International Scholars Journals

Advanced Journal of Microbiology Research ISSN 2736-1756 Vol. 14 (12), pp. 001-006, December, 2020. Available online at www.internationalscholarsjournals.org © International Scholars Journals

Full Length Research Paper

A study of the genome sequence of food-and-mouth disease virus WFL strain

Linzhu Ren[1], Lin Wang[3], Yunzhi Ma[1*] and Xinglong Wang[2]

[1]Heping Campus, Jilin University, 5333 Xi'an Road, Changchun, 130062, China.
[2]Department of Animal Food, The 11th Institute, Academy of Military Medical Sciences, 1068 Qinglong Road, Changchun, China.
[3]Department of Pharmacology, School of Pharmacy, Second military medical university, No.800, Xiangyin Road, Shanghai, China.

Accepted 10 September, 2020

The complete genome of the foot-and-mouth disease virus (FMDV) strain WFL was cloned and sequenced. The results showed that the complete genome was 8155 nucleotides (nt) in length (including the poly(C) tract, but excluding poly(A) tail) and was composed of a 1059-nt 5'-untranslated region (UTR), a 6969-nt open reading frame, and a 127-nt 3'- UTR. *cre* region of 5'UTR was 55nt with 45.5% of G/C, and had a stem-loop. The stem -loops region of 3'UTR can fold into two stem -loops, SL1 and SL2. A phylogenetic tree was constructed based on complete amino acids sequences of WFL strain and reference strains. The strains were divided into 4 clusters. O/ES/2001, HKN/2002, LZ and WFL strain can be divided into one group. It was obvious that WFL strain had a close relationship to LZ strain, which indicated that the WFL strain was of serotype O. There were 16 different deduced amino acid residues between the WFL strain and the LZ strain.

Key words: Food-and-mouth disease virus, sequence, complete genome, untranslated region (UTR).

INTRODUCTION

Foot and mouth disease (FMD) is a highly contagious and economically devastating disease of cloven-hoofed livestock, characterized by the appearance of vesicles on the feet and mouth (Marvin and Barry, 2004; Salguero et al., 2005; Sobrino et al., 2001).

The foot-and-mouth disease virus (FMDV) is a member of the family *Picornaviridae*, genus *Aphthovirus*. Seven serotypes (A, O, C, Asia 1, and South African Territories 1, 2, and 3) have been identified serologically based on their geographic origin (topotypes), e.g., the serotype O can be grouped into eight topotypes [Cathay, Middle East-South Asia (ME-SA), South-East Asia (SEA), Europe-South America (Euro-SA), Indonesia-1 (ISA-1), Indonesia-2 (ISA-2), East Africa (EA), and West Africa (WA)] based on nucleotide differences of up to 15% (Feng, 2004), and multiple subtypes occur within each serotype (Marvin and Barry, 2004). Viral infection or immunization with one serotype does not confer protection against the other serotypes (Grubman and Baxt, 2004). FMDV consists of a single- stranded, plus-sense RNA genome of approximately 8,500 bases surrounded by four structural proteins that form an icosahedral capsid. The genome contains 5' UTR (untranslated region), 3'UTR and a long open reading frame (ORF). The ORF can be translated into a single polyprotein, that can be cleaved into four structural proteins (VP4, VP2, VP3 and VP1), and 10 non-structural proteins (L, 2A, 2B, 2C, 3A, 3B1, 3B2, 3B3, 3C, and 3D) (Feng, 2004).

The main goal of the present study was to obtain the entire genome sequence of food-and-mouth disease virus WFL strain, including the 3 - and 5 -terminal non-coding regions of the genome.

MATERIALS AND METHODS

Viral isolates, RT-PCR, and sequencing

Foot -and-mouth disease virus WFL strain was isolated from swine host in 1999 in Yunnan province, China, and adapted to BHK-21 cells. Total RNA was extracted using the RNeasy Mini kits (QIAGEN) according to the manufacturer's instructions.

*Corresponding author. E-mail: rodgeren2006@yahoo.com.cn.

Table 1. Primers used for the amplification of the WFL strain.

Designation	Sequence of primers	Length	Location
A	5’ TTGAAAGGGGGCGTTAGGGTCTC3’	46	1-19
B	5’ TTGGCGCGCCACTAGTTTACCTCAGGGTACCT3’	27	921-941
C	5’TGCCCTTTAGGTACCCTGA3’	19	931-950
F	5’ TGCGCGGCCGCATGCATGACAGGCGGCTC3’	26	4119-4139
G	5’ TGCGAATTCTGTCATGCATGGCCGCTGT 3’	25	4124-4144
H	5’ TTGCGGCCGCACTAGTCATATGTT3’	22	8140-8155
I_{RT}	5’TTGCGGCCGCACTAGTCATATGTTTTTTTTTTTTTT3	34	

Table 2. Information on the foot-and-mouth disease virus referenced in this study.

