleotide chain termination method (15). We sequenced the overlapping RT-PCR products crossing the boundary of these cDNA fragments. The sequence determined encompassed p58 (NS5A) and part of p76 (NS5B) coding regions (Fig. 1c). The deduced amino acid sequences of the DNA were obtained with the TRANSLATE program of Wisconsin GCG package. The sequence of the 3′-terminal 3431 nucleotides of the LPC strain is shown in frame 1 (Fig. 2). The nucleotide sequence contains a long open reading frame that presumedly encodes a polypeptide of 1143 amino acids in frame 2 (Fig. 2). The sequence has been submitted to GenBank with an accession number AF001986.

The manipulated region of the proximal 3′ end of CSFV LPC vaccine strain included non-structural proteins p58 and part of p76. The p58 was a protein with a hydrophobic property and its function was unclear (2,3). The p76 was identified as RNA-dependent RNA polymerase (RdRp) which contained three conserved domains (the methyltransferase domain, the NTP binding domain of the RNA helicase and the core binding domain of the polymerase) (19). These three domains were found to be conserved in many other single-stranded (+) RNA viruses. Comparisons of deduced amino acid sequences from the polymerase region of different RNA viruses revealed a highly conserved tripeptide Gly-Asp-Asp motif (Fig. 2), which is a characteristic of RdRp (2,19). Viruses with RNA genomes have been shown to mutate more frequently than those with DNA genomes. It is generally assumed that high mutation rate results in rapid evolution of RNA viral proteins. Despite the rapid mutational change that is typical of (+) RNA viruses, enzymes mediating the replication and expression of virus genomes contain arrays of conserved sequence motifs. Proteins with such motifs include RdRp, putative RNA helicase, chymotrypsin-like proteases, papain-like proteases, and methyltransferases. Based on the phylogeny for the RdRp that is universally conserved protein of (+) RNA viruses, it has been proposed that (+) RNA viruses can be classified based on the respective methyltransferase, helicase and polymerase domains of the

Table 1. Nucleotide sequence and location on the CSFV genome of the synthetic oligonucleotide primers used

Primer	Sequence	Location *
P1	5′GTTGAACGGGCTTTGGCTAC3′	7859–7978
P2	5′GACTTGTATACCACCTTTATGTA3′	8261–8283
P3	5′AACTGGGTGATACAAGACGAC3′	9907–9927
P4	5′GAAAATAAACGGGGCAACCTG3′	9926–9946
C1	5′GCTTACCAGCTTCATATAGG3′	11329–11310
C2	5′GCGGTCCAGTCATCATTGAC3′	10728–10709
C3	5′CCCTTATTGCTTGGTGGAAG3′	10414–10395
C4	5′GTTCCTCTACCAACTCCTTTAAC3′	10147–10125
C5	5′TTCCCTTGAGCTAGTTGGTA3′	10041–10022
C6	5′TACTCCCAGTAGTTCCACAG3′	8385–8366

* The location of cDNA sequences is according to the position within the Alfort strain sequence.

