

Original Article

Sequence and phylogenetic analysis of the non-structural 3A and 3B protein-coding regions of foot-and-mouth disease virus subtype A Iran 05

Saber Jelokhani-Niaraki¹, Majid Esmaelizad^{2,*}, Morteza Daliri⁴, Rasoul Vaez-Torshizi¹, Morteza Kamalzadeh³, Mohsen Lotfi³

¹Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran

²Departments of Biotechnology, and ³Quality Control, Razi Vaccine & Serum Research Institute, Karaj, Iran

⁴Department of Animal Biotechnology, National Institute for Genetic Engineering and Biotechnology, Tehran, Iran

The A Iran 05 foot-and-mouth disease virus (FMDV) subtype was detected in Iran during 2005 and has proven to be highly virulent. This study was undertaken to focus on molecular and phylogenetic analysis of 3A and 3B coding-regions in the A Iran 05 field isolate. To assess the genetic relatedness of A Iran 05 isolate the nucleotide and predicted amino acid sequences of the 3AB region of type A FMDV isolates were compared with twenty previously described type A FMDV isolates. The phylogenetic tree based on the 672 bp 3AB gene sequences of type A FMDV from thirteen different locations clustered them into five distinct lineages. The A Iran 05 isolate clustered in lineage A along with four type A variants and was closely matched with viruses isolated in Turkey and Pakistan during 2005~2006. The number of protein sequence differences exhibited by each of the isolates revealed that A Iran 05 isolate contains three amino acid substitutions at positions 47 and 119 of 3A and 27 of the 3B coding region. The nucleotide identity between A Iran 05 and the other four isolates of lineage A was estimated to be 98%.

Keywords: A Iran 05, FMDV, non-structural protein 3A, non-structural protein 3B, sequencing

Introduction

Foot-and-mouth disease (FMD) is a highly contagious viral disease of cloven-hoofed animals, both domestic and wild [8,18]. The disease is caused by FMD virus (FMDV), which is one of two species in the genus *Aphthovirus* in the family of Picornaviridae, an important group of mammalian RNA viruses [3,11]. FMD has been endemic in many continents, including Africa, South America and Asia. It has had a

dramatic impact on the farming industry leading to tremendous economic losses particularly in countries which are naturally FMD-free [20]. Previous studies have shown that Iran has one of the highest reported rates of FMD cases per year. The high incidence of FMD (including serotypes A, Asia 1 and O) has allowed the identification of new variants of the virus over the last 7 years. [16]. Furthermore, due to high error rates in the process of genome replication, FMDV is genetically highly variable. It has been grouped into seven serological serotypes (O, A, C, Asia 1, SAT 1, SAT 2 and SAT 3) and multiple subtypes, with more than 60 subtypes of the virus now characterized [13]. This large antigenic variation is considered one of the major obstacles for the control of FMD by vaccination in endemic areas [1,4,9]. Non-enveloped FMDV contains a single-stranded, positive-sense RNA genome with approximately 8,500 nucleotides. The open reading frame (ORF) encodes a single polyprotein which can be cleaved into four structural proteins (VP4, VP2, VP3, and VP1) and eight non-structural proteins (L, 2A, 2B, 2C, 3A, 3B, 3C, and 3D) [6,17]. In general, structural proteins are more variable than non-structural proteins. Mutations or deletions in structural proteins may help FMDV to evade an immune response produced by the host, whereas mutations or deletions in the non-structural proteins could be detrimental to viral replication and protein processing [7]. In the non-structural protein regions, the 2A, 2B, 2C, 3C, and 3D proteins are highly conserved, while the 3A and 3B proteins are less conserved among serotypes, probably due to their functions or actions with host factors [12]. The wide variation of serotypes across different geographical areas points to the need for urgent identification of regional serotypes for outbreak control and vaccine development.

The purpose of this study was to determine the full-length nucleotide sequence of genes encoding 3A and 3B proteins of A Iran 05 subtype, and to compare it with available corresponding sequences deposited in the GenBank database.