GenBank ID	Serotype	Isolate	Genome size(bp)	5’ NCR	3’ NCR	Amino acids
EF149009	Asia 1	Asia1/Jiangsu/China/2005	8189	1-1092	8083-8183	2329
EF149010	Asia 1	Asia 1/HNK/CHA/05	8187	1-1090	8080-8187	2329
FJ906802	Asia 1	Asia1/WHN/CHA/06	8239	1-1090	8080-8239	2329
DQ533483	Asia 1	ZB/CHA/58(att)	8193	1-1093	8084-8193	2329
AY390432	Asia1	YNBS/58	8163	1..1060	8051..8163	2329
AY686687	O	O/ES/2001	8163	1-1084	8054-8163	2322
AF511039	O	Akesu/58	8147	1..1039	8039..8147	2332
HQ009509	O	China/5/99(Fujian)	8231	1..1101	8101..8231	2332
DQ478937	O	OGBF15 derivative	8166	1-1058	8058-8166	2332
DQ248888	O	LZ	8104	1..1041	8011..8104	2322
AJ539138	O	Tibet/CHA/99	8183	1..1091	8088..8183	2332
AY317098	O	HKN/2002	8104	1..1042	8012..8104	2322
AF506822	O	China/1/99(Tibet)	8173	1..1081	8081..8173	2332
AY359854	O	OMIII	8083	1..1008	7963..8083	2317
AY333431	O	O/NY00	7731	1..712	7712..7731	2332
EF175732	O	WFL	8155	1-1059	8029-8155	2322

Subsequently, RNA was reverse-transcribed into cDNA using the primer I_{RT} and *SuperScript*® III Reverse Transcriptase (*Invitrogen*). The first-strand cDNA was then subjected to PCR amplification using primer pairs, A-B, C- F, and G- H (Table 1), to amplify 3 separate overlapping PCR products containing the complete genome of FMDV using *LA Taq* polymerase (*Takara Biotechnology* (*Dalian*) CO., LTD). The PCR products were purified and sequenced (Sangon Biological Engineering Technology and Service, Shanghai, China). The primers were designed based on the complete reference sequence obtained from GenBank.

Sequence analysis

The RNA structure was depicted according to the RNA-fold prediction program (Gruber et al., 2008). The reference sequences included in the analysis were obtained from GenBank (Table 2). Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 4.1 (Tamura et al., 2007).

RESULTS AND DISCUSSION

Full-length genomic sequence of WFL

Here, we obtained the full-length genome of the WFL strain by RT- PCR. Using a total of 7 primers (Table 1), the complete genome sequence of the WFL strain was amplified as 3 separate overlapping PCR products. The result showed that the complete genome sequence of the WFL strain was 8155 nucleotides (nt) in length, including poly(C). The full-length sequence was submitted to GenBank (GenBank ID: EF175732).

The genomic organization of WFL was shown in Table 3. The complete sequence was divided into sixteen fragments except poly (A). 5’ UTR (non-translated region) and 3’UTR were located in 1-1059 and 8029-8155, respectively (Table 3).

Characteristics of UTR

5’UTR played an important role in replication and selective translation of the viral RNA. The FMDV 5’UTR contains a short fragment called S-fragment, a poly (C) tract of variable length, followed by a large fragment (LF) of over 700 bases in length (LF-5’ UTR) that can form a number of highly conserved secondary structures that include randomly repeated pseudoknots (PKs), a cis-

Table 3. Architecture of the complete genome sequence of the WFL strain.

