Virology 399 (2010) 98–108

Contents lists available at ScienceDirect

Virology

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

Ngaingan virus, a macropod-associated rhabdovirus, contains a second glycoprotein gene and seven novel open reading frames

Aneta Gubala [a,b,c,d,*], Steven Davis [e], Richard Weir [e,d], Lorna Melville [e,d], Chris Cowled [a,d], Peter Walker [a], David Boyle [a,d]

[a] CSIRO Livestock Industries, Australian Animal Health Laboratory, Geelong, Victoria, Australia
[b] School of Chemistry and Molecular Sciences, University of Queensland, St. Lucia, Queensland, Australia
[c] Human Protection and Performance Division, Defence Science and Technology Organisation, Melbourne, Australia
[d] Australian Biosecurity Cooperative Research Centre for Emerging Infectious Disease, Brisbane, Queensland, Australia
[e] Northern Territory Department of Regional Development, Primary Industry, Fisheries and Resources, Berrimah Veterinary Laboratories, Berrimah, Northern Territory, Australia

ARTICLE INFO

Article history:
Received 16 October 2009
Returned to author for revision
16 November 2009
Accepted 14 December 2009
Available online 20 January 2010

Keywords:
Rhabdovirus
Ngaingan
Rhabdoviridae
Arbovirus
Evolution
Genome
Characterization
(-) ssRNA
Virus
Serology

ABSTRACT

Ngaingan virus (NGAV) was isolated from a pool of biting midges that were collected in the tropics of northern Australia. Reported here is the full-length sequence of the NGAV genome, which, at over 15.7 kb, is the largest in any rhabdovirus described to date and contains 13 genes, the highest number of genes observed in any (−) ssRNA virus. Seven of these putative genes show no significant homology to known proteins. Like viruses in the genus *Ephemerovirus*, NGAV possesses a second glycoprotein gene (G_{NS}). Phylogenetic analyses, however, place NGAV within the yet to be classified "Hart Park" group containing Wongabel and Flanders viruses, which do not contain a second glycoprotein gene. Screening of various animal sera from northern Australia has indicated that NGAV is currently circulating in macropods (wallabies, wallaroos and kangaroos), highlighting the need for further studies to determine its potential to cause disease in these species.

Introduction

Ngaingan virus (NGAV) was isolated in 1970 from a pool of biting midges that were collected for surveillance for arboviruses at the low-lying plains of the Mitchell River Aboriginal community (Kowanyama; lat/long −15° 28′ S, 141° 44′ E), Gulf of Carpentaria, northern Queensland (Doherty et al., 1973). At the time, the midge pool was thought to consist of only *Culicoides brevitarsis* (which do not favour feeding on macropods), but later studies of *Culicoides* spp. from the region suggested that *C. actoni* (which have a wide feeding range including macropods and cattle) were probably also present (Kay et al., 1978). Another yet uncharacterised rhabdovirus, Almpiwar virus, was also isolated from this area a number of times from skinks. Morphological examination of NGAV revealed bullet-shaped virions typical of the *Rhabdoviridae* (Karabatsos, 1985). Experimental infection of mosquitoes showed that it can multiply in *Aedes aegypti*, a characteristic common of arboviruses (Carley et al., 1973). Initial characterisation suggested that NGAV was related antigenically to Tibrogargan virus, another uncharacterised Australian rhabdovirus (Calisher et al., 1989). However, recent phylogenetic analysis of a short L gene fragment indicated that it is a member of the Hart Park group, which also includes Wongabel (WONV) and Parry Creek viruses, both from Australia, and Flanders virus (FLAV) isolated from mosquitoes and birds in the USA (Bourhy et al., 2005).

Early serologic surveys suggested that NGAV infects wallabies, kangaroos and possibly cattle (Doherty et al., 1973). Whilst a recent outbreak of blindness in kangaroos and wallabies has been associated with the Wallal and Warrego orbiviruses (Hooper et al., 1999; Reddacliff et al., 1999), viral diseases in these species are poorly studied and it remains to be determined whether NGAV also causes disease.

