Virus Research 145 (2009) 187–199

Contents lists available at ScienceDirect

Virus Research

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

Genetic and phylogenetic analysis of the core proteins VP1, VP3, VP4, VP6 and VP7 of epizootic haemorrhagic disease virus (EHDV)

S.J. Anthony[a,b,*], N. Maan[a], S. Maan[a], G. Sutton[b], H. Attoui[a], P.P.C. Mertens[a]

[a] Vector-borne Diseases Program, Institute for Animal Health, Ash Road, Pirbright, Surrey GU24 0NF, United Kingdom
[b] Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom

ARTICLE INFO

Article history:
Received 9 March 2009
Received in revised form 15 July 2009
Accepted 16 July 2009
Available online 24 July 2009

Keywords:
EHDV
Core proteins
VP1
VP3
VP4
VP6
VP7
East/west split
Electropherotype

ABSTRACT

The core proteins of *epizootic haemorrhagic disease virus* (EHDV) have important roles to perform in maintaining the structure and function of the virus. A complete genetic and phylogenetic analysis was therefore performed on these proteins (and the genes that code for them) to allow comparison of the selective pressures acting on each. Accession numbers, gene and protein sizes, ORF positions, G + C contents, terminal hexanucleotides, start and stop codons and phylogenetic relationships are all presented. The inner core proteins (VP1, VP3, VP4 and VP6) were characterised by high levels of sequence conservation, and the ability to topotype isolates very strongly into eastern or western groups. This is particularly evident in genome segment 9 (VP6) which exists as two different sized homologues. VP7 did not topotype, but rather exhibited a more random, radial phylogeny suggestive of genetic drift. With the exception of VP6, all of the core proteins also showed high numbers of synonymous mutations in the third base position, suggesting they have been evolving for a long period of time. Interestingly, VP6 did not show this, and possible reasons for this are discussed.

1. Introduction

Epizootic haemorrhagic disease (EHD) is an infectious, non-contagious insect transmitted disease of ruminants caused by viruses belonging to the species *Epizootic haemorrhagic disease virus* (EHDV). Along with *Bluetongue virus* (BTV), EHDV is classified within the genus *Orbivirus* of the family *Reoviridae* (Mertens et al., 2005). It was first formally identified in 1955 by Richard Shope and colleagues in New Jersey (USA) following reports of a fatal epizootic in white-tailed deer (*Odocoileus virginianus*) (Shope et al., 1955).

EHDV and BTV are both icosahedral viruses, ~80 nm in diameter, with 10-segmented, double-stranded RNA genomes (Huismans and Van Dijk, 1990; Mertens, 1999; Verwoerd et al., 1972; Verwoerd et al., 1970). Each of these 10 segments codes for 1 of 10 distinct viral proteins, 7 of which are structural components of the virus particle, and three of which are non-structural (Huismans et al., 1979; Mecham and Dean, 1988; Mertens et al., 1984). As demonstrated for BTV, the seven structural proteins form the viral capsid and are arranged as three concentric capsid shells surrounding the viral dsRNA (Grimes et al., 1998; Hewat et al., 1992a,b; Huismans and Van Dijk, 1990; Roy, 1992; Verwoerd et al., 1972).

The viral core is made up of VP1, VP3, VP4, VP6 and VP7, encoded by genome segments 1, 3, 4, 9 and 7, respectively (Huismans et al., 1979; Mecham and Dean, 1988). VP3 forms the innermost subcore capsid shell, and its self-assembly controls the overall size and organization of the capsid structure (Grimes et al., 1998; Mertens and Diprose, 2004). VP3 is also an RNA binding protein and interacts with both the viral RNA genome and with the minor proteins VP1, VP4 and VP6 (Mertens et al., 1992; Ramadevi et al., 1998; Ramadevi and Roy, 1998). VP7 is the main immuno-dominant viral protein (Gumm and Newman, 1982; Huismans and Erasmus, 1981; Thevasagayam et al., 1996), can bind dsRNA and can be involved in cell entry, particularly insect cells for which cores have a high specific infectivity—as demonstrated for BTV (Diprose et al., 2002; Mertens et al., 1996; Roy et al., 1990). VP7 is antigenically specific to each of the *Orbivirus* species, and therefore represents a suitable target for the design of group-specific diagnostic assays (Gumm and Newman, 1982; Luo and Sabara, 2005; Mecham and Jochim, 2000; White et al., 1991).

