COMPLETE NUCLEOTIDE SEQUENCE ANALYSIS OF A VACCINE STRAIN AND A FIELD ISOLATE OF FOOT-AND-MOUTH DISEASE VIRUS SEROTYPE ASIA1 WITH AN INSERTION IN VP1 GENOMIC REGION

A. SANYAL*, J.K. MOHAPATRA, R. MANOJ KUMAR, S. BISWAS, D. HEMADRI, C. TOSH, G.P. SABARINATH, S.K. GUPTA, M. MITTAL, P. GIRIDHARAN, S.K. BANDYOPADHYAY

Project Directorate on Foot-and-Mouth Disease, Indian Veterinary Research Institute Campus, Mukteswar-Kumaon, Nainital 263 138, Uttaranchal, India

Received March 10, 2004; accepted August 17, 2004

Summary. – Complete nucleotide sequences except the poly (C) tract and poly (A) tail of a vaccine strain (IND 491/97) and an atypical field isolate (IND 321/01) of Foot-and-mouth disease virus (FMDV) serotype Asia1 are described. Amino acid (aa) sequence analysis of the VP1 protein of the field isolate revealed that the latter has 212 instead of 210 or 211 aa found in the so far available sequences of other FMDV isolates of Asia1 serotype. The insertion was localized in the hypervariable region of aa 130–160 of VP1 protein. Nucleotide sequencing of the entire genome was therefore carried out to detect changes in other parts of the genome, if any, besides VP1, which could contribute to its fitness. An 8.16 kb sequence of IND 491/97 and an 8.162 kb sequence of IND 321/01 were compared with each other and also with the known sequence of IND 63/72, another vaccine strain of serotype Asia1. Comparison of the entire polyprotein coding (L to 3D) region of IND 321/01 with those of the two Asia1 vaccine strains (IND 63/72 and IND 491/97) revealed no significant differences. A similar comparison of IND 491/97 with IND 63/72 revealed variability across the entire length of the genome. In addition to the capsid-coding region, sequence variability was also observed in non-structural proteins albeit to different extent. This study shows that in the gene pool of serotype Asia1 at least three groups of isolates/strains are present with respect to the length of VP1 protein.

Key words: field isolate IND 321/01; Foot-and-mouth disease virus; genome; serotype Asia1; vaccine strain IND 491/97; nucleotide sequence

Introduction

Foot-and-mouth disease (FMD) continues to be a global problem in international trade with wide-ranging impacts on livestock production and economic development of a country. FMDV exhibits an extensive genetic variation as evident from the prevalence of multiple genotypes/topotypes/lineages within each serotype (Samuel and Knowles, 2001; Gurumurthy *et al*., 2002; Tosh *et al*., 2002). In India FMD is endemic and incidence of the disease caused by FMDV of serotype Asia1 is encountered throughout the year. This serotype is also prevalent in many other countries of South Asia and Middle East and was found to be responsible for the outbreaks in Greece (Knowles and Davies, 2000). Phylogenetic analyses of Asia1 isolates from India and other countries have revealed the prevalence of more than one genetic lineage (Gurumurthy *et al*., 2002; Mohapatra *et al*., 2002) and recently, a divergent group has been identified within the widely circulating lineage of the Asia1 virus (Sanyal *et al*., 2004).

Of two vaccine strains of serotype Asia1 (IND 63/72 and IND 491/97) used in India a full-length sequence of the former (IND 63/72) has recently been determined (GenBank Acc. No. AY304994). These vaccine strains happen to be the representatives of the two circulating genetic lineages

*E-mail: asanyal68@email2me.net; fax: +915942-286307.

Abbreviations: FMD = Foot-and-mouth disease; FMDV = FMD virus; aa = amino acid; PBS = phosphate-buffered saline; UTR = untranslated region; nt = nucleotide

in India as evident from the 1D region-based phylogeny (Gurumurthy *et al*., 2002). Recently, complete nucleotide sequences have been reported for a serotype O vaccine strain (Pereda *et al*., 2002) and several PanAsia strains (Kanno *et al*., 2002; Kweon *et al*., 2002; Mason *et al.,* 2003). It has been established that changes in different regions (2C, 3A and 3B) of the FMDV genome are responsible for altered viral properties (Escarmis *et al*., 1992; Beard *et al*., 2000; Nunez *et al*., 2001).

A complete molecular characterization of vaccine strains, both genetic and antigenic, and further elucidation of their genetic and evolutionary relationships with the circulating field isolates is necessary in terms of molecular epidemiology and control programs.

During routine nucleotide sequencing aimed at epdemiological investigation of FMD, one isolate (IND 321/01) from the state of Madhya Pradesh was found to harbor an insertion of a codon in the hypervariable region of 1D region (Acharya *et al*., 1989; Bittle *et al.*, 1982). This insertion prompted us to investigate the whole genome sequence for other noteworthy changes, which might have occurred in the course of evolution.

Here we report (i) a complete nucleotide sequence except the poly(C) tract and the poly(A) tail of a FMDV vaccine strain (IND 491/97) and a field isolate (IND 321/01) of serotype Asia1 and (ii) a comparison of these viruses with the vaccine strain IND 63/72 in relation to their nucleotide and amino acid sequences.

