Site-specific substitution (Q172R) in the VP1 protein of FMDV isolates collected from asymptomatic carrier ruminants in Vietnam


1. Introduction

Foot-and-mouth disease (FMD) is a highly contagious disease of cloven-hoofed animals. FMD is endemic in much of Asia and Africa and is of significant importance to the transboundary trade in animals and products. The clinical presentation of FMD is characterized by fever, loss of appetite, lameness, and vesicular lesions in, on, or around the feet, mouth and teats (reviewed in Alexandersen and Mowat, 2005). FMD outbreaks can result in devastating economic losses owing to decreased productivity, culling and trade limitations (Knight-Jones and Rushton, 2013). There are seven distinct serotypes of FMD virus (FMDV) (family Picornaviridae, genus Aphthovirus) which may be further subdivided into topotypes, genotypes and lineages, each with varying levels of antigenic variation (reviewed in Jamal and Belsham, 2013; Brito et al., 2015). Phylogenetic analysis of one of the four
strucutral proteins of FMDV (VP1) has been extensively used to characterize viral evolution, identify epidemiological trends and trace outbreaks (Knowles and Samuel, 2003; de Carvalho Ferreira et al., 2015).

In addition to the numerous other challenges associated with controlling outbreaks of FMDV, the virus can persist in the nasopharynx of infected ruminants even years after clinical signs have disappeared (reviewed in Arzt et al., 2011; Alexandersen et al., 2002), while suids do not become long-term carriers of infectious FMDV (Stenfeldt et al., 2016a). Persistently infected FMDV carriers are defined as asymptomatic animals from which FMDV can be recovered ≥28 days after infection (OIE, 2012). The length of the carrier state is believed to be influenced by an undetermined combination of host and viral factors (Moonen and Schrijver, 2000; Gebauer et al., 1988; Salt, 1993), however the identification of these influential factors remains elusive. Circumstantial data have linked carrier cattle to outbreaks (Salt, 2004), but recent work exploring their possible role in seeding new infections in naïve animals showed that this risk is very low under controlled conditions (Parthiban et al., 2015; Tenzin et al., 2008). Currently, the African buffalo (Syncerus caffer) is the only species that has been demonstrated to initiate new FMDV infections from a persistently infected carrier animal (Dawe et al., 1994; Vosloo et al., 1996; Hedger and Condy, 1985; Bengis et al., 1986; Sutmoller et al., 2000; Thomson et al., 2003). While the role of carriers in the evolution and epidemiology of FMDV is undetermined, they are still considered a possible source of infection and this forms the basis of current FMD control policies (Salt, 2004).

Several studies have examined possible viral determinants associated with persistent infection of FMDV under experimental conditions (Parthiban et al., 2015; Horsington and Zhang, 2007; O’Donnell et al., 2014; Barros et al., 2007). One study examined approximately 850 nucleotides of the genomic region that encodes the viral structural proteins and observed a tyrosine (Y) to histidine (H) substitution in the B-C loop of VP2 in four out of six carrier cattle 28 days after experimental infection with FMDV type O UKG/34/2001 (Horsington and Zhang, 2007). Another study examined the complete genomic sequences of viruses from carrier cattle experimentally infected with the same virus (UKG/34/2001) and found the same VP2 Y to H substitution in two of the four carrier cattle and in four of the seven viral sequences obtained from these four cattle (Parthiban et al., 2015). Among the six full-length sequences of viruses from three persistently infected cattle in the latter study, no single substitution relative to the inoculum was observed in all of the viruses, leading to the conclusion that there was no evidence of viral determinants influencing persistent infection (Parthiban et al., 2015). To our knowledge, no previous studies have reported specific substitutions within the FMDV genome associated with persistent infection under field conditions.