Genome segments	5'UTR			L	P1			P2				P3				3'UTR
	S	poly C	LF		VP4	VP2	VP3	VP1	2A	2B	2C	3A	3B	3C	3D	
Nucleotide	371	17	671	603	255	660	654	639	48	462	954	429	213	639	1413	127
Amino acid				201	85	220	218	213	16	154	318	143	71	213	471	

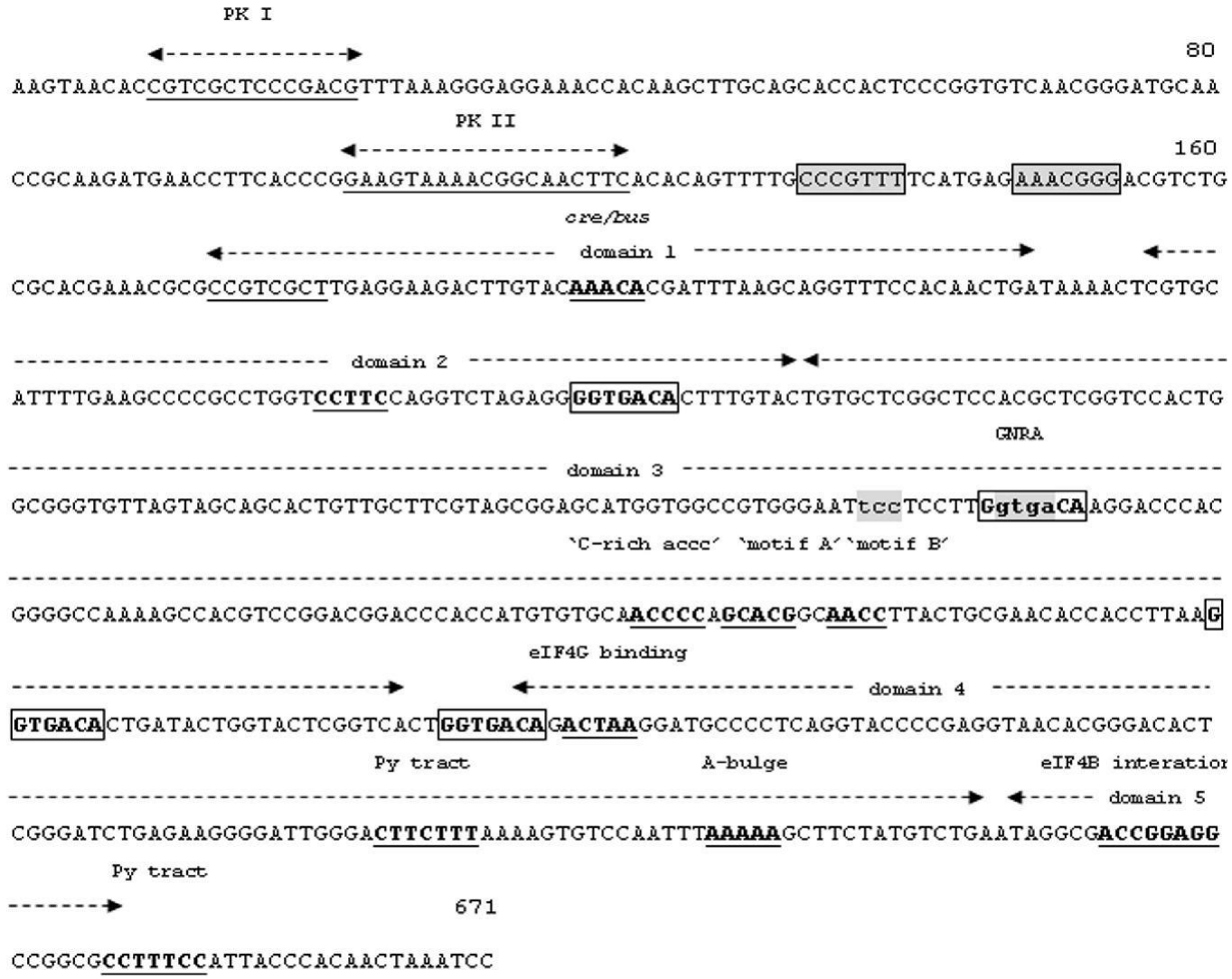

Figure 1. Architecture of the large fragment-5'UTR of FMDV transcriptional control region. Boundaries of different domains have been marked with dashed lines and arrows above the sequence. Conserved critical motifs were depicted in bold faced letter and underlined. Direct repeat motifs were depicted in bold faced letter and frame. Inverted repeats were depicted in frame and shadow. GNRA and $T_{377}CC$ were in low case and shadow eIF4C binding domain was GACTAA, and eIF4B interation domain was ACCGGAGG.

acting replication element (*cre*) and an internal ribosome entry site (IRES) (Mohapatra et al., 2009; He et al., 2011). RNA helicase A (RHA) and $3C^{pro}$ specifically bind the FMDV S fragment. RHA interacts with the S fragment of the FMDV 5 NTR (Lawrence and Rieder, 2009). The result showed (Figure 1) that the IRES element of the WFL strain was about 458 nt in length and had five domains, which participated in the viral protein translation in a cap independent manner. There were two PKs followed by an inverted repeats CCCGTTT/AAACGGG

