A new lineage of foot-and-mouth disease virus serotype O in India

Saravanan Subramaniam, Punam Bisht, Jajati K. Mohapatra, Aniket Sanyal & Bramhadev Pattnaik*

ICAR-Project Directorate on Foot-and-mouth disease, Mukteswar-Kumaon, Nainital-263138, Uttarakhand, India.

** Corresponding author at: ICAR-Project Directorate on Foot and Mouth Disease, Mukteswar, Nainital-263 138, Uttarakhand, India. Tel.: +91 5942 286004, e-mail: pattnaikb@gmail.com.*

Veterinaria Italiana 2015, **51** (2), 145-149. doi: 10.12834/VetIt.106.296.2
Accepted: 18.03.2015 | Available on line: 30.06.2015

Keywords
Foot and mouth disease virus,
Ind2011,
PanAsia,
Serotype O.

Summary
The complete nucleotide sequence of a new lineage of foot and mouth disease virus (FMDV) serotype O was determined. The lineage designated as Ind2011 first appeared during 2011 in the Southern region of India. Excluding the poly C tract and poly A tail, the genome of Ind2011 ranged from 8,169 to 8,172 nucleotides. Variation in the genome length was due to insertions/deletions in LF-UTR. The lineage had a higher sequence identity with lineage PanAsia-1 at P1 and P2 regions, and with lineage PanAsia-2 at P3 and L regions. Phylogenetically, the isolates were placed closely to both PanAsia-1 and 2 lineages, and appear to be a novel variant of the PanAsian lineage.

Sequenziamento nucleotidico completo di un nuovo lineage del virus dell'afta epizootica, sierotipo O, in India

Parole chiave
Afta epizootica,
Ind2011,
PanAsia,
Sierotipo O,
Virus.

Riassunto
La comunicazione descrive il sequenziamento nucleotidico completo di un nuovo lineage del virus dell'afta epizootica, sierotipo O. Ind2011 è apparso per la prima volta nel corso del 2011 nella regione meridionale dell'India. Escludendo il tratto poly C e la coda del tratto poly A, il genoma di Ind2011 varia tra 8169-8172 nucleotidi. La variabilità nella lunghezza genomica è dovuta ad inserzioni/cancellazioni in LF-UTR. Ind2011 mostra una sequenza simile al lineage PanAsia1 nelle regioni P1 e P2, e al lineage PanAsia-2 nelle regioni P3 e L. Filogeneticamente, il nuovo lineage è simile a PanAsia-1 e 2, e sembra essere una nuova variante del lineage PanAsian.

Foot-and-mouth disease virus (FMDV) belongs to the family of *Picornaviridae* and is classified within genus *Aphthovirusgenus*. The genome of FMDV consists of a single-stranded positive-sense RNA of approximately 8,500 nucleotides. The viral genome contains a 5'and 3' untranslated regions (UTRs), which are essential for viral RNA replication. The 5' UTR consists of SF-UTR and LF-UTR. The genome is translated as a single large polyprotein that is cleaved into 4 structural proteins, VP1, VP2, VP3, and VP4, and 10 nonstructural proteins, L, 2A, 2B, 2C, 3A, 3B1-3, 3C, and 3D. The genome is encapsulated in an icosahedral capsid composed of 60 copies of 4 structural proteins; VP1, VP2, and VP3 are surface exposed, while VP4 is entirely internal. (Acharya *et al.* 1989). The P1 region encodes structural proteins VP1 to VP4. The P2 (2A, 2B, 2C) and P3 (3A, 3B, 3C, and 3D) regions encode nonstructural proteins involved in viral replication. Foot-and-mouth disease viruses are categorised in 7 serotypes [A, O, C, Asia1 and Southern African Territories (SAT1-3)] on the basis of the antigenic and nucleotide differences. For each serotype, different genotypes and lineages can be identified. Three serotypes (A, O and Asia1) are currently prevalent in India. Serotype O accounts for about 85% of the FMD incidence in the country (Subramaniam *et al.* 2012). The isolates collected to date belong to Middle East-South Asia topotype. During the last 10-12 years, 2 prominent lineages have been circulating in India; PanAsia and Ind2001. A new genetic group (named Ind2011) in serotype O appeared in 2011 in India (Subramaniam *et al.* 2013). Geographically, the Ind2011 lineage was restricted to Southern regions in the states of Karnataka,

Tamilnadu, Andhra Pradesh, and Kerala. Four FMDV strains representing Ind2011 lineage were selected for genetic characterization at complete coding and non-coding region. This short communication describes the degree of genetic variations in these strains at the whole genome level with respect to other lineages circulating in India.

