Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 308 (2003) 899–905
BBRC
www.elsevier.com/locate/ybbrc

RNA-dependent RNA polymerase gene sequence from foot-and-mouth disease virus in Hong Kong

Xi Chen,[a,1] Qian Feng,[a,1] Zhiyin Wu,[a,b] Yingying Liu,[a] Kun Huang,[a] Rui Shi,[a] Sidi Chen,[a] Wen Lu,[a] Mingxiao Ding,[a] Richard A. Collins,[c] Yin-Wan Wendy Fung,[c] Lok-Ting Lau,[c,d] Albert Cheung-Hoi Yu,[c,d] and Jianguo Chen[a,b,*]

[a] *Department of Cell Biology and Genetics, College of Life Sciences, Peking University, Beijing 100871, PR China*
[b] *Center for Theoretical Biology, Peking University, Beijing 100871, PR China*
[c] *Hong Kong DNA Chips Limited, 1/F, Cosmos Centre, 108 Soy Street, Kowloon, Hong Kong SAR, PR China*
[d] *Neuroscience Research Institute, Department of Neurobiology, Peking University Health Science Center, Peking University, 38 Xue Yuan Road, Beijing 100083, PR China*

Received 18 July 2003

Abstract

A foot-and-mouth disease virus (FMDV, HKN/2002) was isolated in Hong Kong in 2002. The nucleotide sequence of the $3D^{pol}$ gene encoding the viral RNA-dependent RNA polymerase was determined and compared with that of the same gene from other FMDVs. The $3D^{pol}$ gene was 1410 nucleotides in length encoding a protein of 470 amino acid residues. Sequence comparisons indicated that HKN/2002 belonged to serotype O. An evolutionary tree based on the $3D^{pol}$ sequences of 20 FMDV isolates revealed that the nucleotide sequence of the HKN/2002 $3D^{pol}$ gene was most similar to those of isolates found in Taiwan in 1997, suggesting that they share a common ancestor. The amino acid sequence of the HKN/2002 $3D^{pol}$ gene was determined and aligned with those of representative isolates from seven other *Picornaviridae* genera. Eight highly conserved regions were detected, indicating a conserved functional relevance for these motifs. Alignment of 20 FMDV $3D^{pol}$ amino acid sequences revealed a hypermutation region near the N-terminus that may help the virus evade host immune systems.

Keywords: Foot-and-mouth disease virus; Hypermutation; RNA-dependent RNA polymerase; Sequence alignment; Serotype; $3D^{pol}$ gene

Foot-and-mouth disease (FMD) is a highly contagious viral disease affecting cloven-hoofed animals [1]. The disease is widespread and causes significant economic losses due to the death of young animals, decreased productivity and failure to return to prime condition following infection, and trade sanctions against livestock and animal products originating from infected regions. In addition, certain animals continue to shed the virus in their secretions for considerable periods of time after recovery and so act as a reservoir for further cycles of infection. The only viable option under these circumstances is mass slaughter of infected and at risk animals.

Foot-and-mouth disease virus (FMDV) is a positive-sense single-stranded RNA virus, of the family *Picornaviridae*, genus *Aphthovirus*. Seven serotypes of FMDV have been classified (A, Asia-1, C, O, SAT-1, SAT-2, and SAT-3). The viral genome, 7.6–8.4 kb in length, is translated into a single polypeptide before subsequent hydrolytic cleavage into four structural proteins and 10 non-structural proteins. Of particular importance to viral replication is the $3D^{pol}$ gene encoding the RNA-dependent RNA polymerase. The $3D^{pol}$ product is located at the C-terminus of the polypeptide, following the 3C gene product. The $3D^{pol}$ gene is 1410 nucleotides in length and encodes a 470-amino-acid protein with a molecular mass of 55 kDa [2]. In a mechanism catalyzed by two bivalent metal ions, the $3D^{pol}$ enzyme elongates a primer, copying the viral RNA template (plus strand). The newly synthesized minus strand snaps back on itself

* Corresponding author. Fax: +86-10-6276-7044.
E-mail address: chenjg@pku.edu.cn (J. Chen).
[1] These authors contributed equally to this study.

0006-291X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0006-291X(03)01511-0

to generate a template-primer structure which is elongated by the $3D^{pol}$ gene product to form covalently linked dimeric RNA molecules [3,4].

Studies on the $3D^{pol}$ gene product from poliovirus revealed that the protein could be divided into three subdomains, termed fingers, thumb, and palm, respectively [5]. The palm subdomain houses the catalytic center and contains four of the amino acid sequence motifs characteristic of RNA-dependent RNA polymerases [5]. The functional relevance of each motif has been examined by site-directed mutation analysis. These motifs play essential roles in template-binding, substrate-selection, and catalytic processing [6–8]. Due to its significance in viral replication, the $3D^{pol}$ gene is highly conserved, especially within the functional motifs. Meanwhile, the stringency and fidelity of the RNA-dependent RNA polymerase determines the rate of mutation of the virus and its ability to evolve and adapt to its environment, which is in part affected by the variation within the non-conserved regions. As a result, the $3D^{pol}$ protein sequence serves as an ideal target for virus detection and classification. In fact, the RNA-dependent RNA polymerase may be more appropriate for phylogenetic analysis of the positive-strand RNA viruses than the capsid protein [9].