Materials and Methods

Virus. A suspected tongue epithel sample (a 10% suspension in PBS pH 7.4) was confirmed to contain FMDV serotype Asia1 by sandwich ELISA (Bhattacharya *et al*., 1996). Subsequently, this material designated as IND 321/01 was passaged in BHK-21 monolayer cell cultures. The isolate IND 491/97 (WBN 117/85 isolated in 1985 in the state of West Bengal), a vaccine strain was taken from the national repository maintained at this laboratory as BHK-21 cell culture-adapted virus and used in the study.

RT-PCR. Viral RNA was extracted from infected cell culture supernatants using the Total RNA isolation kit (Qiagen). The whole genome was amplified in six overlapping fragments (Table 1) in one-step RT-PCR (Tosh *et al*., 1997) using the Taq DNA polymerase/AMV reverse transcriptase enzyme of Access RT-PCR kit (Promega). Initially, the RT step was done at 48°C for 45 mins. Then the PCR was carried out in 40 cycles; the denaturation step consisted of 94°C/30 secs, the annealing step lasted 30 secs and the extension step proceeded at 68°C. The annealing temperature and extension time varied depending on the primer combinations used (Table 1). A final extension step at 68°C for 10 mins was given for all the combinations. For the amplification of VP1 region, the RT step was carried out with 10 µl of RNA in 25 µl of final reaction volume using AMV reverse transcriptase and 20 pmoles of NK61 primer (Knowles and Samuel, 1995). The reaction mixture was incubated at 48°C for 1 hr. Additional incubation at 95°C for 5 mins was done to inactivate the enzyme. The PCR amplification from this cDNA was carried out using Taq DNA polymearase/pfu DNA polymerase and 20 pmoles of each primer (NK61 and 1C505) (Knowles and Samuel, 1995) as per the recommendation of the manufacturer. The PCR products were purified using QIAquick gel extraction kit (Qiagen).

Table 1. Primer combinations used to generate the overlapping PCR fragments

Primer combination	Annealing temperature (°C)	Extension time (mins)	Length of amplicon (bp)
SF-1F/SF-370R	55	30 secs	300
LF-1F/MG39	50	2.0	1629
L01F/NK61	55	3.0	3005
1C505/MG11	55	3.0	2937
3AB1/CTLV12	50	3.0	2813
3D1081/oligo $d(T)_A$	50	30 secs	426

Nucleotide sequencing of different templates was performed using 21 different 5'-Cy5-labeled primers (Table 2) and the *fmol* DNA cycle sequencing kit (Promega). The reaction was stopped by the addition of 4 µl of a stop solution and was resolved on ALF Express II automated DNA sequencer (Amersham Pharmacia Biotech). Analysis of data was carried out on ALFwin™ Sequence Analyser software version 2.10.

Analysis of nucleotide sequence. Multiple alignment and all the subsequent analyses of the sequences from the viruses corresponding to the full-length nucleotide sequence were performed. The nucleotide and amino acid sequences were aligned using the CLUSTAL W algorithm (Thompson *et al*., 1994), available in the OMIGA 2.0 (Oxford Molecular Ltd., UK). The sequence of IND 63/72 (AY 304994) was taken from the GenBank. The nucleotide sequences generated in the present study have been submitted to GenBank. Their Acc. Nos. are: AY687333 (IND 321/01) and AY687334 (IND 491/97).

Results and Discussion

Existence of field isolates with alteration in the region coding for non-structural proteins (deletion in the 3A gene) has already been reported in case of the FMDV serotype O. Detection of a naturally occurring viable virus (IND 321/01) with a unique insertion in the region coding for structural proteins aroused the idea of sequencing the whole genome of this virus to determine other noteworthy change(s), if any, which might have contributed to the survivability of this field isolate.

In order to rule out possible sequencing artifacts, RNA of IND 321/01 was extracted from four successive passages; the PCR amplification was carried out in different *in vitro* amplification experiments using both *Taq* and *Pfu* DNA polymerase. Each amplicon harboring the 1D genomic region was sequenced for both the strands to make sure that

Table 2. Synthetic oligonucleotides used as primers for amplification and sequencing

