Genome and Molecular Characterization of a CSFV Strain Isolated from a CSF Outbreak in South China

Hai-Yan Shen\textsuperscript{a,b} Jia-Ying Wang\textsuperscript{a} Xiao-Ying Dong\textsuperscript{a} Ming-Qiu Zhao\textsuperscript{a}
Yanmei Kang\textsuperscript{a} Yin-Guang Li\textsuperscript{a} Jing-Jing Pei\textsuperscript{a} Ming Liao\textsuperscript{a} Chun-Mei Ju\textsuperscript{a}
Lin Yi\textsuperscript{a} Yongming Hu\textsuperscript{a} Jin-Ding Chen\textsuperscript{a}

\textsuperscript{a}College of Veterinary Medicine, South China Agricultural University, and \textsuperscript{b}Institute of Veterinary Medicine, Guangdong Academy of Agricultural Sciences, Guangzhou, PR China

Key Words
Classical swine fever virus • Phylogenetic tree • Genome comparison • Homologous recombinant • Virulence

Abstract
In the present study, the full-length nucleotide sequences of the CSFV-GZ-2009 strain of classical swine fever virus (CSFV) isolated from a hog pen in Guangdong province in China was determined. Results demonstrated that the genome of CSFV-GZ-2009 is 12,298 nucleotides (nt) in length, is composed of a 373-nt 5'-untranslated region (UTR), has an 11,697-nt open reading frame encoding a polyprotein of 3,898 amino acids, and has a 228-nt 3'-UTR. Genome comparison of the CSFV-GZ-2009 isolate (GenBank accession No. HQ380231) with other CSFV strains was also analyzed. Gene regions from CSFV-GZ-2009 and other known strains were shown to share 92.7–96.7% identity at the nucleotide level and 94.7–99.2% identity at the amino acid level. Phylogenetic analysis of the full-length genome and the following regions E\textsubscript{NS}, E2 and NS5B revealed that the CSFV-GZ-2009 isolate was classified within subgroup 1.1 of group I and closely related to the highly virulent strain JL1 (06), CF114, Shimen and SWH with pairwise distances of 0.0037, 0.0043, 0.0058 and 0.0107, respectively. Analysis of recombination with the SimPlot program demonstrated that strain CSFV-GZ-2009 was not a naturally homologous recombinant. Furthermore, the change of clinical signs of pigs after infection of CSFV-GZ-2009 isolates showed typical symptoms such as diarrhea, persistent fever, and mononuclear lymphocytosis after CSFV infection. Based on phylogenetic analysis and an animal infection test, we could conclude that the CSFV-GZ-2009 isolate belonged to subgroup 1.1 of group I and was of high virulence.

Introduction
Classical swine fever (CSF) is an economically important, highly contagious disease affecting pigs in various parts of the world [1]. The causative agent of this disease is CSF virus (CSFV), a member of the Pestivirus genus within the Flaviviridae family [2]. CSF epidemic usually...
results in high fever, hemorrhagic syndrome and high mortality, and causes considerable economic loss of the pig industry worldwide.

Nucleotide sequence diversity is usually analyzed for tracing virus spread and for developing disease control strategies. In Taiwan, by analyzing the E2 sequences of CSFV from field outbreaks during 1993–2001, CSFVs have been classified into two subgroups – 3.4 and 2.1 [3, 4]. Also, the Taiwanese strains of 2.1 were further divided into two different genotypes termed 2.1a and 2.1b which might be introduced from different origins [4]. Another finding suggested that all CSFV strains were subgenotype 3.4 prior to 1996. However, genotype 2 strains of CSFV, identified as subgenotype 2.1 or 2.2, have been isolated since 1994 and have replaced gradually the subgenotype 3.4 strain [5]. 110 clinical specimens representing 109 epizootic sites from 1986 to 1999 in China were also analyzed. A phylogenetic tree showed that 103 of the 110 field isolates were clustered within group 2 and subdivided into three subgroups, while the remaining 7 isolates were clustered into subgroup 1.1 indicating that sub-groups 2.1 and 2.2 predominated in the more recent epizootics in China, while group 1 viruses have caused only limited epizootics [6].

Guangdong province, located in the extreme south of China and bordering Hong Kong, is one of the most economically developed provinces. It has an extensive commercial and pig movement population, with most of its breeding stocks introduced from overseas or from inner Chinese provinces. This high mobility allows transmission of CSFVs. In 2009, a virulent strain from a farm with CSF outbreak in Guangdong province, South China, was isolated. In this study, the complete genome and molecular characterization of this virulent strain named CSFV-GZ-2009 was sequenced and characterized. The genome analyses provided insights into the phylogenetic origin of the new CSFV isolates, which would be useful for the further study in viral pathogenesis of CSFV.