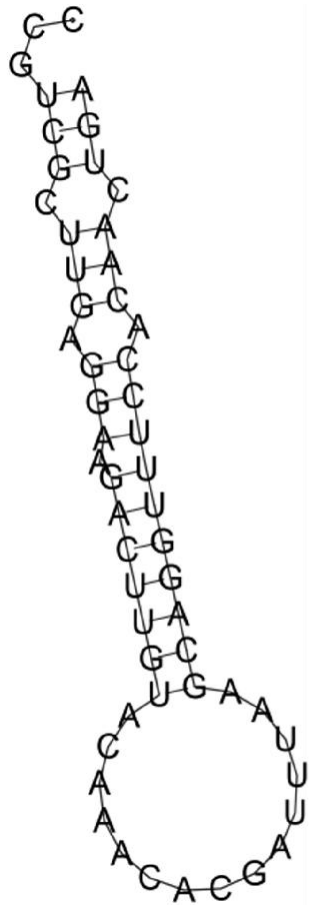

Figure 2. Structure analysis of cre motif. The RNA structure was depicted according to the RNAfold prediction program (MFE plain structure).

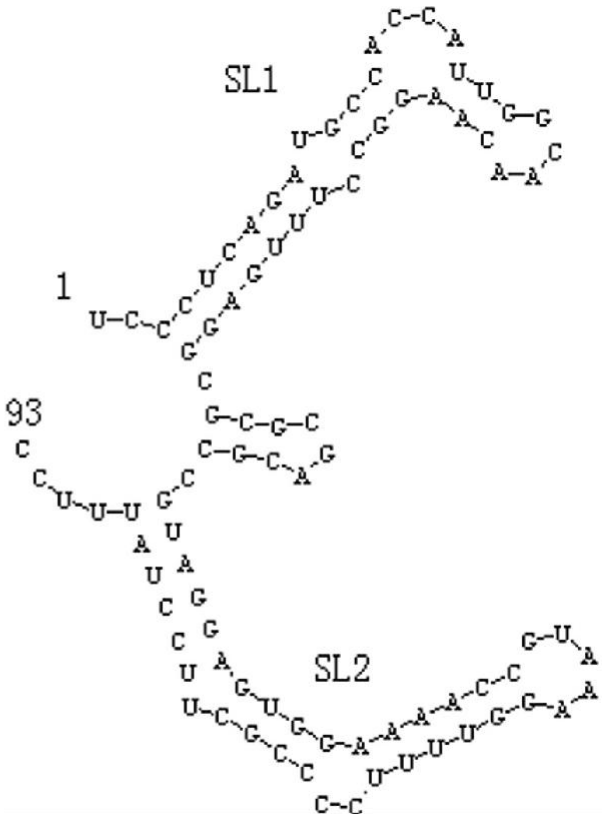

Figure 3. Secondary structure of the FMDV 3'UTR. The RNA structure was depicted according to the RNA-fold prediction program (MFE plain structure). SL1 and SL2 denoted conserved stem-loop structures.

and *cre* (Figure 2).

The *cre* region was essential for RNA genome replication (Marvin and Barry, 2004) . And a conserved 'AAACA' motif in the *cre/bus* region has been recently shown to be involved in VPg uridylylation (López et al., 2001; He et al., 2011) . In this study, *cre* was 55nt with 45.5% of G/C, and had a stem-loop (Figure 2).

IRES including domain 2 to domain 5, and four direct repeat motifs, GGTGACA, were located in IRES region. It was reported that the conserved motifs and the structural domains in the IRES interact with an array of cellular factors involved in host translation initiation (Ramos and Martínez-Salas, 1999; Pacheco and Martinez-Salas, 2010) and some motifs were also crucial in maintenance of the tertiary structure of the IRES through RNA-RNA interaction (Fernández et al., 2006). Domain 4 followed by domain 5 in the IRES displayed highest degree of conservation (Mohapatra et al., 2009). The GNRA tetraloop was a thermostable tetraloop which can exist within a RNA structure solely on its own, or take place in an interaction with a receptor.