The nucleotide sequence and predicted amino acid sequence of the $3D^{pol}$ gene from a novel strain of FMDV, HKN/2002, isolated in Hong Kong in February 2002 have been determined. Alignment of the nucleotide sequence with 19 similar sequences from other FMDV isolates indicated the serotype and geographic origins of the new strain. Alignment of the predicted amino acid

```
1       GGGTTGATCGTTGATACCAGAGATGTGGAGGAACGCGTCCACGTGATGCGCAAAACCAAGCTCGCGCCCACCGTAGCACACGGTGTGTTC
1        G  L  I  V  D  T  R  D  V  E  E  R  V  H  V  M  R  K  T  K  L  A  P  T  V  A  H  G  V  F
91      AATCCTGAGTTCGGGCCTGCTGCTCTGTCCAACAAGGACCCGCGTCTGAATGAAGGGGTCGTCCTCGACGACGTCATTTTCTCAAAACAC
31       N  P  E  F  G  P  A  A  L  S  N  K  D  P  R  L  N  E  G  V  V  L  D  D  V  I  F  S  K  H
181     AAGGGAGACACGAGGATGTCTGAGGAAGACAAAGCGCTGTTCCGGCGCTGTGCTGCCGACTACGCGTCGCGTCTACACAGCGTGCTAGGG
61       K  G  D  T  R  M  S  E  E  D  K  A  L  F  R  R  C  A  A  D  Y  A  S  R  L  H  S  V  L  G
271     ACGGCAAACGCCCCACTGAGTGTATACGAAGCCATCAAAGGCGTCGATGGACTTGACGCCATGGAGCCAGACACCGCACCCGGTCTCCCC
91       T  A  N  A  P  L  S  V  Y  E  A  I  K  G  V  D  G  L  D  A  M  E  P  D  T  A  P  G  L  P
361     GGGGCTCTCCAAGGAAAACGCCGAGGTGCCCTGATCGACTTCGAAAACGGTACTGTCGGGCCCGAGGTTGAAGCAGCACTCAAGCTCATG
121      G  A  L  Q  G  K  R  R  G  A  L  I  D  F  E  N  G  T  V  G  P  E  V  E  A  A  L  K  L  M
451     GAAAGCCGTGAGTATAAATTCGTCTGCCAAACCTTTCTGAAAGACGAAATTCGGCCGCTAGAGAAGGTGCGCGCCGGTAAGACACACATT
151      E  S  R  E  Y  K  F  V  C  Q  T  F  L  K  D  E  I  R  P  L  E  K  V  R  A  G  K  T  H  I
541     GTCGACGTTTTGCCTGTTGAACACATTCTCTATACCAGAATGATGATTGGTAGATTCTGTGCTCAGATGCACTCAAACAACGGACCGCAA
181      V  D  V  L  P  V  E  H  I  L  Y  T  R  M  M  I  G  R  F  C  A  Q  M  H  S  N  N  G  P  Q
631     ATTGGCTCAGCGGTCGGTTGCAACCCTGATGTTGATTGGCAAAGATTTGGCACACATTTCGCCCAGTACAAGAACGTGTGGGATGTGGAC
211      I  G  S  A  V  G  C  N  P  D  V  D  W  Q  R  F  G  T  H  F  A  Q  Y  K  N  V  W  D  V  D
721     TACTCAGCCTTCGATGCAAACCACTGCAGCGATGCGATGAACATCATCTTCGAAGAGGTGTTCCGCACGGAGTTTGGGTTCCACCCGAAC
241      Y  S  A  F  D  A  N  H  C  S  D  A  M  N  I  I  F  E  E  V  F  R  T  E  F  G  F  H  P  N
811     GCCGAGTGGATTCTGAAGACTCTGGTGAACACGGAGCACGCTTACGAGAACAAGCGCATCACTGTGGAGGGTGGAATGCCGTCCGGTTGT
271      A  E  W  I  L  K  T  L  V  N  T  E  H  A  Y  E  N  K  R  I  T  V  E  G  G  M  P  S  G  C
901     TCCGCAACAAGCATCATCAACACAATTTTGAACAACATCTACGTGCTCTACGCTCTGCGTAGGCACTATGAAGGAGTTGAGCTGGACACC
301      S  A  T  S  I  I  N  T  I  L  N  N  I  Y  V  L  Y  A  L  R  R  H  Y  E  G  V  E  L  D  T
991     TACACAATGATCTCCTATGGAGACGACATCGTGGTGGCTAGTGACTACGACCTGGACTTCGAGGCTCTCAAGCCCCACTTCAAGTCCCTC
331      Y  T  M  I  S  Y  G  D  D  I  V  V  A  S  D  Y  D  L  D  F  E  A  L  K  P  H  F  K  S  L
1081    GGTCAGACCATCACTCCAGCCGACAAAAGCGACAAAGGTTTTGTTCTTGGTCACTCCATAACCGATGTCACTTTCCTCAAAAGACACTTC
361      G  Q  T  I  T  P  A  D  K  S  D  K  G  F  V  L  G  H  S  I  T  D  V  T  F  L  K  R  H  F
1171    CACATGGACTACGGAACTGGGTTTTACAAACCTGTGATGGCCTCGAAGACCCTCGAGGCTATCCTCTCCTTTGCACGCCGTGGGACCATA
391      H  M  D  Y  G  T  G  F  Y  K  P  V  M  A  S  K  T  L  E  A  I  L  S  F  A  R  R  G  T  I
1261    CAGGAGAAGTTGATCTCCGTGGCAGGACTCGCCGTCCACTCCGGACCTGACGAGTACCGGCGTCTCTTTGAACCTTTCCAAGGTCTCTTC
421      Q  E  K  L  I  S  V  A  G  L  A  V  H  S  G  P  D  E  Y  R  R  L  F  E  P  F  Q  G  L  F
1351    GAGATTCCAAGCTACAGATCACTTTACCTGCGATGGGTGAACGCCGTGTGCGGTGACGCATAA
451      E  I  P  S  Y  R  S  L  Y  L  R  W  V  N  A  V  C  G  D  A
```