Materials and Methods

Cell Culture and Virus Isolation

Samples of spleen, kidney and lymph node were collected from a hog pen in Guangdong province, South China, during the outbreak of CSF in 2009. A permissive cell line of porcine kidney (PK-15) was used to isolate CSFV. Prior to inoculating onto PK-15 cells, the samples from the infected pigs were ground and then grounded in Dulbecco’s modified Eagle medium (DMEM; Gibco-BRL, USA) (pH 7.2) and filtered through a 0.22-μm filter. The infected PK-15 cells were incubated at 37° for 1 h. Fresh DMEM containing 2% (v/v) fetal calf serum and the antibiotics penicillin (100 U/ml) and streptomycin (0.1 mg/ml) were then added and the cells were incubated for 72 h. Five blind passage cultures were also done for CSFV isolation. The replication of the CSFV was monitored in cultured PK-15 cells by indirect immunofluorescence assay using monoclonal antibody WH303 against the E2 glycoprotein of CSFV (a kind gift by China Institute of Veterinary Drug Control, FITC-conjugated anti-mouse rabbit IgG (Santa Cruz Biotech, USA). After washing, the cultured PK-15 cells were observed under a fluorescence microscope. Intracellular brilliant green fluorescence was indicative of viral infection.

Cloning of the Full-Length Genomic Sequence of the CSFV-GZ-2009

Nine pairs of primers (table 1) were designed using the software package Primer 5.0. The primers were based on the sequences and the distribution of restriction endonuclease sites of CSFV strains (GenBank No. AF333000) and synthesized by Sangon Biotech Co., Ltd (Shanghai, China). Using these primers, nine cDNA fragments corresponding to nucleotides (nt) 1–910, 911–2,463, 2,441–5,140, 5,071–7,408, 7,382–8,422, 8,413–9,913, 9,720–11,671, 10,500–11,810, and 11,653–12,298 were amplified by RT-PCR.

Total RNA from CSFV-infected PK-15 cells was extracted from virus stock using TRIzol reagent (TaKaRa, Japan), according to the manufacturer’s instructions. Reverse transcription (RT) was carried out using SuperScript™ III reverse transcriptase (Invitrogen, USA) and the reverse primer (table 1) for priming and the extracted viral RNA as template. PCR was carried out in a total volume of 50 μl using PluUltra™ High-Fidelity DNA Polymerase (Stratagene, USA). PCR amplification was performed according to the following thermocycling profiles: 3 min of initial denaturation at 95°, 30 cycles of 45 s of denaturing at 95°, 50 s of annealing at 48.0–57.0°, and 1–3 min extension at 72°. A final step of 10 min extension at 72° was added after the 30-cycle amplification. The PCR product was verified by electrophoresis on 1% agarose gels and purified with a PCR Purification Kit (Omega, Germany). The purified PCR fragment was cloned into pJET1.2 vector (Fermentas, Canada). Recombinant clones were identified by PCR. At least five recombinant clones for each fragment were randomly selected and sequenced with an ABI PRISM 3730 sequencer (Applied Biosystems, USA).

Sequence Analysis

The complete genome sequence and deduced amino acid sequence of the CSFV-GZ-2009 strain were analyzed with DNASTAR (DNASTAR Inc., Madison, Wisc., USA) and Sequencher (Gene Codes Corp., USA). The homologous analysis of the nucleotide sequence and deduced amino acid of CSFV-GZ-2009 and other reported 28 CSFV strains was performed with DNASTAR software. Phylogenetic tree analyses were conducted using the full-length of the open reading frame (ORF), E2, E™ and N5B, respectively, by a neighbor-joining (NJ) method and the bootstrap test with MEGA version 4.0. All sequences of CSFV strains including CSFV-GZ-2009 used in alignment and phylogenetic tree analyses were obtained from GenBank (table 2). The RNA secondary structures of CSFV 3’-untranslated region (UTR) and 5’-UTR were predicted and analyzed using RNA structure 4.4 software (http://rna.chem.rochester.edu/). The functional domains of pre-polyprotein of CSFV were analyzed with Expert Protein Analysis System (ExPASy; http://ca.expasy.org).
Analysis of Recombination

Twenty-eight CSFV genomes were downloaded from GenBank and aligned with Clustal W35. Among these genomes, the redundant sequences were excluded from further analysis (criteria: ≥99% similarity). Putative recombinant sequences were identified with the SimPlot program [7]. Bootscanning [8] was also carried out employing software from the SimPlot program [7], using the putative recombinant sequence as a query. Mosaicism is suggested when high levels of phylogenetic relatedness between the query sequence and more than one reference sequence in different genomic regions is obtained.

