Veterinary Microbiology 177 (2015) 106–122

Contents lists available at ScienceDirect

Veterinary Microbiology

journal homepage: www.elsevier.com/locate/vetmic

Determination of common genetic variants within the non-structural proteins of foot-and-mouth disease viruses isolated in sub-Saharan Africa

P. Nsamba [a,b], T.A.P. de Beer [c,1], M. Chitray [a], K. Scott [a], W. Vosloo [b,d], F.F. Maree [a,e,*]

[a] Agricultural Research Council, Onderstepoort Veterinary Institute, Transboundary Animal Diseases Programme, Private Bag X05, Onderstepoort 0110, South Africa
[b] Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria, Private Bag X04, Onderstepoort 0110, South Africa
[c] European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, United Kingdom
[d] CSIRO-Australian Animal Health Laboratory, Private Bag 24, Geelong 3220, VIC, Australia
[e] Department of Microbiology and Plant Pathology, Faculty of Agricultural and Natural Sciences, University of Pretoria, Pretoria 0002, South Africa

ARTICLE INFO

Article history:
Received 10 April 2014
Received in revised form 3 March 2015
Accepted 4 March 2015

Keywords:
Foot-and-mouth disease virus
Non-structural proteins
Leader-protease, 2C/AAA+ ATPase
3A-protein
3B/Vpg-protein
3C-protease, RNA-dependent RNA polymerase

ABSTRACT

The non-structural proteins of foot-and-mouth disease virus (FMDV) are responsible for RNA replication, proteolytic processing of the viral polyprotein precursor, folding and assembly of the structural proteins and modification of the cellular translation apparatus. Investigation of the amino acid heterogeneity of the non-structural proteins of seventy-nine FMDV isolates of SAT1, SAT2, SAT3, A and O serotypes revealed between 29 and 62% amino acid variability. The Leader protease (L^{pro}) and 3A proteins were the most variable whilst the RNA-dependent RNA polymerase ($3D^{pol}$) the most conserved. Phylogeny based on the non-structural protein-coding regions showed separate clusters for southern African viruses for both the L^{pro} and 3C protease ($3C^{pro}$) and sequences unique to this group of viruses, e.g. in the 2C and $3C^{pro}$ proteins. These groupings were unlike serotype groupings based on structural protein-coding regions. The amino acid substitutions and the nature of the naturally occurring substitutions provide insight into the functional domains and regions of the non-structural proteins that are critical for structure–function. The L^{pro} of southern African SAT type isolates differed from A, O and SAT isolates in northern Africa, particularly in the auto-processing region. Three-dimensional structures of the 3C protease ($3C^{pro}$) and $3D^{pol}$ showed that the observed variation does not affect the enzymatic active sites or substrate binding sites. Variation in the $3C^{pro}$ cleavage sites demonstrates broad substrate specificity.

1. Introduction

Foot-and-mouth disease (FMD) is widely considered the most economically important disease of livestock, and is a notifiable disease in many countries due to its highly contagious nature and associated productivity losses among even-toed ungulates (*Artiodactyla*). The disease is

* Corresponding author at: Transboundary Animal Diseases Programme, Onderstepoort Veterinary Institute, Private Bag X05, Onderstepoort 0110, South Africa. Tel.: +27 12 529 9584/93; fax: +27 12 529 9595/05.
E-mail address: mareef@arc.agric.za (F.F. Maree).
[1] Current address: Biozentrum & SIB Swiss Institute of Bioinformatics, University of Basel, Klingelbergstrasse 50/70, 4056 Basel, Switzerland.

http://dx.doi.org/10.1016/j.vetmic.2015.03.007
0378-1135/

endemic to large parts of the African continent and an impediment to lucrative export markets for animal products (Vosloo et al., 2002). The different serotypes of FMD virus (FMDV) cause a clinically indistinguishable vesicular disease in cloven-hoofed animals and display different geographical distributions and epidemiology (Bastos et al., 2001, 2003a,b; Bronsvoort et al., 2004; Knowles and Samuel, 2003; Samuel and Knowles, 2001). Of the seven serotypes, the South African Territories (SAT) types 1, 2 and 3 are confined to sub-Saharan Africa, although incursions into the Middle East by SAT1 (1961–1965 and 1970) and SAT2 (1990, 2000 and currently in North Africa) viruses have been recorded (Bastos et al., 2001; Ferris and Donaldson, 1992; Valdazo-González et al., 2012; records of the Office International des Épizooties or OIE). In contrast, serotypes A and O occur globally (Samuel and Knowles, 2001) with the exception of southern Africa (Vosloo et al., 2002).