Fig. 1. Nucleotide and predicted amino acid sequences of the RNA-dependent RNA polymerase $3D^{pol}$ gene of HKN/2002. Amino acids are indicated below the nucleotide sequence by standard one-letter codes.

sequence with sequences derived from typical isolates representing seven other *Picornaviridae* genera identified conserved residues and motifs within this protein. A phylogenetic analysis indicated how the $3D^{pol}$ gene of HKN/2002 was related to other FMDV serotypes.

Materials and methods

Virus culture and RNA extraction. Foot-and-mouth disease virus, isolate HKN/2002, obtained from a pig displaying clinical symptoms of FMD was generously supplied by the Agriculture, Fisheries, and Conservation Department of the Hong Kong SAR, China. The virus was used to infect baby hamster kidney (BHK-21) cells cultured in Dulbecco's modified Eagle's medium (DMEM) containing 10% (v/v) foetal calf serum (FCS) as described previously [10]. Total RNA was extracted using RNAgents Total RNA Isolation System (Promega, Madison, WI) following the manufacturer's instructions.

Reverse transcription-PCR, cloning, and sequencing. Reverse transcription was carried out using Superscript II RT RNase H^- (Invitrogen, Carlsbad, CA) and the antisense primer R1 (5′-GGG GCG GCC GCG GAT TAA GGA AGC GGG AAA AGC CC-3′). The primer included a *Not*I artificial restriction site (underlined). The first-strand cDNA was then subjected to PCR amplification using the sense primer F3D (5′-GGG TTG ATC GTT GAT ACC AGA GA-3′) and R1, to amplify a 1.5-kb fragment containing the 1410 nucleotide 3D gene and the 92 nucleotide 3′-UTR of the virus. The PCR product was recovered using Geneclean II Kit (Qbiogene, Carlsbad, CA) and cloned directly into pGEM-T Easy vector (Promega). Positive clones were sequenced (Shanghai Sangon Biologic Engineering Technology and Service, Shanghai, China).

Nucleotide sequence alignment and comparison. Reference FMDV sequences were obtained from the National Center for Biotechnology Information (NCBI) website (www.ncbi.nlm.nih.gov). Multiple alignments were analyzed using ClustalX multiple sequence alignment program 1.83 [11] A phylogenetic tree was constructed using Mega software, version 2.1 [12].

Results and discussion

Nucleotide sequence of HKN/2002 $3D^{pol}$

The $3D^{pol}$ gene of HKN/2002 was sequenced after molecular cloning. The consensus nucleotide sequence and the predicted amino acid sequence are shown in Fig. 1. The sequence was deposited in GenBank (Accession No.: AY152808). The 92 nucleotide 3′-UTR was also determined. The $3D^{pol}$ gene nucleotide sequence is 1410 nucleotides in length. The gene utilizes TAA as the termination codon and encodes a protein of 470 amino acid residues.

Phylogenetic analysis of $3D^{pol}$ gene sequences

Very few complete FMDV $3D^{pol}$ gene sequences are available for analysis in GenBank. Nineteen sequences (Table 1) were selected and compared with the HKN/2002 sequence. The sequences share significant sequence similarity at the nucleotide and protein levels, ranging from 85.0% to 100.0% (nucleotide) and 92.6% to 100.0% (amino acid), respectively (data not shown). As the polymerase is essential for viral replication the degree of sequence similarity between the $3D^{pol}$ genes is much higher than that of the VP1 surface antigen [10]. From these data, a genetic distance matrix was constructed (Table 2). Of the sequences analyzed, HKN/2002 has a maximum genetic distance from SAT-2/KEN (0.145 substitutions per nucleotide) and a minimum genetic distance from O/YunLin (0.063 substitutions per nucleotide).

