Virology 504 (2017) 152–167

Contents lists available at ScienceDirect

Virology

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

Genetic characterization, molecular epidemiology, and phylogenetic relationships of insect-specific viruses in the taxon Negevirus

Marcio R.T. Nunes[a], María Angélica Contreras-Gutierrez[b,c], Hilda Guzman[d,e,f], Livia C. Martins[f], Mayla Feitoza Barbirato[g], Chelsea Savit[h], Victoria Balta[i], Sandra Uribe[c], Rafael Vivero[b,c], Juan David Suaza[b,c], Hamilton Oliveira[f], Joaquin P. Nunes Neto[f], Valeria L. Carvalho[g], Sandro Patroca da Silva[a], Jedson F. Cardoso[a], Rodrigo Santo de Oliveira[a], Poliana da Silva Lemos[f], Thomas G. Wood[j], Steven G. Widen[j], Pedro F.C. Vasconcelos[f], Durland Fish[k], Nikos Vasilakis[d,e,f,*], Robert B. Tesh[d,e,f,*]

[a] *Center for Technological Innovation, Evandro Chagas Institute, Ministry of Health, Ananindeua, Para, Brazil*
[b] *Programa de Estudio y Control de Enfermedades Tropicales – PECET - SIU-Sede de Investigación Universitaria – Universidad de Antioquia, Medellín, Colombia*
[c] *Grupo de Investigación en Sistemática Molecular-GSM, Facultad de Ciencias,Ciencias, Universidad Nacional de Colombia, sede Medellín, Medellín, Colombia*
[d] *Department of Pathology and Center for Biodefense and Emerging Infectious Diseases, University of Texas Medical Branch, Galveston, TX 77555-0609, United States*
[e] *Center for Tropical Diseases, University of Texas Medical Branch, Galveston, TX 77555-0609, United States*
[f] *Department of Arbovirology and Hemorrhagic Fevers, Evandro Chagas Institute, Ministry of Health, Ananindeua, Para, Brazil*
[g] *Vale S.A. – Projeto Ferro Carajás S11D, Canaã dos Carajás, Para, Brazil*
[h] *Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, United States*
[i] *School of Public Health, University of Washington, Seattle, WA 98195, United States*
[j] *Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX 77555-0645, United States*
[k] *Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT 06520, United States*

ARTICLE INFO

Keywords:
Negevirus
Nelorpivirus genus
Sandewavirus genus
Insect-specific viruses
Virus diversity

ABSTRACT

The recently described taxon *Negevirus* is comprised of a diverse group of insect-specific viruses isolated from mosquitoes and phlebotomine sandflies. In this study, a comprehensive genetic characterization, molecular, epidemiological and evolutionary analyses were conducted on nearly full-length sequences of 91 new negevirus isolates obtained in Brazil, Colombia, Peru, Panama, USA and Nepal. We demonstrated that these arthropod restricted viruses are clustered in two major phylogenetic groups with origins related to three plant virus genera (*Cilevirus, Higrevirus* and *Blunevirus*). Molecular analyses demonstrated that specific host correlations are not present with most negeviruses; instead, high genetic variability, wide host-range, and cross-species transmission were noted. The data presented here also revealed the existence of five novel insect-specific viruses falling into two arthropod-restrictive virus taxa, previously proposed as distinct genera, designated *Nelorpivirus* and *Sandewavirus*. Our results provide a better understanding of the molecular epidemiology, evolution, taxonomy and stability of this group of insect-restricted viruses.

1. Introduction

The taxon Negevirus is a monophyletic taxon of non-segmented, positive-sense ssRNA viruses that have been isolated from mosquitoes and phlebotomine sand flies in the Americas, Europe, Africa and Asia (Vasilakis et al., 2013; Kallies et al., 2014; Auguste et al., 2014; Nabeshima et al., 2014; Carapeta et al., 2015). Negevirus particles are spherical in shape with a size of 45–55 nm in diameter (Vasilakis et al., 2013). Three interesting biological characteristics of the known negeviruses are: (1) their wide geographic distribution; (2) their broad host range among biting *Diptera*; and (3) the relatively high natural infection rates in some mosquito species (Vasilakis et al., 2013; Kallies et al., 2014).

Based on their phylogenetic relationships, Kallies et al. (2014) proposed that the taxon Negevirus can be separated into two major

* Correspondence to: Department of Pathology, University of Texas Medical Branch, 301 University Boulevard, Galveston, TX 77555-0609, United States.
E-mail addresses: nivasila@utmb.edu (N. Vasilakis), rtesh@utmb.edu (R.B. Tesh).

http://dx.doi.org/10.1016/j.virol.2017.01.022
Received 26 October 2016; Received in revised form 25 January 2017; Accepted 27 January 2017
Available online 11 February 2017

Table 1
Virus isolates used for genetic characterization and phylogenetic analyses.