CSFV Infection Test

RT-PCR and antibody assay kit were used to measure CSFV RNA and CSFV-specific antibody and make sure all 50-day-age Landrace pigs used in this experiment are free of CSFV. Three CSFV-negative pigs in each group were then intramuscularly injected with 1 ml CSFV Thiverval strain or CSFV-GZ-2009 strain (1 × 10^5 TCID50/ml), and pigs in the uninfected group were intramuscularly injected with 1 ml physiological saline.

Body temperature and clinical symptoms were recorded every day post-infection. 10 ml peripheral blood was collected from the anterior vena cava of pigs after typical CSF clinical symptoms appeared. A blood sample with EDTA-K2 anticoagulant was mixed with hydroxypropylmethylcellulose, and washed three times with PBS at 4°C, each time followed by centrifugation for 10 min at 1,500 rpm. After centrifugation, 0.1 mol/l hydrochloric acid was added to dilute the leukocyte preparation (1:20), and mononuclear lymphocytes were counted in a hemocytometer using light microscopy.

Results

Isolation of Strain CSFV-GZ-2009

PK-15 cell lines were inoculated with positive specimens, followed by blind passage culture for CSFV isolation. The E2 gene of strain CSFV-GZ-2009 was amplified by RT-PCR and obtained the anticipant length of 1,119 nt (the sequence has been uploaded to the GeneBank No. HQ380231). The antigen of strain CSFV-GZ-2009 was identified by indirect immunofluorescence assay using WH303 antibody resulting in brilliant green fluorescence which was indicative of viral antigen (fig. 1). In a word, the strain CSFV-GZ-2009 has been successfully isolated.

Full-Length Genomic Sequence of CSFV-GZ-2009 Isolate

As expected, nine cDNA fragments of CSFV-GZ-2009 isolate were amplified using nine primers in table 1 by RT-PCR and sequenced. The spliced complete genome is 12,298 nt in length and C+G content is about 46.6%. It includes a 373-nt 5′-UTR (including an Internal Ribosome Entry Site, IRES sequence), an 11,697-nt ORF encoding a polyprotein of 3,898 amino acids, and a 228-nt 3′-UTR. The genome sequence has been submitted to the

<table>
<thead>
<tr>
<th>Primer</th>
<th>Position (nt)</th>
<th>Oligonucleotide sequence (5′–3′)</th>
<th>Annealing temperature, °</th>
<th>Fragment size (nt), bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 and R1</td>
<td>1–24</td>
<td>GTATACTGAGGTTAGTCCATTCTCG CTTCTTTTACCTCAGTCC</td>
<td>48</td>
<td>910</td>
</tr>
<tr>
<td>F2 and R2</td>
<td>911–933</td>
<td>AAGCGCAGTAGGTCAACAAAGG GTAATCTTCTGACGGTACCC</td>
<td>52</td>
<td>1,553</td>
</tr>
<tr>
<td>F3 and R3</td>
<td>2,441–2,463</td>
<td>CGGCTAAGCTGCAAGGAGATTA TCTAGACCCAGGACAGTGGTT</td>
<td>57</td>
<td>2,670</td>
</tr>
<tr>
<td>F4 and R4</td>
<td>5,071–5,089</td>
<td>CAAGTAGCTCTGGTGCCGA CACACATCTCTGACATCCC</td>
<td>48</td>
<td>2,338</td>
</tr>
<tr>
<td>F5 and R5</td>
<td>7,382–7,405</td>
<td>GCTCAGGGGGATGTCAGAGATGT TAGCTGAGGATCTTTTCCACT</td>
<td>56</td>
<td>1,041</td>
</tr>
<tr>
<td>F6 and R6</td>
<td>8,413–8,445</td>
<td>CGCCCAATCTAATACATTACCTAGAGCT CAGTTCTCATGAAATACCTTTTGACCAACCTCC</td>
<td>55</td>
<td>1,501</td>
</tr>
<tr>
<td>F7 and R7</td>
<td>9,720–9,738</td>
<td>AAGCACAAATGGAGAGCT TTGCTTGGGGATATGTTATG</td>
<td>47</td>
<td>1,071</td>
</tr>
<tr>
<td>F8 and R8</td>
<td>10,500–10,518</td>
<td>AAGCACCAGATGGTAGAGTC ACCATATTGACACAGTCTTG</td>
<td>45</td>
<td>1,331</td>
</tr>
<tr>
<td>F9 and R9</td>
<td>11,665–11,681</td>
<td>AGGGAGCCGAATCTCG GGGCGGTTAGGAAATTACC</td>
<td>44</td>
<td>634</td>
</tr>
</tbody>
</table>

Table 1. Nucleotide sequences of primers used for amplification of isolate CSFV-GZ-2009 (forward primers are listed first)
A CSFV Strain Isolated from a CSF Outbreak in South China