There are different selection pressures influencing viral genomic changes in persistent infections relative to viruses obtained from clinical samples. Reports have estimated the nucleotide substitution rate per site per year (s/s/y) to be approximately 0.9 × 10−2 to 7.4 × 10−2 in the VP1-coding region of FMD viruses from carriers (Gebauer et al., 1988; Barros et al., 2007), while the rate across the entire viral genome has been estimated to be approximately 1.9 × 10−2 to 9.2 × 10−2 s/s/y (Parthiban et al., 2015; Barros et al., 2007). While the viral serotype and lineage may influence the substitution rate, in general a lower rate of 6.3 × 10−3 to 7.8 × 10−3 s/y was estimated in the VP1-coding region of FMD viruses from outbreaks (Jamal et al., 2011; Subramaniam et al., 2015) while the rates across the entire viral genome were estimated to be approximately 8 × 10−3 to 9 × 10−3 s/y (Hanada et al., 2004; Valdazo-Gonzalez et al., 2012). Additionally, the nucleotide substitution rate can vary between persistently infected animals (de Carvalho Ferreira et al., 2015; Parthiban et al., 2015; Barros et al., 2007). The sum of these data demonstrates complex and varying influences on the viral genome that may contribute to establishing and maintaining persistent subclinical infections in susceptible animals.

The aim of the current study was to investigate viral determinants of FMDV persistence in naturally infected hosts and to explore the molecular dynamics of within-host evolution of persistent FMDV with regard to time and anatomic loci. This work further supports the hypotheses that specific viral determinants are associated with FMDV persistence and that persistently infected animals may serve as a source of new outbreaks.

2. Methods

A recent study from our group described the phylogenetic relationships of VP1 sequences of FMDVs collected from carrier animals and related outbreak viruses of the serotype O ME-SA/PanAsia lineage in Vietnam between 2010 and 2013 (de Carvalho Ferreira et al., 2015). The present study included the viruses in the previously described dataset and 27 novel viral sequences obtained from either carrier animals (4 sequences, accession numbers KX869894-KX869897) or outbreak (clinically diseased) animals (23 sequences, accession numbers KX944714-KX944736, N.J. Knowles and J. Wadsworth, unpublished data) naturally infected in Vietnam and collected between 2012 and 2014. Carriers were defined as animals from which FMDV was recovered from oropharyngeal fluid and which were situated at farms at which clinical cases of FMD had not been identified within the previous or subsequent 30 days relative to sample collection (de Carvalho Ferreira et al., 2015). The new sequences from carrier animals were obtained from tissues collected at necropsy and sequenced either directly from raw material or from first-passage supernatant as previously described (de Carvalho Ferreira et al., 2015; Pacheco et al., 2010). Additional outbreak sequences were obtained from previously unpublished samples submitted to the World Reference Laboratory for FMD (Pirbright, United Kingdom). The final dataset contained 114 VP1 sequences collected from either outbreaks (n = 83) or carriers (n = 31), collected in either Vietnam (n = 112), Kazakhstan (n = 1) or China (n = 1) between 2010 and 2014.

To identify specific genetic changes associated with FMDV persistence in Asian buffalo (Bubalus bubalis) and study the within-host genetic variability, viral sequences were obtained from two persistently infected animals (V-BU 1 and V-BU 10) sampled at different time points throughout infection and again at necropsy. From V-BU 10 (a bovine), a total of eight viral sequences were obtained from clinical samples. Reports have estimated the nucleotide substitution rate per site per year (s/s/y) to be approximately 0.9 × 10−2 to 7.4 × 10−2 in the VP1-coding region of FMD viruses from carriers (Gebauer et al., 1988; Barros et al., 2007), while the rate across the entire viral genome has been estimated to be approximately 1.9 × 10−2 to 9.2 × 10−2 s/s/y (Parthiban et al., 2015; Barros et al., 2007). While the viral serotype and lineage may influence the substitution rate, in general a lower rate of 6.3 × 10−3 to 7.8 × 10−3 s/y was estimated in the VP1-coding region of FMD viruses from outbreaks (Jamal et al., 2011; Subramaniam et al., 2015) while the rates across the entire viral genome were estimated to be approximately 8 × 10−3 to 9 × 10−3 s/y (Hanada et al., 2004; Valdazo-Gonzalez et al., 2012). Additionally, the nucleotide substitution rate can vary between persistently infected animals (de Carvalho Ferreira et al., 2015; Parthiban et al., 2015; Barros et al., 2007). The sum of these data demonstrates complex and varying influences on the viral genome that may contribute to establishing and maintaining persistent subclinical infections in susceptible animals.