Family	Genus	Virus name	Strain Identification	Genbank number	Host source	Year of Isolation	Country of Isolation	State Province
Unassigned	***Nelorpivirus***	Big Cypress virus	US/BCPV/BCNP_2-24A/2014	KX518750	*Anopheles atropos*	2014	USA	Florida
			US/BCPV/BCNP_2-30A/2014	KX518751	*Culex iolambdis*	2014	USA	Florida
			US/BCPV/BCNP_3-125/2014	KX518754	*Anopheles atropos*	2014	USA	Florida
			US/BCPV/BCNP_3-112A/2014	KX518752	*Aedes taeniorhynchus*	2014	USA	Florida
			US/BCPV/BCNP_3-126B/2014	KX518755	*Anopheles atropos*	2014	USA	Florida
			US/BCPV/BCNP_3-128A/2014	KX518756	*Aedes taeniorhynchus*	2014	USA	Florida
			US/BCPVBCNP_3-124/2014	KX518753	*Anopheles atropos*	2014	USA	Florida
		Brejeira virus	CO/BREJV/CoB_35/2013	KX518759	Mosquito unknown	2013	Colombia	Cordoba
			CO/BREJV/CoB_37A/2013	KX518760	Mosquito unknown	2013	Colombia	Cordoba
			CO/BREJV/CoB_37dA/2013	KX518761	Mosquito unknown	2013	Colombia	Cordoba
			BR/BREJV/BEAR803410/2005	KX518762	*Culex* sp.	2005	Brazil	Para
			BR/BREJV/BEAR803412A/2005	KX518763	*Culex* sp.	2005	Brazil	Para
			BR/BREJV/AR800208B/2013	KM350508	*Culex* sp.	2013	Brazil	Para
			BREJV/BR/AR800212B/2013	KM350509	*Culex* sp.	2013	Brazil	Para
			BR/BREJV/AR800213B/2013	KM350510	*Culex* sp.	2013	Brazil	Para
			BR/BREJV/AR800214/2013	KM350511	*Culex* sp.	2013	Brazil	Para
			BR/BREJV/AR800224/2013	KM350512	*Culex* sp.	2013	Brazil	Para
			BR/BREJV/AR800227/2013	KM350513	*Culex* sp.	2013	Brazil	Para
			BR/BREJV/AR800233B/2013	KM350514	*Culex* sp.	2013	Brazil	Para
			BR/BREJV/AR800239/2013	KM350515	*Culex* sp.	2013	Brazil	Para
		Cordoba virus	NP/CDBV/Nepal_12_3/1984	KX518772	Mosquito unknown	1984	Nepal	na
			NP/CDBV/Nepal_24-3/1984	KX518773	Mosquito unknown	1984	Nepal	na
			US/CDBV/GMC_30/2013	KX518771	*Culex quinquefasciatus*	2013	USA	Texas
			US/CDBV/EVG_7-47B/2013	KX518767	*Culex nigripalpus*	2013	USA	Florida
			US/CDBV/EVG_7-41B/2013	KX518766	*Culex nigripalpus*	2013	USA	Florida
			US/CDBV/EVG_9-196B/2013	KX518768	*Psorophora ferox*	2013	USA	Florida
			US/CDBV/EVG_9-226B/2013	KX518769	*Culex cedecei*	2013	USA	Florida
			US/CDBV/EVG_9-228B/2013	KX518770	*Anopheles crucians*	2013	USA	Florida
			CO/CDBV/CoB_37B/2013	KX518764	Mosquito unknown	2013	Colombia	Cordoba
			CO/CDBV/CoB_37dC/2013	KX518765	Mosquito unknown	2013	Colombia	Cordoba
		Loreto virus	PE/LORV/3940_83_2/1983	KX518775	*Anopheles albimanus*	1983	Peru	Loreto
			PE/LORV/3940-83/1983	JQ675610	*Anopheles albimanus*	1983	Peru	Lima
			PE/LORV/PE_AR_2612_77/1977	JQ675611	*Culex* sp.	1977	Peru	Loreto
			PE/LORV/PE_AR_2617_77/1977	JQ675612	*Lutzomyia* sp.	1977	Peru	Loreto
		Negev virus	US/NEGV/V_05087/2014	KX518776	*Culex quinquefasciatus*	2014	USA	Texas
			PT/NEGV/#730/2010	AB935183	*Culex univitatus*	2010	Portugal	na
			IL/NEGV/EO239/1983	JQ675605	*Anopheles constant*	1983	Israel	na
			US/NEGV/M_30957_58/2008	JQ675608	*Culex coronator*	2008	USA	Texas
			US/NEGV/M_33056/2008	JQ675609	*Culex quinquefasciatus*	2008	USA	Texas
			PT/NEGV/#174/2010	HF913429	*Ochlerotatus caspius*	2010	Portugal	na
		Ngewotan virus	NP/NWTV/Nepal_12/1984	KX518778	Mosquito unknown	1984	Nepal	na
			NP/NWTV/Nepal_12-2/1984	KX518777	Mosquito unknown	1984	Nepal	na
			NP/NWTV/Nepal_22/1984	KX518779	Mosquito unknown	1984	Nepal	na
			ID/NWTV/JKT_9982/1981	JQ686833	*Culex vishnui*	1981	Indonesia	Central Java
		Piura virus	US/PIUV/EVG_1-47/2013	KX518788	*Culex cedecei*	2013	USA	Florida

(continued on next page)