*Corresponding author

Tel: +98-261-4502849; Fax: +98-261-4552194

E-mail: m.esmaelizad@rvsri.ir

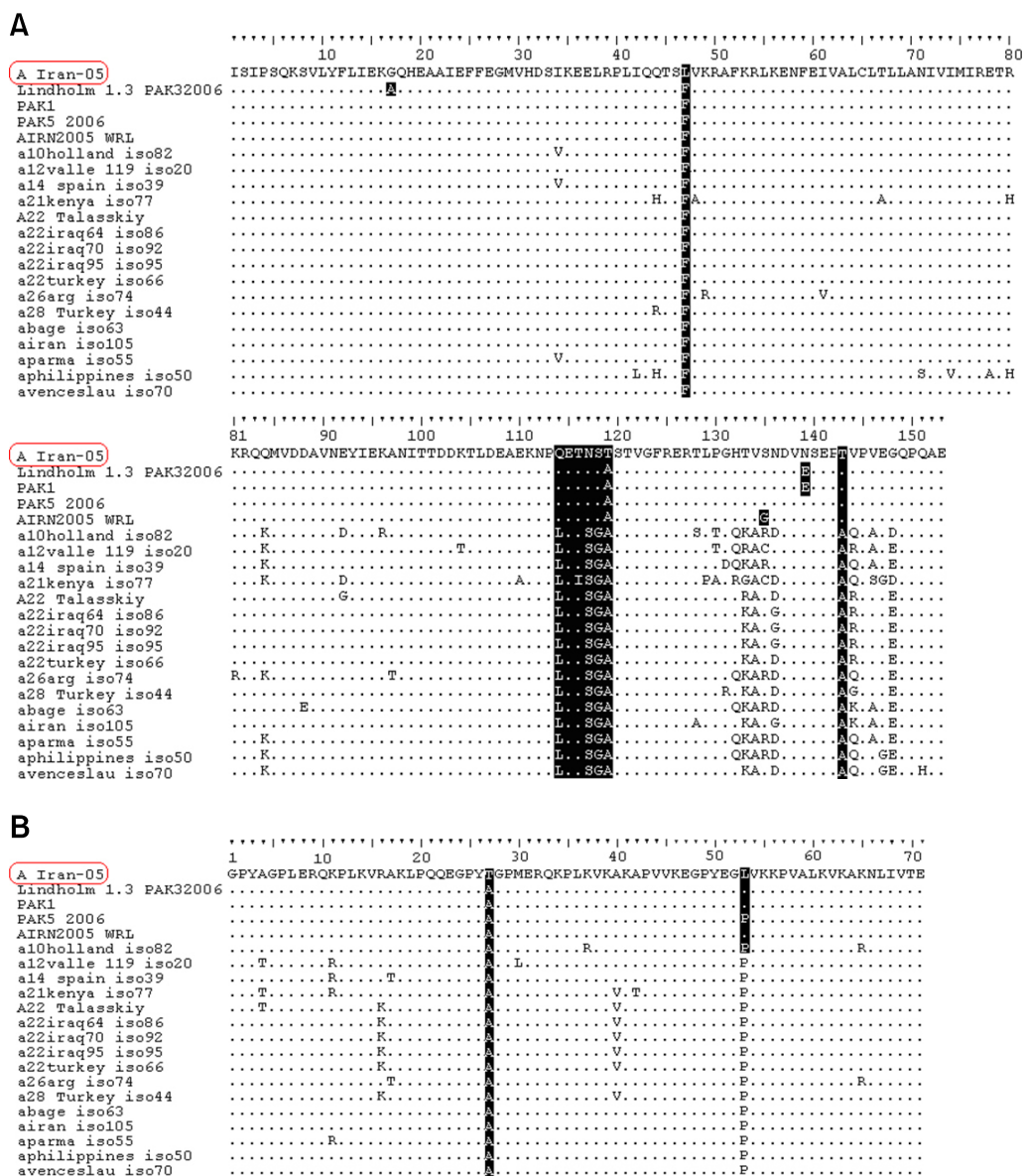
Materials and Methods

Viruses

In 2008, clinical specimens of FMDV which included tongue epithelium tissue were collected from one of the Iranian fields located in Tehran. They were sent immediately to the laboratory of Razi Vaccine and Serum Research Institute in Karaj. Laboratory and serological tests for FMDV confirmed the presence of infected samples. The new virulent isolate was passaged in baby hamster kidney cell monolayer cultures, and the infected cell culture supernatant was clarified and stored at -70°C before use.