The aim of the current study was to investigate viral determinants of FMDV persistence in naturally infected hosts and to explore the molecular dynamics of within-host evolution of persistent FMDV with regard to time and anatomic loci. This work further supports the hypotheses that specific viral determinants are associated with FMDV persistence and that persistently infected animals may serve as a source of new outbreaks.

2. Methods

A recent study from our group described the phylogenetic relationships of VP1 sequences of FMDVs collected from carrier animals and related outbreak viruses of the serotype O ME-SA/PanAsia lineage in Vietnam between 2010 and 2013 (de Carvalho Ferreira et al., 2015). The present study included the viruses in the previously described dataset and 27 novel viral sequences obtained from either carrier animals (4 sequences, accession numbers KX869894-KX869897) or outbreak (clinically diseased) animals (23 sequences, accession numbers KX944714-KX944736, N.J. Knowles and J. Wadsworth, unpublished data) naturally infected in Vietnam and collected between 2012 and 2014. Carriers were defined as animals from which FMDV was recovered from oropharyngeal fluid and which were situated at farms at which clinical cases of FMD had not been identified within the previous or subsequent 30 days relative to sample collection (de Carvalho Ferreira et al., 2015). The new sequences from carrier animals were obtained from tissues collected at necropsy and sequenced either directly from raw material or from first-passage supernatant as previously described (de Carvalho Ferreira et al., 2015; Pacheco et al., 2010). Additional outbreak sequences were obtained from previously unpublished samples submitted to the World Reference Laboratory for FMD (Pirbright, United Kingdom). The final dataset contained 114 VP1 sequences collected from either outbreaks (n = 83) or carriers (n = 31), collected in either Vietnam (n = 112), Kazakhstan (n = 1) or China (n = 1) between 2010 and 2014.

To identify specific genetic changes associated with FMDV persistence in Asian buffalo (Bubalus bubalis) and study the within-host genetic variability, viral sequences were obtained from two persistently infected animals (V-BU 1 and V-BU 10) sampled at different time points throughout infection and again at necropsy. From V-BU 10 (a bovine), a total of eight viral sequences were examined: five from oropharyngeal fluid samples collected by probang cup between April and November 2012 (the last collected on the day of necropsy) and three from tissue samples of the larynx, dorsal soft palate (DSP) and dorsal nasopharynx (DNP),
collected at necropsy in November 2012. From V-BU 1 (a buffalo), three viral sequences were examined: two probang samples collected between May and November 2012 and one tissue sample (from the larynx) collected at necropsy in November 2012.

Sequence alignments (ClustalW) and phylogenetic analysis were conducted using MEGA v6 (Tamura et al., 2013). An analysis of natural selection was conducted in HyPhy (Pond et al., 2005) and MEGA v6, using the joint Maximum Likelihood reconstructions of ancestral states under the Muse-Gaut model (Muse and Gaut, 1994) of codon substitution and the Felsenstein 1981 model (Felsenstein, 1981) of nucleotide substitution. To detect codons that have undergone positive selection, the test statistic dN-dS was calculated for each of the 211 codons in the FMDV VP1 dataset, where dS is the number of synonymous substitutions per synonymous site (s/S) and dN is the number of nonsynonymous substitutions per nonsynonymous site (n/N). A mid-point rooted phylogenetic tree (log likelihood −3390.0360) was inferred using the Maximum Likelihood algorithm based on the GTR + G + I model. These alignments were examined and further analysis focused on sites of consistent amino acid variations. To each individual VP1 sequence, we associated metadata containing (when available): predicted amino acid at persistent infection-associated VP1 sites, clinical or subclinical FMDV status, host species, and time (year) and place (province) of sample collection. This metadata-associated phylogenetic alignment was implemented and analyzed with the CLC genomics workbench 8.5.1 (Qiagen) to visualize possible associations of these factors with the genetic relationship of the viruses.