Total RNA was extracted using RNeasy Mini Kit (Qiagen, Hilden, Germany). Reverse transcription was performed using oligod$(T)_{15}$ primer and M-MLV Reverse transcriptase (Promega, Fitchburg, Wisconsin, USA). The complete genome was amplified in 7 overlapping fragments (SF1F-SF370R, LFIF-DHP2, L463R-NK61, DHP13-DH5, MG33-CTLV10, CTLV2-V4, 3D1081-anchored oligodT) using Pfu DNA polymerase (Fermentas, Opelstrasse, Germany). The details of the primers used in this study are specified in Table I. After the gel purification of the polymerase chain reaction (PCR) products conducted using QIA quick Gel Extraction Kit (Qiagen, Hilden, Germany), cycle sequencing reactions were performed using BigdyeV3.1 terminator kit on 3,130 genetic analyzer (Applied Biosystems, Waltham, Massachusetts, USA). Multiple sequence reads were assembled using Edit seq module of Lasergene core suite 10 (DNASTAR, Inc., Madison, Wisconsin, USA). Percent nucleotide and amino acid identities were calculated through Megalign module of Lasergene core suite 10. Phylogenetic analysis was conducted using MEGA 5.05 software (Tamura *et al*. 2011) employing the best fit nucleotide substitution model, GTR+G +I under Maximum Likelihood (ML) method.

The complete genome sequence was determined from BHK-21 cell culture adapted isolates (n = 4) of Ind2011 lineage at passage level ~2-3 (Table II). The clinical materials were collected from the suspected FMD outbreaks during the year 2011. Frank vesicular/erosive lesions were evident on tongue and feet of affected cattle. Smacking lips and excessive salivation was also observed. Nucleotide and amino acid (aa) sequences were compared with sequences of FMDV isolates available in the genetic database of Project Directorate on FMD, India and GenBank. Excluding the poly C tract and polyA tail, the length of the genome of Ind2011 isolates ranged from

Table I. *The deoxyoligonucleotide primers used in this study with location and polarity.* The details of the primers can be found in the corresponding references.

