Arch Virol (2007) 152: 1457–1465
DOI 10.1007/s00705-007-0971-8
Printed in The Netherlands

Archives of Virology

# Characterization of complete genome sequence of the spring viremia of carp virus isolated from common carp (*Cyprinus carpio*) in China

**Y. Teng**[1,*], **H. Liu**[2,3,*], **J. Q. Lv**[2], **W. H. Fan**[2], **Q. Y. Zhang**[3], and **Q. W. Qin**[1]

[1] State Key Laboratory of Biocontrol, College of Life Sciences, Sun Yat-sen University, Guangzhou, P.R. China
[2] The Key Laboratory of Aquatic Animal Diseases, Shenzhen Exit & Entry Inspection and Quarantine Bureau, Shenzhen, P.R. China
[3] State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, P.R. China

Received November 15, 2006; accepted March 14, 2007; published online April 20, 2007


## Summary

The complete genome of spring viraemia of carp virus (SVCV) strain A-1 isolated from cultured common carp (*Cyprinus carpio*) in China was sequenced and characterized. Reverse transcription-polymerase chain reaction (RT-PCR) derived clones were constructed and the DNA was sequenced. It showed that the entire genome of SVCV A-1 consists of 11,100 nucleotide base pairs, the predicted size of the viral RNA of rhabdoviruses. However, the additional insertions in bp 4633–4676 and bp 4684–4724 of SVCV A-1 were different from the other two published SVCV complete genomes. Five open reading frames (ORFs) of SVCV A-1 were identified and further confirmed by RT-PCR and DNA sequencing of their respective RT-PCR products. The 5 structural proteins encoded by the viral RNA were ordered 3′-N-P-M-G-L-5′. This is the first report of a complete genome sequence of SVCV isolated from cultured carp in China. Phylogenetic analysis indicates that SVCV A-1 is closely related to the members of the genus *Vesiculovirus*, family *Rhabdoviridae*.

## Introduction

Spring viraemia of carp (SVC) is a viral disease that causes outbreaks in common carp (*Cyprinus carpio*) and several other freshwater fish species like koi carp (*C. carpio koi*), crucian carp (*Carassius carassius*), big head (*Hypophthalmichthys molitrix*), silver carp (*Aristichthys nobilis*) and sheatfish (*Silurus glanis*). The disease was initially identified in European countries, and recently has been reported in the Middle East area and America as an emerging viral disease [2, 6, 10, 20].

SVC outbreaks often occur at water temperatures between 10 and 18 °C, and mostly in the spring. The major clinical signs in common carp are hemorrhages of the gills and skin, abdominal distension and an inflamed vent, internal ascites, peritonitis and petechial hemorrhages of the internal swim-

* These authors contributed equally to this work.
Author's address: Dr. Qi Wei Qin, State Key Laboratory of Biocontrol, College of Life Sciences, Sun Yat-sen University, 135 West Xingang Road, Guangzhou 510275, P.R. China. e-mail: lssqinqw@mail.sysu.edu.cn

bladder wall and in skeletal muscle [2]. Mortality caused by SVC in juvenile carp can reach up to 70% during spring outbreaks. It has been reported that the disease has a significant impact on yearling carp in Europe, with estimated losses of 10–15% for 1-year-old carp, equivalent to approximately 4,000 metric tons annually [15].

The causative agent of SVC is spring viraemia of carp virus (SVCV), which exhibits a typical bullet shape with size about 80–180 nm in length and 60–90 nm in diameter and is closely related to members of the genus *Vesiculovirus*, family *Rhabdoviridae* [3, 6]. SVCV contains a linear, negative-sense and single-stranded RNA genome, which encodes 5 structural proteins: nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G), and viral RNA-dependent RNA polymerase (L) in the order 3′-N-P-M-G-L-5′ [5, 18]. To date, only two complete nucleotide sequences, representing the genome of an American Type Culture Collection (ATCC) reference strain of SVCV VR-1390, have been published: SVCV NC_002803 [5] and SVCV AJ318079 [10].

Recently, a novel SVCV strain was isolated from cultured common carp at our laboratory in China, and the virus isolate was able to infect and cause significant mortality in young carp. To our knowledge, this is the first isolation of pathogenic SVCV from fish in Asia, and we tentatively designated it as SVCV Asia-1 (SVCV A-1). In order to understand SVCV A-1 at the molecular level, the entire genome of the virus was cloned and sequenced. The authenticity of the genomic open reading frames (ORFs) was further confirmed by reverse transcription-PCR (RT-PCR). The gene and deduced amino acid sequences of the 5 structural proteins were compared with other known representative fish rhabdoviruses.

## Materials and methods

### *Cell line and virus*

Grass carp ovary (CO) cells were used for propagation of the SVCV A-1 strain. The CO cells were cultured in M199 medium containing 10% fetal bovine serum (FBS) at 24 °C. The viral pathogen SVCV A-1 was originally isolated from virus-infected common carp in China.