Within the orbiviral core, the 10 dsRNA segments of the viral genome form a highly ordered, liquid crystal structure. This gives each genome segment considerable freedom of movement within the central space, allowing them to act as templates for ssRNA synthesis by the viral transcriptase complexes (TC—composed of VP1,

* Corresponding author. Current address: San Diego Zoo: Institute for Conservation Research. 15600 San Pasqual Valley Road, Escondido, CA 92027, USA.
E-mail address: santhony@sandiegozoo.org (S.J. Anthony).

0168-1702/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2009.07.011

VP4 and VP6). The TCs are situated on the inner surface of the sub-core shell at each of the 12 icosahedral vertices, as demonstrated for cypovirus and rotavirus (Miura et al., 1969; Prasad et al., 1988) and studies with other reoviruses, including cytoplasmic polyhedrosis viruses (genus *Cypovirus*) have suggested that each TC may be associated with a separate genome segment (Yazakia and Miura, 1980); perhaps explaining why no member of the *Reoviridae* contains more than 12 RNA segments (Mertens, 2004). Within this TC, VP1 functions as the viral RNA-dependant RNA polymerase (RdRP), and is highly conserved (Boyce et al., 2004; Urakawa et al., 1989; Wehrfritz et al., 2007). VP4 is the capping enzyme (type 1), which has guanylyltransferase and two transmethylase activities (Mertens et al., 1992; Mertens and Diprose, 2004; Ramadevi et al., 1998; Ramadevi and Roy, 1998) and VP6 is believed to be the viral helicase, as demonstrated in both BTV and African horse sickness virus (de Waal and Huismans, 2005; Kar and Roy, 2003).

It is significant that our knowledge of the structure and function of the EHD viral core is largely based on assumption for what we know to be true for BTV, or other related reoviruses. For example the high conservation in VP1, or the ability of VP3 to topotype virus isolates are well documented for BTV—but there is only very limited data currently for EHDV. A full-genome database was recently completed for 11 different strains of EHDV (Anthony et al., submitted for publication-a, submitted for publication-b). The genetic and phylogenetic analyses of the core proteins of EHDV are described in this report.

2. Materials and methods

2.1. Viruses

Eleven strains of EHDV were used in these studies, covering strains from both the eastern and western hemispheres. Western viruses: USA1955/01; CAN1962/01; NIG1967/01; NIG1968/01; BAR1983/01. Eastern viruses: AUS1979/05; AUS1977/01; AUS1981/07; AUS1981/06; AUS1982/06; JAP1959/01 (Ditchfield et al., 1964; Lee, 1979; Mohammed et al., 1996; Omori et al., 1969; Shope et al., 1960; St. George et al., 1983). When used in this manuscript, the term 'Eastern' virus refers to one that has been isolated in Australia or Japan. 'Western' viruses refer to those isolated in the Middle-East, Africa and North America. The strains included here are used as EHDV diagnostic reference strains by the EU Community Reference Centre and OIE Reference Laboratory for Bluetongue, in Pirbright (United Kingdom); vary serologically, temporally and geographically. All of the viruses were obtained from the dsRNA virus collection at the Institute for Animal Health (IAH) in Pirbright and are described in Table 1. All viruses were cultivated in BHK-21 cell monolayers (175 cm^3) exhibiting ~90% confluence, until complete cytopathic effect (CPE) was observed.