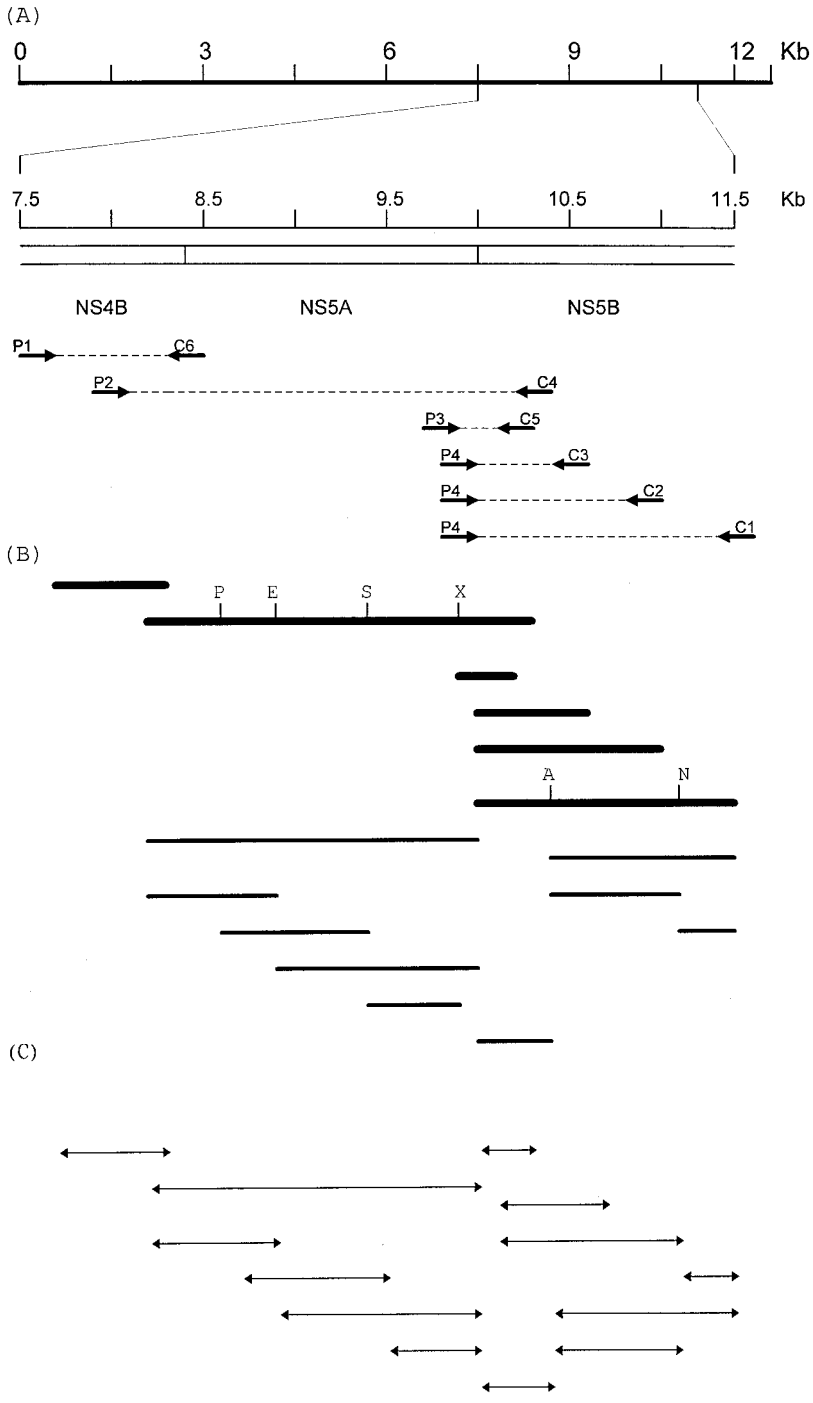

Fig. 1. Schematic representation of the LPC strain of CSFV genome. The optimum conditions for PCR were performed with an initial denaturation (94°C, 5 min) and then 35 cycles of amplification [94°C, 30 sec for denaturation; 1 min of annealing at 47°C (P1/C6), 50°C (P2/C4, P4/C1), 45°C (P3/C5), and 55°C (P4/C2, P4/C3), respectively; 72°C, 2 min for extension], and a final extension of 72°C 7 min. (A) Genome organization with the scale. Primers' location and orientation are indicated by an arrowhead. (B) Partial restriction map and cDNA clones selected for sequencing. PCR products are presented as thick black lines. Abbreviations for restriction sites are: A: *Apa* I; E: *Eco* RV; N: *Nsi* I; P: *Pst* I; S: *Sma* I; X: *Xba* I. Subclones are shown by thin black lines. (C) Sequencing strategy. Double arrows indicate regions sequenced using cDNA clones and their deletion subclones in both strands.