### *Virus propagation, purification and RNA extraction*

SVCV A-1 virus was propagated and purified as described by Hoffmann et al. [10] with minor modifications. Briefly, confluent monolayers of CO cells were inoculated with the virus at a MOI of approximately 0.1 for 1 h. The inoculum was removed and the monolayer cells were incubated in fresh culture medium supplemented with 10% FBS. When CPE became significant, virus-infected cultures were harvested and treated by 3 cycles of rapid freezing/thawing and ultrasonication, and then centrifuged at $4000 \times g$ for 20 min at 4 °C. The supernatant was collected and layered onto a 5-ml cushion of 10% sucrose in STE (0.15 M NaCl, 0.01 M Tris–HCl, pH 8.3, 0.01 M EDTA, pH 7.0), and then centrifuged in a Beckman SW41 rotor at 24,000 rpm for 120 min at 12 °C. The virus pellet was loaded onto 10–40% (w/v) sucrose gradients and centrifuged at 21,000 rpm for 60 min at 12 °C in a SW40 rotor. Resulting virus bands were collected by puncturing the centrifuge tubes with a needle, diluted with STE buffer, and repelleted in a SW41 rotor at 42,000 rpm for 60 min at 12 °C. The purified virus pellet was resuspended in 200 µl STE buffer and stored at −20 °C. The genomic RNA was extracted with phenol:chlorophorm:isoamyl alcohol and precipitated in 70% ethanol. The precipitated RNA was resuspended in diethypyrocarbonate (DEPC)-treated water.

### *Cloning of the SVCV genome*

In order to clone the 3′-terminal region of the genomic RNA, the 3′ poly(A) tail was synthesized with poly(A) polymerase. Tailing was catalyzed by 10 µl of poly(A)-tailing mixture, which contained 2 µl of 5 × reaction buffer, 0.5 µl of 100 mM ATP, 0.5 µl of RNasin (40 U/µl), 1 µl of yeast poly(A) polymerase (Promega, USA), and 10 ng of SVCV A-1 genomic RNA. The cDNA synthesis and polymerase chain reaction were conducted by the method of Kima et al. [13], using the ODT-M13 primer (5′-ACG ACG GCC AGT TTT TTT TTT TT-3′) for the first-strand synthesis, in combination with the SVCV N primer (5′-ATG AGT GTC ATT CGG ATC AA-3′) for PCR.

The 5′-end of the genomic RNA was cloned by poly(G)- and poly(C)-tailing of the cDNA, followed by polymerase chain reaction. The first strand was synthesized using the Access RT-PCR System (Promega, USA), with 1 µg of purified genomic RNA and 50 pM of SVCV-L8 (5′-TCT GGG ACA TGG GTA GAA TAG-3′) primer located at nucleotides 11041–11061 on the genomic RNA. The cDNA was treated with RNase H, extracted with phenol/chloroform, and precipitated with ethanol. The cDNA was divided into two aliquots, which were tailed separately with poly(G) or poly(C) as described previously [1, 7], using terminal deoxynucleotidyl transferase (TdT; TaKaRa, Japan). The tailed cDNA was amplified with the internal primer SVCV-t2 (5′-GTA AAC ATC CAG TTT TTT TC-3′), together with the SVCV-Rt (5′-ATC AAT AGA AGG GGG GGG GG-3′) and SVCV-Ft (5′-CCT GTC AGT ACC CCC CCC CC-3′)

**Table 1.** Oligonucleotide primers used for genome sequencing of SVCV A-1

| Clone | Primer name | Sense | Sequence (5′–3′) | Remark |
|---|---|---|---|---|
| pGT-GL | SVCV-C | Genomic | AGAGTAAACGGCCTGCAACA | *Ava*I |
| | SVCV-N | Antigenomic | CCAAAATGCTTCAATGTGAG | *Eco*RI |
| pGT-L1 | SVCV-FL1 | Genomic | GAGAGCCATACCATCGTTAT | *Nco*I |
| | SVCV-RL1 | Antigenomic | TCAGTATTGTTCACTAGTGA | |
| pGT-L2 | SVCV-FL2 | Genomic | AGATATACAGCTGATATTGACC | *Pst*I |
| | SVCV-RL2 | Antigenomic | AGATGAGACTCGACTCAAAG | *Eco*RI |
| pGT-L3 | SVCV-FL3 | Genomic | ACTACAACGGGAGGCCAGAC | |
| | SVCV-RL3 | Antigenomic | CAATCATCATAAGGCAAAAA | |
| pGT-L4 | SVCV-FL4 | Genomic | CAATGCACAATCCTGCCATT | *Eco*RI |
| | SVCV-RL4 | Antigenomic | TGCTAGAGCATTCCCAGATG | |
| pGT-L5 | SVCV-FL5 | Genomic | CTCAAGCAGATTTGCTTCGA | |
| | SVCV-RL5 | Antigenomic | TCTCAATTCTCCAGTTGTCA | |
| pGT-L6 | SVCV-FL6 | Genomic | GAGAATTCGACAAATTCTACAG | *Eco*RI |
| | SVCV-RL6 | Antigenomic | AATTTCTGATCAGGTACCCC | *Eco*RI |
| pGT-L7 | SVCV-FL7 | Genomic | ATCTGAAAAGATCTTTGAAA | *Ava*I |
| | SVCV-RL7 | Antigenomic | GACACGACAGTGTTTATTTT | *Hin*dIII |
| pGT-L8 | SVCV-FL8 | Genomic | ATCTTTAACAGTCTTTTGGAGC | *Eco*RI |
| | SVCV-RL8 | Antigenomic | CTATTCTACCCATGTCCCAG | *Cla*I |

primers, which are complementary to the poly(G) and poly(C) tails, respectively.

The primers for the N, P, M, G, and L genes of SVCV A-1 were derived from the published sequences of SVCV AJ318079 and are listed in Table 1. All PCR products including 3′ and 5′ regions were cloned into the pGEM-T vector (Promega, USA) and sequenced using an ABI Prism 3730XL DNA Analyzer (ABI, USA).