Table 1
FMDV RNA-dependent RNA polymerase nucleotide sequences compared in this study

Serotype	Virus	GenBank Accession No.	Origin	Outbreak year
O	O/Chu-pei	AF026168	Taiwan	1997
	O/Tibet	AF506822	Tibet/China	1999
	HKN/2002	AY152808	Hong Kong	2002
	O/Japan/2000	AB079061	Japan	2000
	O/YunLin	AF308157	Taiwan	2000
	O/Tao-Yuan	AF154271	Taiwan	1997
	O/SKR/2000	AF377945	Korea	2000
	O/Akesu/58	AF511039	China	1958
	O1/Geshure	AF189157	Israel	1999
	O1/Campos	AJ320488	Brazil	2002
	O1/K	X00871	Germany	1984
C	C-MARLS	AF274010	Spain	1998
	C-rp99	AJ133358	Spain	1999
	C-s8c1	AJ133357	Spain	1999
	C-1/Santa Pau	M11027	N/A	1985
A	A10-61	X00429	Argentina	1984
	A12	M10975	United Kingdom	1985
	A22/550	X74812	Azerbaijan	1993
Asia 1	Asia-1/IND	AF207520	India	1999
SAT-2	SAT-2/KEN	NC_003992	Kenya	1999

N/A, not available.

Table 2
Distance matrix for 20 FMDV RNA-dependent RNA polymerase $3D^{pol}$ gene sequences generated by ClustalX

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 C-rp99	0	0	0.004	0.021	0.064	0.066	0.072	0.072	0.077	0.079	0.089	0.091	0.082	0.072	0.084	0.116	0.114	0.116	0.116	0.126
2 C-s8c1	0	0	0.004	0.021	0.064	0.066	0.072	0.072	0.077	0.079	0.089	0.091	0.082	0.072	0.084	0.116	0.114	0.116	0.116	0.126
3 C-MARLS	0.004	0.004	0	0.023	0.066	0.067	0.073	0.074	0.079	0.081	0.09	0.093	0.082	0.072	0.086	0.115	0.113	0.115	0.118	0.126
4 C-1/Santa Pau	0.021	0.021	0.023	0	0.074	0.081	0.087	0.083	0.092	0.094	0.101	0.103	0.094	0.087	0.102	0.123	0.121	0.123	0.128	0.138
5 A10-61	0.064	0.064	0.066	0.074	0	0.065	0.068	0.062	0.08	0.081	0.09	0.09	0.084	0.07	0.085	0.118	0.117	0.119	0.119	0.135
6 O1/Campos	0.066	0.066	0.067	0.081	0.065	0	0.011	0.058	0.074	0.072	0.082	0.082	0.079	0.07	0.071	0.112	0.11	0.113	0.113	0.125
7 O1/K	0.072	0.072	0.073	0.087	0.068	0.011	0	0.061	0.078	0.077	0.087	0.089	0.082	0.072	0.074	0.114	0.112	0.115	0.113	0.128
8 A12	0.072	0.072	0.074	0.083	0.062	0.058	0.061	0	0.084	0.084	0.092	0.087	0.079	0.071	0.094	0.118	0.116	0.118	0.122	0.129
9 O/Tibet	0.077	0.077	0.079	0.092	0.08	0.074	0.078	0.084	0	0.011	0.026	0.085	0.078	0.08	0.079	0.114	0.112	0.115	0.115	0.134
10 O/Japan/	0.079	0.079	0.081	0.094	0.081	0.072	0.077	0.084	0.011	0	0.029	0.085	0.078	0.081	0.076	0.115	0.113	0.116	0.118	0.134
11 O/SKR/2000	0.089	0.089	0.09	0.101	0.09	0.082	0.087	0.092	0.026	0.029	0	0.091	0.086	0.087	0.09	0.123	0.121	0.123	0.124	0.138
12 A22/550	0.091	0.091	0.093	0.103	0.09	0.082	0.089	0.087	0.085	0.085	0.091	0	0.097	0.085	0.095	0.131	0.129	0.132	0.132	0.142
13 O/Akesu/58	0.082	0.082	0.082	0.094	0.084	0.079	0.082	0.079	0.078	0.078	0.086	0.097	0	0.086	0.101	0.12	0.119	0.121	0.121	0.131
14 Aisa-1/IND	0.072	0.072	0.072	0.087	0.07	0.07	0.072	0.071	0.08	0.081	0.087	0.085	0.086	0	0.067	0.109	0.106	0.109	0.118	0.128
15 O1/Geshure	0.084	0.084	0.086	0.102	0.085	0.071	0.074	0.094	0.079	0.076	0.09	0.095	0.101	0.067	0	0.121	0.118	0.121	0.127	0.132
16 O/Tao-Yuan	0.116	0.116	0.115	0.123	0.118	0.112	0.114	0.118	0.114	0.115	0.123	0.131	0.12	0.109	0.121	0	0.002	0.005	0.063	0.15
17 O/Yunlin	0.114	0.114	0.113	0.121	0.117	0.11	0.112	0.116	0.112	0.113	0.121	0.129	0.119	0.106	0.118	0.002	0	0.004	0.062	0.148
18 O/Chu-pei	0.116	0.116	0.115	0.123	0.119	0.113	0.115	0.118	0.115	0.116	0.123	0.132	0.121	0.109	0.121	0.005	0.004	0	0.063	0.15
19 HKN/2002	0.116	0.116	0.118	0.128	0.119	0.113	0.113	0.122	0.115	0.118	0.124	0.132	0.121	0.118	0.127	0.063	0.062	0.063	0	0.145
20 SAT-2/KEN	0.126	0.126	0.126	0.138	0.135	0.125	0.128	0.129	0.134	0.134	0.138	0.142	0.131	0.128	0.132	0.15	0.148	0.15	0.145	0