Table 2. Information of the CSFV strains used in this study

<table>
<thead>
<tr>
<th>Geographic location</th>
<th>Defined isolate name</th>
<th>GenBank accession number</th>
<th>Total length (5'-NTR,ORF,3'-NTR), nt</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>CSFV/GZ/2009</td>
<td>HQ380231</td>
<td>12,298 (373,11697,228)</td>
</tr>
<tr>
<td></td>
<td>Shimen/CA/1999</td>
<td>AF092448</td>
<td>12,298 (373,11697,228)</td>
</tr>
<tr>
<td></td>
<td>cF114/CA/2001</td>
<td>AF333000</td>
<td>12,297 (373,11697,227)</td>
</tr>
<tr>
<td></td>
<td>C/CA/2003</td>
<td>AY382481</td>
<td>12,310 (374,11697,239)</td>
</tr>
<tr>
<td></td>
<td>CWH/CA/2003</td>
<td>AY663656</td>
<td>12,310 (374,11697,239)</td>
</tr>
<tr>
<td></td>
<td>HCLV/CA/2002</td>
<td>AF531433</td>
<td>12,310 (374,11697,239)</td>
</tr>
<tr>
<td></td>
<td>GXWZ02/CA/2003</td>
<td>AY367767</td>
<td>12,296 (373,11697,226)</td>
</tr>
<tr>
<td></td>
<td>CSFV39/CA/2001</td>
<td>AF407339</td>
<td>12,297 (373,11697,227)</td>
</tr>
<tr>
<td></td>
<td>P97/CA(TW)/1995c</td>
<td>L49347</td>
<td>12,144 (372,11697,181)</td>
</tr>
<tr>
<td></td>
<td>0406/CH/01/TWN/2004</td>
<td>AY568569</td>
<td>12,296 (372,11697,225)</td>
</tr>
<tr>
<td></td>
<td>LPC/CA(TW)/2001</td>
<td>AF352565</td>
<td>12,344 (373,11694,277)</td>
</tr>
<tr>
<td></td>
<td>SWH/CA/2004</td>
<td>DQ127910</td>
<td>12,296 (373,11697,226)</td>
</tr>
<tr>
<td></td>
<td>JL1(06)/2008</td>
<td>EU497410</td>
<td>12,298 (373,11697,228)</td>
</tr>
<tr>
<td></td>
<td>C-ZJ/2008</td>
<td>HM175885</td>
<td>12,311 (374,11697,240)</td>
</tr>
<tr>
<td></td>
<td>Thiverval/2008</td>
<td>EU490425</td>
<td>12,321 (373,11697,251)</td>
</tr>
<tr>
<td>Japan</td>
<td>ALD/JP/1999</td>
<td>D49532</td>
<td>12,298 (373,11697,228)</td>
</tr>
<tr>
<td></td>
<td>GPE.-/JP/1999</td>
<td>D49533</td>
<td>12,298 (373,11697,228)</td>
</tr>
<tr>
<td>Russia</td>
<td>CS/RA/2000</td>
<td>AF099102</td>
<td>12,310 (372,11697,241)</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Brescia/SL/1998</td>
<td>AF091661</td>
<td>12,297 (373,11697,227)</td>
</tr>
<tr>
<td></td>
<td>Allort/187/SL/2003</td>
<td>X87939</td>
<td>12,298 (373,11697,228)</td>
</tr>
<tr>
<td></td>
<td>CAP/SL/1996</td>
<td>X96550</td>
<td>12,297 (373,11697,227)</td>
</tr>
<tr>
<td></td>
<td>Eystrup/SL/2003</td>
<td>AF326963</td>
<td>12,301 (373,11697,231)</td>
</tr>
<tr>
<td>Germany</td>
<td>Riems/GM/1996</td>
<td>U45477</td>
<td>12,298 (373,11697,228)</td>
</tr>
<tr>
<td></td>
<td>Glentorf/GM/1996</td>
<td>U45478</td>
<td>12,278 (353,11697,228)</td>
</tr>
<tr>
<td></td>
<td>Allort/Tuebingen</td>
<td>J04358</td>
<td>12,297 (372,11697,228)</td>
</tr>
<tr>
<td></td>
<td>CSF857/2006</td>
<td>GU233731</td>
<td>12,297 (374,11697,226)</td>
</tr>
<tr>
<td>Denmark</td>
<td>Paderborn/DM/2002</td>
<td>AY072924</td>
<td>12,229 (372,11697,160)</td>
</tr>
<tr>
<td>Korea</td>
<td>LOM/2008</td>
<td>EU789580</td>
<td>12,298 (373,11697,228)</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>C strain/NL/1996</td>
<td>Z46258</td>
<td>12,311 (373,11697,241)</td>
</tr>
</tbody>
</table>