*Assembly and analysis of the SVCV genome*

After sequencing of the viral fragments, Vector NTI Suite 9.0 (InforMax Inc., Frederick, Mass.) was applied to create the contigs, assemble the genome, and draw the genomic map. Putative ORFs were predicted by submitting to http://www.softberry.com (Softberry Inc., Mount Kisco, N.Y.). In all cases, an ORF starting from the ATG codon with a minimum of 40 codons was considered to be a putative gene. These ORFs were searched in the NCBI nucleotide database for identity analysis. The nucleotide and deduced amino acid sequences were analyzed using Vector NTI Suite 9.0, and functional analysis of proteins encoded by the 5 ORFs was performed by using TMHMM-2.0 and OMIGA 2.0 software (Oxford Molecular Group Inc., Oxford, UK).

*RT-PCR confirmation*

Reverse transcription-polymerase chain reaction (RT-PCR) was used for confirmation of potential ORFs. One-step RT-PCR was performed by the protocol of the One Step RNA PCR Kit (AMV) (TaKaRa, Japan). The respective primers for RT-PCR designed from 5 ORFs of SVCV A-1 are listed in Table 2.

**Table 2.** The primers of five genes used for identification of the five SVCV A-1 ORFs

| Genes | Primer names | Sense | Sequence (5′–3′) | Product length |
|---|---|---|---|---|
| N | SVCV N1 | Genomic | GCCGATTATCCTTCCACT | 370 bp |
| | SVCV N2 | Antigenomic | TTGCTCTTCCCACTCTGT | |
| P | SVCV P1 | Genomic | TTTGGACCTGGGATAGTG | 528 bp |
| | SVCV P2 | Antigenomic | CTCGCTCTGTTTGGCTGT | |
| M | SVCV M1 | Genomic | CCCACTTACGAGGAGACAC | 343 bp |
| | SVCV M2 | Antigenomic | TGAAAGGAGGCATTTGACC | |
| G | SVCV G1 | Genomic | TCTGTTCATTTGGAGCCG | 850 bp |
| | SVCV G2 | Antigenomic | GGATTTCATCGTCGTCAT | |
| L | SVCV L1 | Genomic | CGATCCCGAGCAAGAAGGTTTTAG | 402 bp |
| | SVCV L2 | Antigenomic | CAGGCCCATTGGATTCTTTTCGTT | |

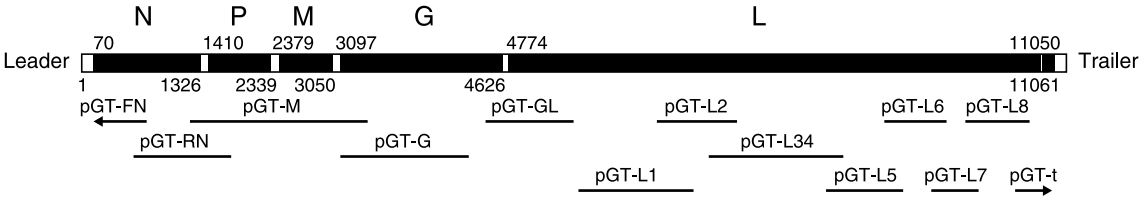



**Fig. 1.** Physical map of the SVCV A-1 genome. Location and relative size of the 5 genes are shown on the map, the numbers indicate starts and ends of the respective ORFs. The locations of cDNA clones are presented at the bottom

*Multiple sequence alignment and phylogenetic analysis*

The sequences of SVCV A-1 and other rhabdoviruses were used to determine the taxonomic position of SVCV A-1. Both nucleotide and amino acid sequence alignments were established using the default options in Clustal W [21], and the phylogenetic relationships among species were determined using the neighbor-joining (NJ) [19] and maximum parsimony methods as implemented in MEGA3 [14]. GenBank accession numbers for the sequences used are as follows: SVCV NC, **NC_002803**; SVCV AJ, **AJ318079**; bovine ephemeral fever virus (BEFV), **NC_002526**; Hirame rhabdovirus (HIRRV), **NC_000855**; infectious hematopoietic necrosis virus (IHNV), **NC_001652**; rabies virus (RABV), **AB009663**; snakehead rhabdovirus (SHRV), **NC_005093**; Tupaia rhabdovirus (TPRV), **AF147498**; viral haemorrhagic septicemia virus (VHSV), **NC_007020**; vesicular stomatitis Indian virus (VSIV), **NC_001560**.

## Results

*Determination of the complete SVCV genomic sequence*

The entire genome of SVCV A-1 was amplified and cloned into 14 overlapping clones (Fig. 1). The assembled SVCV A-1 genome is a negative single-strand RNA of 11,100 bp in length, similar to the 2 published SVCV genomes. The percentage of G+C content of SVCV A-1 (42%) is the same as in SVCV AJ and NC, and comparable with other rhabdoviruses: HIRRV (51%), IHNV (51%), VHSV (50%), SHRV (48%), RABV (45%), TPRV (43%), VSIV (41%) and BEFV (33%) (shown in Table 3). The complete genome sequence of SVCV A-1 reported here was deposited in the GenBank sequence database under accession no. **DQ097384**.