An evolutionary tree was constructed to examine the phylogenetic relationship between the FMDV serotypes (Fig. 2). The 20 isolates could be segregated into six clusters (designated A, B, C, D, E, and F) based on the predicted degree of relatedness. Most of the isolates clustered according to their serotype and geographic origins. HKN/2002 grouped closely with FMDV-O serotype isolates in cluster A. From an analysis of the $3D^{pol}$ gene sequences, it is predicted that HKN/2002 belongs to serotype O. This confirms our previous report based on the VP1 sequence of the same isolate [10]. However, virus-neutralization tests are needed to identify conclusively the serotype of HKN/2002. Furthermore, the isolates clustering most strongly with HKN/2002, i.e., O/Chu-pei, O/Tau-Yuan, and O/YunLin, were isolated in Taiwan province of China in 1997, suggesting a possible infection path between Taiwan and Hong Kong, or a putative common ancestor for these isolates. Isolate SAT 2/KEN was isolated in Africa, a considerable geographic distance from the location of the other isolates, and it is not surprising that it clustered separately with the largest genetic distance from the other isolates (0.125–0.150 substitutions per nucleotide). Isolates O/Tibet, O/Akesu/58, O/SKR2000, and O/Japan/

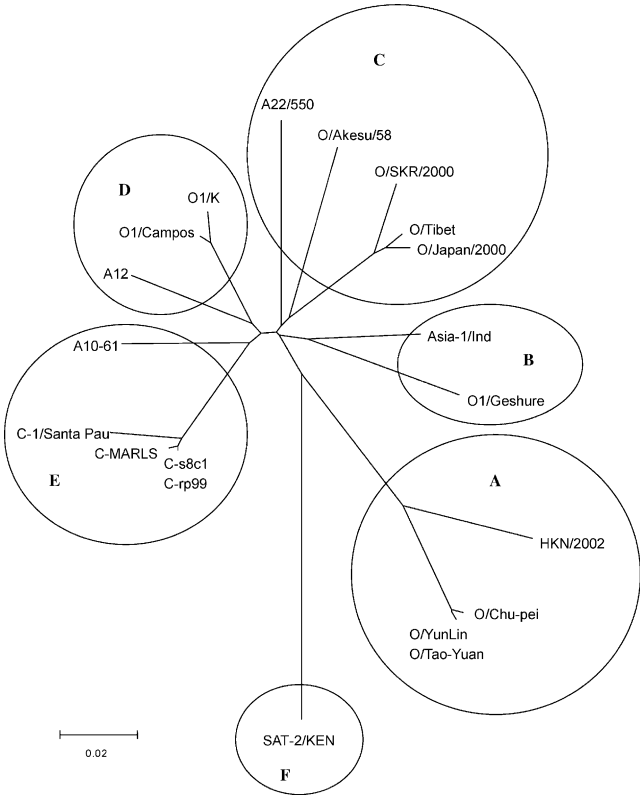


Fig. 2. Neighbor-joining tree based on sequence analysis of 20 RNA-dependent RNA polymerase $3D^{pol}$ genes from various FMDV serotypes. The analysis was performed using MEGA software version 2.1. Scale bar represents substitutions per nucleotide position. Six genetic clusters (designated A, B, C, D, E, and F) are shown. Isolate abbreviations are listed in Table 1.

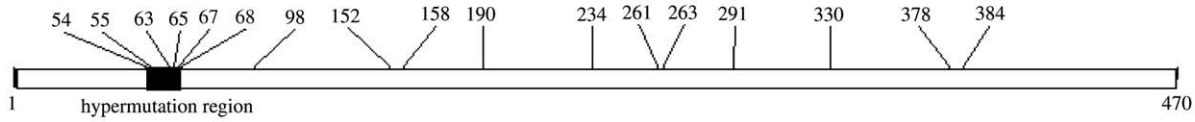
2000 also belong to serotype O, but aggregate separately from HKN/2002, O/Chu-Pei, O/Tau-Yuan, and O/YunLin, suggesting a lower degree of relatedness.

