Virology 474 (2015) 154–162

Contents lists available at ScienceDirect

Virology

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

Evolutionary genetics and vector adaptation of recombinant viruses of the western equine encephalitis antigenic complex provides new insights into alphavirus diversity and host switching

Andrew B. Allison [a,b,*], David E. Stallknecht [b,c], Edward C. Holmes [d]

[a] Baker Institute for Animal Health, Department of Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
[b] Southeastern Cooperative Wildlife Disease Study, College of Veterinary Medicine, University of Georgia, Athens, GA 30602, USA
[c] Department of Population Health, College of Veterinary Medicine, University of Georgia, Athens, GA 30602, USA
[d] Marie Bashir Institute for Infectious Diseases and Biosecurity, Charles Perkins Centre, School of Biological Sciences and Sydney Medical School, University of Sydney, New South Wales 2006, Australia

ARTICLE INFO

Article history:
Received 6 August 2014
Returned to author for revisions
28 August 2014
Accepted 23 October 2014

Keywords:
Fort Morgan virus
Highlands J virus
Western equine encephalitis virus
Western equine encephalitis antigenic complex
Alphavirus
Togavirus
Recombination
Host range
Virus emergence

ABSTRACT

Western equine encephalitis virus (WEEV), *Highlands J virus* (HJV), and *Fort Morgan virus* (FMV) are the sole representatives of the WEE antigenic complex of the genus *Alphavirus*, family *Togaviridae*, that are endemic to North America. All three viruses have their ancestry in a recombination event involving *eastern equine encephalitis virus* (EEEV) and a *Sindbis* (SIN)-like virus that gave rise to a chimeric alphavirus that subsequently diversified into the present-day WEEV, HJV, and FMV. Here, we present a comparative analysis of the genetic, ecological, and evolutionary relationships among these recombinant-origin viruses, including the description of a nsP4 polymerase mutation in FMV that allows it to circumvent the host range barrier to Asian tiger mosquito cells, a vector species that is normally refractory to infection. Notably, we also provide evidence that the recombination event that gave rise to these three WEEV antigenic complex viruses may have occurred in North America.

Introduction

The genus *Alphavirus* within the family *Togaviridae* is comprised of 31 arthropod-borne viruses with a worldwide distribution (Powers et al., 2012; Nasar et al., 2012). Of these viruses, four are endemic to North America: *eastern equine encephalitis virus* (EEEV), *western equine encephalitis virus* (WEEV), *Highlands J virus* (HJV), and *Fort Morgan virus* (FMV) (Weaver et al., 1997). Two additional alphaviruses found in the western United States, Buggy Creek virus and the recently described Stone Lakes virus, are generally regarded as variants of FMV (Hopla et al., 1993; Powers et al., 2001; Brault et al., 2009). WEEV, HJV, and FMV are notable in that they are descendants of a recombination event between a *Sindbis* (SIN)-like virus and EEEV that is believed to have occurred in the neotropics of South America (Hahn et al., 1988; Strauss and Strauss, 1997). As the major surface glycoproteins (E1 and E2) of the ancestral recombinant were obtained from the SIN-like virus, WEEV, HJV, and FMV are antigenically related (i.e., in neutralization tests) to SINV rather than to EEEV (Calisher et al., 1988). Along with *Aura virus* (AURAV) and *Whataroa virus* (WHAV), WEEV, HJV, FMV, and SINV collectively constitute the WEE antigenic complex (Weaver et al., 1997).

Other than the three recombinants, AURAV is the only WEE antigenic complex member found in the New World, having been isolated from *Culex* and *Aedes* species of mosquitoes in Brazil and Argentina (Causey et al., 1963; Rümenapf et al., 1994, 1995). As AURAV is endemic to South America and is related to SINV, it initially provided an attractive candidate as the putative SIN-like parental virus of the recombination event. However, as WEEV is more closely related to SINV than it is to AURAV (Rumenapf et al., 1995), and AURAV and SINV are believed to have diverged prior to the recombination event (Weaver et al., 1997), it is likely that EEEV recombined with a virus more closely related to the present-day SINV. SINV (or a SIN-like virus other than AURAV) has not been detected in North or South America, suggesting that after its introduction into the New World (presuming a non-New World