Table 1 *(continued)*

Family	Genus	Virus name	Strain Identification	Genbank number	Host source	Year of Isolation	Country of Isolation	State Province
			US/PIUV/EVG_1-50/2013	KX518789	*Aedes infirmatus*	2013	USA	Florida
			US/PIUV/EVG_1-76/2013	KX518790	*Culex nigripalpus*	2013	USA	Florida
			US/PIUV/EVG_2-7/2013	KX518791	*Culex nigripalpus*	2013	USA	Florida
			US/PIUV/EVG_2-86/2013	KX518793	*Aedes atlanticus*	2013	USA	Florida
			US/PIUV/EVG_2-8/2013	KX518792	*Culex nigripalpus*	2013	USA	Florida
			US/PIUV/EVG_3-129/2013	KX518795	*Culex nigripalpus*	2013	USA	Florida
			US/PIUV/EVG_3-97/2013	KX518794	*Anopheles crucians*	2013	USA	Florida
			US/PIUV/EVG_4-9A/2013	KX518796	*Culex nigripalpus*	2013	USA	Florida
			US/PIUV/EVG_4-9B/2013	KX518797	*Culex nigripalpus*	2013	USA	Florida
			US/PIUV/EVG_7-47A/2013	KX518804	*Culex nigripalpus*	2013	USA	Florida
			US/PIUV/EVG_5-53/2013	KX518798	*Anopheles quadrimaculatus*	2013	USA	Florida
			US/PIUV/EVG_7-35/2013	KX518802	Psorophora ferox	2013	USA	Florida
			US/PIUV/EVG_7−19/2013	KX518800	*Aedes atlanticus*	2013	USA	Florida
			US/PIUV/EVG_7−1/2013	KX518799	*Aedes taeniorhynchus*	2013	USA	Florida
			US/PIUV/EVG_7-94/2013	KX518806	*Anopheles crucians*	2013	USA	Florida
			US/PIUV/EVG_7-95/2013	KX518807	*Anopheles crucians*	2013	USA	Florida
			US/PIUV/EVG_7-78/2013	KX518805	*Anopheles crucians*	2013	USA	Florida
			US/PIUV/EVG_7-29/2013	KX518801	*Culex cedecei*	2013	USA	Florida
			US/PIUV/EVG_7-41 A/2013	KX518803	*Culex nigripalpus*	2013	USA	Florida
			US/PIUV/EVG_9-196A/2013	KX518816	*Psorophora ferox*	2013	USA	Florida
			US/PIUV/EVG_7-110/2013	KX518808	*Aedes infirmatus*	2013	USA	Florida
			US/PIUV/EVG_7-111/2013	KX518809	*Anopheles crucians*	2013	USA	Florida
			US/PIUV/EVG_7-112/2013	KX518810	*Aedes atlanticus*	2013	USA	Florida
			US/PIUV/EVG_7-114/2013	KX518811	*Uranotaeneia iowii*	2013	USA	Florida
			US/PIUV/EVG_9-226A/2013	KX518817	*Culex cedecei*	2013	USA	Florida
			US/PIUV/EVG_9-228A/2013	KX518818	*Anopheles crucians*	2013	USA	Florida
			US/PIUV/EVG_9-151/2013	KX518813	*Culex cedecei*	2013	USA	Florida
			US/PIUV/EVG_9-92/2013	KX518812	*Culex cedecei*	2013	USA	Florida
			US/PIUV/EVG_9-162/2013	KX518814	*Culex cedecei*	2013	USA	Florida
			US/PIUV/EVG_9-183/2013	KX518815	*Anopheles crucians*	2013	USA	Florida
			US/PIUV/V_08449/2013	KX518819	*Culex quinquefasciatus*	2013	USA	Texas
			US/PIUV/V_08458/2013	KX518820	*Culex quinquefasciatus*	2013	USA	Texas
			US/PIUV/BCNP_2-60/2014	KX518780	*Culex nigripalpus*	2014	USA	Florida
			US/PIUV/BCNP_3-126A/2014	KX518781	*Anopheles atropos*	2014	USA	Florida
			US/PIUV/BCNP_3-148/2014	KX518784	*Culex nigripalpus*	2014	USA	Florida
			US/PIUV/BCNP_4-73/2014	KX518785	*Aedes infirmatus*	2014	USA	Florida
			US/PIUV/BCNP_3-141/2014	KX518783	*Aedes infirmatus*	2014	USA	Florida
			US/PIUV/BCNP_3-134/2014	KX518782	*Aedes taeniorhynchus*	2014	USA	Florida
			CO/PIUV/CoR_10/2013	KX527894	*Lutzomyia evansi*	2013	Colombia	Sucre
			CO/PIUV/CoR_29/2013	KX518786	*Culex* sp.	2013	Colombia	Sucre
			CO/PIUV/CoR_33/2013	KX518787	*Culex* sp.	2013	Colombia	Sucre
			PE/PIUV/JQ675607/1996	JQ675607	*Culex* sp.	1996	Peru	Piura
			MX/PIUV/C6.7-MX-2008/2008	KM249340	*Culicidae* sp.	2008	Mexico	Chiapas
		San Bernardo virus	CO/SBDV/CoB_52/2013	KX518821	Mosquito unknown	2013	Colombia	Cordoba
			CO/SBDV/CoG_72_b/2013	KX518822	*Deinocerites atlanticus*	2013	Colombia	Cordoba
	Sandewavirus	Biratnagar virus	NP/BIRV/Nepal_12/1984	KX518757	Mosquito unknown	1984	Nepal	na
			NP/BIRV/Nepal_24_2/1984	KX518758	Mosquito unknown	1984	Nepal	na
		Dezidougou virus	CI/DEZV/ARA_20086/1984	JQ675604	*Aedes aegypti*	1984	Ivory Coast	
		Goutanap virus	CI/GANV/F33_CI_2004/2004	KF588035	*Culex nebulosus*	2004	Ivory Coast	
			CI/GANV/F35_CI_2004/2004	KF588036	*Culex nebulosus*	2004	Ivory Coast	
			CI/GANV/F47_CI_2004/2004	KF588037	*Culicinae* sp.	2004	Ivory Coast	

(continued on next page)