The genomes of rhabdoviruses always contain genes encoding the five structural proteins, nucleoprotein (N), phosphoprotein (P), matrix protein (M), transmembrane glycoprotein (G) and RNA-dependent RNA polymerase (L). The complete genome sequences

* Corresponding author. Human Protection and Performance Division, Defence Science and Technology Organisation, PO Box 4331, Melbourne 3001, Victoria, Australia. Fax: +61 3 9626 8410.
E-mail address: ania.gubala@dsto.defence.gov.au (A. Gubala).

0042-6822/$ – see front matter.
doi:10.1016/j.virol.2009.12.013

obtained for a number of rhabdoviruses demonstrate that some possess numerous additional genes or ORFs that putatively encode small proteins with unknown functions (Basurco and Benmansour, 1995; Dietzgen et al., 2006; Gubala et al., 2008; Huang et al., 2003; McWilliam et al., 1997; Reed et al., 2005; Revill et al., 2005; Scholthof et al., 1994; Springfeld et al., 2005; Tanno et al., 2000; Walker et al., 1991; Walker et al., 1992; Wang et al., 1994; Wang and Walker, 1993). In addition, the two characterised ephemeroviruses, bovine ephemeral fever virus (BEFV) and Adelaide River virus (ARV), are known to contain a second glycoprotein (G_{NS}) with an unknown function (Walker et al., 1991; Walker et al., 1992; Wang and Walker, 1993). Due to the large number of evolutionarily divergent rhabdoviruses for which little or no genome sequence information is available, many remain unassigned to any of the six genera (*Lyssavirus*, *Vesiculovirus*, *Ephemerovirus*, *Novirhabdovirus*, *Cytorhabdovirus* and *Nucleorhabdovirus*) recognised currently.

This paper describes the sequence of the complete genome of NGAV, which at 15,764 nt is currently the largest known within the *Rhabdoviridae*. The NGAV genome possesses 13 discrete genes predicted to encode proteins, which is the most of any known (−) ssRNA virus and highlights the levels of complexity rhabdoviruses are capable of reaching. In addition to the five genes coding for structural proteins typical of rhabdoviruses, NGAV possesses a gene encoding a second glycoprotein, which has previously only been observed in members of the genus *Ephemerovirus*, and seven encoding novel proteins. Phylogenetic analyses support assignment of a seventh genus within the *Rhabdoviridae* comprising of NGAV, WONV and FLAV. Serologic data indicate that NGAV is still circulating in some animal species in northern Australia.

Results

Complete genomic sequence of NGAV

The complete genome of NGAV is 15,764 nt in length and contains the five typical rhabdovirus genes N, P, M, G and L (Fig. 1A). In addition, it contains seven relatively short ORFs predicted to encode novel proteins with no significant similarity at the amino acid level to any proteins currently in GenBank, as well as a longer ORF consistent in genome position and sequence to a second glycoprotein gene, a feature previously only observed in the ephemeroviruses BEFV and ARV (Walker et al., 1991; Walker et al., 1992; Wang and Walker, 1993). Each of the ORFs is bounded by recognisable transcription control sequences (with the exception of U1/U2 and U5/U6 which appear to be bicistronic, as described later), suggesting that all are transcribed. Details of the genes, ORFs, putative proteins and untranslated regions (UTRs) are collated in Table 1.

Although a consensus sequence of high confidence was generated for the whole genome, four nucleotide positions remained ambiguous despite repeated sequencing. Conflicts at positions 5842, 8687 and 12,853 within the G, U6 and L genes, respectively, were nonsynonymous and resulted in amino acid transitions, and the conflict at position 3153 introduced a stop codon within the U3 gene. These conflicts are noted in the NGAV GenBank entry (accession no. FJ715959).