Name	Purpose	Sequence	Position	Polarity	Reference
SF-1F	P & S	5'TTGAAAGGGGGCGCTAGGGTC	SF-UTR	1, positive	Toja *et al.,* 1999
SF-370R	P & S	5'CGGTAAAACTTAGGGGGGATGAAAG GCGGGCGCCGGGTG	SF-UTR	350, negative	This study
LF-1F	P	5'CCCCCCTAAGTTTTACCGTCGTTCCCG	LF-UTR	1, positive	This study
DH11	S	5'CCGTCGTTCCCGACGTAAAA	LF-UTR	6, positive	This study
L01F	P & S	5'GTGCCCCAGTTTAAAAAGCTT	LF-UTR	632, positive	Roberts and Belsham, 1995
DH13	S	5'TGTAGACCCAGTCGAAG	L	201, negative	This study
DH6	S	5'TTGTTCTGAGTGTTGGTTGTGTG	VP4	172, negative	Mohapatra *et al.,* 2002
DH1	S	5'AACAACTACTACATGCA	VP4	67, positive	Sabarinath, 2001
MG39	P	5'GGTGGTGAGGATGCGGTCTTC	VP2	31, negative	George, 2000
MG41	S	5'GTAGGTGTTGGACATGTGCCCCGC	VP3	244, negative	Gurumurthy, 2000
MG45	S	5'GGCAACATGGTGACCACAGACCC	VP3	34, positive	Gurumurthy, 2000
1C505	P & S	5'TACACTGCTTCTGACGTGGC	VP3	508, positive	Knowles and Samuel, 1995
NK72	S	5'GAAGGGCCCAGGGTTGGACTC	2A	34, negative	Knowles and Samuel, 1995
NK72F	S	GAGTCCAACCCTGGGCCCTTC	2A	34, positive	This study
NK61	P & S	5'GACATGTCCTCCTGCATCTG	2B	58, negative	Knowles and Samuel, 1995
2B325	S	5'GACTCGCTCTCCAGTCTCTTT	2B	325, positive	This study
MG32	S	5'TGTCTCCGTGGCAAATCCGGCCA	2C	319, positive	George, 2000
MG33	S	5'ATGCAACAAGATATGTTTAAGCC	2C	835, positive	George, 2000
3AB1	P	5'CGTGGGATCCCAATCTCAATTCCTTCTCAA	3A	1, positive	Diego *et al.,* 1997
3A402F	S	5'GAGTGATGACGTGAACTCCGAGCC	3A	402, positive	This study
MG11	P	5'CGGTGCGGCCGCTCACTCAGTGACAATCAG	3B3	58, negative	George, 2000
3D26R	S	5'ACATCTCTGGTGTCAACAATCAACCCCTCGTG	3C	634, negative	This study
3D331	S	5'AGGCGCGGTGTCTGGCTCCAT	3D	331, negative	George, 2000
3D786	S	5'GCGGAACACCTCCTCAAACAT	3D	766, negative	George, 2000
3D1081	P & S	5'GGCCAAACCATCACTCCAGCTGA	3D	1081, positive	George, 2000
CTLV2	S	5'TGATCTGTAGCTTGGTATCT	3D	1352, negative	Pattnaik *et al.,* 1997
CTLV12	P	5'GGAAGCGGGAAAAGCTCTTT	3'UTR	61, negative	Pattnaik *et al.,* 1997
OLIGO $d(T)_A$	P	23 MER	poly (A)	This study	

S = sequencing primer, P = PCR primer.

Table 3. Pairwise nucleotide homology and divergence observed between the vaccine strains IND 63/72 and IND 491/97 and the field isolate IND 321/01 at different genomic regions

Region	Length[a]	Divergence/homology (%)		
		IND 321/01 & IND 63/72	IND 63/72 & IND 491/97	IND 321/01 & IND 491/97
5'-SF-UTR[b]	367	12.9/86.5	16.3/83.5	20.0/80.3
5'-LF-UTR	710[c]	9.3/89.8	9.5/88.5	8.3/89.8
L	603	13.8/86.9	14.5/86.6	17.5/84.2
VP4	255	7.9/92.5	10.3/90.6	11.3/89.8
VP2	654	14.0/86.9	14.2/86.5	8.2/92.0
VP3	657	14.6/86.0	13.2/87.5	9.6/90.9
VP1	633[c]	17.6/80.9	15.2/83.7	10.2/89.3
2A	48	6.6/93.8	8.9/89.6	6.6/91.7
2B	462	6.8/93.5	8.3/92.2	6.4/93.9
2C	954	12.5/87.5	10.9/89.5	8.1/92.0
3A	459	8.6/91.7	10.1/89.8	5.9/94.1
3B	213	9.1/90.6	8.0/92.0	7.0/93.4
3C	639	9.4/90.6	8.3/91.7	8.9/91.2
3D	1410	7.9/92.3	7.8/92.5	7.7/92.1
3'-UTR	96[c]	14.3/78.3	7.7/89.6	6.8/87.2

[a]The length is shown for IND 491/97.
[b]The sequence of the first three nucleotides could not be deduced.
[c]The length varies in dependence on the isolate.

the insertion was not due to mis-incorporation by *Taq* polymerase in the process of synthesis of DNA and allowed the determination of the consensus sequence of the virus population in this region.

Sequencing of IND 321/01 and subsequent comparison with IND 63/72 and IND 491/97 revealed only insertion of an amino acid in the VP1 protein. Phylogenetic analysis of the 1D genomic region of published Asia1 viruses and IND 321/01 revealed that the latter belongs to the commonly circulating lineage (Mohapatra *et al*., 2002) in the country (data not shown). The two vaccine strains IND 63/72 and IND 491/97 currently in use are obviously members of two different lineages of which IND 491/97 belongs to the B lineage (Mohapatra *et al.,* 2002). Comparison of the amino acid sequence at the coding region revealed 82 positions where IND 321/01 showed similarity with IND 491/97 but not with IND 63/72. Likewise, at 32 positions IND 321/01 showed homology with IND 63/72 but not with IND 491/97.