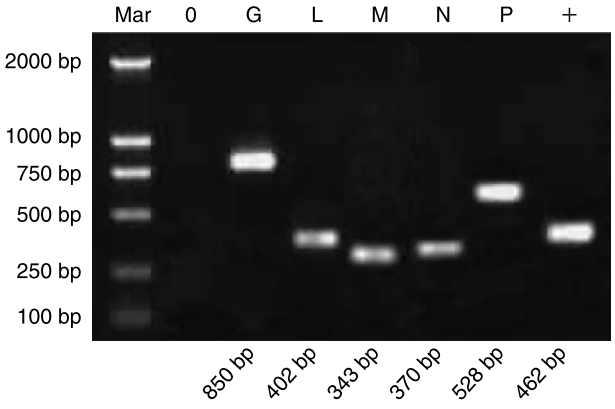



**Fig. 2.** RT-PCR identification of 5 ORFs of SVCV A-1. *Mar*, molecular markers; 0, negative control; *G–P*, show the 5 ORFs; +, positive control

*RT-PCR confirmation*

The RT-PCR confirmation result is shown in Fig. 2. The PCR products of the 5 genes were consistent with the designed length, confirmed the RNA transcription of the 5 ORFs and the successful sequencing of the SVCV A-1 genome.

*Analysis of the SVCV genome*

Five putative ORFs of the SVCV A-1 genome were identified and analyzed:

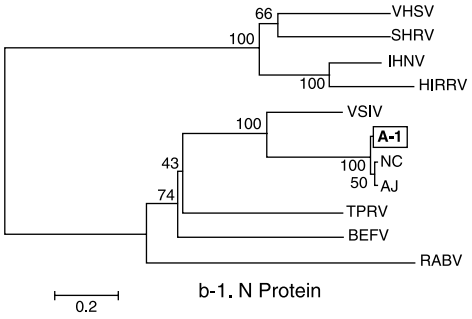
(i) ORF 1: nucleoprotein (N) gene

ORF1 is located between nt 70 and 1326, encoding an N protein of 418 amino acids, with a molecular mass of 47 kDa. The start codon ATG is localized downstream of the motif AACAG, and the transcription stop sequence TAA is located at nt 1326. The amino acid identities of N proteins

**Fig. 3.** Identity comparison and phylogenetic relationship of SVCV A-1 to representative rhabdoviruses. The analysis was based on multiple alignments of the 5 proteins of rhabdoviruses. *A-1*, SVCV A-1; *AJ*, SVCV AJ; *NC*, SVCV NC

| | A-1 | NC | AJ | BEFV | HRRV | IHNV | RABV | SHRV | TPRV | VHSV |
|---|---|---|---|---|---|---|---|---|---|---|
| NC | 97.6% | | | | | | | | | |
| AJ | 97.4% | 99.0% | | | | | | | | |
| BEFV | 31.0% | 31.0% | 31.3% | | | | | | | |
| HRRV | 11.4% | 12.6% | 12.4% | 11.6% | | | | | | |
| IHNV | 11.7% | 11.5% | 11.5% | 11.2% | 61.6% | | | | | |
| RABV | 19.4% | 19.8% | 19.8% | 21.6% | 11.4% | 11.6% | | | | |
| SHRV | 9.2% | 11.1% | 11.1% | 12.1% | 41.3% | 43.4% | 11.1% | | | |
| TPRV | 30.2% | 30.4% | 30.5% | 30.5% | 13.7% | 13.7% | 20.5% | 11.9% | | |
| VHSV | 12.0% | 11.1% | 11.1% | 10.7% | 38.6% | 39.8% | 12.7% | 45.8% | 11.4% | |
| VSIV | 52.8% | 51.9% | 52.1% | 28.8% | 12.0% | 13.3% | 18.8% | 12.5% | 30.8% | 10.0% |

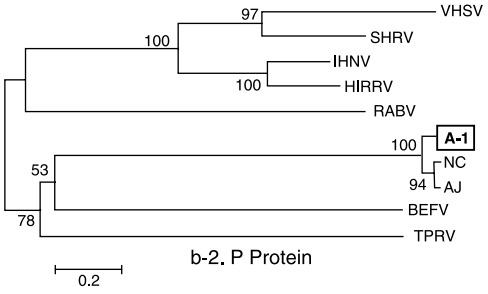
a-1. N Protein




b-1. N Protein

| | A-1 | NC | AJ | BEFV | HRRV | IHNV | RABV | SHRV | TPRV | VHSV |
|---|---|---|---|---|---|---|---|---|---|---|
| NC | 90.3% | | | | | | | | | |
| AJ | 89.7% | 98.7% | | | | | | | | |
| BEFV | 11.7% | 12.1% | 12.4% | | | | | | | |
| HRRV | 8.9% | 10.7% | 10.0% | 13.6% | | | | | | |
| IHNV | 10.6% | 10.6% | 11.5% | 12.2% | 63.5% | | | | | |
| RABV | 12.1% | 14.6% | 14.6% | 11.6% | 9.5% | 13.4% | | | | |
| SHRV | 11.6% | 11.3% | 12.9% | 12.2% | 27.0% | 29.3% | 13.1% | | | |
| TPRV | 14.9% | 15.2% | 15.2% | 11.5% | 9.5% | 10.4% | 11.9% | 12.2% | | |
| VHSV | 12.5% | 12.3% | 12.3% | 10.7% | 32.6% | 33.9% | 13.5% | 41.7% | 9.2% | |
| VSIV | | | | | | | | | | |