Alignment of the predicted $3D^{pol}$ protein sequences of 20 FMDV isolates was performed using an appropriate program. To eliminate possible errors generated during DNA sequencing, only variants present in at least two independent sequences are shown (Fig. 3.) Seventeen variant loci were identified. In addition, a hypermutant region between amino acid 54 and 68 was identified (Fig. 3). Due to the highly conserved enzymic activity of the $3D^{pol}$ protein, mutations at the indicated positions are predicted to cause little or no significant structural or functional change to the enzyme. The hypermutant region may serve to evade the immune system while maintaining the core structure of the functional protein.

FMDV $3D^{pol}$ and the Picornaviridae

The predicted amino acid sequences of the $3D^{pol}$ gene from representative isolates of eight genera of *Picornaviridae* were compared. The isolates selected are shown in Table 3. The largest difference in the length of the $3D^{pol}$ gene product sequence is 37 amino acid residues between hepatitis A virus (HAV) and porcine teschovirus 1. The genetic distance between each pair of isolates was calculated (data not shown). The maximum distance of 0.731 substitutions per nucleotide was between bovine kobuvirus and HAV. The minimum distance between pairs of isolates was 0.395 substitutions per nucleotide between poliovirus and simian picornavirus 1. The FMDV isolate HKN/2002 showed a maximum distance of 0.700 substitutions per nucleotide from HAV and a minimum distance of 0.462

Position / Isolates	54	55	63	65	67	68	98	152	158	190	234	261	263	291	330	378	384
C-1-Santa Pau	E	V	D	K	S	A	I	K	A	M	R	R	E	T	T	H	I
A12	E	V	D	K	S	A	I	K	V	M	R	R	D	T	T	H	T
A22-550	E	V	D	K	T	E	I	K	T	M	K	R	E	T	T	Q	I
Asia-1/Ind	E	V	D	K	S	A	I	K	A	M	R	R	E	T	T	H	I
C/Marls	E	V	D	K	S	A	I	K	A	M	R	R	E	T	T	H	I
C/rp99	E	V	D	K	S	A	I	K	A	M	R	R	E	T	T	H	I
A10-61	D	V	D	K	T	E	I	K	A	M	R	R	D	T	T	Q	I
SAT-2/KEN	E	V	D	K	S	E	I	N	T	I	K	R	E	T	S	Q	I
O/Akesu/58	E	V	D	K	S	E	I	K	A	M	R	N	E	T	T	H	I
O/Chu-pei	D	V	D	R	S	E	V	S	V	I	K	R	E	V	T	H	I
HKN/2002	D	V	D	R	S	E	V	S	V	I	K	R	E	T	T	H	I
O/SKR2000	E	A	N	K	S	E	I	K	V	M	R	N	D	S	S	H	I
O/Tau-Yuan	D	V	D	R	S	E	V	S	V	I	K	R	E	V	T	H	I
O/Tibet	E	A	N	K	S	E	T	K	V	M	R	N	D	T	S	H	I
O/YunLin	D	V	D	R	S	E	V	S	V	I	K	R	E	V	T	H	I
O/Japan/2000	E	V	N	K	S	E	A	K	A	M	R	N	D	T	S	H	T
O1/Geshure	E	V	D	K	S	G	I	K	A	M	R	R	E	T	T	H	I
O1/Campos	E	V	D	K	S	E	I	K	A	M	R	R	E	T	T	H	I
O1/K	E	V	D	K	S	E	I	K	V	I	R	R	E	T	T	H	I
C-s8c1	E	V	D	K	S	A	I	K	A	M	R	R	E	T	T	H	I
Variation frequency							4	3	3	2	2	2	2	3	2	2	2



3D RdRp Variations

Fig. 3. Amino acid variations among 20 FMDV RNA-dependent RNA polymerases (RdRp). The variations are labelled in the schematic diagram of the 3D gene below the table. The black box represents the hypermutant region. The hypermutant residues are shaded in the table.

Table 3
Picornavirus RNA-dependent RNA polymerase amino acid sequences compared in this study

Abbreviation	Species	Genus	Length (aa)	GenBank Accession No.
BKV	Bovine kobuvirus	Kobuvirus	469	AB084788
EMCV	Encephalomyocarditis virus	Cardiovirus	460	NC_001479
ERV	Equine rhinovirus 3	Rhinovirus	468	NC_003077
HAV	Hepatitis A virus	Hepatovirus	489	M14707
PV	Poliovirus	Enterovirus	461	V01149
SPV	Simian picornavirus 1	New, unnamed	466	AY064708
HKN/2002	Foot-and-mouth disease virus	Aphthovirus	470	AY152808
PTV	Porcine teschovirus 1	Teschovirus	452	NC_003985