* Data already published since August 2010 were collected from GenBank. The represented sequences are arranged by the origin of country.
* All sequences were identified by the name of CSFV isolate and GenBank accession number.
* A top codon was found approximately in the middle of the polyprotein ORF in P97 strain. But we still consider the polyprotein ORF of P97 to be as long as 11,697 nt.
NCBI and the GenBank accession number is HQ380231. The 5′-UTR of CSFV-GZ-2009 is 99.4% homology with the Shimen strain, and two nucleotide mutations at positions 305 (C→T) and 365 (T→A) occurred. The ORF nucleotide sequence of the CSFV-GZ-2009 strain shows 99.4% and 98.9% identity to that of the Shimen strain and SWH strain, while the identity of the deduced amino acid sequences of CSFV-GZ-2009 shows 99.3% and 98.87% homology, respectively.

**CSFV-GZ-2009 Isolate Is Closely Related to Highly Virulent Strains**

Sequence analysis of the CSFV-GZ-2009 genome with 28 CSFV strains that were reserved in the GenBank database indicated that CSFV-GZ-2009 was a highly virulent strain. As shown in figure 2, a NJ tree constructed by the bootstrap test based on the sequence alignment of these 29 CSFV genomes demonstrated that these CSFV strains were distinctly divided into three gene clusters, namely group I (highly virulent strains), group II (avirulent strains) and group III (moderately virulent strains) (fig. 2). CSFV-GZ-2009 is clustered as a highly virulent type and closely related to four CSFV isolates, JL1 (06), cF114, Shimen and SWH, with pairwise distances of 0.0037, 0.0043, 0.0058 and 0.0107, respectively. The close
A CSFV Strain Isolated from a CSF Outbreak in South China

Intervirology 2013;56:122–133

the non-structural protein regions, N\textsuperscript{pro}, NS3, NS4A and NS5A genes are divergent, and N\textsuperscript{pro} has the lowest conservation (88.1–95.6%). The structural protein, nucleocapsid protein C, is highly conserved (91.9–96.8% identity). Among all the NJ trees constructed by each region of the CSFV genome, only E\textsuperscript{rms}, E2 and NS5B-based trees (fig. 3) are distinctly divided into three clusters based on different virulence, possibly corresponding to their roles as antigenic determinants (E\textsuperscript{rms} and E2). The CSFV-GZ-2009 strain belongs to group I. For the NJ trees of other regions, although the branches are generally divided according to virulence, a few overlaps are evident.

Sequence Alignment of Regions Reveals Mutations in 5'-UTR, E\textsuperscript{rms}, E2 and 3'-UTR

Similar to the majority of CSFV viruses, the sequence of 5'-UTR of the CSFV-GZ-2009 strain is 373 nt. Adenine

Fig. 3. The NJ trees based on the three proteins of CSFV, E\textsuperscript{rms} (a), E2 (b), and NS5B (c), respectively.

Characteristics of Sequence Conservation in Different Regions of the CSFV-GZ-2009 Genome

The similarity of nucleotide and amino acid sequences of each region (5'-UTR, N\textsuperscript{pro}, C, E\textsuperscript{rms}, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, NS5B, and 3'-UTR) of CSFV-GZ-2009 and other CSFV strains is compared in table 3. The most highly conserved region in CSFV genome is 5'-UTR (89.5–96.7% identity in nucleotide sequences, 96.7% is the average value), whereas p7 shows the lowest sequence identity (79.2–92.7%). In the coding region, the non-structural proteins are more conserved than the structural proteins, NS4A exhibiting the highest conservation (93.8–99.2% identity in amino acid sequences). In

relationship between CSFV-GZ-2009 and these four highly virulent isolates shows they may belong to the same topotype.

Characteristics of Sequence Conservation in Different Regions of the CSFV-GZ-2009 Genome

The similarity of nucleotide and amino acid sequences of each region (5'-UTR, N\textsuperscript{pro}, C, E\textsuperscript{rms}, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, NS5B, and 3'-UTR) of CSFV-GZ-2009 and other CSFV strains is compared in table 3. The most highly conserved region in CSFV genome is 5'-UTR (89.5–96.7% identity in nucleotide sequences, 96.7% is the average value), whereas p7 shows the lowest sequence identity (79.2–92.7%). In the coding region, the non-structural proteins are more conserved than the structural proteins, NS4A exhibiting the highest conservation (93.8–99.2% identity in amino acid sequences). In

relationship between CSFV-GZ-2009 and these four highly virulent isolates shows they may belong to the same topotype.
Fig. 4. Alignment of the 5′-end sequences from different subgroups of CSFVs. Sources of the sequence data are CSFV-GZ-2009 (HQ380231), HCLV (AF091507), Brescia (AF091661), GXWZ02 (AY367767), Strain39 (AF407339) and Alfort-Tubingen (J04358). Dots (.........) indicate the identical nucleotides as the consensus sequence. The remarked secondary structure containing the various domains (hairpins Ia, Ib, II, IIIa, IIb, IIc, IIId1, IIId2, IIIe) and the pseudoknot, which was designated according to the nomenclature in current usage.