Table 1 *(continued)*

Family	Genus	Virus name	Strain Identification	Genbank number	Host source	Year of Isolation	Country of Isolation	State Province
			CI/GANV/F54_CI_2004/2004	KF588038	*Culex antennatus*	2004	Ivory Coast	
			CI/GANV/F55_CI_2004/2004	KF588039	*Culex antennatus*	2004	Ivory Coast	
			CI/GANV/C68_CI_2004/2004	KM249339	*Culicidae* sp.	2004	Ivory Coast	
		Santana virus	BR/SANV/Be_Ar_517449/1992	JQ675606	*Culex* sp.	1992	Brazil	Amapa
		Tanay virus	PH/TANAV/11_3/2005	KF425262	*Culex* sp.	2005	Philippines	Rizal
			PH/TANAV/11_4/2005	KF425263	*Armigeres* sp.	2005	Philippines	Rizal
			PH/TANAV/11_5/2005	KF425264	*Armigeres* sp.	2005	Philippines	Rizal
			PH/TANAV/11_2/2005	NC_024071	*Culex quinquefasciatus*	2005	Philippines	Rizal
		Wallerfield virus	CO/WALV/CoB_92/2013	KX518836	*Deinocerites* sp.	2013	Colombia	Cordoba
			CO/WALV/CoB_37dB/2013	KX518835	Mosquito unknown	2013	Colombia	Cordoba
			PA/WALV/GAM_058A/2013	KX518837	Mosquito unknown	2013	Panama	Panama
			PA/WALV/GAM_195/2013	KX518839	Mosquito unknown	2013	Panama	Panama
			PA/WALV/GAM_058A_2/2013	KX518838	Mosquito unknown	2013	Panama	Panama
			US/WALV/BCNP_2_24B/2014	KX518827	*Anopheles atropos*	2014	USA	Florida
			US/WALV/BCNP_2_15/2014	KX518826	*Anopheles punctipennis*	2014	USA	Florida
			US/WALV/BCNP_2_5/2014	KX518825	*Anopheles crucians*	2014	USA	Florida
			US/WALV/BCNP_2_4/2014	KX518824	*Anopheles crucians*	2014	USA	Florida
			US/WALV/BCNP_2_30B/2014	KX518828	*Culex iolambdis*	2014	USA	Florida
			US/WALV/BCNP_2_3/2014	KX518823	*Anopheles crucians*	2014	USA	Florida
			US/WALV/BCNP_3_112B/2014	KX518829	*Aedes taeniorhynchus*	2014	USA	Florida
			US/WALV/BCNP_3_128B/2014	KX518830	*Aedes taeniorhynchus*	2014	USA	Florida
			US/WALV/BCNP_4_205/2014	KX518831	*Anopheles atropos*	2014	USA	Florida
			US/WALV/BEAR803412B/2005	KX518832	*Culex* sp.	2005	Brazil	Para
			US/WALV/BEAR803418/2005	KX518833	*Culex* sp.	2005	Brazil	Para
			BR/WALV/BEAR804033/2005	KX518834	*Culex* sp.	2005	Brazil	Para
			BR/WALV/AR800208_A/2013	KM350504	*Culex* sp.	2013	Brazil	Para
			BR/WALV/AR800212_A/2013	KM350505	*Culex* sp.	2013	Brazil	Para
			BR/WALV/AR800213_A/2013	KM350506	*Culex* sp.	2013	Brazil	Para
			BR/WALV/AR800233_A/2013	KM350507	*Culex* sp.	2013	Brazil	Para
			TT/WALV/TR7904/2009	NC_023440	*Culex declarator*	2009	Trinidad and Tobago	na
	Blunevirus	Blueberry necrotic ring blotch virus (RNA 1)	US/BNRBV/2009	NC_016084	na	2009	USA	na
	Cilevirus	Citrus leprosis virus C	BR/CiLV-C/NA/2006	DQ157466	na	2006	Brazil	Sao Paulo
			BR/CiLV-C/2006	DQ352194	*Citrus sinesis var*	2006	Brazil	Sao Paulo
			PA/CiLV-C/NA/2006	DQ388512	*Citrus* spp.	2006	Panama	na
			BR/CiLV-C/SJRP/2011	KP336746	*Citrus sinesis var*	2011	Brazil	Sao Paulo
	Hifocrovirus	Hibiscus Fort Crockett virus	US/HFCKV/ACMB-1/2015	KX518774	*Paracoccus marginatus*	2015	USA	Texas
	Higrevirus	Hibiscus green spot virus	US/HGSV/ NC_016141/2009	NC_016141	*Citrus volkameriana*	2009	USA	na
Virgaviridae	***Furovirus***	Chinese wheat mosaic virus	CN/CWMV/ NC_002359/1997	NC_002359	wheat	1997	China	na
		Oat golden stripe virus	UK/OGSV/ NC_002358/1998	NC_002358	*Avena sativa* (oat)	1998	United Kingdom	na
		Soil-borne cereal mosaic virus	FR/EWMV/ NC_002351/1999	NC_002351	*Triticum aestivum* (wheat)	1999	France	na
		Soil-borne wheat mosaic virus	US/SBWMV/ NC_002041/1981	NC_002041	wheat	1981	USA	na
	Hordovirus	Barley stripe mosaic virus (RNA 1)	US/BSMV	NC_003469	barley	na	USA	na
	Peculovirus	Indian peanut clump virus	IPCV/ NC_004729	NC_004729	Groundnut	na	na	na
		Peanut clump virus	PCV/ NC_003672	NC_003672	na	na	na	na
	Pomovirus	Beet soil-borne virus	BSBV/ NC_003520	NC_003520	na	na	na	na
		Beet virus Q	BVQ/ NC_003510	NC_003510	*Beta vulgaris*	na	Germany	na
		Potato mop-top virus	PMTV/ NC_003723	NC_003723	na	na	Sweden	na

(continued on next page)

Table 1 *(continued)*

Family	Genus	Virus name	Strain Identification	Genbank number	Host source	Year of Isolation	Country of Isolation	State Province
	Tobravirus	Pea early browning virus	PEBV/NC_002036	NC_002036	na	na	United Kingdom	na
		Pepper ringspot virus	PepRSV/NC_003669	NC_003669	na	na	na	na
		Tobacco rattle virus	TRV/NC_003805	NC_003805	na	na	na	na
	Tomovirus	Cucumber green mottle mosaic virus	CGMMV/NC_001801	NC_001801	na	na	na	na
		Odontoglossum ringspot virus	ORSV/NC_001728	NC_001728	na	na	na	na
		Tobacco mosaic virus	TMV/NC_001367	NC_001367	na	na	na	na

na – not available.