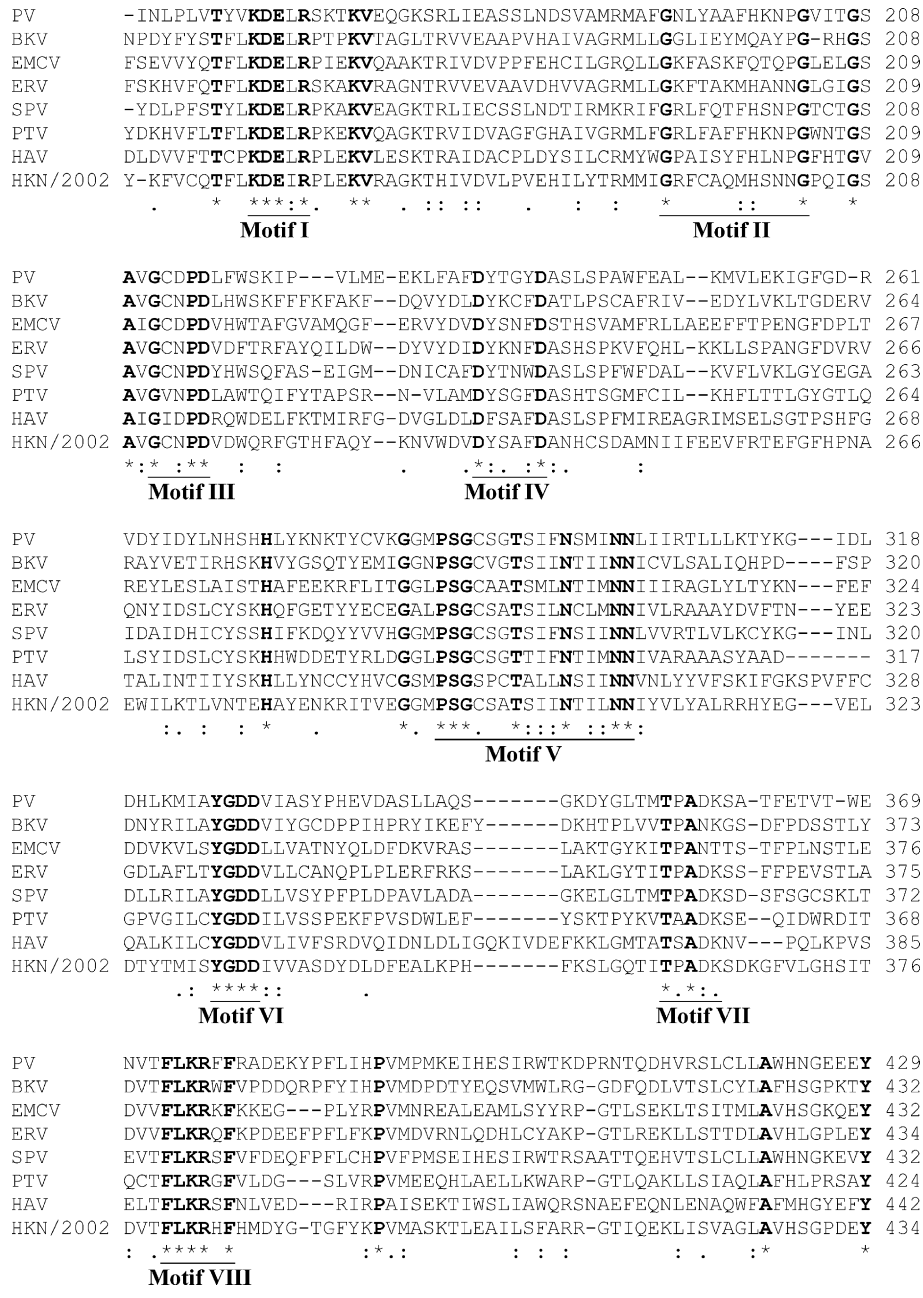


Fig. 4. Amino acid sequence alignment of RNA-dependent RNA polymerases from eight *Picornaviridae* genera. The alignment was produced using ClustalX 1.83. An underline indicates motifs containing regions of highly conserved amino acid residues. (*) Residues conserved across all compared sequences. Such identical residues are also highlighted in bold. (:) Residues across all the compared sequences fall into in one of the following groups: STA; NEQK; NHQK; NDEQ; QHRK; MILV; MILF; HY; and FYW. (.) Residues across all the compared sequences fall into in one of the following groups: CSA; ATV; SAG; STNK; STPA; SGND; SNDEQK; NDEQHK; NEQHRK; FVLIM; HFY. Isolate abbreviations are described in Table 3.

substitutions per nucleotide from equine rhinovirus 3. A distance matrix based on the amino acid sequences was also calculated, revealing a generally lower degree of sequence similarity (18.9–45.1%) than that observed at the nucleotide level (30.2–53.8%), indicating that the majority of the nucleotide changes between the species were non-synonymous mutations (data not shown). The regions containing the most sequence variation are the N-terminus (about 150 amino acid residues in length) and C-terminus (about 70 amino acid residues in length), respectively. The conserved residues (about 250 amino acid residues in length) cluster at the center, forming the core structure of the enzyme.