We determined a complete nucleotide sequence of FMDV Asia1 vaccine strain IND 491/97. Excluding the poly(C) tract and poly(A) tail, the overall length of the genome was about 8.16 kb which comprised of 367 nucleotides of the short fragment of 5'-untranslated region (SF-UTR) (Fig. 1a), 710 nucleotides of the large fragment of 5'-untranslated region (LF-UTR) (Fig. 1b), 6987 nucleotides of the polyprotein coding region, and 96 nucleotides of 3'-UTR (Fig. 1c). Comparison of the nucleotide sequences of all the regions in IND 491/97 and in another vaccine strain IND 63/72 revealed that the short fragment was most divergent (16.7%). Amino acid sequence comparison between the two vaccine strains showed that 5.6% of the amino acids of coding region are variable. Pair-wise comparison revealed a maximum sequence divergence of 20% in the SF-UTR region between IND 321/01 and IND 491/97 while a minimum sequence divergence of 6.6% in the 2A region between IND 321/01 and IND 63/72. Among the structural proteins, VP1 was most heterogeneous, exhibiting nucleotide divergences of 17.6% (between IND 321/01 and IND 63/72), 15.2% (IND 491/97 and IND 63/72) and 10.2% (IND 321/01 and IND 491/97) (Table 3). Insertion of one amino acid was observed in VP1 protein at position 44 in IND 491/97 in relation to IND 63/72. Nucleotide sequence variation was observed throughout the genome of the two viruses which further supported the previous finding of their belonging to separate lineages of the serotype Asia1 (Gurumurthy *et al*., 2002).

The 8,162 kb sequence of the genome of IND 321/01 was determined. In comparing IND 63/72 with IND 491/97 the sequences of 5'- and 3'-UTRs of IND 321/01revealed a divergence of 12.9–20.0% at the nucleotide level. The LF-UTR of IND 321/01 was found to be 711 nucleotides long as per alignment with IND 63/72 and IND 491/97 and it has an extra stretch of 43 nucleotides at the 5'-end of LF-UTR as compared to IND 63/72 (Fig. 1b). The isolate had an 8.9–9.3% divergence in the nucleotide sequence in relation to IND 63/72 and IND 491/97 in this region. The poly-pyrimidine tract (Kuhn *et al*., 1990) was 10 nucleotides long and identical to IND 63/72 and IND 491/97 (Fig. 1b). This tract was followed by a sequence of 17 nucleotides, which represented a variable region of FMDV and differed from that of IND 63/72 and IND 491/97. Comparison of the nucleotide sequence at the 5'-end of 3'-UTR of IND 321/01 with those of IND 491/97 and IND 63/72 revealed a divergence of 6.8% and 14.3%, respectively. The 3'-UTR of IND 321/01 was 94 nt long and was shorter by 3 nucleotides than that of IND 63/72 (Fig. 1c).

Whereas the VP1 sequences of Asia1 PAK 1/54 and IND 63/72 comprise of 210 amino acids, isolates from Bangladesh, Israel, Saudi Arabia, Thailand and most of the isolates from India have 211 amino acids (one extra amino acid at position 44) (Marquardt *et al*., 2000; Gurumurthy *et al*., 2002, Mohapatra *et al*., 2002). The isolate IND 321/01 had two extra amino acids in VP1 at positions 44 and 140 as compared to PAK 1/54, and thus had total 212 amino acids. It should be noted that in the gene pool of FMDV serotype Asia1, there are three groups of viruses concerning the length of VP1 protein. Comparison of the amino acid sequences of the VP1 protein of various isolates from India (Sanyal *et al*., 2000; Gurumurthy *et al*., 2002; Mohapatra *et al*., 2002) has revealed the existence of two groups of viruses; one having a total of 210 amino acids and the motifs T-Q-P-T and V-S-N-R at the positions 139–142 and 154–157, respectively, and another having a total of 211 amino acids and the motifs E-T-T-S and L-S-G-Q at the same positions. IND 321/01 had also the motif L-S-G-Q at the positions 154–157 (Fig. 2); however, at the position 140, one extra amino acid resulting in the motif E-T-I-T-S was found.

In addition, IND 321/01 revealed a total of 32 changes in the genome. These changes were unique to this isolate and were not be detected in any of the two vaccine viruses. The maximum number of the aforesaid changes (seven) were observed in the L protease out of which 3 substitutions, leucine 19, threonine 23 and glutamic acid 129 were never found in any of the Indian field isolates of FMDV Asia1 sequenced so far (Mohapatra *et al*., 2002). However, these changes did not encompass any of the established motifs (catalytic dyad and substrate-binding pocket) of the leader protein. Similar isolate-specific changes were also noticed at four positions (valine 24, serine 34, valine 50 and alanine 209) in the VP1 protein (Fig. 1). A total of six changes were detected in the RNA polymerase but not involving the identified major functional motifs of this protein (Koonin, 1991). The 3A protein of IND 491/97 and IND 321/01 was found to contain no deletion in relation to O/CAM/3/98 (GenBank Acc. No. AJ320488) (11 amino acid deletions at positions 133 to 143) or O/YUN/TAW/97 (GenBank Ac. No. AF308157) (10 amino acid deletions at positions 93 to 102).