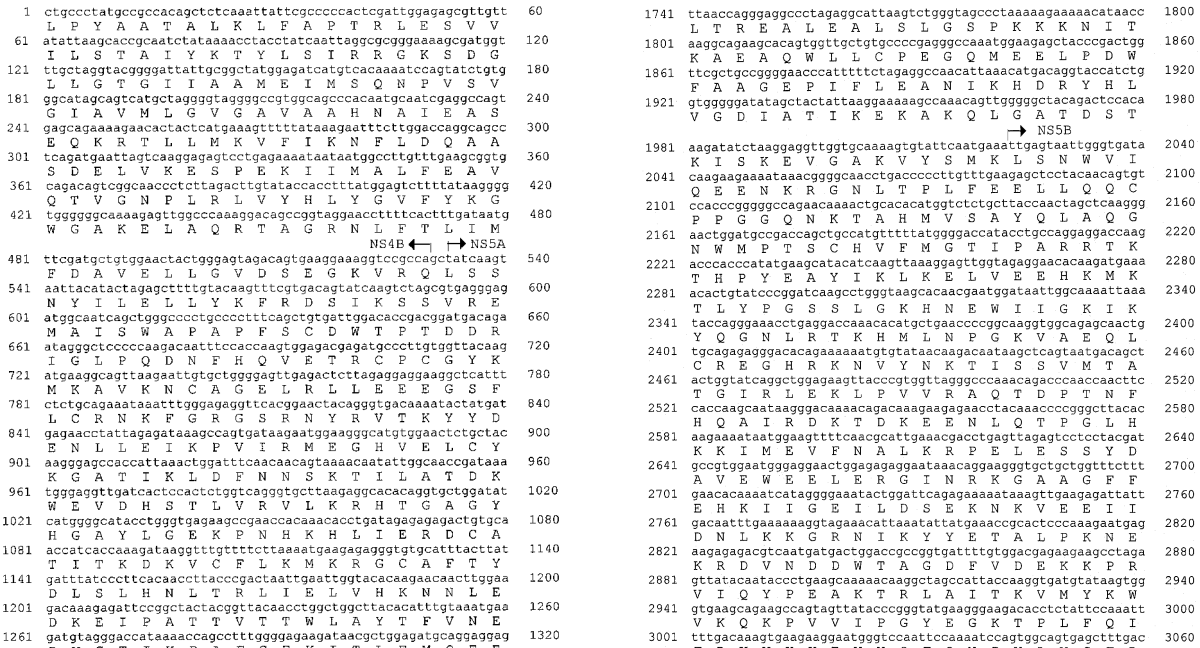

Fig. 2. Nucleotide sequence of the 3′-terminal 3431 nucleotides of LPC vaccine strain. Below the DNA sequence is the deduced amino acid sequence of the N terminal protein in standard capital letter code. The arrowheads indicate the locations of non-structural genes (NS4B, NS5A, and NS5B).

RdRp protein, and that they could have evolved from a common ancestor virus (6).

Comparing nucleotide sequences of the LPC strain with previously reported sequences, molecular phylogenetic trees were constructed to delineate the relationship among different pestiviruses. The nucleotide sequence similarity is 70% between LPC and BDV, and 66% between LPC and BVDV. These results are similar to the previous study by Meyers et al. (7). The LPC strain showed a 85% similarity with the Alfort strain and a 92% similarity with the Brescia strain. However, a higher level of the amino acid sequence identity was found among these strains of CSFV, ranging from 95% to 98%. According to the nucleotide sequence, a polypeptide of 1143 amino acids was obtained, and was same as the other CSFV published sequences.