RNA extraction, reverse transcription-polymerase chain reaction (RT-PCR) and sequencing

The infected cell culture supernatants were used for the extraction of viral RNA using the Total RNA Isolation kit (Roche, Swiss) following the recommendations of the suppliers. The 3AB coding region of FMDV is 672 nucleotides long and was amplified by standard methods in a one-step RT-PCR, using the primer combination forward G10 (5'CCAAGGACGGGTACAAAGTTAAC3') and reverse G11 (5'ACCATCTTTTGCAAGTCG/AGTC3'). The amplified PCR products (672 bp) of the expected length were subjected to electrophoresis in a 1% agarose gel and visualized by staining



with ethidium bromide under UV transilluminator. After the successful amplification of the target DNA sequence, fragments were purified using gel extraction kit (Roche, Swiss) in accordance with the manufacturer's instructions. Purified PCR product was cloned into PNTZ57T vector (Fermentas, Germany) according to the manufacturer's instructions and sequencing was carried out using T7 promoter primer (MWG Biotech, Germany).

Sequence analysis and phylogenetic tree construction

A fragment of 672 nucleotides containing the full length 3A and 3B regions (edited manually) enabled us to perform multiple sequence alignments using the BioEdit software, version 7 [10]. To determine the degree to which genetic diversity is reflected in the 3A and 3B proteins, alignment and comparison of the deduced amino acid sequences of isolates were carried out (Fig. 1). The published sequences of 20 FMDV type A isolates recovered from different parts of the world were included in this analysis and compared with the corresponding sequence of A Iran 05 isolate. The phylogenetic tree was constructed using the Neighbour Joining method in the program Clustal X, version 1.8 [19]. The nucleotide sequence homology/divergence was calculated using the MegAlign project of DNASTar software package version 5.1 (DNASTar, USA). A description of type A FMDV isolates has been given in Table 1.

Results

The sequences for the non-structural protein coding regions (3A and 3B) were determined by PCR amplification and sequencing. A nucleotide sequence comparison based upon the alignment of complete nucleotide sequence of the 3AB region indicated that A Iran 05 had the greatest sequence similarity with four isolates (AIRN2005_WRL, Lindholm 1.3 PAK32006, PAK1, and PAK5 2006) with a nucleotide identity of approximately 98% (data not shown). Fig. 2 shows a phylogenetic tree constructed based on the sequence alignment of 21 genomes, which are distinctly divided into five lineages. The isolate A Iran 05 clustered with a Turkish and three Pakistan isolates into a separate branch from other type A isolates (lineage A). The A Iran 05 isolate was closely matched with viruses isolated in Turkey and Pakistan during 2005 ~ 2006. All the field isolates (except for isolates of lineage A) shared comparatively lesser homology (81 ~ 89%) with the Iranian isolate (A Iran 05). The topology of the phylogenetic tree indicated that the isolates airan iso105 and a21kenya iso77 differed from the Iranian subtype and could be placed in the lineages B and D, respectively (Fig. 2). All a22 type A isolates and lone isolate a28turkey iso44 were grouped closely in lineage C (Fig. 2).