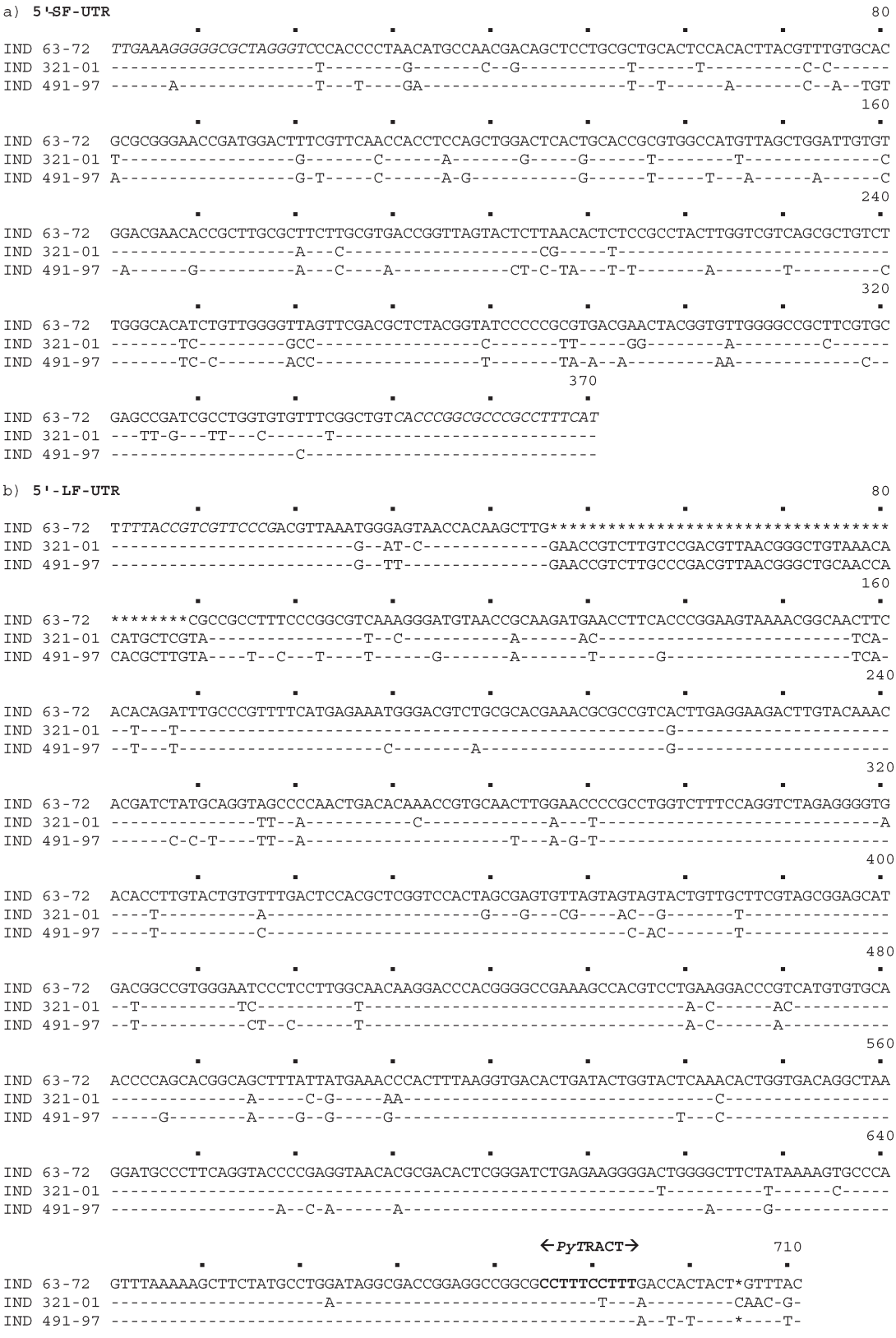
a) 5'-SF-UTR
IND 63-72 TTGAAAGGGGGCGCTAGGGTCCCACCCCTAACATGCCAACGACAGCTCCTGCGCTGCACTCCACACTTACGTTTGTGCAC
IND 321-01 ----------T--------G-------C--G----------T------T----------C-C------
IND 491-97 ------A----------T---T----GA----------T--T------A-------C--A--TGT
IND 63-72 GCGCGGGAACCGATGGACTTTCGTTCAACCACCTCCAGCTGGACTCACTGCACCGCGTGGCCATGTTAGCTGGATTGTGT
IND 321-01 T----------G-------C------A-------G-----G------T--------T----------C
IND 491-97 A----------G-T-----C------A-G----------G------T-----T---A------A------C
IND 63-72 GGACGAACACCGCTTGCGCTTCTTGCGTGACCGGTTAGTACTCTTAACACTCTCCGCCTACTTGGTCGTCAGCGCTGTCT
IND 321-01 ----------A---C----------CG-----T----------
IND 491-97 -A------G----------A---C----A----------CT-C-TA---T-T-------A-------T---------C
IND 63-72 TGGGCACATCTGTTGGGGTTAGTTCGACGCTCTACGGTATCCCCCGCGTGACGAACTACGGTGTTGGGGCCGCTTCGTGC
IND 321-01 -------TC---------GCC----------C-------TT-----GG--------A---------C------
IND 491-97 -------TC-C-------ACC----------T-------TA-A--A---------AA----------C--
IND 63-72 GAGCCGATCGCCTGGTGTGTTTCGGCTGTCACCCGGCGCCCGCCTTTCAT
IND 321-01 ---TT-G---TT---C------T----------
IND 491-97 ----------C----------
b) 5'-LF-UTR
IND 63-72 TTTTACCGTCGTTCCCGACGTTAAATGGGAGTAACCACAAGCTTG**********
IND 321-01 ----------G--AT-C----------GAACCGTCTTGTCCGACGTTAACGGGCTGTAAACA
IND 491-97 ----------G--TT----------GAACCGTCTTGCCCGACGTTAACGGGCTGCAACCA
IND 63-72 ********CGCCGCCTTTCCCGGCGTCAAAGGGATGTAACCGCAAGATGAACCTTCACCCGGAAGTAAAACGGCAACTTC
IND 321-01 CATGCTCGTA----------T--C----------A------AC----------TCA-
IND 491-97 CACGCTTGTA----T--C---T----T------G-------A-------T------G----------TCA-
IND 63-72 ACACAGATTTGCCCGTTTTCATGAGAAATGGGACGTCTGCGCACGAAACGCGCCGTCACTTGAGGAAGACTTGTACAAAC
IND 321-01 --T---T----------G----------
IND 491-97 --T---T----------C--------A----------G----------
IND 63-72 ACGATCTATGCAGGTAGCCCCAACTGACACAAACCGTGCAACTTGGAACCCCGCCTGGTCTTTCCAGGTCTAGAGGGGTG
IND 321-01 ----------TT--A----------C----------A---T----------A
IND 491-97 ------C-C-T----TT--A----------T---A-G-T----------
IND 63-72 ACACCTTGTACTGTGTTTGACTCCACGCTCGGTCCACTAGCGAGTGTTAGTAGTAGTACTGTTGCTTCGTAGCGGAGCAT
IND 321-01 ----T----------A----------G---G---CG----AC--G-------T----------
IND 491-97 ----T----------C----------C-AC-------T----------
IND 63-72 GACGGCCGTGGGAATCCCTCCTTGGCAACAAGGACCCACGGGGCCGAAAGCCACGTCCTGAAGGACCCGTCATGTGTGCA
IND 321-01 --T----------TC----------T----------A-C------AC----------
IND 491-97 --T----------CT--C------T----------A-C------A----------
IND 63-72 ACCCCAGCACGGCAGCTTTATTATGAAACCCACTTTAAGGTGACACTGATACTGGTACTCAAACACTGGTGACAGGCTAA
IND 321-01 ----------A-----C-G-----AA----------C----------
IND 491-97 -----G--------A----G--G-----G----------T---C----------
IND 63-72 GGATGCCCTTCAGGTACCCCGAGGTAACACGCGACACTCGGGATCTGAGAAGGGGACTGGGGCTTCTATAAAAGTGCCCA
IND 321-01 ----------T----------T------C-----
IND 491-97 ----------A--C-A------A----------A-----G----------
←PyTRACT→
IND 63-72 GTTTAAAAAGCTTCTATGCCTGGATAGGCGACCGGAGGCCGGCGCCTTTCCTTTGACCACTACT*GTTTAC
IND 321-01 ----------A----------T---A---------CAAC-G-
IND 491-97 ----------A--T-T----*----T-