The nucleotide sequence was further analyzed in order to establish the phylogenetic relationships between several strains of pestivirus and to construct an evolutionary tree based on the sequence data from the 3′-terminal 3431 nucleotides. Phylogenetic analysis was performed using multiple programs of the PHYLIP package. Comparison of nucleic acid identities varied from 84%–95% between LPC and the other nine strains of CSFV. Fig. 3 illustrates the phylogenetic tree from maximum likelihood analysis of each set of data. Statistical analysis using bootstrapping showed that the confidence was high. The dendrogram clearly demonstrated a close relationship between Riems, Chinese and LPC. Moreover, highly conserved nucleotide sequence with a similarity of 95% was the region containing RNA polymerase in the 3′-end.

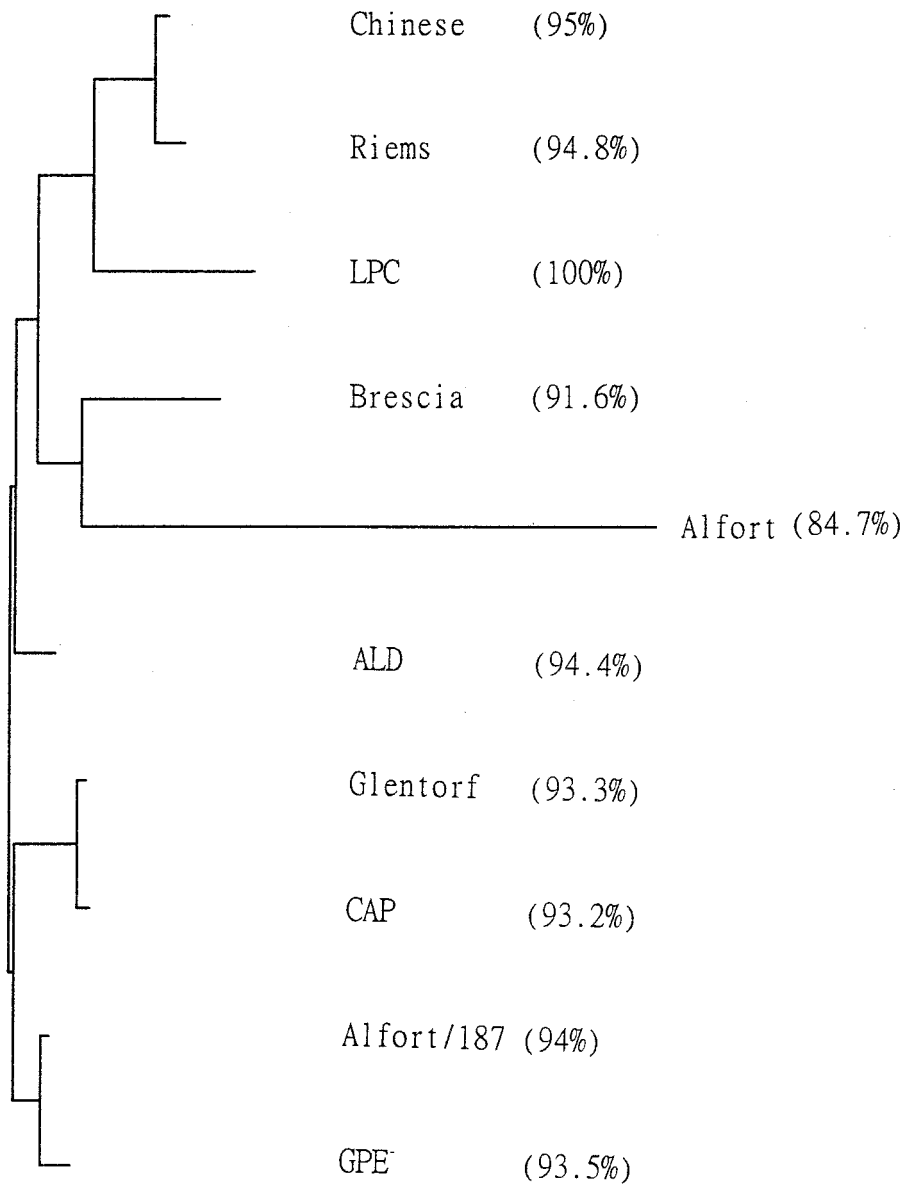

Fig. 3. Phylogenetic relationship of the 3′-terminal 3431 nucleotides of various isolates of pestiviruses. EMBL/GenBank accession numbers of the clones in the phylogenetic tree are: AF001986 (LPC); Z46258 (Chinese); U45477 (Riems); D49532 (ALD); X87939 (Alfort/187); D49533 (GPE⁻); U45478 (Glentorf); X96550 (CAP); M31768 (Brescia); J04358 (Alfort). The similarity between LPC and other strains are represented by percentages.

In order to assess the validity of classification of RNA viruses. In our study, the same size of 3′-terminal 3431 nucleotides was very uniform among strains of CSFV with sequence similarity ranges from 84% to 95%. Phylogenetic analysis of the sequence of p58 and RdRp (Fig. 3) seems to have the potential use for sequence-based molecular epidemiology of CSFV infections and diagnosis. In summary, this work demonstrated that the 3431 nucleotide sequences of the 3′-terminal region of LPC viral genome were determined and can be used as a tool for the epidemiological study of CSFV strains.