Fig. 5. Alignment of the 3′-end sequences from viruses located in different subgroups. Sources of the sequence data are CSFV-GZ-2009 (HQ380231), HCLV (AF091507), Brescia (AF091661), GXWZ02 (AY367767), Strain39 (AF407339) and Alfort-Tubingen (J04358). Dots (.........) indicate the identical nucleotides as in the consensus lane, and the dashes (--------) indicate a gap. The 3′-UTR consists of a variable region of 61 nt, three repeat AT-rich stretch (box) and a highly conserved region with approximately 120 nt. A 12-nt insertion found in vaccine strains (HCLV) was underlined.
Fig. 6. The amino acid sequence comparison of E\textsuperscript{trans}-p7 (859 aa) with 12 representative CSFV isolates. Some motifs, T- and B-lymphocyte epitopes mainly, and only one transmembrane structure (TMS) are indicated in panels. 15 Cys of glycoprotein E2, absolutely conservative aa residues, are identified in the shadow. Surprisingly, the amino acid deletion occurred at 124 aa in the E\textsuperscript{trans} of the LPC/CA (TW)/2001 strain.
at 48 nt in HCLV, CSFV-GZ-2009, Brescia, Strain39 and GXWZ02 strains was deleted in Alfort-Tubingen virus, while adenine at 123 nt in HCLV was deleted in CSFV-GZ-2009, Brescia, Alfort-Tubingen, Strain39 and GXWZ02 strains (fig. 4). Analysis of the 5'-UTR from various groups of viruses including HCLV (subgroup 1.1), Brescia (subgroup 1.2), GXWZ02 (subgroup 2.1), Strain39 (subgroup 2.2) and Alfort-Tubingen (subgroup 2.3), a highly variable region in CSFV-GZ-2009 locates at the region of 1–70 nt with 0.0–26.1% divergence and its highly conserved region locates at 71–373 nt with only 0.3–4.1% divergence (fig. 4).

The 3'-UTR of the CSFV-GZ-2009 virus contains 227 nucleotides. At the 3'-end, the CCC nucleotides were similar to other CSFVs, type I bovine viral diarrhea viruses (BVDV-1) and border disease viruses (BDVs) [9, 10]. Comparison of the 3'-UTR among various groups of viruses, including subgroups 1.1, 1.2, 2.1, 2.2 and 2.3, showed that the 3'-end of CSFV-GZ-2009 contains a highly variable region (positioning 1–60 nt in the 3'-UTR region) with 0.0–53.6% divergence, a 3-repeat AT-rich stretch and a highly conserved region (positioning 74–238 nt in the 3'-UTR region) with only 1.8–9.1% divergence. Different from other CSFV strains, the vaccine strains such as HCLV contains a 12-nt (CTTTTTTCTTTT) insertion at 62–73 nt in the 3'-UTR region (fig. 5) [11].

The sequences of two B-cell epitopes in the enveloped glycoprotein E2 (residues 829–837 and 995–998) for the CSFV-specific murine monoclonal antibody (mAb) WH303 and the pan-pestivirus-reactive murine mAb 4-9D4 have been previously reported [12]. We also confirmed these epitopes in the CSFV-GZ-2009 genome. Furthermore, 15 cysteine residues in E2 are conserved among different strains. Some studies have revealed that four antigenic determinant regions of E2 are evidently affected by the deletion or replacement of these Cys. The sequence of the Ribonuclease T2 family histidine-active site is located at 292–299 aa and 338–349 aa in the E\textsuperscript{\text{ns}} region (fig. 6).

Secondary Structure Prediction of Pre-Polyprotein of CSFV-GZ-2009

The polyprotein of CSFV-GZ-2009 is composed of 3,898 aa. Secondary structure analyses are performed in the PROSITE database. There is a specific lysine-rich region at 177–249 aa in the middle of structural protein C (169–267 aa), where there is no analogous region in the CSFV avirulent strains. There are 27 N-glycosylation sites between 107 and 3,797 aa. 118 phosphorylation sites are observed, including 51 protein kinase C, 54 casein kinase II, 8 tyrosine kinase, 5 CAMP- and cGMP-dependent protein kinase. The sequence of the Ribonuclease T2 family histidine-active site is located at 338–349 aa in the E\textsuperscript{\text{ns}} region, and plays an important role in replication, expression and regulation in the process of CSFV infection (fig. 6).