Fig. 1

Fig.1. Contd.,

```
c) 3'-UTR                                                                    80
                ?        ?        ?        ?        ?        ?        ?        ?
IND 63-72  TAATCCCTCAGATGTCCCAATTGGCAGAAAGACTCTGAGGCGAGCGACACCGCAGGAGTGAAAGGCCGGAAACGGCTTTT
IND 321-01 ---------------T**-----------GA-----------C-----G---T-----C----A---C-G--A-------
IND 491-97 ----------------A------------------------------G---T----------A---C----A-------

                     97
                ?     ?
IND 63-72  CCCGCTTCCCTATTTCC
IND 321-01 ---------T--*----
IND 491-97 ---------*--A----
```

Fig. 1

Alignment of SF-UTR (a) and LF-UTR R (b) and 3'-UTR (c) of the vaccine strains IND 63-72 and IND 491-97 and the field isolate IND 321-01 of FMDV Asia1 serotype

*Py*TRACT denotes a polypyrimidine tract. Dash (-), indicates no change in relation to IND 63/72. Asterisk (*) indicates absence of a nucleotide as per alignment. PCR primer-binding end sequences are italicized.

```
                    ▪         ▪         ▪         ▪         ▪        60
IND 63-72  TTTAGESADPVTTTVENYGGETQSARRLHTDVAFVLDRFVKLT*PKNTQILDLMQIPSHT
IND 491-97 ---T-------------------T-------------------T--S--T----------
IND 321-01 ---T-------------------V---------S---------A-----V----------

                    ▪         ▪         ▪         ▪         ▪       120
IND 63-72  LVGALLRSATYYFSDLEVALVHTGSVTWVPNGAPKDALDNHTNPTAYQKKPITRLALPYT
IND 491-97 ------------------------P-------S-------Q------R-R----------
IND 321-01 ------------------------P-------S-------Q--------Q----------

                    ▪         ▪         ▪         ▪         ▪       180
IND 63-72  APHRVLATVYNGKTTYGTQ*PTRRGDLAVLAQRVSNRLPTSFNYGAVKADTITELLIRMT
IND 491-97 -----------------ET*TS----M-A----L-G------------E---------K
IND 321-01 -----------------ETITP----M-A----L-G----------------------K

                    ▪         ▪       210
IND 63-72  RAETYCPRPLLALDTTHDRRKQEIIAPEKQVL
IND 491-97 ----------------Q---------------
IND 321-01 ----------------Q----------A--MM
```