clades (genera), designated Nelorpivirus and Sandewavirus. In the current study, 91 new negevirus isolates from mosquitos and phlebotomine sand flies collected in Brazil, Colombia, Nepal, Panama, Peru and the United States were sequenced. Four novel negevirus species were identified; these viruses cluster into two major phylogenetic groups, corresponding to the two proposed genera noted above. In addition, a fifth novel negevirus was isolated from a pool of 12 mealy bugs (*Paracoccus marginatus*), a hemipteran pseudococcid feeding on hibiscus plants. This virus is genetically distinct from the nelorpiviruses and sandewaviruses and is phylogenetically closer to plant viruses in the genera *Higrevirus*, *Blunevirus*, and *Cilevirus*. These findings confirm previous reports (Vasilakis et al., 2013; Kallies et al., 2014; Carapeta et al., 2015) that the negeviruses are genetically and evolutionarily related to several taxa of plant viruses.

2. Materials and methods

A total of 149 viruses or viral sequences were used in this study. Their taxonomic assignments, names and strain identifications, GenBank accession numbers, host source, and year and geographic locality of isolation are shown in Table 1. Fifty-eight of the viruses were previously characterized agents; information about them was obtained from GenBank and from earlier publications (Vasilakis et al., 2013; Kallies et al., 2014; Nabeshima et al., 2014). The remaining 91 viruses listed in Table 1 were obtained from virus repositories located and were sequenced at the University of Texas Medical Branch (UTMB) or at the Instituto Evandro Chagas (IEC), Ministry of Health, Ananindeua, Para, Brazil. Viruses from the IEC were isolated in that laboratory from insects collected during fieldwork in northern Brazil (Amazon Basin). Viruses from UTMB were obtained from samples collected during field studies in the USA or from material (viruses or insect pools) sent from other laboratories for diagnostic studies by staff of the World Reference Center for Emerging Viruses and Arboviruses. The methods used for isolation and preliminary identification of the latter 91 viruses were described in previous publications (Vasilakis et al., 2013; Auguste et al., 2014; Nunes et al., 2015a, 2015b).

2.1. RNA extraction

Virus stocks of the 91 viruses sequenced in our study were prepared in cultures of C6/36 cells and RNA was extracted as follows. For viruses sequenced at UTMB, RNA samples were prepared using the Qiamp RNA mini kit (Qiagen). The Brazilian viruses were prepared for RNA extraction using the Trizol reagent RNA purification method (Rio et al., 2010).

2.2. Comparative studies of negevirus stability

To compare the stability of a representative negevirus (Piura, strain EVG 7–47) with that of an alphavirus (Sindbis strain Eg 339) and a flavivirus (West Nile strain NY305-99), 1.0 ml of a high-titered stock of each virus was added to 9.0 ml of a 10% sucrose solution. The virus-sucrose solutions were held in an incubator at 28 °C and periodically sampled for infectivity. Samples (0.5 ml) of the virus-sucrose solutions were taken initially (day 0) and then daily for 7 consecutive days for the Sindbis and West Nile virus solutions. Aliquots of the Piura sample were taken for 7 consecutive days and then intermittently until day 76.

Viability (infectivity) of each virus was determined by titration in quadruplicate in 24-well microplate cultures of Vero (Sindbis and West Nile) or C6/36 cells (Piura). Serial ten-fold dilutions of each daily sample were prepared in phosphate-buffered (pH 7.4) saline solution, containing 10% fetal bovine serum. After mixing, four wells were inoculated with 100 uL or each dilution. After incubation for 2 h, maintenance medium was added to the cultures and each plate was

Table 2.
Genomic characteristics of novel insect-specific viruses based on each genome length, terminal UTR sequences (5′ and 3′), ORFs, Intergenic region, GC content and Poly A tail.