3. Results and discussion

Amino acid changes in the VP1 protein that were present in 10 or more of the 114 VP1 sequences analyzed are presented in Table 1. Among substitutions that occurred 10 or more times, only a single specific substitution occurred in all 31 viral sequences from subclinical carrier animals. These 31 viral sequences all had an arginine (R) at VP1 residue 172, whereas all temporally antecedent outbreak-derived viruses had a glutamine (Q) at this site. Given that 18 of these 31 sequences were derived directly from original material, this substitution was not an artefact of passage in cell culture. Furthermore, we analyzed the 211 codons in the VP1 dataset for evidence of positive selection using the test statistic dN-dS where a positive value indicates an overabundance of nonsynonymous substitutions (i.e. positive selection). The codon for residue 172 had the highest dN-dS value (3.15) indicating this position is under the highest amount of positive selection force although the value was not statistically significant (P-value = 0.14). A phylogenetic tree of the 114 VP1 sequences is shown in Fig. 1. Four clades are indicated on this tree: i) 2013-2014 outbreak/arginine clade, which contains 24 viral sequences collected from outbreaks between 2013 and 2014 which all possess an arginine (R) at VP1 residue 172; ii) persistently infected individual cattle V-BU 10 sequences (n = 8) (see methods for details); iii) persistently infected individual buffalo V-BU 1 sequences (n = 3) (see methods for details) and; iv) 2011-2014 porcine-derived clade, which contains 18 viral sequences collected from outbreaks between 2011 and 2014, 17 of which were from pigs and possess either a glutamine (Q) or proline (P) at VP1 residue 172.

To further analyze the Q172R substitution present in all 31 viral sequences obtained from persistently infected animals, a metadata-associated cladogram was constructed which allowed simultaneous visualization of the distinct attributes: sample type (outbreak or carrier), VP1 residue 172 predicted amino acid, year, host species, and location (Fig. 2). Unlike phylograms, cladograms only show branching order and not branch length, however with the associated metadata they present a robust platform to further analyze the phylodynamic properties of this dataset. The 31 subclinal isolates all possessed the Q172R substitution, regardless of their host species (cattle or buffalo), province of origin (Son La, Lang Son or Long An) or relative genetic distances (Figs. 1 and 2, Table 1). The presence of this Q172R substitution across all carrier isolates is more noteworthy considering their relatively high sequence heterogeneity within a single lineage (range: 0.0%-5.6%; mean: 3.1%) relative to the entire data set (range: 0.0%-8.4%; mean: 2.9%) and their disparate placements across the corresponding phylogenetic trees (Figs. 1 and 2). Other substitutions were found with lower frequency, and no other substitution was present in a majority of the carrier-derived isolates (Table 1).

The monophyletic group referred to as the 2013–2014 outbreak/arginine clade (Figs. 1 and 2) shares a recent ancestor with a carrier (probang) isolate from a persistently infected bovine in Long An Province in 2012 (O/VIT/362/2012_pro, accession #KT153143). The 24 outbreak isolates comprising the 2013–2014 outbreak/arginine clade were collected after that carrier isolate and all possess the Q172R substitution present in all carrier isolates in the dataset. Another isolate (O/VIT/413/2012_pro, accession #KX763147) occurred early in the 2013/2014 outbreak outbreak/arginine clade and all except one isolate (O/VIT/362/2012_pro) were collected after this isolate.