Fig. 2

Alignment of the sequences of VP1 protein of the vaccine strains IND 63-72 and IND 491-97 and the field isolate IND 321-01 of FMDV Asia1 serotype

Amino acid positions marked with asterisk represent the locations of insertion as per alignment. Dash (-) indicates no change in relation to IND 63/72.

Since the bB-bC and bG-bH loops in the VP1 protein are antigenically important (Bittle *et al*., 1982; Pfaff *et al*., 1982; Strohmaier *et al*., 1982), insertion of amino acids would be expected to affect the antigenicity of the protein. The bG-bH loop of VP1 is disordered in the crystal structure of FMDV (Acharya *et al*., 1989). Substitution at the bB-bC loop of VP1 in FMDV serotype O has been reported to affect antibody binding to the bG-bH loop by changing the loop conformation (Parry *et al*., 1990). Variation in the length of bG-bH loop has also been observed for FMDV of serotypes A and C. In the isolate under study, heterogeneity in the length was observed in both the bB-bC and bG-bH loops. The observed insertion could be due to the spatial proximity of certain loop/region of the genome (coding for the inserted

isoleucine) to the E-T-T-S coding region of VP1. This might have favored an intra-molecular strand switching by the 3D polymearse resulting in the origin of a viable progeny virion with an E-T-I-T-S motif in place of the naturally observed E-T-T-S motif. However, it is difficult to pinpoint the exact mechanism of the insertion observed in the isolate under study. In general, the insertion or deletion of amino acids modifies the function of a protein to some extent. This can happen in the loop regions (such as the bB-bC and bG-bH loops) which are structurally diverse and can accept changes with minimum loss of function, and being highly antigenic they are prone to change under positive selection pressure.

Acknowledgements. The authors thank Indian Council of Agricultural Research for providing necessary facilities to carry out this work and Dr. A.M.Q. King for his helpful comments to improve the manuscript.

References

Acharya R, Fry E, Stuart D, Fox G, Rowlands D, Brown F (1989): The three-dimensional structure of foot-and-mouth disease virus at 2.9 Å resolution. *Nature* **337**, 709–716.

Beard CW, Mason PW (2000): Genetic determinants of altered virulence of Taiwanese foot-and-mouth disease virus. *J. Virol.* **74**, 987–991.

Bhattacharya S, Pattnaik B, Venkataramanan R (1996): Development and application of sandwich enzyme-linked immunosorbent assay (ELISA) for type identification of foot-and-mouth disease (FMD) virus in direct field materials. *Indian J. Anim. Sci.* **66**, 1–9.

Bittle JL, Houghten RA, Alexander H, Shinnick TM, Sutcliffe JG, Lerner RA, Rowlands DJ, Brown F (1982): Protection against foot-and-mouth disease by immunization with chemically synthesized peptide predicted from the viral nucleotide sequence. *Nature* **298**, 30–33.

Diego M de, Brocchi E, Mackay D, Simone F de (1997): The non-structural polyprotein 3ABC of foot-and-mouth disease virus as a diagnostic antigen in ELISA to differentiate infected from vaccinated cattle. *Arch. Virol.* **142**, 2021–2033.

Escarmis C, Toja M, Medina M, Domingo E (1992): Modifications of the 5' untranslated region of foot-and-mouth disease virus after prolonged persistence in cell culture. *Virus Res.* **26**, 113–125.

George M (2000): Molecular cloning and expression of the nonstructural proteins of foot-and-mouth disease virus serotype Asia1. *Ph.D. Thesis* submitted to the Deemed University, Indian Veterinary Research Institute, Izatnagar/Mukteswar, Uttar Pradesh, India.

Gurumurthy CB (2000): Antigenic sites of foot-and-mouth disease virus type Asia1. *Ph.D. Thesis* submitted to the Deemed University, Indian Veterinary Research Institute, Izatnagar/Mukteswar, Uttar Pradesh, India.

Gurumurthy CB, Sanyal A, Venkataramanan R, Tosh C, George M, Hemadri D (2002): Genetic diversity in the VP1 gene of foot-and-mouth disease virus serotype Asia1. *Arch. Virol.* **147**, 85–102.

Kanno T, Yamakawa M, Yoshida K, Sakamoto K (2002): The complete nucleotide sequence of PanAsia strain of foot-and- mouth disease virus isolated in Japan. *Virus Genes* **25**, 119–125.