Virus and Strain	GenBank accesion no.	Genome length (bp)	Genomic region										
			5-UTR	ORF1 nt (aa)	%GC	Intergenic region 1	ORF2 nt (aa)	%GC	Intergenic region 2	ORF3 nt (aa)	%GC	3´- UTR	Poly A tail
Big Cypress													
BCNP 2–24	KX518750	9506	308	7002 (2333)	44.7%	NA	1386 (461)	43.3%	29	561 (186)	49.2%	360	Yes
BCNP 2–30	KX518751	9504	308	7002 (2333)	44.7%	NA	1386 (461)	43.3%	29	561 (186)	49.2%	359	Yes
BCNP 3–125	KX518754	9481	295	7002 (2333)	44.7%	NA	1386 (461)	43.3%	29	561 (186)	49%	348	Yes
BCNP 3–112	KX518752	9504	308	7002 (2333)	44.7%	NA	1386 (461)	43.3%	29	561 (186)	49%	358	Yes
BCNP 3–126	KX518755	9501	306	7002 (2333)	44.7%	NA	1386 (461)	43.3%	29	561 (186)	49%	357	Yes
BCNP 3–128	KX518756	9502	307	7002 (2333)	44.7%	NA	1386 (461)	43.3%	29	561 (186)	49%	357	Yes
BCNP 3–124	KX518753	9503	307	7002 (2333)	44.7%	NA	1386 (461)	43.3%	29	561 (186)	49%	358	Yes
Biratnagar													
Nepal 12-1	KX518757	9111	212	6798 (2265)	39.4%	18	1149 (382)	39.3%	162	585 (194)	44.6%	188	Yes
Nepal 24-2	KX518758	9104	203	6798 (2265)	39.4%	18	1149 (382)	39.1%	162	585 (194)	44.8%	189	Yes
Brejeira													
CoB 35	KX518759	9744	432	7014 (2337)	43%	29	1203 (400)	41.6%	141	624 (207)	50.3%	301	Yes
CoB 37	KX518760	9738	419	7014 (2337)	43%	29	1203 (400)	41.4%	141	624 (207)	50.3%	308	Yes
CoB 37d	KX518761	9748	431	7014 (2337)	43%	29	1203 (400)	41.4%	141	624 (207)	50.3%	306	Yes
BEAR803410	KX518762	9754	455	7011 (2336)	40.9%	28	1203 (400)	38.1%	117	624 (207)	46.2%	292	Not Found
BEAR803412	KX518763	9781	459	7011 (2336)	40.9%	28	1203 (400)	38.1%	117	624 (207)	46.2%	315	Yes
Cordoba													
Nepal 12-3	KX518772	7537	32	7035 (2344)	37.5%	NA	NA	NA	NA	NA	NA	NA	NA
Nepal 24-3	KX518773	7538	30	7035 (2344)	37.5%	NA	NA	NA	NA	NA	NA	NA	NA
GMC 30	KX518771	7515	29	7023 (2340)	36.7%	NA	NA	NA	NA	NA	NA	NA	NA
EVG 7-47	KX518767	7615	20	7023 (2340)	36.4%	NA	NA	NA	NA	NA	NA	NA	NA
EVG 7-41	KX518766	7584	28	7023 (2340)	36.4%	NA	NA	NA	NA	NA	NA	NA	NA
EVG 9-196	KX518768	7606	34	7023 (2340)	36.5%	NA	NA	NA	NA	NA	NA	NA	NA
EVG 9-226	KX518769	7601	28	7023 (2340)	36.5%	NA	NA	NA	NA	NA	NA	NA	NA
EVG 9-228	KX518770	7594	30	7023 (2340)	36.5%	NA	NA	NA	NA	NA	NA	NA	NA
CoB 37	KX518764	7598	28	7023 (2340)	35.9%	NA	NA	NA	NA	NA	NA	NA	NA
CoB 37d	KX518765	7501	28	7023 (2340)	35.9%	NA	NA	NA	NA	NA	NA	NA	NA
Hibiscus Fort Crockett													
ACMB-1	KX518774	7903	86	6957 (2318)	36.2%	NA	NA	NA	NA	NA	NA	NA	Not Found
Loreto													
3940-83	KX518775	9136	151	7014 (2337)	40%	33	1206 (401)	32.6%	19	642 (213)	39.3%	87	Not Found
Negev													
V 05087	KX518776	9522	230	7107 (2368)	48.4%	31	1203 (400)	46.8%	38	627 (208)	54.9%	276	Yes
Negwotan													
Nepal 12	KX518778	9424	144	6849 (2282)	48.1%	34	1197 (398)	47.1%	49	624 (207)	55.3%	237	Not Found
Nepal 12-2	KX518777	9546	238	6849 (2282)	48.3%	34	1202 (399)	47.2%	49	624 (207)	55.3%	265	Yes
Nepal 22	KX518779	9530	227	6849 (2282)	48.3%	34	1197 (398)	47.1%	49			260	Yes
Piura													
EVG 1-47	KX518788	10054	743	7011 (2336)	48.3%	42	1203 (400)	49.1%	140	618 (205)	53.6%	311	Yes
EVG 1-50	KX518789	10053	728	7011 (2336)	48.1%	42	1203 (400)	49.2%	140	618 (205)	54.2%	311	Yes
EVG 1-76	KX518790	10038	726	7011 (2336)	48.3%	42	1203 (400)	49.1%	140	618 (205)	53.6%	298	Yes

(continued on next page)