Designation	Sequence (5'→3')	Location and polarity	Reference
SF1F[a, b]	TTGAAAGGGGGCGCTAGGGTC	SFUTR1 (+)	Toja *et al.*1999
SF370R[a]	CGGTAAAACTTAGGGGGGATGAAAGGCGGGCGCCGGGTG	SFUTR370 (-)	Sanyal *et al.* 2004
SF370Rs[b]	CGGTAAAACTTAGGGGGGATG	SFUTR370 (-)	Sanyal *et al.* 2004
LF1F[a]	CCCCCCTAAGTTTTACCGTCGTTCCCG	LFUTR1 (+)	Sanyal *et al.* 2004
DHP2[a]	CCGATTCCGGTGTTGAGCAGGTG	L 325 (-)	Sabarinath 2005
DH13[b]	TGTAGACCCAGTCGAAG	L217 (-)	Sanyal *et al.* 2004
L463F[a, b]	ACCTCCRACGGGTGGTACGC	L463 (+)	George 2000
NK61[a, b]	GACATGTCCTCCTGCATCTG	2B77 (-)	Knowles & Samuel 1995
ARS4[b]	ACCAACCTCCTTGATGTGGCT	VP3 124 (+)	Knowles & Samuel 1995
DHP4[a, b]	CACAAACAAAAGATTGTGGCA	1D600 (+)	Sabarinath 2005
DHP5[a, b]	GTGTTGTACTTTCTCATTGAAAAA	3A25 (-)	Sabarinath 2005
MG33[a, b]	ATGCAACAAGATATGTTTAAGCC	2C835 (+)	George 2000
CTLV10[a]	CATCGACAATGCGAGTCTTGCC	3D547 (-)	Pattnaik *et al.* 1997
CTLV2[a, b]	TGATCTGTAGCTTGGTATCT	3D1371 (-)	Pattnaik *et al.* 1997
CTLV4[a, b]	TGACCCTGAACCACAACACG	3C 618 (+)	Pattnaik *et al.* 1997
3D1081[a, b]	GGCCAAACCATCACTCCAGCTGA	3D1081 (+)	George 2000
3D26R[b]	ACATCTCTGGTGTCAACAATCAACCCCTCGTG	3D26 (-)	Sanyal *et al.* 2004
2C540R	GGTSGARACCATYTGGGCAAARTA	2C540 (-)	This study
DHP13	GTGACTGAACTGCTTTACCGCAT	VP1 516 (+)	Sabarinath 2005
3D331R[b]	AGGCGCGGTGTCTGGCTCCAT	3D351 (-)	George 2000
NK72F[b]	GAGTCCAACCCTGGGCCCTTC	2A34 (+)	Sanyal *et al.* 2004
01C(S)237R	CATTGCTTTGCTGCCAAAGAC	1C237(-)	This study
2B325[b]	GACTCGCTCTCCAGTCTCTTT	2B325 (+)	Sanyal *et al.* 2004
3D786R[b]	GCGGAACACCTCCTCAAACAT	3D786 (-)	George 2000

a = primers used for PCR amplification of the complete genome; b = primers used for sequencing of the complete genome.

Table II. *History of the field isolates used in the study.*

Isolate ID	Place	Date of Collection	Host	Source
PD367/2011	Srirangapura, Karnataka	07-10-2011	Cow	Tongue epithelium
PD369/2011	Thammanapalli, Karnataka	07-10-2011	Cow	Tongue epithelium
PD411/2011	Kattur, Kerala	14-10-2011	Cattle	Tongue epithelium
PD422/2011	Palakkad, Kerala	18-11-2011	Cattle	Buccal mucosa

8,169 to 8,172 nucleotides. Variability in the genome length was due to insertions/deletions in LF-UTR. The ORF of all the 4 isolates encodes a polyprotein consisting of 2,332 aa and ends in stop codon TAA. The individual proteins were identical in size to that of other lineages. The level of nucleotides and the aa sequence identity varied among individual proteins. The highest degree of sequence divergence from the other lineages circulating in India was observed at P1 region (7.3-12.2%) followed by P2 (6.8-8.2%) and P3 (6.5-8.6%) regions. The Ind2011 lineage had higher sequence identity with lineage PanAsia-1 at P1 (92.5-92.7%) and P2 (92.9-93.2%) regions, and with lineage PanAsia-2 at P3 (92.3-93.5%) and L (89.4-90.5%) regions. At the level of complete ORF, this lineage had close genetic homology with lineage PanAsia-1 (92-92.3%) followed by PanAsia-2 (91.6-92%), Ind2001 lineage (91-91.4%) and vaccine strain, INDR2/1975 (89.6-90.1%). Similar trend was observed at aa level. The SF-UTR and LF-UTR of Ind2011 lineage had close sequence homology with PanAsia-2 lineage. Among the 4 strains of Ind2011 lineage, a maximum divergence of 1 and 1.1% at nucleotide and aa level respectively, was observed between PD367/2011 and PD411/2011. Though the strains PD367/2011 and PD369/2011 were collected on the same day from adjacent villages, they showed 0.5 and 0.8% divergence at nucleotide and aa level, respectively. Irrespective of the genetic differences from the vaccine strain, the Ind2011 lineage isolates tested in micro-neutralization test were found to be antigenically related to the currently used Indian vaccine strain INDR2/1975 and antigenic relationships (*r*-value) for this 4 isolates ranged from 0.49 to 0.86.