* Corresponding author at: Baker Institute for Animal Health, Department of Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA.
E-mail address: aba75@cornell.edu (A.B. Allison).

http://dx.doi.org/10.1016/j.virol.2014.10.024
0042-6822/

origin of SINV), this virus either went extinct or that it may possibly still be circulating endemically (e.g., in the neotropics) but has not been detected. Conversely, EEEV is found only in the New World and represents the sole species constituting the EEE antigenic complex, although South American lineages (lineages II–IV) of EEEV have recently been reclassified as *Madaraiga virus* (MADV) based on genetic, ecological, epidemiological, and pathogenic differences between North and South American strains (Arrigo et al., 2010).

In North America, EEEV, WEEV, HJV, and FMV all circulate in transmission cycles involving passerine birds as amplifying hosts and hematophagous arthropods as vectors. However, FMV is unique among these viruses in that the normal invertebrate vector is the cimicid swallow bug (*Oeciacus vicarius*), rather than a mosquito species [i.e., *Culiseta* (*Cs.*) *melanura* for EEEV and HJV; *Culex* (*Cx.*) *tarsalis* for WEEV] (Calisher et al., 1980). The distribution of each of the recombinant alphaviruses in the United States is, for the most part, spatially discrete, and is essentially defined by the geographical range of their respective enzootic vectors. HJV is primarily confined to eastern states along the Gulf and Atlantic seaboard and some inland foci around the Great Lakes region (Cilnis et al., 1996), while WEEV and FMV are endemic throughout most of the western United States (Reisen and Monath, 1988; Pfeffer et al., 2006), although WEEV appears to be declining in North America (Forrester et al., 2008; Bergren et al., 2014). Additionally, FMV is more focally distributed than WEEV due to its unique transmission cycle (outlined below). Interestingly, HJV and EEEV share apparently identical transmission cycles in North America and thus nearly identical geographical ranges (Scott and Weaver, 1989). Although lineages of WEEV and EEEV exist in Central and South America (Srihongse and Galindo, 1967; Mitchell et al., 1987; Weaver et al., 1994, 1997; Brault et al., 1999), FMV and HJV have not been isolated outside of the United States. Additionally, unlike EEEV and WEEV, neither FMV nor HJV are normally associated with disease in mammals (Hayes and Wallis, 1977; Calisher et al., 1980; Englund et al., 1986; Karabatsos et al., 1988; Przelomski et al., 1988), although all four viruses are avian pathogens to varying degrees (Scott et al., 1984; Ficken et al., 1993; Randolph et al., 1994; Huyvaert et al., 2008).

Fort Morgan virus was first isolated by Hayes et al. (1977) from *O. vicarius* in eastern Colorado. Since its initial description, FMV (along with Buggy Creek virus and Stone Lakes virus) has been reported from a number of additional western and central states including Nebraska, Oklahoma, Texas, North Dakota, South Dakota, Washington, and California (Calisher et al., 1980; Hopla et al., 1993; Pfeffer et al., 2006; Padhi et al., 2008; Brault et al., 2009; Brown et al., 2009). The primary vertebrate amplifying hosts for FMV are cliff swallows (*Petrochelidon pyrrhonata*), and to a lesser extent, house sparrows (*Passer domesticus*), with the latter being inadvertently involved in transmission as they opportunistically occupy cliff swallow nests (Calisher et al., 1980; Scott et al., 1984; O'Brien et al., 2011). *O. vicarius* is a sedentary ectoparasite that strictly blood-feeds on cliff swallows (and house sparrows which parasitize cliff swallow nests) (Brown et al., 2010a, 2010b). This, in part, likely restricts the geographical range of FMV to more discrete endemic foci compared to that of other North American alphaviruses that are vectored endemically by motile mosquito species that may feed on multiple avian species (*Cs. melanura*) or are more catholic in host preference (*Cx. tarsalis*) (Loye, 1985). That *Cx. tarsalis* and *Cx. pipiens* have been demonstrated to be refractory to FMV infection following intra-thoracic inoculation, and *Aedes (Ae.) albopictus* cell cultures do not support replication of FMV (Calisher et al., 1980), suggests that FMV is exclusively adapted to *O. vicarius*. However, it was also demonstrated that, in vivo, FMV could infect *Cs. melanura*, the vector for EEEV and HJV (Calisher et al., 1980).