Table 2. *(continued)*

Virus and Strain	GenBank accesion no.	Genome length (bp)	Genomic region										
			5-UTR	ORF1 nt (aa)	%GC	Intergenic region 1	ORF2 nt (aa)	%GC	Intergenic region 2	ORF3 nt (aa)	%GC	3´- UTR	Poly A tail
EVG 2-7	KX518791	10036	728	7011 (2336)	48.2%	42	1203 (400)	49.1%	140	618 (205)	54.2%	294	Yes
EVG 2-86	KX518793	10047	724	7011 (2336)	48.2%	42	1203 (400)	48.9%	140	618 (205)	53.9%	309	Yes
EVG 2-8	KX518792	10034	723	7011 (2336)	48.2%	42	1203 (400)	48.9%	140	618 (205)	53.9%	297	Yes
EVG 3-129	KX518795	10040	727	7011 (2336)	48.3%	42	1203 (400)	49.2%	140	618 (205)	53.9%	299	Yes
EVG 3-97	KX518794	10039	726	7011 (2336)	48.2%	42	1203 (400)	49%	140	618 (205)	53.6%	299	Yes
EVG 4-9 (1)	KX518796	9930	620	7011 (2336)	48.3%	42	1203 (400)	49.1%	140	618 (205)	54%	296	Yes
EVG 4-9 (2)	KX518797	10037	729	7011 (2336)	48.3%	42	1203 (400)	49.1%	140	618 (205)	54%	294	Yes
EVG 7-47	KX518804	10039	728	7011 (2336)	48.2%	42	1203 (400)	49%	140	618 (205)	54%	297	Yes
EVG 5-53	KX518798	10054	740	7011 (2336)	48.3%	42	1203 (400)	49.3%	140	618 (205)	54%	300	Yes
EVG 7-35	KX518802	10042	729	7011 (2336)	48.4%	42	1203 (400)	49.1%	140	618 (205)	53.9%	299	Yes
EVG 7-19	KX518800	10040	724	7011 (2336)	48.2%	42	1203 (400)	49.1%	140	618 (205)	53.7%	302	Yes
EVG 7-1	KX518799	10041	726	7011 (2336)	48.2%	42	1203 (400)	49%	140	618 (205)	53.7%	301	Yes
EVG 7-94	KX518806	10040	724	7011 (2336)	48.2%	42	1203 (400)	49.1%	140	618 (205)	54.2%	302	Yes
EVG 7-95	KX518807	10046	728	7011 (2336)	48.3%	42	1203 (400)	49%	140	618 (205)	53.9%	304	Yes
EVG 7-78	KX518805	10039	727	7011 (2336)	48.2%	42	1203 (400)	49.1%	140	618 (205)	53.6%	301	Yes
EVG 7-29	KX518801	10039	728	7011 (2336)	48.4%	42	1203 (400)	49%	140	618 (205)	53.9%	297	Yes
EVG 7-41	KX518803	10041	728	7011 (2336)	48.2%	42	1203 (400)	49%	140	618 (205)	54%	299	Yes
EVG 9-196	KX518816	9942	724	7011 (2336)	48.2%	42	1203 (400)	49.1%	140	618 (205)	53.9%	206	Not Found
EVG 7-110	KX518808	7039	727	7011 (2336)	48.2%	42	1203 (400)	49.1%	140	618 (205)	54%	298	Yes
EVG 7-111	KX518809	10050	733	7011 (2336)	48.2%	42	1203 (400)	49%	140	618 (205)	54.2%	303	Yes
EVG 7-112	KX518810	1041	727	7011 (2336)	48.2%	42	1203 (400)	48.9%	140	618 (205)	54%	300	Yes
EVG 7-114	KX518811	10039	726	7011 (2336)	48.2%	42	1203 (400)	49%	140	618 (205)	53.7%	299	Yes
EVG 9-226	KX518817	10038	724	7011 (2336)	48.2%	42	1203 (400)	49%	140	618 (205)	53.7%	300	Yes
EVG 9-228	KX518818	10042	728	7011 (2336)	48.2%	42	1203 (400)	49%	140	618 (205)	53.7%	300	Yes
EVG 9-151	KX518813	10031	726	7011 (2336)	48.1%	42	1203 (400)	49.1%	140	618 (205)	54.2%	291	Yes
EVG 9-92	KX518812	10039	727	7011 (2336)	48.3%	42	1203 (400)	49.2%	140	618 (205)	53.9%	298	Yes
EVG 9-162	KX518814	10040	726	7011 (2336)	48.3%	42	1203 (400)	49.1%	140	618 (205)	53.9%	300	Yes
EVG 9-183	KX518815	10043	729	7011 (2336)	48.1%	42	1203 (400)	49.4%	140	618 (205)	54%	299	Yes
V 08449	KX518819	10028	719	7008 (2335)	48.2%	42	1203 (400)	48.5%	140	618 (205)	52.9%	298	Yes
V 08458	KX518820	10010	706	7008 (2335)	48.2%	42	1203 (400)	48.5%	140	618 (205)	52.9%	293	Yes
BCNP 2-60	KX518780	10048	730	7011 (2336)	48.2%	42	1203 (400)	49.1%	140	618 (205)	54%	304	Yes
BCNP 3-126	KX518781	10030	715	7011 (2336)	48.3%	42	1203 (400)	48.9%	140	618 (205)	54%	301	Yes
BCNP 3-148	KX518785	10045	726	7011 (2336)	48.2%	42	1203 (400)	49.1%	140	618 (205)	53.9%	305	Yes
BCNP 4-73	KX518783	10043	726	7011 (2336)	48.3%	42	1203 (400)	49.2%	140	618 (205)	54.2%	303	Yes
BCNP 3-141	KX518782	10046	726	7011 (2336)	48.3%	42	1203 (400)	49.2%	140	618 (205)	53.9%	306	Yes
BCNP 3-134	KX527894	10046	726	7011 (2336)	48.3%	42	1203 (400)	49.2%	140	618 (205)	53.9%	306	Yes
CoR 29	KX518786	10050	730	7011 (2336)	48.7%	42	1203 (400)	48%	140	618 (205)	53.9%	306	Yes
CoR 33	KX518787	10040	730	7011 (2336)	48.7%	42	1203 (400)	48%	140	618 (205)	53.9%	296	Yes
San Bernado													
CoB 52	KX518821	9651	412	7116 (2371)	46.9%	24	1191 (396)	45.9%	64	579 (192)	50.4%	265	Yes
CoG 72 b	KX518822	9657	414	7116 (2371)	46.9%	24	1191 (396)	45.9%	64	579 (192)	50.6%	269	Yes
Wallerfield													
CoB 92	KX518836	8989	68	6723 (2240)	34.7%	29	1269 (422)	32.2%	121	601 (199)	42.4%	182	Yes
CoB 37d	KX518835	8995	65	6723 (2240)	34.7%	29	1269 (422)	32.5%	121	601 (199)	42.2%	191	Yes
GAM 058A_2	KX518837	8995	63	6723 (2240)	34.3%	29	1269 (422)	32.7%	121	601 (199)	42.6%	188	Yes
GAM 195	KX518839	8978	61	6723 (2240)	34.3%	29	1269 (422)	32.7%	121	601 (199)	42.8%	175	Yes
GAM 058A	KX518838	8994	63	6723 (2240)	34.3%	29	1269 (422)	32.7%	121	601 (199)	42.6%	189	Yes
BCNP 2-24	KX518827	8923	59	6723 (2240)	34.5%	29	1266 (421)	32.9%	121	601 (199)	42.3%	135	Not Found
BCNP 2-15	KX518826	8995	68	6723 (2240)	34.5%	29	1266 (421)	32.9%	121	601 (199)	42.3%	188	Not Found
BCNP 2-5	KX518825	8940	66	6723 (2240)	34.5%	29	1266 (421)	32.9%	121	601 (199)	42.3%	135	Not Found

(continued on next page)