The SMART database was utilized to analyze the structural and functional domains of the CSFV-GZ-2009 polyprotein. Two motifs, DEXDc (DEAD-like helicases superfamily) located at 1,777–1,977 aa and HELICc (helicase superfamily C-terminal domain) located at 2,001–2,100 aa, are both included in the NS3 protein. Furthermore, ten transmembrane domains are concentrated within the segment at 804–1,384 aa, six of which are in the NS2 protein (1,127–1,589 aa), two in the E2 protein and two in the p7 protein (fig. 6).

CSFV-GZ-2009 Infection

To further determine that CSFV-GZ-2009 is a highly virulent strain, the investigation was undertaken in pigs after infection of the CSFV-GZ-2009 strain and Thiverval strain (moderately virulent strains). Results showed that there was little change in body temperature, but no typical clinical symptoms were observed in the Thiverval-infected group, while in the CSFV-GZ-2009-infected group, all pigs showed typical CSFV symptoms by day 3 post-infection, including elevated body temperatures (fig. 7), varied degrees of anorexia and diarrhea, and an ocular discharge. By day 7 post-infection, deterioration continued, body temperature was elevated to 41.5°, hind
limbs were weak with the typical staggering gait and a cyanotic discoloration of the skin of the abdomen, ears and legs was also observed. By day 15 post-infection, body temperatures decreased to 39.0%; all inoculated pigs died.

A rapid onset of mononuclear lymphocytopenia was also detected at day 7 post-infection in inoculated pigs, levels of mononuclear lymphocytes in CSFV-GZ-2009-infected pigs declining from $2.83 \times 10^7$ to $1.3 \times 10^7$ cells/ml and from $2.95 \times 10^7$ to $1.63 \times 10^7$ cells/ml in Thiver-val-infected animals. This result confirms the depletion of mononuclear lymphocytes in the infected pigs.

**Discussion**

CSF is one of the most devastating diseases for the pig industry throughout the world, from both the economic and sanitary points of view [13]. Eradication programs against CSF are based on rapid identification of outbreaks, slaughtering of infected herds, control of pig and pork movements within infected regions, and forwards and backwards tracing of disease spread. In recent years, molecular typing has proven to be a potent method to find relations between different outbreaks of CSF and in combination with epidemiological surveys has helped to trace the spread of the disease [14].

In this study, the whole genome of the CSFV-GZ-2009 strain was cloned and its complete genome was sequenced and analyzed in comparison with other CSFV strains. By analyzing with 28 other reported CSFV genomes, the novel CSFV-GZ-2009 strain is placed together with Shimen, JL1-06, cF114 and SWH strains and shows 99.4, 99.6, 99.5 and 98.9% identities at the nucleotide, respectively. The novel CSFV-GZ-2009 strain is clustered with the gene cluster I (highly virulent strains) and separated from the avirulent Chinese C strain, with a pairwise distance of 0.149. It is suggested that the novel CSFV-GZ-2009 strain is of high virulence. It also indicates that the CSFV-GZ-2009 strain is unlikely a revertible form of the vaccine virus.

A previous study revealed that homologous recombination occurred in natural CSFV populations, generating genetic diversity, which is an important mechanism in CSFV evolution [15]. It also suggested that recombination may result in the change of the virus epidemiologic character. In China, a nation-wide policy of biannual vaccinations of pigs (spring and autumn) has always been performed, and large-scale outbreaks of CSF have rarely been seen. However, this does not mean that the Shimen strain has disappeared in the field. In this study, the novel CSFV-GZ-2009 strain shares 99.4% genome sequence identity with the Shimen strain (AF092448). It suggested that the CSFV-GZ-2009 strain was possibly derived from the Shimen strain. In view of the applicability of RNA recombination, the new recombinant CSFV vaccine with a desired phenotype can be designed to protect pigs against CSFV infection. Nevertheless, by ignoring the presence of recombination, phylogenetic analysis may be severely compromised [16]. As the molecular epidemiological study was limited by the phylogenetic analysis that was based on the several subgenomic regions, it may lead to erroneous conclusions. Thus, the availability of new sequence data covering the complete CSFV genome will help establish a more consistent genetic relatedness among CSFV isolates and the role of recombination in its evolution.