Acknowledgments

We thank Professor Poa-Chun Chang (National Chung Hsing University) for invaluable help in the phylogenetic analysis. The financial support of this work was provided by grants from Council of Agriculture, Taiwan.

References

1. Chomczynski P. and Sacchi N., Anal Biochem *162*, 156–159, 1987.
2. Collett M.S., Anderson D.K., and Retzel E., J Gen Virol *69*, 2637–2643, 1988a.
3. Collett M.S., Larson R., Belzer S.K., and Retzel E., Virology *165*, 200–208, 1988b.
4. Hofmann M.A., Brechtbuhl K., and Stauber N., Arch Virol *139*, 217–229, 1994.
5. Koonin E.V., J Gen Virol *72*, 2197–2206, 1991.
6. Koonin E.V. and Dolja V.V., Crit Rev Biochem Mol Biol *28*, 375–430, 1993.
7. Meyers G., Rumenapf T., and Thiel H.-J., Virology *171*, 555–567, 1989.
8. Meyers G. and Thiel H.-J., Adv Virus Res *47*, 53–118, 1996.
9. Meyers G., Thiel H.-J., and Rumenapf T., J Virol *70*, 1588–1595, 1996.
10. Moormann R.J.M., Warmerdam P.A.M., van der Meer B., and Schaaper W.M., Virology *177*, 184–198, 1990.
11. Poch O., Sauvaget I., Delarue M., and Tordo N., EMBO J *8*, 3867–3874, 1989.
12. Ruggli N., Tretschin J.D., Mittelholzer C., and Hofmann M.A., J Virol *70*, 3478–3487, 1996.
13. Rumenapf T., Meyers G., Stark R., and Thiel H.-J., Virology *171*, 18–27, 1989.
14. Rumenapf T., Unger G., Strauss R., and Thiel H.-J., J Virol *67*, 3288–3294, 1993.
15. Sanger F., Nicklen S., and Coulson A.R., Proc Natl Acad Sci USA *74*, 5436–5467, 1977.
16. Stark R., Rumenapf T., Meyers G., and Thiel H.-J., Virology *174*, 286–289, 1990.
17. Thiel H.-J., Stark R., Weiland., Rumenapf T., and Meyers G., J Virol *65*, 4705–4712, 1991.
18. Vilcek S., Stadejek T., Ballagi-Pordany A., Lowings J.P., Paton D.J., and Belak S., Virus Res *43*, 137–147, 1996.
19. Wong S.M., Mahtani P.H., Lee K.C., Yu H.H., Neo K.K., Chan Y., Wu M., and Chng C.G., Arch Virol *142*, 383–391, 1997.