The remaining eight isolates of the present study originated from different geographical areas and formed a single group (lineage E). A minimum homology of 81% in nucleotide

Table 1. Description of type A field isolates used to determine the sequence variability in the non-structural protein 3A and 3B coding regions

Serial No.	Isolate	Country	Date of collection	Accession No.
1	A Iran 05	Iran	2005	Not given
2	PAK5 2006	Pakistan	2006	EF494488
3	PAK1	Pakistan	2006	EF494487
4	Lindholm 1.3 PAK32006	Pakistan	2006	EF117837
5	AIRN2005_WRL	Turkey	2005	EF494486
6	a28turkey iso44	Turkey	1972	AY593772
7	a22turkey iso66	Turkey	1965	AY593765
8	airan iso105	Iran	1998	AY593791
9	avenceslau iso70	Brazil	1979	AY593803
10	abage iso63	Brazil	1977	AY593787
11	a21kenya iso77	Kenya	1964	AY593761
12	a14 spain iso39	Spain	1959	AY593754
13	a10holland iso82	Netherlands	1942	AY593751
14	a12valle 119 iso20	England	1932	AY593752
15	a22iraq64 iso86	Iraq	1964	AY593763
16	a22iraq70 iso92	Iraq	1970	AY593764
17	a22iraq95 iso95	Iraq	1964	AY593762
18	a26arg iso74	Argentina	1966	AY593770
19	A22 Talasskiy	Kazakhstan	1999	FJ623456
20	aphilippines iso50	Philippin	1975	AY593793
21	aparma iso55	Italy	1962	AY593792

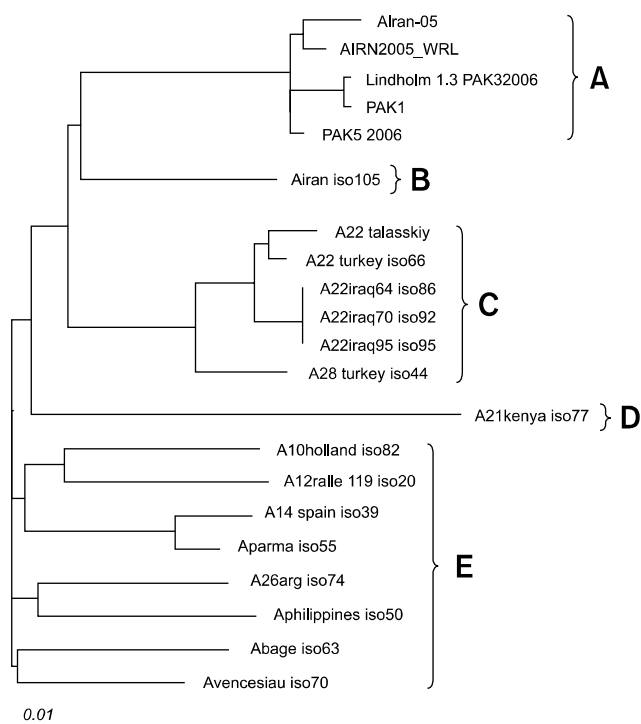


Fig. 2. Nucleotide sequence similarity tree based on a comparison of A Iran 05 isolate to other type A isolates, established with 3AB coding sequences (672 bp).

sequence was noted between the A Iran 05 isolate and the Kenyan isolate a21kenya iso77. Analysis of sequence data revealed that the Iranian isolate (airan iso105) described earlier has a low similarity to A Iran 05 with a sequence identity of 89%. Comparison of amino acid sequences identified amino acid changes in both the non-structural proteins. The number of sequence differences exhibited by each of the isolates revealed that A Iran 05 contains two amino acid substitutions at positions 47 and 119 of the 3A coding region (Fig. 1). Comparison of 3B protein among the different variants showed that in relation to the A Iran 05 the entire isolates showed a change at position 27. Apart from this, no variation leading to amino acid substitution was found in 3B protein. In 3A protein, the N-terminus (1 ~ 33) amino acids were highly conserved (Fig. 1A) while the C-terminus amino acids were highly variable. In comparison to the 3A protein, the N-terminus (1 ~ 17) amino acids of the 3B protein were highly variable (Fig. 1B). In general, the 3A and 3B coding regions were prone to amino acid alterations. As can be visualized in Fig. 1, there is obvious evidence of amino acid substitutions in lineage A, during 2005 from Iran up to the 2006 from Turkey and Pakistan. From the results highlighted in Fig. 1, it becomes evident that in relation to the A Iran 05, four other isolates of lineage A had changes at residues 3A (Gly₁₇→Ala, Ser₁₃₅→Gly and Asn₁₃₉→Glu) and 3B (Leu₅₃→Pro).