Table 1

<table>
<thead>
<tr>
<th>Position</th>
<th>Change</th>
<th>Change/Total</th>
<th>Change/Outbreak</th>
<th>Change/Carrier</th>
<th>Change/Porcine</th>
<th>Change/Cattle</th>
<th>Change/Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>D-E</td>
<td>10/114</td>
<td>2/83</td>
<td>8/31</td>
<td>2/32</td>
<td>8/60</td>
<td>0/18</td>
</tr>
<tr>
<td>25</td>
<td>Q-R</td>
<td>50/114</td>
<td>46/83</td>
<td>4/31</td>
<td>26/32</td>
<td>9/60</td>
<td>14/18</td>
</tr>
<tr>
<td>133</td>
<td>N-D</td>
<td>15/114</td>
<td>15/83</td>
<td>0/31</td>
<td>14/32</td>
<td>1/60</td>
<td>0/18</td>
</tr>
<tr>
<td>135</td>
<td>R-K</td>
<td>11/114</td>
<td>4/83</td>
<td>7/31</td>
<td>0/32</td>
<td>10/60</td>
<td>1/18</td>
</tr>
<tr>
<td>141</td>
<td>V-A</td>
<td>14/114</td>
<td>1/83</td>
<td>13/31</td>
<td>0/32</td>
<td>12/60</td>
<td>2/18</td>
</tr>
<tr>
<td>152</td>
<td>A-T</td>
<td>11/114</td>
<td>11/83</td>
<td>0/31</td>
<td>11/32</td>
<td>0/60</td>
<td>0/18</td>
</tr>
<tr>
<td>153</td>
<td>Q-P</td>
<td>13/114</td>
<td>13/83</td>
<td>0/31</td>
<td>13/32</td>
<td>0/60</td>
<td>0/18</td>
</tr>
<tr>
<td>172</td>
<td>Q-R</td>
<td>60/114</td>
<td>29/83</td>
<td>31/31</td>
<td>4/32</td>
<td>48/60</td>
<td>7/18</td>
</tr>
<tr>
<td>201</td>
<td>H-R</td>
<td>10/114</td>
<td>10/83</td>
<td>0/31</td>
<td>10/32</td>
<td>0/60</td>
<td>0/18</td>
</tr>
</tbody>
</table>

- The codon with the highest dN-dS estimate and lowest P-value (3.15 and 0.14, respectively).
- All of these outbreak sequences were collected in 2013 or 2014.
Fig. 1. Mid-point rooted, maximum-likelihood phylogenetic tree of 114 FMDV O/ME-SA/PanAsia VP1 sequences. Standard nomenclature for the sequences is used, including whenever available: serotype, country of isolation, isolate number, year and Vietnamese province of isolation, as well as host species and GenBank accession number. In addition, the deduced amino acid residue for VP1 position 172 is shown in parenthesis before the accession number using the standard one-letter amino acid code. Four subtrees or clades are demarcated: i) 2013–2014 outbreak/arginine clade; ii) animal V-BU 10; iii) animal V-BU 1 and iv) 2011–2014 porcine-derived clade. For the serial probang samples taken from V-BU 1 and V-BU 10, the month of collection is given as the standard three letter abbreviation. An asterisk (*) denotes a clade described in the text.

#KT153152) from a different persistently infected bovine in Long An Province forms a monophyletic group with outbreak isolates that were collected after the carrier isolate and also possess the Q172R substitution (Fig. 1). Collectively, these findings suggest...
that persistently infected animals may play a role in the ecology and evolution of FMDV in Vietnam. However, there are limitations intrinsic to analyses of field samples that make definitively establishing the directionality of infection all but impossible. Potential confounding factors that could not be accounted for include sampling bias, lack of access to spatially and temporally co-circulating outbreak viruses, animal movement and virus introduction from outside sources. Additionally, it is noteworthy that arginine at VP1 codon 172 has previously been described in other outbreak viruses (Jamal et al., 2011; Muthuchelvan et al., 2001).