Phylogenetic trees were constructed based on the complete ORF and individual protein coding regions. In the ORF based tree (Figure 1), the Ind2011 isolates were clustered distinctly and shared branching with PanAsia-1 similar to that of VP1 coding region-based tree. But phylogenetic analysis based on individual protein coding regions yielded different results. In VP2, VP3, 2C, 3A, 3C, and LF-UTR based tree, the Ind2011 strains were positioned more proximally to PanAsia-1. Whereas in L, 2B, and 3D based phylogeny, the isolates were placed closely with PanAsia-2 than PanAsia-1. The results of phylogenetic analysis are consistent with the percentage identity observed

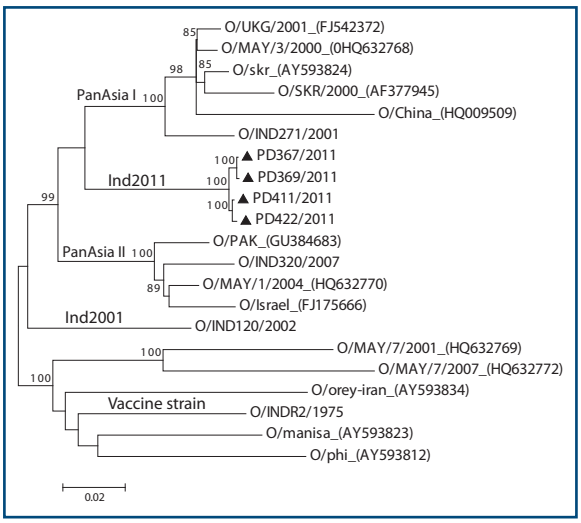

Figure 1. *Maximum Likelihood tree showing the phylogenetic relationship of FMDV serotype O isolates at complete coding region.* The numbers at each node represent the percentage bootstrap scores (10,000 replicates). Isolates sequenced in this study are marked with filled diamond.

between these strains and PanAsian lineages. Incongruent tree topologies for different genomic regions indicate possibilities of recombination; such events occur frequently in the non-structural protein coding region and less often in the structural protein coding region. Inter and intratypic recombination was shown to play an essential role in genome evolution and to the generation of FMDV genetic and population diversity (Carrillo *et al.* 2005).

The length of SF-UTR of the new lineage isolates was 371 nucleotides, and for Indian isolates of serotype O, the length varied between 348 and 397 nucleotides (Mathapati 2012). Mean length of SF-UTR of Euroasiatic isolates was reported to be 373 nucleotides (Carrillo *et al.* 2005). The secondary structures (minimum free energy and thermodynamically stable) predicted using m-fold web server (http://bioweb.pasteur.fr/seqanal/interfaces/mfold-simple.html) revealed a single stem loop structure similar to other lineages. The length of LF-UTR ranged from 707 to 710 nucleotides for Ind2011 isolates, whereas for serotype O Indian isolates, the length varied from 626 to 716 nucleotides (Mathapati 2012). All the 4 Pseudoknots (PK 1-4) could be predicted without

any junk deletion; presence of such deletions was most commonly observed to disrupt the formation of PK-I and PK-II. Secondary structure predicted here for Ind2011 isolates possessed all the 5 domains and well defined motifs. The A238AACA motif in cre (Domain I) essential for virus replication and VPg uridylylation (Nayak *et al.* 2006) was found to be fully conserved. The GNRA motif in domain 3 essential for maintenance of tertiary structure of IRES and long range RNA-RNA interaction was GTAA in the Ind2011 isolates and was found to be totally conserved. The 3'UTR excluding poly A tail was 92 nucleotides in length. The secondary structure predicted with minimum free energy showed the typical Y shaped structure observed in picornaviruses (Carrillo *et al.* 2005). Motif TCCTCAGATGT which follow stop codon TAA was fully conserved across all the isolates.