Discussion

Analysis of the viral genome sequence is of importance to monitor the field isolates in areas where the disease is endemic. Continuous co-circulation of FMDV in the field and lack of proofreading activity of the viral polymerase in the replication process results in the emergence of many genetic variants [2,5]. Within the FMDV serotypes, serotype A displays the greatest number of newly occurring subtypes [13]. The Middle East has been severely affected by separate type A (A Iran 05) epidemics, which initially emerged in Iran in 2005 and moved westwards into Turkey (including the European part of Thrace). It has continued to spread in 2006, circulating in Turkey and Iran, and has been also detected in Pakistan, Saudi Arabia and in the first months of 2009 in Iraq, Kuwait and Bahrain, and was recently identified in samples sent by Lebanon and Libya to the World Reference Laboratory [18]. This particular FMDV subtype has proven to be highly virulent and has caused severe disease in all ages of cattle. Variation in virulence could be due to the quasispecies structure of FMDV populations and the error-prone replication process. Together with this mechanism, recombination within the non-structural genome regions, potentially modifying the virulence of the virus, may be involved in the success of this subtype [13]. Previous studies have shown that both regions (3A and 3B) appear to influence the virus's pathogenic potential [14].

It is thought that virus circulation is limited to some areas of western Asia and southern Asia. Iran is bordered on the east by Pakistan and on the west by Turkey and Iraq. Permeable borders and live animal trade in west Asia are likely reasons for virus circulation. Throughout the period of virus circulation, amino acid substitutions accumulated at the various sites of protein sequence. The genetic polymorphism amongst different subtypes of type A demonstrated that in the 3A coding region of isolate A Iran 05, two nucleotides 139 and 355 were changed from T to C and G to A, respectively, resulting in exchange in amino acids from Phe₄₇→Leu and Ala₁₁₉→Thr. In addition, there is a substitution at nucleotide 79 (G to A) of 3B region, which leads to an amino acid exchange of Ala₂₇→Thr. Surprisingly, in contrast to A Iran 05 isolate the remaining twenty isolates showed no corresponding changes in the mentioned positions. This trend leads us to suggest that evolutionary events are occurring. The low conservation of the non-structural protein 3A could be due to its putative role in the adaptability of the virus to the host [12]. The phylogenetic analysis based on the 3A and 3B genes is in complete agreement with a previous report documenting that four isolates of lineage A have significant sequence similarity [13]. Nucleotide sequence analysis of the field isolates showed that virus circulation during 2005 ~ 2006 in Iran, Pakistan and Turkey may have originated in Iran.

Analysis of the amino acid sequences revealed that five isolates of lineage A had four conserved regions (Leu₁₁₄→Gln), (Ser₁₁₇→Asn), (Gly₁₁₈→Ser) and (Ala₁₄₃→Thr) in the 3A protein. The other viruses revealed no such changes. Therefore, these residues might be looked upon as signature amino acids for the group of lineage A field isolates.

Acknowledgments

This work was supported financially by the Razi Vaccine and Serum Research Institute (Iran). The authors wish to thank Miss Hasheminejad for her collaboration.