By comparing the CSFV-GZ-2009 isolate and other virulent and avirulent strains, a highly conservative Lys-rich region between CSFV-GZ-2009 and other virulent CSFV strains were found, but not among avirulent strains. We presume that the Lys-rich region may be closely related to the virulence of CSFV strain CSFV-GZ-2009. In this study, CSFV-GZ-2009 E2 glycoprotein contains a discrete epitope, TAVSPTTLR, recognized by mAb WH303, used to differentiate CSFV from related ruminant pestiviruses and is routinely used for CSF diagnostics. Research has demonstrated that the CSFV-specific sequences in the WH303 epitope of E2 have been identified as a CSFV virulence determinant, involving a mechanism specific and critical for CSFV infection [17]. We therefore presume that E2 residue TAVSPTTLR plays a significant role in GZ-2009-CSFV virulence.

Virus glycoproteins are crucial in key steps of the virus cycle such as attachment to host cell receptors, entry, assembly of newly produced viral progeny, and exit. Research has observed that rescue of viable virus was completely impaired by removal of all putative glycosylation sites in E2 but restored when mutation N185A reverted to wild-type asparagine-produced viable virus that was attenuated in pigs. Single mutations of each of the E2 glycosylation sites showed that amino acid N116 (N1v virus) was responsible for Brescia infectious clone (BICv) attenuation [18]. Recent studies have also demonstrated that the E'^rs' glycosylation can influence CSFV pathogenesis [19]. The CSFV E'^rs' could directly bind to exogenous dsRNA and inhibit dsRNA-induced IFN-β; importantly, the N-glycan of CSFV E'^rs' is essential for E'^mr' blocking of IFN-β induction [20]. E1 along with E'^rs' and E2 [21] is
also implicated in CSFV adsorption to host cells [22]. It has been demonstrated that residue N594 of CSFV E1 is critical for virus viability and removal of all three putative glycosylation sites in E1 yielded nonviable progeny, while single or dual site mutants excluding N594 were viable. However, individual N594 (E1.N3 virus) or combined N500/N513 (E1.N1N2 virus) substitutions resulted in viable progeny. In this study, there are seven putative glycosylation sites in the CSFV-GZ-2009 E2 glycoprotein. The secondary structure analysis of CSFV-GZ-2009 displayed that E ℓns contains seven putative N-linked glycosylation sites and CSFV strain CSFV-GZ-2009 E1 glycoprotein contains three N-linked putative glycosylation sites at CSFV amino acid residue position N500, N513 and N594. These characteristics of the CSFV strain CSFV-GZ-2009 may possibly play an important role in the CSFV infectious cycle and virus immune evasion. Moreover, in view of the previously observed data of E ℓns, E2 and E1, it is suggested that strain CSFV-GZ-2009 of CSFV may be the best candidate for developing CSFV live-attenuated vaccines when glycosylation patterns have been modified. Previous research showed that there were distinct variants found causing CSF epizootics in Guangdong province, 14 field viruses belonging to group 2, which could be further subdivided into three subgroups. Guangdong province, located in the extreme south of China bordering Hong Kong, has an extensive commercial and pig movement population which allows transmission of divergent CSFVs. However, the CSFV strain CSFV-GZ-2009 isolated from Guangdong province was classified within subgroup 1.1 of group 1. We presumed that as follows: China has the largest population of pigs in the world, and factors such as the coexistence of various animal husbandry methods, a vaccination-based disease eradication policy, highly mobile commercial herds and a complex ecological environment have possibly had an impact on virus spread, variation and evolution [6], making homologous recombination occur in CSFV natural populations. Recombination may also result in the change of the CSFV epidemiologic character. Under the pressure of massive vaccination in China, large-scale outbreaks of CSF have rarely been seen. This does not mean that classical highly virulent subgroup 1.1 strains have disappeared in the field. The highly virulent subgroup 1.1 CSFV-GZ-2009 isolate therefore caused CSF outbreaks in pig herds in Guangdong province.

In conclusion, the CSFV-GZ-2009 strain belonged to subgroup 1.1 of group 1. The secondary structure of the non-translated region was predicted and some functional domains that are crucial for viral replication and regulation were indicated. To our knowledge, this is the first report of a complete genome analysis of CSFV which was isolated in Guangdong. It provides a basic understanding of the molecular epidemiology of CSFV in this region and is also useful for understanding the processes involved in CSFV evolution.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (Nos. 31072137, 31172321 and 30771611), the Key Project of National Natural Science Foundation of Guangdong province, China (No. S2011020001037) and Special Fund for Agro-Scientific Research in the Public Interest (No. 201203056) and Research Fund for the Doctoral Program of Higher Education of China (No. 20114404110015).

References

A CSFV Strain Isolated from a CSF Outbreak in South China


Hulst MM, Moormann RJ: Inhibition of pestivirus infection in cell culture by envelope proteins E\textsuperscript{rmn} and E2 of classical swine fever virus: E\textsuperscript{rmn} and E2 interact with different receptors. J Gen Virol 1997;78:2779–2787.