FMDV is a small non-enveloped virus, a member of the *Aphthovirus* genus within the family *Picornaviridae*. The icosahedral capsid consists of 60 copies of four structural proteins, VP1 to 4, arranged in a pseudo $T = 3$ composition. The three surface-exposed proteins, VP1 (1D), VP2 (1B) and VP3 (1C), assemble into a protomeric subunit, with the smaller VP4 (1A) located internally (Acharya et al., 1989; Curry et al., 1995; Sobrino et al., 2001). Despite the high levels of genetic and antigenic variation (Vosloo et al., 1995; Reeve et al., 2010; Maree et al., 2011), a consequence of the high mutation rate of the virus, the structural arrangement of the capsid is remarkably conserved, indicating plasticity within the three-dimensional structure of the capsid proteins (Acharya et al., 1989; Curry et al., 1995; Fry et al., 1999; Lea et al., 1994). The capsid encloses a ca. 8.5 kilobase, positive-sense, single-stranded RNA genome with a single open reading frame and two in-frame translation-initiation codons. Covalently linked to the 5′ end of the genome is the small viral protein 3B (or VPg), while the 3′ end is poly-adenylated (Carrillo et al., 2005). Upon virus infection, the interactions between VP2 and VP3 at the pentameric interfaces are disrupted by acidification within cellular endosomes, thereby releasing the viral RNA (Ellard et al., 1999; Knipe et al., 1997). The viral genome is rapidly translated into a polyprotein which is co- and post-translationally cleaved by viral proteinases into several partially cleaved intermediates and ultimately into 12 mature proteins (Pereira, 1981; Rueckert, 1996).

In addition to the capsid proteins, the ORF of the viral RNA genome encodes eight non-structural proteins, each with its unique function within the replication cycle of FMDV (Belsham, 1993, 2005). The non-structural proteins include three proteases, i.e. L^{pro}, 2A and $3C^{pro}$, responsible for cleavage of the viral polyprotein and shut-down of the host cap-dependent translational system (Bablanian and Grubman, 1993; Martinez-Salas et al., 1996). Although several of the picornavirus proteins involved in RNA replication (2B, 2C and 3A) have membrane binding properties and disrupt protein trafficking in the cell (Moffat et al., 2005, 2007), their particular functions during viral replication are still unknown. The 2B protein has been implicated in virus-induced cytopathic effect (CPE) (van Kuppeveld et al., 1997), while the 2C protein has recently been classified as an AAA+ ATPase enzyme that may act as an RNA helicase (Sweeney et al., 2010). The 3D gene encodes the viral RNA-dependent RNA polymerase (RDRP), and together with the 3A co-localizes with ER membrane-associated replication complexes (Lama et al., 1994).

Based on the genetic variability of the VP1-coding region, the FMDV strains that exist among the serotypes, group into topotypes that are geographically specific (Knowles and Samuel, 2003; Samuel and Knowles, 2001). Serotype A has three topotypes, of which the Africa topotype is endemic to sub-Saharan Africa. Of the eleven topotypes defined for serotype O, five are endemic in Africa; the East Africa (EA1–EA4) and West Africa (WA) topotypes (Di Nardo et al., 2011). On the other hand, the SAT serotypes are genetically more diverse, and nine, fourteen and five topotypes have been defined for SAT1, SAT2 and SAT3 respectively (Bastos et al., 2001, 2003a,b; Knowles et al., 2010).

A few studies have looked at genome comparisons mainly focusing on serotypes A, O, C and Asia-1 with a geographic distribution in Euro-Asia and South America (Carrillo et al., 2005; Mason et al., 2003; Pereda et al., 2002). However, a limited number of complete non-structural protein analyses for viruses belonging to the SAT serotypes have been described (Carrillo et al., 2005). Here we have compared the non-structural proteins and their coding regions for the three SAT serotypes and viruses from serotype A and O found in sub-Saharan Africa from 1974 to 2006. Additionally, the natural variation found within the non-structural protein sequences of 79 viruses to identify structurally and possibly functionally conserved regions were mapped. Variation in the $3C^{pro}$ and $3D^{pol}$ was also mapped onto the protein structures to improve understanding of the plasticity of these enzymes. The deduced amino acid sequences of the non-structural proteins of two closely related SAT2 viruses, causing numerous of outbreaks in North Africa and the Middle East in 2012, were also included (Valdazo-González et al., 2012).

2. Materials and methods

2.1. Virus isolates

The viruses included in this study were either supplied by the World Reference Laboratory (WRL) for FMD at the Pirbright Institute (United Kingdom) or form part of the virus collection at the Transboundary Animal Diseases Programme, Onderstepoort (South Africa). The SAT1 ($n = 30$), SAT2 ($n = 26$), SAT3 ($n = 7$), serotype A ($n = 7$) and serotype O ($n = 10$) FMDV isolates from 17 countries in sub-Saharan Africa were selected for genomic characterization. The isolates span a 32 year time period and represent various geographic locations and animal species. The viruses were propagated in IB-RS-2 cells prior to RNA extraction, cDNA synthesis and amplification of the relevant genome regions by PCR (Maree et al., 2011). A description of the passage histories, host species and representative topotypes can be found in Table 1.