The 'GNRA' tetraloop in domain 3, which plays critical role in determining the tertiary structural conformation of the IRES element (Mohapatra et al., 2009), was found to be 'GTGA' in the WFL strain. The cleavage site for RNase P within the 'GNRA' stem- loop was 'T_{377}CC' motif. In this study the conserved 'motif A', which interacts with 'GNRA' motif to maintain structural organization of the central domain of IRES (Nayak et al., 2006), was found to be 'G_{448}CACG' (Figure 1). The eIF4C binding domain was GACTAA, and the eIF4B interation domain was ACCGGAGG.

The 3' UTR, composed of two stem-loops and a poly(A) tract, was required for viral infectivity and stimulates IRES activity (Serrano et al., 2006). The 3' end established two distinct strand- specific, long-range RNA-RNA interactions, one with the S region and another with the IRES element (Serrano et al., 2006). The S region was recognized by each of the separate stem-loops. S-3'UTR interaction was dependent on a structural conformation induced by the presence of the poly(A) tract (Serrano et al., 2006). Here, it was found that 3' UTR of the WFL strain was 127nt, including 93nt stem -loops and poly (A). The 93nt stem-loops region can fold into two stem-loops, SL1 and SL2 (Figure 3).

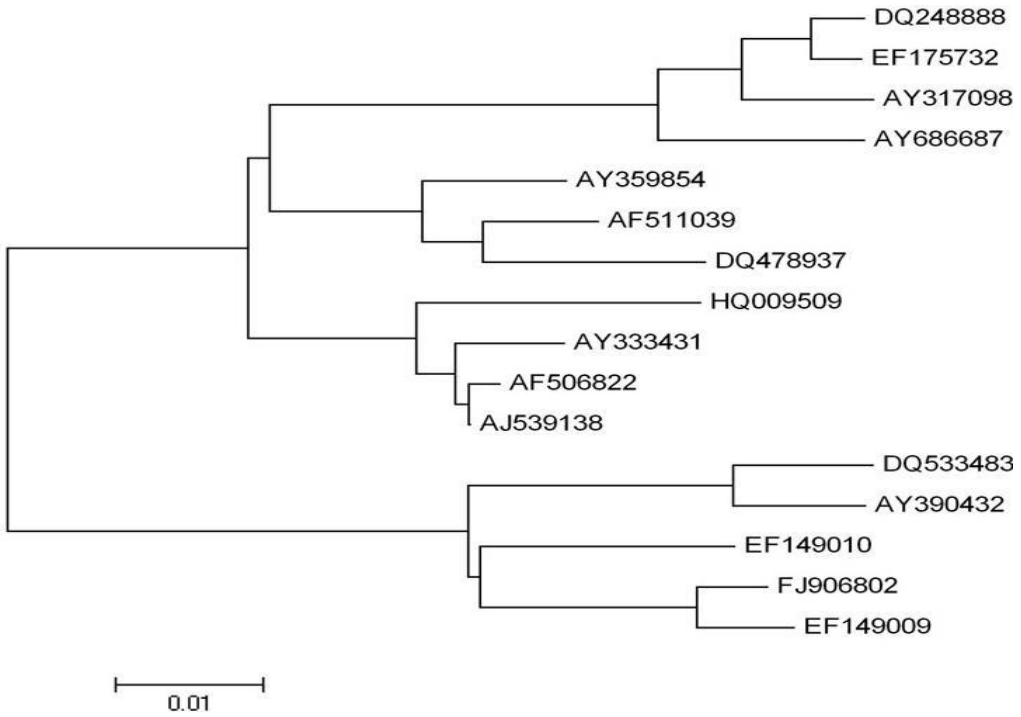

Figure 4. Phylogenetic tree constructed using the complete amino acids sequence of FMDV. A phylogenetic tree (Neighbor-Joining) was constructed using the software MEGA version 4.1 [15]. Four groups were categorized according to their evolutionary relationships. O/ES/2001, HKN/2002, LZ and WFL strain can be divided into one group.

Table 4. Different deduced amino acid residues between the WFL strain and the LZ strain.

Amino acid site	DQ248888	EF175732
37	K	R
57	R	Q
131	M	V
626	V	A
743	G	S
768	Q	K
1048	Q	E
1217	R	K
1405	E	K
1453	K	M
1533	V	A
1653	N	D
1667	A	V
2219	G	D
2258	A	T
2292	R	G

WFL belonged to serotype O

The open reading frame (ORF) of the WFL strain was 6969 nt, encoding 2322 amino acids (Table 2). A phylogenetic tree was constructed based on the deduced, complete amino acid sequences of the WFL strain and reference strains (Table 3) using MEGA 4.1 software. The four groups were categorized (Figure 4) according to their evolutionary relationships. The strains O/ES/2001, HKN/2002, LZ and WFL strain were grouped together. It was reported that both HKN/2002 and LZ were isolated from swine hosts in China, belonging to serotype O (Feng, 2004; Ma et al., 2006). And the strain O/ES/2001 was a recombinant of serotype O and Asia 1 (Wu et al., 2009). It was obvious that WFL strain had a close relationship to LZ strain (GenBank ID: DQ248888), indicating that the WFL strain belonged to serotype O. The close link between WFL and these three isolates was consistent with our previous finding using the VP1 sequence (data not shown).