To analyze the FMDV evolutionary dynamics within naturally infected carrier animals, we identified and characterized specific nucleotide substitutions in the viral genome at different time points and anatomic sites (Fig. 1 and Table 2). Successive samples of oropharyngeal fluid (probang) were compared with mucosal tissues collected at necropsy from mucosal sites previously shown to maintain persistent FMDV infection (Burrows, 1966; Pacheco et al., 2015; Stenfeldt et al., 2016b). Viral sequences from probangs taken from a persistently infected buffalo (V-BU 1) in Lang Son Province in May and November 2012 were compared to a viral sequence retrieved from laryngeal tissue of the same animal collected at necropsy in November 2012, identifying the larynx as a site of FMDV persistence in Asian buffalo (Fig. 1). At the nucleotide level, the larynx-derived sequence (O/VIT/2012/Buff(V-BU1)/Phar(A)) was identical to the probang sequence collected on the same day of the necropsy and all three formed a
monophyletic group. Additionally, five viral sequences from probangs taken from a persistently infected bovine (V-BU 10) in Son La Province in April, May, July, October and November 2012 were compared to viral sequences retrieved from three tissues (larynx, dorsal soft palate, and dorsal nasopharynx) which were collected at necropsy from this animal in November 2012 (Fig. 1 and Table 2). The viral sequence obtained from the dorsal soft palate (O/VIT/2012/Catt(V-BU10)/Phar(A)) was identical to the viral sequences from two probangs collected one month prior and at the time of necropsy, while the viral sequence obtained from the dorsal nasopharynx (O/VIT/2012/Catt(V-BU10)/Phar(B)) was highly similar to all three sequences, differing by only two synonymous substitutions. These eight sequences, separated by up to seven months and from different anatomic locations, differ by only a single amino acid in the usually highly variable VP1 protein. Even more remarkable is that all of the samples from this animal collected for 6 months prior to necropsy were identical at the protein level. Further analysis, including complete capsid or full-length sequence analysis of these isolates, is necessary to determine what evolutionary forces (such as codon usage bias, nucleotide changes in other viral proteins, recombination or a decrease in viral replication) may influence the FMDV genetic changes in a persistently infected animal.

In a recent experimental study, three arginines (R) and three glutamines (Q) were predicted at VP1 residue 172 from six full-length genomes of FMD viruses collected during the early phase of the carrier state (Parthiban et al., 2015) while the inoculum used in this controlled study coded for an arginine at this VP1 residue (Accession number KR265075). In another experimental study that analyzed FMDV VP1 sequences from Asian (water) buffalo persistently infected with O1Campos, the authors found a Q172R substitution at least once in three of the five persistently infected animals and in five of the 19 VP1 sequences analyzed (Barros et al., 2007). Additionally, viruses causing outbreaks in Vietnam from 2002 (Accession number DQ165026) and 2005 (Accession number HQ116283) both code for an R at VP1 residue 172. Thus it seems likely that this residue by itself or in combination with additional substitutions in the viral capsid outside of VP1 may be involved in virus-host co-adaptation during the carrier phase. This theory is supported by the findings of Borley et al., who predict that this residue is part of a conformational epitope (Borley et al., 2013). Overall, these data suggest that there may be viral determinants for establishing and/or maintaining persistent infection and that these animals may play a role in the ecology and evolution of FMDV. To the authors’ knowledge, this is the first report of a distinct anatomic site of FMDV persistence in Asian buffalo and FMDV within-host evolution with regard to time and anatomic site in a naturally infected carrier animal. Overall, our analyses provide novel insights into the dynamics of persistent FMDV infection under natural conditions.

Acknowledgements

This research was funded in part by ARS-CRIS Project 1940-32000-057-00D and through an interagency agreement with the Science and Technology Directorate of the U.S. Department of Homeland Security under Award Number HSHQDC-12-X-0060. Additional funding was provided by the U.S. Department of State, Biosecurity Engagement Program through the USDA, ARS Office of International Research Programs (agreement# 58-1940-0-070F) and the Cooperative Biological Engagement Program of the U.S. Department of Defense, Defense Threat Reduction Agency (interagency agreement# 7 U.S.C. 3318(b)-13113). Michael Eschbaumer, Barbara Brito and Helena C. de Carvalho are the recipients of a Plum Island Animal Disease Center Research Participation Program fellowship, administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and the U.S. Department of Agriculture (USDA). All opinions expressed in this paper are the authors’ and do not necessarily reflect the policies and views of the USDA-ARS FADRU, Vietnam DAH or ORAU/ORISE. The Vietnam DAH and USDA-ARS FADRU are members of the Global Foot-and-Mouth Disease Research Alliance (GFRA). The authors would like to thank Nick Knowles and Jemma Wadsworth for kindly providing sequences, Ethan Hartwig and George Smoliga for their technical support and Penny Rempe and Paula Disabella for administrative assistance.
References