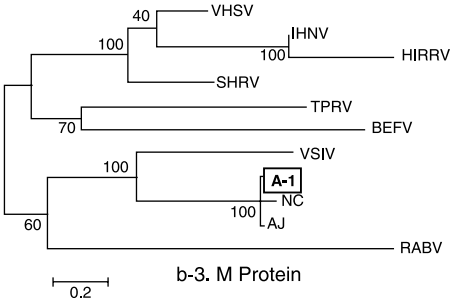
a-2. P Protein




b-2. P Protein

| | A-1 | NC | AJ | BEFV | HRRV | IHNV | RABV | SHRV | TPRV | VHSV |
|---|---|---|---|---|---|---|---|---|---|---|
| NC | 97.8% | | | | | | | | | |
| AJ | 98.2% | 99.6% | | | | | | | | |
| BEFV | 15.2% | 14.9% | 14.9% | | | | | | | |
| HRRV | 10.2% | 9.9% | 9.0% | 13.0% | | | | | | |
| IHNV | 12.8% | 12.0% | 12.0% | 9.0% | 74.4% | | | | | |
| RABV | 17.0% | 15.2% | 15.2% | 13.8% | 12.0% | 12.0% | | | | |
| SHRV | 10.7% | 11.2% | 9.9% | 11.6% | 34.7% | 33.3% | 9.9% | | | |
| TPRV | 11.4% | 10.1% | 10.1% | 14.8% | 9.5% | 7.8% | 12.9% | 11.0% | | |
| VHSV | 10.9% | 11.3% | 10.0% | 13.1% | 36.1% | 36.6% | 10.1% | 41.6% | 8.3% | |
| VSIV | 25.8% | 25.9% | 25.9% | 14.4% | 10.4% | 11.4% | 12.2% | 8.2% | 10.5% | 10.5% |

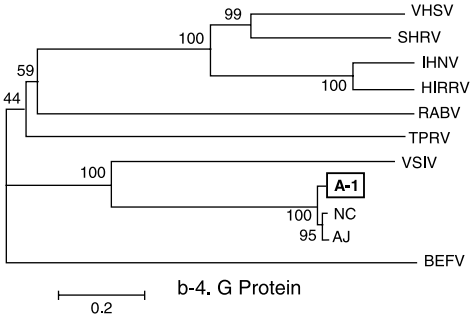
a-3. M Protein




b-3. M Protein

| | A-1 | NC | AJ | BEFV | HRRV | IHNV | RABV | SHRV | TPRV | VHSV |
|---|---|---|---|---|---|---|---|---|---|---|
| NC | 94.9% | | | | | | | | | |
| AJ | 95.1% | 99.4% | | | | | | | | |
| BEFV | 14.7% | 14.2% | 14.2% | | | | | | | |
| HRRV | 17.8% | 17.1% | 17.9% | 12.7% | | | | | | |
| IHNV | 16.1% | 16.3% | 16.3% | 14.7% | 73.6% | | | | | |
| RABV | 17.6% | 18.1% | 18.9% | 13.1% | 16.1% | 16.9% | | | | |
| SHRV | 18.3% | 18.6% | 18.6% | 13.6% | 39.0% | 38.2% | 18.4% | | | |
| TPRV | 19.6% | 20.0% | 20.2% | 16.0% | 15.5% | 16.5% | 17.4% | 18.1% | | |
| VHSV | 18.7% | 19.1% | 19.1% | 13.1% | 37.1% | 37.5% | 14.8% | 47.2% | 15.5% | |
| VSIV | 31.1% | 30.5% | 30.7% | 15.8% | 13.6% | 16.7% | 19.3% | 16.8% | 18.5% | 14.7% |

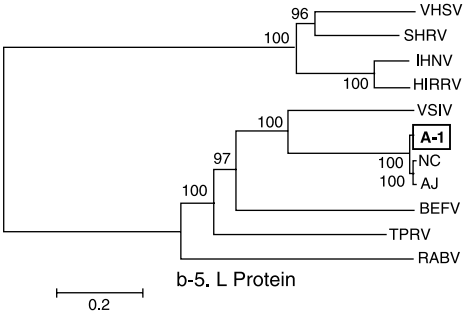
a-4. G Protein




b-4. G Protein

| | A-1 | NC | AJ | BEFV | HRRV | IHNV | RABV | SHRV | TPRV | VHSV |
|---|---|---|---|---|---|---|---|---|---|---|
| NC | 97.5% | | | | | | | | | |
| AJ | 97.6% | 99.6% | | | | | | | | |
| BEFV | 39.7% | 39.6% | 39.7% | | | | | | | |
| HRRV | 14.8% | 14.3% | 14.5% | 14.6% | | | | | | |
| IHNV | 14.8% | 14.8% | 14.7% | 14.2% | 84.4% | | | | | |
| RABV | 32.9% | 32.7% | 32.8% | 30.4% | 13.9% | 14.7% | | | | |
| SHRV | 14.8% | 14.7% | 14.5% | 15.1% | 58.7% | 58.2% | 14.2% | | | |
| TPRV | 39.5% | 39.2% | 39.2% | 36.1% | 14.5% | 15.6% | 30.7% | 14.9% | | |
| VHSV | 14.5% | 14.3% | 14.3% | 15.1% | 59.3% | 59.6% | 14.1% | 65.2% | 15.5% | |
| VSIV | 52.1% | 52.4% | 52.4% | 39.7% | 14.4% | 14.6% | 32.3% | 14.6% | 40.7% | 14.8% |

a-5. L Protein




b-5. L Protein

encoded by all known rhabdovirus are shown in Fig. 3 a-1. The N protein of SVCV A-1 was highly identical both to SVCV AJ, with 97.4%, and SVCV NC, with 97.6%.

(ii) ORF2: phosphoprotein (P) gene

The region of ORF2 between nt 1410 and 2339 of SVCV A-1 encodes a P protein of 309 amino acids with a molecular mass of 36 kDa. The ORF2 transcription starts at nt 1500 with the consensus start sequence AACAG and ends at nt 2339 with the consensus transcription stop sequence TAA. The P protein of SVCV A-1 is highly identical to those of SVCV AJ (89.7%) and SVCV NC (90.3%; Fig. 3 a-2).