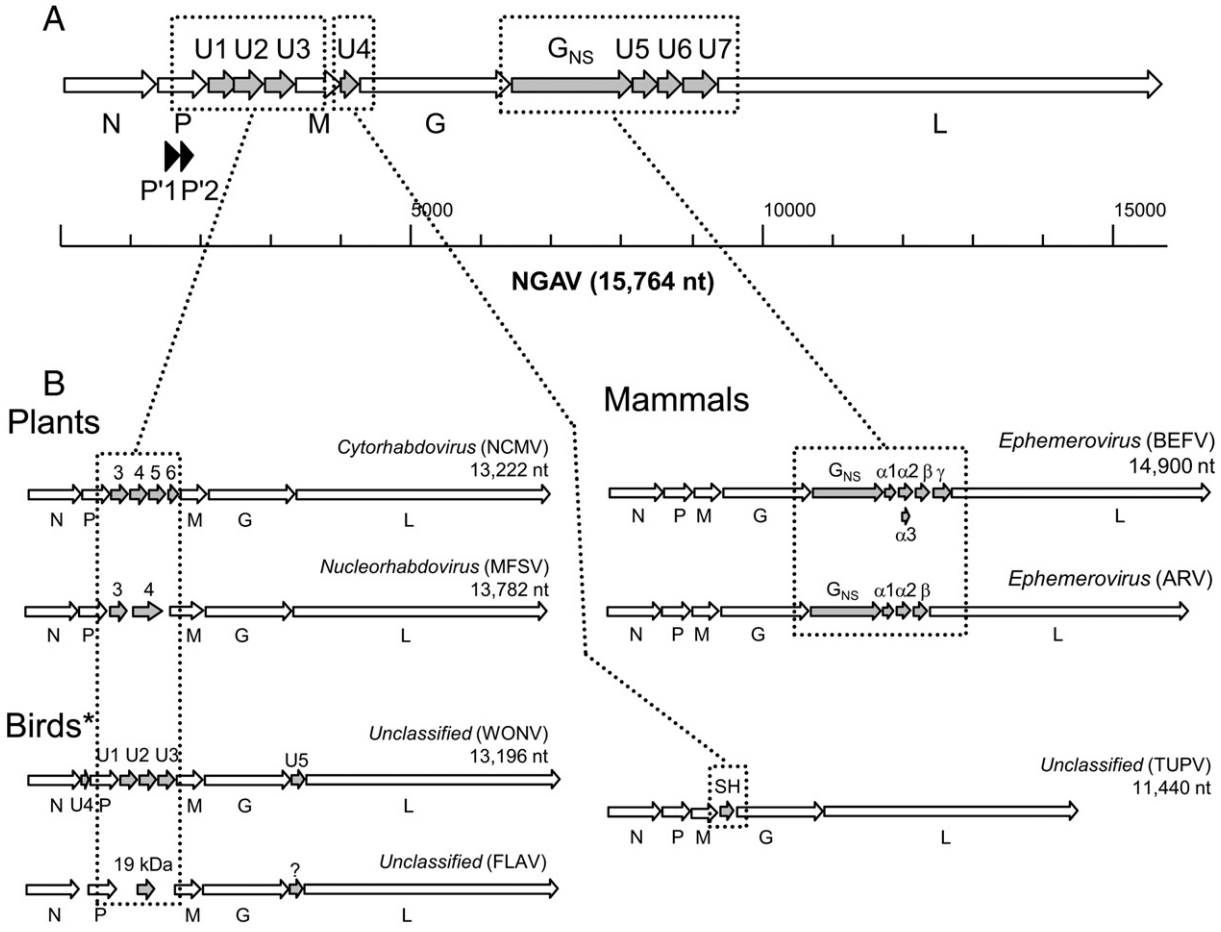

Fig. 1. Comparison of NGAV genome organisation with other rhabdoviruses. Common rhabdovirus proteins are depicted by hollow arrows and novel ORFs by shaded arrows. (A) Organisation of the NGAV genome. ORFs P'1 and P'2 (black arrow heads) overlap with P in different frames. (B) Genomic locations of novel NGAV ORFs compared with other rhabdoviruses. *Serologic evidence suggests that WONV infects birds although it has never been isolated from this host. FLAV has been isolated from birds and mosquitoes.