References

1. **Araújo JP Jr, Montassier HJ, Pinto AA.** Extensive antigenic and genetic variation among foot-and-mouth disease type A viruses isolated from the 1994 and 1995 foci in São Paulo, Brazil. *Vet Microbiol* 2002, **84**, 15-27.
2. **Biswas S, Sanyal A, Hemadri D, Tosh C, Mohapatra JK, Manoj Kumar R, Bandyopadhyay SK.** Sequence analysis of the non-structural 3A and 3C protein-coding regions of foot-and-mouth disease virus serotype Asia1 field isolates from an endemic country. *Vet Microbiol* 2006, **116**, 187-193.
3. **Curry S, Roqué-Rosell N, Zunszain PA, Leatherbarrow RJ.** Foot-and-mouth disease virus 3C protease: recent structural and functional insights into an antiviral target. *Int J Biochem Cell Biol* 2007, **39**, 1-6.
4. **Domingo E, Baranowski E, Escarmís C, Sobrino F.** Foot-and-mouth disease virus. *Comp Immunol Microbiol Infect Dis* 2002, **25**, 297-308.
5. **Dopazo J, Sobrino F, Palma EL, Domingo E, Moya A.** Gene encoding capsid protein VP1 of foot-and-mouth disease virus: a quasispecies model of molecular evolution. *Proc Natl Acad Sci USA* 1988, **85**, 6811-6815.
6. **Du J, Chang H, Cong G, Shao J, Lin T, Shang Y, Liu Z, Liu X, Cai X, Xie Q.** Complete nucleotide sequence of a Chinese serotype Asia1 vaccine strain of foot-and-mouth disease virus. *Virus Genes* 2007, **35**, 635-642.
7. **Feng Q, Yu H, Liu Y, He C, Hu J, Sang H, Ding N, Ding M, Fung YWW, Lau LT, Yu ACH, Chen J.** Genome comparison of a novel foot-and-mouth disease virus with other FMDV strains. *Biochem Biophys Res Commun* 2004, **323**, 254-263.
8. **Gerner W, Denyer MS, Takamatsu HH, Wileman TE, Wiesmüller KH, Pfaff E, Saalmüller A.** Identification of novel foot-and-mouth disease virus specific T-cell epitopes in c/c and d/d haplotype miniature swine. *Virus Res* 2006, **121**, 223-228.
9. **Grubman MJ.** Development of novel strategies to control foot-and-mouth disease: marker vaccines and antivirals. *Biologicals* 2005, **33**, 227-234.
10. **Hall TA.** BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 1999, **41**, 95-98.
11. **Haydon DT, Bastos AD, Knowles NJ, Samuel AR.** Evidence for Positive Selection in foot-and-mouth disease virus capsid genes from field isolates. *Genetics* 2001, **157**, 7-15.
12. **Kamstrup S, Frimann TH, Barfoed AM.** Protection of Balb/c mice against infection with FMDV by immunostimulation with CpG oligonucleotides. *Antiviral Res* 2006, **72**, 42-48.
13. **Klein J, Hussain M, Ahmad M, Normann P, Afzal M, Alexandersen S.** Genetic characterisation of the recent foot-and-mouth disease virus subtype A/IRN/2005. *Virology* 2007, **4**, 122.
14. **OIE.** Situation of Foot and Mouth Disease (FMD) in the Middle East 2008-2009. OIE, Paris, 2009.
15. **Pacheco JM, Henry TM, O'Donnell VK, Gregory JB, Mason PW.** Role of nonstructural proteins 3A and 3B in host range and pathogenicity of foot-and-mouth disease virus. *J Virol* 2003, **77**, 13017-13027.
16. **Perez AM, Thurmond MC, Grant PW, Carpenter TE.** Use of the scan statistic on disaggregated province-based data: foot-and-mouth disease in Iran. *Prev Vet Med* 2005, **71**, 197-207.
17. **Sangar DV, Bryant J, Harris TJR, Brown F, Rowlands DJ.** Removal of the Genome-Linked Protein of foot-and-mouth disease virus by rabbit reticulocyte lysate. *J Virol* 1981, **39**, 67-74.
18. **Sellers R, Gloster J.** Foot-and-mouth disease: A review of intranasal infection of cattle, sheep and pigs. *Vet J* 2008, **177**, 159-168.
19. **Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG.** The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997, **24**, 4876-4882.
20. **Wernery U, Kaaden OR.** Foot-and-mouth disease in camelids: a review. *Vet J* 2004, **168**, 134-142.