Complete amino acids of the WFL strain and the LZ strain were compared. The results showed that there were 16 different deduced amino acid residues between WFL and LZ (Table 4). Detailed comparison of WFL with other strains is still doing.

ACKNOWLEDGEMENT

This project was founded by scientific and technical

development project of Jilin Province (20050549), China.

REFERENCES

Feng Q, Yu H, Liu Y, He C, Hu J, Sang H, Ding N, Ding M, Fung YW, Lau LT, Yu AC, Chen J (2004). Genome comparison of a novel foot-and-mouth disease virus with other FMDV strains. Biochem. Biophys. Res. Commun., 323: 254-263.

Fernández-Miragall O, Ramos R, Ramajo J, Martínez-Salas E (2006). Evidence of reciprocal tertiary interactions between conserved motifs involved in organizing RNA structure essential for internal initiation of translation. RNA, 12: 223-234.

Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL (2008). The Vienna RNA websuite. Nucleic Acids Res., 36: 70-74.

Grubman MJ, Baxt B (2004). Foot-and-mouth disease. Clin. Microbiol. Rev., 17: 465-493.

He DS, Li KN, Lin XM, Lin SR, Su DP, Liao M (2011). Genomic comparison of foot-and-mouth disease virus R strain and its chick-passaged attenuated strain. Vet. Microbiol., 150: 185-190.

Lawrence P, Rieder E (2009). Identification of RNA helicase A as a new host factor in the replication cycle of foot-and-mouth disease virus. J. Virol., 83: 11356-11366.

López de QS, Lafuente E, Martínez-Salas E (2001). IRES interaction with translation initiation factors: functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII. RNA, 7: 1213-1226.

Ma MX, Jin NY, Li C, Liu HJ (2006). Molecular characterization of foot-and-mouth disease virus O/LZ. Acta Veterinaria Et Zootechnica Sinica, 37: 1027-1035.

Marvin JG, Barry B (2004). Foot-and-Mouth Disease. Clin. Microbiol. Rev., 17: 465-493.

Mohapatra JK, Sahu A, Barik SK, Sanyal A, Pattnaik B (2009). Comparative analysis of the large fragment of the 5 untranslated region (LF-5 UTR) of serotype a foot-and-mouth disease virus field isolates from India. Virus Genes, 39: 81-89.

Nayak A, Goodfellow IG, Woolaway KE, Birtley J, Curry S, Belsham GJ (2006). Role of RNA structure and RNA binding activity of foot-and-mouth disease virus 3C protein in VPg uridylylation and virus replication. J. Virol., 80: 9865-9875.

Pacheco A, Martinez-Salas E (2010). Insights into the biology of IRES elements through riboproteomic approaches. J. Biomed. Biotechnol., 458927.

Ramos R, Martínez-Salas E (1999). Long-range RNA interactions between structural domains of the aphthovirus internal ribosome entry site (IRES). RNA, 5: 1374-1383.

Salguero FJ, Sánchez-Martín MA, Díaz-San SF, de Avila A, Sevilla N (2005). Foot-and-mouth disease virus (FMDV) causes an acute disease that can be lethal for adult laboratory mice. Virology, 332: 384-396.

Serrano P, Pulido MR, Sáiz M, Martínez-Salas E (2006). The 3' end of the foot-and-mouth disease virus genome establishes two distinct long-range RNA -RNA interactions with the 5' end region. J. Gen. Virol., 87(Pt 10): 3013-3022.

Sobrino F, Sáiz M, Jiménez-Clavero MA, Núñez JI, Rosas MF, Baranowski E, Ley V (2001). Foot-and-mouth disease virus: a long known virus, but a current threat. Vet. Res., 32: 1-30.

Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mole. Biol. Evol., 24: 1596-1599.

Wu ZY, He CQ, Liu YY, Feng Q, Teng JL, Chen JG (2009). A study of homologous recombination in foot-and-mouth disease virus in China. Progress biochem. Biophys., 36: 689-695.