(iii) ORF 3: matrix protein (M) gene

ORF3 for the M gene (nt 2379–3050) encodes a protein of 223 amino acids with a molecular mass of 26 kDa. The identities of the SVCV A-1 M protein with those of all other known rhabdoviruses are listed in Table 3 and shown in Fig. 3 a-3.

The highly conserved PPXY motif (P, Proline; X, any amino acid; Y, tyrosine), which is typical for the M proteins, is located between amino acids 17 and 20 of the M protein.

(iv) ORF 4: glycoprotein (G) gene

The region between nt 3097 and 4626 encodes the G gene. Transcription starts at nt 3097 and ends at nt 4626. The G protein consists of 509 amino acids with a calculated molecular mass of 57 kDa. The identities of the G protein with those of all known rhabdoviruses are listed in Table 3 and shown in Fig. 3 a-4.

The functional analysis of G protein showed one hydrophobic region between aa 1 and 50, one possible antigenic site between aa 50 and 450 and one transmembrane domain between aa 468 and 487. Two Arg-GLY-Asp (RGD) sites between aa 212 and 252 and one TYR-PHOSPHO site at aa 83 were also found in the SVCV A-1 G protein.

(v) ORF 5: RNA polymerase (L) protein

The fifth ORF on the SVCV A-1 genome encodes the viral RNA polymerase protein. ORF5 is the largest and is located between nt 4774 and 11,061. The resulting protein consists of 2095 amino acids with a calculated molecular mass of 121 kDa. Transcription starts at nt 4774 and ends at nt 11,061.

**Table 3.** Summary of genomic sequence data for members of nine virus species within the family *Rhabdoviridae*

| Characteristic | SVCV | | | BEFV | HIRRV | IHNV | RABV | SHRV | TPRV | VHSV | VSIV |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | A-1 | AJ | NC | | | | | | | | |
| Genome size (bp) | 11,100 | 11,019 | 11,019 | 14,900 | 11,034 | 11,158 | 11,926 | 11,550 | 11,440 | 11,131 | 11,161 |
| GC (%) | 42% | 42% | 42% | 33% | 51% | 50% | 45% | 48% | 43% | 51% | 41% |
| Molecular Weight (kDa) | 3607.50 | 3581.18 | 3581.18 | 4843.50 | 3586.05 | 3626.35 | 3875.95 | 3753.75 | 3718.00 | 3617.57 | 3627.32 |
| No. of putative ORFs | 5 | 5 | 5 | 9 | 6 | 6 | 5 | 6 | 7 | 6 | 5 |
| No. of structural protein | 5 | 5 | 5 | 11 | 6 | 6 | 5 | 6 | 7 | 6 | 5 |
| ORF size (no. of AA) | 223–2095 | 223–2095 | 223–2095 | 48–2144 | 122–1984 | 111–1986 | 202–2127 | 111–1986 | 122–1983 | 93–2107 | 229–2109 |
| GenBank accession no. | DQ097384 | AJ318079 | NC_002803 | NC_002526 | NC_000855 | NC_001652 | AB009663 | NC_005093 | AF147498 | NC_007020 | NC_001560 |

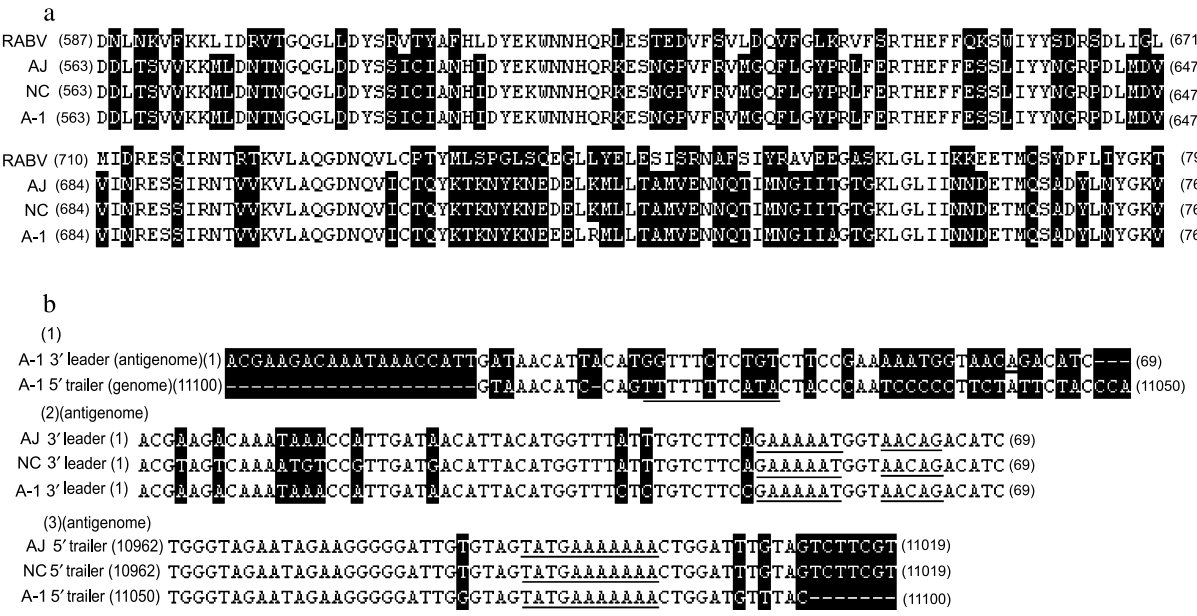



**Fig. 4.** **a**. Partial comparison of L protein conserved catalytic domains among rabies virus (RABV), SVCV A-1, SVCV AJ and SVCV NC. **b**. Analysis of the determined 3′ and 5′ terminal nucleotide sequences of SVCV A-1 genome. The leader RNA transcription stop signal, the transcription stop/polyadenylation signals, and the transcription initiation sequences are underlined. Different nucleotides and amino acids are shaded in black

The L protein encoded by the SVCV A-1 ORF5 is 97.6 and 97.5% identical to those of SVCV AJ and SVCV NC, respectively (Fig. 3 a-5).