2.2. Full-length amplification of cDNAs (FLAC) and sequencing

Double-stranded RNA was extracted from an infected monolayer showing 100% CPE, as described by Attoui et al. (2000). Full-length cDNA copies of EHDV genome segments 2 and 6 were synthesised and sequenced as described by Maan et al. (2007b).

2.3. Analysis of sequence data

Sequence data was analysed and aligned using BioEdit Sequence Alignment Editor (version 6.0) Mega (version 3.1) (Tamura et al., 2007), Clustal W and Clustal X alignment programs (Higgins and Sharp, 1988; Thompson et al., 1997; Thompson et al., 1994).

Phylogenetic comparisons were made using Neighbour-Joining (NJ), Maximum Likelihood (ML) and Bayesian algorithms. NJ trees were constructed in Mega 3.1 using full-length nucleotide sequences in a pair-wise deletion, *p*-distance algorithm, and bootstrapped using 2000 replicates. ML trees were generated in PAUP* and bootstrapped using 1000 replicates. Alignments were first analysed with ModelTest3.7 to determine the model most appropriate for the ML analysis. For genome segments 1 and 9, the TrN + I + G model was selected. For segments 3 and 4, the TrN + I model was selected. For genome segment 7, the HKY + I + G model was selected. Bayesian trees were created with MrBayes (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) using a GTR + inv gamma model. Probabilities were calculated using the MCMC with 100,000 generations and were sampled every 100th generation. A phylogenetic tree is presented that represents a consensus between these different methods (NJ, ML and MrBayes). Confidence levels (>50%) are presented for each algorithm individually.

Sequences were analysed for evidence of recombination within a segment using the Recombination Detection Programme (RDP), version 2Beta08 (http://darwin.uvigo.es/rdp/rdp.html). Sequence identities were calculated for each segment using a pair-wise deletion, *p*-distance algorithm in MEGA 3.1. Genetic drift was investigated by assessing neutrality with a standard Tajima's *D*-test, calculated using the total number of segregating sites.

3. Results

Full-genome sequences were generated for 11 different EHDV isolates (Table 1). Only genome segments 1, 3, 4, 7 and 9 are discussed in this report, representing the core virus particle. Analysis

Table 1
List of EHDV isolates used in this study.

IAH-P[a] reference collection number	Serotype[b] (strain)	Location (year)	Species isolated from	Reference
USA1955/01	1 (New Jersey)	USA (1955)	White-tailed deer	Shope et al. (1960)
NIG1967/01	1 (Ib Ar 22619)	Nigeria (1967)	Culicoides spp	Lee (1979)
CAN1962/01	2 (Alberta)	Canada (1962)	White-tailed deer	Ditchfield et al. (1964)
AUS1979/05	2 (CSIRO 439)	Australia (1979)	Sentinel Cattle	St. George et al. (1983)
JAP1959/01	2 (Ibaraki virus)	Japan (1959)	Cattle	Omori et al. (1969)
NIG1968/01	4 (Ib Ar 33853)	Nigeria (1968)	Culicoides spp	Lee (1979)
AUS1977/01	5 (CSIRO 157)	Australia (1977)	Sentinel Cattle	St. George et al. (1983)
AUS1981/07	6 (CSIRO 753)	Australia (1981)	Sentinel Cattle	St. George et al. (1983)
BAR1983/01	6 (318)	Bahrain (1983)	Cattle	Mohammed et al. (1996)
AUS1981/06	7 (CSIRO 775)	Australia (1981)	Sentinel Cattle	St. George et al. (1983)
AUS1982/06	8 (DPP59)	Australia (1982)	Sentinel Cattle	St. George et al. (1983)

Note: NIG1967/01 was previously classified as the reference serotype 3 strain, but has since been re-classified as a type 1 virus. Consequently, there are no type 3 viruses listed.

[a] Institute for Animal Health, Pirbright.

[b] Serotypes listed refer to the new classifications, as described in Anthony et al. (submitted for publication-b).