Four conserved motifs in the amino acid sequences of RNA-dependent RNA polymerases have been described previously [5,8,13]. Structural analysis indicated three

subdomains (termed fingers, thumb, and palm) within the $3D^{pol}$ protein. The palm subdomain represents the core structure of the polymerase [9]. Eight conserved motifs present in the palm subdomain were identified after amino acid sequence alignment (Fig. 4), corresponding to those identified previously [9]. The eight motifs, I ^{164}KDE[I/L]R^{168}, II ^{197}G(X)$_9$[P/N]G^{208}, III ^{216}GXXPD^{220}, IV ^{240}D[Y/F]XXXD^{245}, V ^{297}PSG[C/S](X)$_2$T(X)$_3$NX[M/I]XNN^{312}, VI ^{336}YGDD^{339}, VII ^{365}T[P/S]A[D/N][K/T]^{369}, and VIII ^{385}FLKRXF^{390}, are highly conserved across the seven $3D^{pol}$ sequences of the picornavirus isolates, suggesting a shared functional significance. Essential catalytic residues include Asp^{240} in motif IV and Asp^{338} in motif VI, which are proposed to coordinate magnesium ions during nucleotidyltransfer catalysis [4]. Asp^{245} in motif IV and Ser^{298}, Thr^{303}, and Asn^{307} in motif V are proposed to interact to discriminate between NTP and dNTP as the catalytic substrate, defining the enzyme as an RNA polymerase instead of a DNA polymerase [8]. The critical role of these residues in mediating enzymatic activity correlates well with their degree of conservation among the RNA polymerases and these strictly conserved residues have been predicted to occupy key positions in the secondary and tertiary structures of picornavirus $3D^{pol}$ gene products [5,6].

In summary, the $3D^{pol}$ gene of a novel isolate of FMDV was sequenced. The isolate probably belongs to serotype O, supporting previous studies on the VP1 surface protein of the same isolate. The new sequence is an important addition to the sparse data on *Aphthovirus* $3D^{pol}$ sequences available publicly. Important structural and functional motifs at the nucleotide and amino acid levels are conserved both within the *Aphthovirus* genus and across *Picornaviridae* genera.

Acknowledgments

The authors thank Trevor Ellis and Pamela Li Chui Har (Agriculture, Fisheries, and Conservation Department, Veterinary Laboratory Division, Tai Lung Experimental Station, Lin Tong Mei, Sheung Shui, Hong Kong SAR, China) for providing virus samples. This work was supported by Grant G1999011904 from the Special Funds for Major State Basic Research of China to J.C. The Jun-Zheng Foundation at PKU supported X.C.

References

[1] Infectious Diseases in Livestock: Policy Document 15/02, The Royal Society, London, 2002.
[2] M. George, R. Venkataramanan, B. Pattnaik, A. Sanyal, C.B. Gurumurthy, D. Hemadri, C. Tosh, Sequence analysis of the RNA polymerase gene of foot-and-mouth disease virus serotype Asia1, Virus Genes 22 (2001) 21–26.
[3] J.M. Lubinski, L.J. Ransone, A. Dasgupta, Primer-dependent synthesis of covalently linked dimeric RNA molecules by poliovirus replicase, J. Virol. 61 (1987) 2997–3003.
[4] T.A. Steitz, A mechanism for all polymerases, Nature 391 (1998) 231–232.
[5] J.L. Hansen, A.M. Long, S.C. Schultz, Structure of the RNA dependant RNA polymerase of poliovirus, Structure 5 (1997) 1109–1122.
[6] D.W. Gohara, S. Crotty, J.J. Arnold, J.D. Yoder, R. Andino, C.E. Cameron, Poliovirus RNA-dependent RNA polymerase ($3D^{pol}$): Structural, biochemical, and biological analysis of conserved structural motifs A and B, J. Biol. Chem. 275 (2000) 25523–25532.
[7] S.A. Jablonski, C.D. Morrow, Mutation of the aspartic acid residues of the GDD sequence motif of poliovirus RNA-dependent RNA polymerase results in enzymes with altered metal ion requirements for activity, J. Virol. 69 (1995) 1532–1539.
[8] E.K. O'Reilly, C.C. Kao, Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure, Virology 252 (1998) 287–303.
[9] E.V. Koonin, The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses, J. Gen. Virol. 72 (1991) 2197–2206.
[10] Q. Feng, X. Chen, O. Ma, Y. Liu, M. Ding, R.A. Collins, L.S. Ko, J. Xing, L.T. Lau, A.C.H. Yu, J. Chen, Serotype and VP1 gene sequence of a foot-and-mouth disease virus from Hong Kong (2002), Biochem. Biophys. Res. Commun. 302 (2003) 715–721.
[11] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 25 (1997) 4876–4882.
[12] S. Kumar, K. Tamura, I.B. Jakobsen, M. Nei, MEGA2: Molecular evolutionary genetics analysis software, Bioinformatics 17 (2001) 1244–1245.
[13] O. Poch, I. Sauvaget, M. Delarue, N. Tordo, Identification of four conserved motifs among the RNA-dependent polymerase encoding elements, EMBO J. 8 (1989) 3867–3874.