The highly conserved catalytic domain is located between amino acids 563 and 786 of the L protein in SVCV A-1 (Fig. 4a).

(vi) Gene junctions, leader and trailer region

The SVCV A-1 genomic order of 3′ Leader-N-P-M-G-L-Trailer 5′ is identical to the genetic organization of members of species in the genera *Lyssavirus* and *Vesiculovirus* of family *Rhabdoviridae*. In the genome of SVCV A-1, distances between translation stop and start codons range from 39 (P-M) to 238 (N-P) nucleotides.

SVCV A-1 has a putative leader region of 69 base pairs at the 3′ terminus. The 3′-leader of SVCV A-1 is highly identical to those of published SVCVs (AJ and NC) with 95.7 and 84.1% (Fig. 4 b-2). The promoter is located between 1 and 19 of 69 nucleotides, the leader RNA transcription stop signal GAAAAAT is contained in 50–56 of 69 nucleotides and shows 100% identity to the corresponding sequences of published SVCVs (Fig. 4 b-2). The transcription start signal AACAG is located between nt 60 and 64. The 5′-trailer region of SVCV A-1 is located between nt 11,050 and 11,100, and exhibits an inverse complementarity (Fig. 4 b-1). The transcription stop/polyadenylation signal $TATG(A)_7$ is included in this region and shows 100% identity to those of published SVCVs (Fig. 4 b-3).

*Phylogenetic analysis*

The deduced amino acid sequences of the SVCV A-1 N, P, M, G and L proteins were compared with those of 2 published SVCVs and several other rhabdoviruses. Phylogenetic trees for all 5 proteins were constructed based on the multiple alignments (Fig. 3 b1-5). The phylogenetic analysis indicated that SVCV A-1 belongs to the genus *Vesiculovirus*, family *Rhabdoviridae*.

## Discussion

This paper firstly presents the complete nucleotide sequence of SVCV A-1 isolated from cultured common carp in China. The complete SVCV A-1 genome consists of 11,100 nucleotides with the predicted

size for the viral RNA of rhabdoviruses. The deduced genome organization of SVCV A-1 contains 5 ORFs encoding N, P, M, G and L, identical to the two published SVCV genomes. The protein sequences of the SVCV genome described here are more than 89% homologous to the published SVCV sequences for the leader region, the trailer region, and the 5 ORFs, especially in the M, G and L proteins. However, SVCV A-1 has two additional insertions at bp 4633–4676 and bp 4684–4724, and lacks 7 nucleotides in the 5′-trailer region.

The highly conserved domain, the PPXY motif, which is typical for the M protein, was also found in SVCV A-1. This domain might be related to virus budding [8, 11]. The highly conserved catalytic domains, typical for the L proteins of RABV, a member of the order *Mononegavirales*, are located between amino acids 587 and 811 [17, 22]. Based on the result of amino acid sequence alignment, the highly conserved catalytic domain was found to be located between amino acids 563 and 786 of the L protein in SVCV A-1.

The G protein of SVCV forms trimeric peplomers or spikes on the virus surface that bind to cellular receptors and induce viral endocytosis. The surface G protein acts as the most important viral antigen and determines the serological properties of rhabdoviruses [4, 9, 12]. The functional sites searched in G protein of SVCV A-1 include two RGD sites and one TYR-PHOSPHO site. RGD is a recognition site in the receptors of cell-adhesion protein families, reacting on cell adhesion. The TYR-PHOSPHO site often acts on disturbing early signaling.

The G gene has been used in the study of genogroup analysis of SVCV isolates [16, 20, 23]. Based on phylogenetic analysis of the nucleotide sequences of G proteins, SVCV isolates have been separated into four subgenogoups: genogroup Ia, originally from Asia; genogroups Ib and Ic, originally from Moldova, Ukraine and Russia; and genogroup Id, originally from the UK, with the exception of a few virus isolates from the former USSR [20]. The new sequence reported in this study indicates that our virus isolate belongs to the SVCV Ia cluster.

Four gene junctions with a TATG$(A)_7$ transcription stop polyadenylation signal at the end of each gene, and an AACAG transcription start signal for the following genes were strictly conserved in the SVCV genome [5]. The intergenic region between the P and M genes (GT) and two additional nucleotides in the intergenic region between the G and L genes (CTAT) were identified in the SVCV A-1 genome. In addition, the inverse complementarity of the 5′-trailer was found to be located between nt 11,050 and 11,100.

The deduced amino acid sequences of all five proteins of the known rhabdoviruses were used in comparative analyses to elucidate the phylogentic relationships of SVCV A-1 to the other rhabdoviruses. The highest identity, of more than 89%, was found in the 5 proteins (N, P, M, G and L) of SVCV A-1 to the corresponding proteins of SVCV AJ and NC. Higher homology of more than 50% was also obtained in the N and L proteins of SVCV A-1 to the respective proteins of VSIV, a member of genus *Vesiculovirus*. However, the identity of SVCV A-1 is less than 20% to IHNV, VHSV, HIRRV and SHRV, the members of the genus *Novirhabdovirus*. This suggests that SVCV A-1 is a novel strain of the genus *Vesiculovirus* within the family *Rhabdoviridae*.