Knowles NJ, Davies PR (2000): Origin of recent outbreaks of foot-and-mouth disease in North Africa, the Middle East and Europe. *Rpt. Sess. Res. Gp. Stand. Tech. Comm. Eur. Comm.* Control of FMD (FAO), Borovets, Bulgaria, pp. 39–45.

Knowles NJ, Samuel AR (1995): Polymerase chain reaction amplification and cycle sequencing of the 1D (VP1) gene of foot-and-mouth disease viruses. *Rpt. Sess. Res. Gp. Stand. Tech. Comm. Eur. Comm.* Control of FMD (FAO), Vienna, Austria, pp. 45–53.

Kweon CH, Ko YJ, Kim WI, Kwon BJ, Hyun BH, Sohn HJ, Choi KS, Shin JH (2002): Molecular characterization of foot-and-mouth disease virus O/SKR/2000. *Virus Res.* **90**, 15–22.

Marquardt O, Rahman MM, Freiberg B (2000): Genetic and antigenic variances of foot-and-mouth disease virus type Asia1. *Arch. Virol.* **145**, 149–157.

Mason PW, Pacheco JM, Zhao QZ, Knowles NJ (2003): Comparisons of the complete genomes of Asian, African and European isolates of a recent foot-and-mouth disease virus type O pandemic strain (PanAsia). *J. Gen. Virol.* **84**, 1583–1593.

Mohapatra JK, Sanyal A, Hemadri D, Tosh C, Sabarinath GP, Venkataramanan R (2002): Sequence and Phylogenetic analysis of the L and VP1 genes of Foot-and-mouth disease virus serotype Asia1. *Virus Res.* **87**, 107–118.

Nunez JI, Baranowski E, Molina N, Ruiz-Jarabo CM, Sanchez C, Domingo E, Sobrino F (2001): A single amino acid substitution in non-structural protein 3A can mediate adaptation of foot-and-mouth disease virus to the guinea pig. *J. Virol.* **75**, 3977–3983.

Parry NR, Fox G, Rowlands DJ, Brown F, Fry E, Acharya R, Logan D, Stuart D (1990): Structural and serological evidence for a novel mechanism of antigenic variation in foot-and-mouth disease virus. *Nature* **347**, 569–572.

Pattnaik B, Sanyal A, George M, Tosh C, Hemadri D, Venkataramanan R (1997): Evaluation of primers for PCR amplification of RNA polymerase gene of foot-and-mouth disease virus. *Acta Virol.* **41**, 333–336.

Pereda AJ, Konig GA, Chimeno Zoth SA, Borca M, Palma EL, Piccone ME (2002): Full length nucleotide sequence of foot-and- mouth disease virus strain O_1 Campos/Bra/58. *Arch. Virol.* **147**, 2225–2230.

Pfaff E, Mussgay M, Bohm HO, Schulz GE, Schaller H (1982): Antibodies against a preselected peptide recognize and neutralize foot-and-mouth disease virus. *EMBO J.* **7**, 869–874.

Roberts PJ, Belsham G (1995): Identification of critical amino acids within the foot-and-mouth disease virus leader protein, a cysteine protease. *Virology* **213**, 140–160.

Sabarinath GP (2001): Comparison of nucleotide sequence of structural protein encoding (P1) region of foot-and-mouth disease virus serotype O isolates. *M.V.Sc. Thesis* submitted

to the Deemed University, Indian Veterinary Research Institute, Izatnagar/Mukteswar, Uttar Pradesh, India.

Samuel AR, Knowles NJ (2001): Foot-and-mouth disease type O viruses exhibit genetically and geographically distinct evolutionary lineages (topotypes). *J. Gen. Virol.* **82**, 609–621.

Sanyal A, Gurumurthy CB, Venkataramanan R, Hemadri D, Tosh C (2000): Comparison of amino acid sequences at the amino acid 130-160 region of VP1 polypeptide on Indian field isolates of foot-and-mouth disease virus serotype Asia1. *Acta Virol.* **44**, 85–90.

Sanyal A, Hemadri D, Tosh C, Bandyopadhyay SK (2004): Emergence of a novel subgroup within the widely circulating lineage of Foot-and-Mouth Disease virus serotype Asia1 in India. *Res. Vet. Sci.* **76**, 151–156.

Strohmaier K, Franze R, Adam KH (1982): Location and characterization of the antigenic portion of the foot-and-mouth disease virus immunizing protein. *J. Gen. Virol.* **59**, 295–306.

Thompson JD, Higgins DG, Gibson TJ (1994): CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.

Toja M, Escarmis C, Domingo E (1999): Genomic nucleotide sequence of a foot-and-mouth disease virus clone and its persistent derivatives. Implications for the evolution of viral quasispecies during a persistent infection. *Virus Res*. **64**, 161–172.

Tosh C, Hemadri D, Sanyal A, Pattnaik B, Venkataramanan R (1997): One-tube and one-buffer system of RT-PCR amplification of 1D gene of foot-and-mouth disease virus field isolates. *Acta Virol.* **41**, 153–155.

Tosh C, Sanyal A, Hemadri D, Venkataramanan R (2002): Phylogenetic analysis of serotype A foot-and-mouth disease virus isolated in India between 1977 and 2000. *Arch. Virol.* **147**, 493–513.