The genomes of two of the recombinants, WEEV and HJV, along with representative progenitors of the parental viruses, EEEV and SINV, have been described previously (Shirako et al., 1991; Weaver et al., 1993; Netolitzky et al., 2000; Allison and Stallknecht, 2009). However, the lack of the FMV genome has precluded a comprehensive comparative evolutionary analysis of the three recombinant viruses, although the genome of the variant Buggy Creek virus was recently sequenced (Forrester et al., 2012). Additionally, the mechanism(s) underlying the inability of FMV to infect mosquito cells remains unknown and represents a novel host range barrier, as most other alphaviruses (including HJV and WEEV) are vectored by mosquitoes. Herein, we present the genome of FMV and analyze the genetic, ecological, and evolutionary relationships of the WEE antigenic complex viruses, including the origin of the ancestral recombinant and other possible recombination events, as well as the in vitro host range adaptation of FMV to mosquito cells.

Results

Genetic relationships of the recombinant WEEV antigenic complex viruses

Genome sequencing of FMV was undertaken to perform genetic and phylogenetic comparisons of the full-length genomes among the three viruses (FMV, WEEV, HJV) known to be derived from the recombination event between EEEV and a SIN-like virus. As the structural polyprotein (C-E3-E2-6K-E1) and 3′ UTR of FMV CM4-146, accounting for ~3.9 kB, have previously been sequenced (Pfeffer et al., 1998), these genomic regions will not be discussed in detail here. Excluding the 5′ cap nucleotide and the 3′ poly(A) tail, the genome of FMV is 11,381 nt in length. The FMV genome is the shortest among the recombinants, being 127 or 145 nt less than either WEEV or HJV, respectively, primarily due to the short 3′ UTR. Whereas HJV and WEEV share a 75% nucleotide and 87% amino acid identity over their entire genomes, FMV is more divergent, with 69% nucleotide and 78% amino acid identity to both HJV and WEEV (Fig. S1). nsP4 was the most conserved (87–88%) protein among the recombinants and the parental EEEV, followed by the capsid (84–88%), nsP1 (80–86%), nsP2 (81–84%), and nsP3 (61–69%) (Fig. S1).

All of the nonstructural proteins of FMV were the same length as in HJV, WEEV, and EEEV, with the exception that the nsP3 protein of FMV (522aa) was of intermediate length between HJV (515aa) and WEEV (532aa), with the nsP3 protein of all the recombinants shorter than that of the parental EEEV (559aa). Although the C-terminal region of nsP3 is the most highly variable coding region of the alphavirus genome, a number of short, conserved amino acid motifs were observed upstream of the nsP3/nsP4 cleavage site sequence in FMV, HJV, WEEV, and EEEV (^{349}-IPSP-^{352}, ^{401}-WSIPS-^{405}, ^{447}-QFLS-^{450}, ^{456}-PAPR-^{459}; numbering based on FMV) (Fig. S2), suggesting that these motifs may have some structural and/or functional role in the C-terminal domain. Similar to HJV, WEEV, and EEEV, the nsP3 gene of FMV contains an opal termination codon (UGA) at genomic positions 5573–5575 followed by a C nucleotide at position 5576 (Strauss et al., 1983). The nsP1/nsP2, nsP2/nsP3, and nsP3/nsP4 cleavage site sequences for FMV were EAGA/GSVE, EAGR/APAY, and RYEAGA/YIFS, respectively, which were identical to those of EEEV, WEEV, and HJV (Strauss and Strauss, 1994; Allison and Stallknecht, 2009). Based on the cleavage site motifs, the theoretical isoelectric point and molecular weight of each of the FMV nonstructural proteins were nsP1 (6.00; 59.8 kDa), nsP2 (8.80; 89.0 kDa), nsP3 (5.54; 57.4 kDa), and nsP4 (6.32; 68.1 kDa).