Although the catalytic triad C_{51}, H_{148} and D_{163} (Guarne *et al.*2000) of the L protein were found to be fully conserved, the residue 76 implicated in autocatalysis revealed E→K substitution in 2 isolates (PD367/2011 and PD369/2011). Residues E_{147} and G_{98} involved in hydrogen bonding with L_{200} and K_{201} were found to be conserved totally except K→R substitution at position 201 in 2 isolates (PD369/2011 and PD422/2011). Apart from this, the critical residues in the non-structural proteins and UTRs were found to be fully conserved. The antigenically critical residues and receptor-binding motif in the structural protein were also found to be conserved and identical to that of vaccine strain INDR2/1975. Lineage specific substitutions were observed in the L region at positions 14 (F/L→I) and 24 (R→Q), and in 3D region at positions 47 (N→D), 68 (E→K), 72 (A-E), and 291 (T→V).

In summary, we describe the full genome analysis of a new lineage of FMDV serotype O identified in the Southern region of India. Phylogenetic analyses based on complete genome sequences suggest that the Ind2011 lineage could be a variant of PanAsia lineage that caused a worldwide pandemic in 2001. The Ind2011 lineage, which emerged in 2011 was not detected in the subsequent years.

Acknowledgements

This work was supported by the Indian Council of Agricultural Research, New Delhi, India. We are sincerely grateful to all the scientific and technical staff associated with the network laboratories for providing clinical materials and all the required data.

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.

Carrillo C., Tulman E.R., Delhon G., Lu Z., Carreno A., Vagnozzi A., Kutish G.F. & Rock D.L. 2005. Comparative genomics of foot-and-mouth disease virus. *J Virol*, **79**, 6487-6504.

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

Guarne A., Hampoelz B., Glaser W., Carpena X., Tormo J., Fita I. & Skern T. 2000. Structural and biochemical features distinguish the foot-and-mouth disease virus leader proteinase from other papain-like enzymes. *J Mol Biol*, 302, 1227-1240.

Knowles N.J. & Samuel A.R. 1995. Polymerase chain reaction amplification and cycle sequencing of the 1D (VP1) gene of foot-and-mouth disease viruses. Report of the Session of the Research Group of the Standing Technical Committee of the European Commission for the Control of Foot and Mouth Disease, Vienna, Austria. Appendix 8, FAO, Rome, 45-53.

Mathapati B.S. 2012. Assessment of antigenic and genetic variation in serotype O Foot and mouth disease virus in India: antigenic cartography and complete genome analysis. Ph.D. Thesis, Deemed University, IVRI, Mukteswar/Izatnagar, India.

Nayak A., Goodfellow I.G., Woolaway K.E., Birtley J., Curry S. & Belsham G.J. 2006. Role of RNA structure and RNA binding activity of foot-and-mouth disease virus 3C protein in VPguridylylation and virus replication. *J Virol*, **80**, 9865-9875.

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

Sabarinath G.P. 2005. Evaluation of Foot-and-Mouth disease virus serotype O isolates for use as candidate vaccine strains. Ph.D. Thesis, Deemed University, IVRI, Mukteswar/Izatnagar, India.

Sanyal A., Mohapatra J.K., Manoj Kumar R., Biswas S., Hemadri D., Tosh C., Sabarinath G.P., Gupta S.K., Mittal M., Giridharan P. & Bandyopadhyay S.K. 2004. Complete nucleotide sequence analysis of a vaccine strain (IND491/97) and a field isolate of foot-and-mouth disease virus serotype Asia1 with an insertion in VP1 genomic region. *Acta Virol*, **48**, 65-72.

Subramaniam S., Pattnaik B., Sanyal A., Mohapatra J.K., Pawar S.S., Sharma G.K., Das B. & Dash B.B. 2012. Status of Foot-and-Mouth Disease in India. *Transbound Emerg Dis*, **60**, 197-203.

Subramaniam S., Sanyal A., Mohapatra J.K., Sharma, G.K., Biswal J.K., Ranjan R., Rout M., Das B., Bisht P., Mathapati B.S., Dash B.B. & Pattnaik B. 2013. Emergence of a novel lineage genetically divergent from the predominant Ind2001 lineage of serotype O foot-and-mouth disease virus in India. *Infect Genet Evol*, **18**, 1-7. doi: 10.1016/j.meegid.2013.04.027.

Tamura K.,Peterson D., Peterson N., Stecher G., Nei M. & Kumar S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*, **28**, 2731-2739.

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.