In general, we report a complete sequence of SVCV A-1. The genomic and phylogenetic analyses of SVCV A-1 provide basic knowledge of the viral structure and functions. The important information from this work will be of benefit to further studies on the molecular mechanisms of viral pathogenesis and will facilitate the development of effective vaccines or specific diagnostics to control this important pathogen.

## Acknowledgements

The authors sincerely thank Dr. X.Y. Xie (Medical College of Georgia, USA) for critical comments and helpful discussion. This work was supported in part by 973 Program of China (2006CB101802), 863 High-Tech Program of China (2006AA100306), Guangdong Science & Technology Program (2006B50104003), Natural Science Foundation of Guangdong, China (06104920, 04205408), China's General Administration for Quality Supervision, Inspection and Quarantine (AQSIQ) (2005IK034) and Institute of Hydrobiology of Chinese Academy of Sciences (2005FB14).

## References


1. Ahlquist P, Luckow V, Kaesberg P (1981) Complete nucleotide sequence of brome mosaic virus RNA3. J Mol Biol 153: 23–38
2. Ahne W, Bjorklund HV, Essbauer S (2002) Spring Viremia of Carp (SVC). Dis Aquat Org 52: 261–272
3. Bachmann PA, Ahne W (1973) Isolation and characterization of agent causing swim bladder inflammation in carp. Nature 244: 235–237
4. Bishop DHL, Smith MS (1977) Rhabdoviruses. In: Nayak DP (ed) The molecular biology of animal viruses. Marcel Dekker, New York, pp 167–280
5. Bjorklund HV, Higman KH, Kurath G (1996) The glycoprotein genes and gene junctions of the fish rhabdoviruses Spring viremia of carp virus and hirame rhabdovirus: analysis of relationship with other rhabdoviruses. Virus Res 42: 65–80
6. Fijan N, Petrinec Z, Sulimanović D, Zwillenberg LO (1971) Isolation of the viral causative agent from the acute form of infectious dropsy of carp. Vet Arh 41: 125–138
7. Gustafson G, Hunter B, Hanau R, Armour SL, Jackson AO (1987) Nucleotide sequence and genetic organization of barley stripe mosaic virus RNA gamma. Virology 158: 394–406
8. Harty RN, Paragas J, Sudol M (1999) A proline-rich motif within the matrix protein of vesicular stomatitis virus and rabies virus interacts with WW domains of cellular proteins: implications for viral budding. J Virol 73: 2921–2929
9. Hill BJ, Underwood BO, Smale CJ, Brown F (1975) Physicochemical and serological characterization of 5 rhabdoviruses infecting fish. J Gen Virol 27: 369–378
10. Hoffmann B, Schutze H, Mettenleiter TC (2002) Determination of the complete genomic sequence and analysis of the gene products of the virus of Spring Viremia of Carp, a fish rhabdovirus. Virus Res 84: 89–100
11. Johansson T, Nylund S, Olesen NJ (2001) Molecular characterization of the nucleocapsid protein gene, glycoprotein gene and gene junctions of rhabdovirus 903/87, a novel fish pathogenic rhabdovirus. Virus Res 80: 11–22
12. Jorgensen PEV, Olesen NJ, Ahne W, Lorenzen N (1989) SVCV and PFR viruses: serological examination of 22 isolates indicates close relationship between the two fish rhabdoviruses. In: Ahne W, Kurstak E (eds) Viruses of lower vertebrates. Springer Berlin New York Tokyo Heidelberg, pp 349–366
13. Kima DH, Oha HK, Eoua JI, Seob HJ, Kimb SK, Ohc MJ, Namd SW, Choia TJ (2005) Complete nucleotide sequence of the hirame rhabdovirus, a pathogen of marine fish. Virus Res 107: 1–9
14. Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5: 150–163
15. Lehmann J, Mock D, Sturenberg FJ (1986) Das Auftreten von erregerbedingten Frischkrankheiten in Nordrhein-Westfalen in den Jahren 1982 bis einschliesslich 1985. Fischwirt 36: 49–50
16. Liu H, Gao L, Shi X, Gu T, Jiang Y, Chen H (2004) Isolation of spring of carp virus (SVCV) from cultured koi (*Cyprinus carpio koi*) and common carp (*Cyprinus carpio carpio*) in P.R. China. Bull Eur Ass Fish Pathol 24 (4): 194–198
17. Poch O, Blumberg BM, Bougueleret L, Tordo N (1990) Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. J Gen Virol 71: 1153–1162
18. Roy P, Clewley JP (1978) Spring viremia of carp virus RNA and virion-associated transcriptase activity. J Virol 25: 912–916
19. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425
20. Stone DM, Ahne W, Denham KL, Dixon PF, Liu CT, Sheppard AM, Taylor GR, Way K (2003) Nucleotide sequence analysis of the glycoprotein gene of putative spring viraemia of carp virus and pike fry rhabdovirus isolates reveals four genogroups. Dis Aquat Org 53: 203–210
21. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680
22. Tordo N, Poch O, Ermine A, Keith G, Rougeon F (1988) Completion of the rabies virus genome sequence determination: highly conserved domains among the L (polymerase) proteins of unsegmented negative-strand RNA viruses. Virology 165: 565–576
23. Warg JV, Dikkeboom AL, Goodwin AE, Snekvik K, Whitney J (2006) Comparison of multiple genes of spring viremia of carp viruses isolated in the United States. Virus Genes Oct 18, 2006 (Epub a head of print)