Table 2. *(continued)*

Virus and Strain	GenBank accesion no.	Genome length (bp)	Genomic region										
			5-UTR	ORF1 nt (aa)	%GC	Intergenic region 1	ORF2 nt (aa)	%GC	Intergenic region 2	ORF3 nt (aa)	%GC	3´- UTR	Poly A tail
BCNP 2-4	KX518824	8934	60	6723 (2240)	34.5%	29	1266 (421)	32.9%	121	601 (199)	42.3%	135	Not Found
BCNP 2-30	KX518828	8934	60	6723 (2240)	34.5%	29	1266 (421)	32.9%	121	601 (199)	42.3%	135	Not Found
BCNP 2-3	KX518823	8996	67	6723 (2240)	34.5%	29	1266 (421)	32.9%	121	601 (199)	42.3%	190	Yes
BCNP 3-112	KX518829	8989	61	6723 (2240)	34.5%	29	1266 (421)	33%	121	601 (199)	42.3%	189	Yes
BCNP 3-128	KX518830	8989	61	6723 (2240)	34.5%	29	1266 (421)	32.9%	121	601 (199)	42.3%	189	Yes
BCNP 4-205	KX518831	8986	60	6723 (2240)	34.5%	29	1266 (421)	32.9%	121	601 (199)	42.3%	187	Yes
BEAR803412_2	KX518832	8965	53	6723 (2240)	34.7%	29	1269 (422)	31.9%	121	601 (199)	42.8%	170	Not Found
BEAR803418	KX518833	8970	53	6723 (2240)	34.7%	29	1269 (422)	31.8%	121	601 (199)	43.1%	175	Yes
BEAR804033	KX518834	8970	53	6723 (2240)	34.7%	29	1269 (422)	32.1%	121	601 (199)	43.2%	175	Yes

examined daily for evidence of viral cytopathic effect (CPE) under an inverted tissue culture microscope for 7 days. The tissue culture infectious dose$_{50}$ endpoint (TCID$_{50}$) for each sample was calculated, using the Reed and Muench formula, as described by Hsiung (Hsiung, 1994).

2.3. Genome sequencing and assembling

Genome sequences were obtained for the virus isolates, using the parallel sequencing and the Ion Torrent semi-conduction methods implemented in the Illumina HiSeq, MiSeq and Ion Torrent instruments, respectively. Briefly, for the Ion Torrent, a complementary DNA was generated from the viral RNA, using random hexamer primers, followed by library preparation, emulsion PCR, and sequencing as previously described (Rothberg et al., 2011). For the HiSeq Illumina method, viral RNA (~0.9 μg) was fragmented by incubation at 94 °C for eight minutes in 19.5 μl of fragmentation buffer (Illumina 15016648). A sequencing library was prepared from the sample RNA using an Illumina TruSeq RNA v2 kit following the manufacturer's protocol using the 2×50 paired-end protocol. For the MiSeq illumine method. A highly efficient protocol was made including the rRNA removal (Ribo-Zero technology) followed by a rapid, ligation-free cDNA synthesis procedure for preparing directional RNA-seq libraries (ScriptSeq v2 technology) as previously described (Pease and Sooknanan, 2012). The data generated by MiSeq and Ion Torrent were assembled using a *de novo* strategy implemented in the Mira Software (Chevreux et al., 2004). The resulting contigs were compared to the Genbank database using Blastx (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web & PAGE_TYPE=BlastHome) to identify candidate genomes to be used as references. Afterward, the raw data (reads) were mapped against the chosen sequence using the Newbler v.2.6 software. Relevant parameters for the mapping assembly were set as follows: minimum length =20 bp, all contig threshold =100, large contig threshold =200, minimum overlap length =40, minimum overlap identity =70%, k-mer =12 (seed step), and k-mer =16 (seed length). Visual inspection and annotation of the resulting assembled sequence were performed using the Geneious v 9 (http://www.geneious.com, Kearse et al., 2012). Reads generated by the HiSeq platform in fastq format were quality-filtered, and any adapter sequences were removed, using Trimmomatic software (Lohse et al., 2012). The *de novo* assembly program ABySS (Simpson et al., 2009) was used to assemble the reads into contigs, using several different sets of reads, and *k* values from 20 to 40. The longest contigs expressing open reading frames were selected and reads were mapped back to the contigs using bowtie2 (Langmead and Salzberg, 2012), and visualized with the Integrated Genomics Viewer (Robinson et al., 2011) to verify that the assembled contigs were correct. Pre-filtering of reads to remove host sequence enhanced the assembly process. Assembly was carried out using a fasta file of *Aedes albopictus* sequences to remove host DNA from the assembly thus reducing the number of contigs present.

2.4. Genome characterization

Virus genomes were characterized based on their genetic traits such as genome size, genome organization (potential Open Reading Frames-ORF) and encoded proteins. Briefly, all genomes were inspected for the existence of ORFs and annotated, using Geneious v 9. Protein domains were inspected by the InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan5/) and implemented in Geneious work package. Conserved motifs, cysteine residues, and N potential glycosylation sites were also assessed using both Geneious and the NetNGlyc V 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/).

2.5. Phylogenetic analyses

The genomes of our 91 negevirus isolates were compared with those

of other genetically related viruses associated with arthropods and plants (Table 2). The amino acid sequences referred to the virus methyl transferase (VMTase), viral helicase (VHEL) and RdRp domains and were individually aligned using the Promals3D webserver (Pei et al., 2008). The sequence ambiguities were removed, using the Gblock software v. 0.91b (Talavera and Gastresana, 2007). The ProtTest v3.4 software (Darriba et al., 2011) was used to determine the NNI+SPR branch swapping as the best-fit model of protein evolution for the dataset. Outcomes alignments were visually inspected using the Geneious software v. 9, concatenated and used to construct a maximum likelihood phylogenetic tree with the PhyML v. 3.3 (Guidon et al., 2010). A bootstrap analysis was performed using 1,000 replicates (reliability value of 95%) to provide confidence on phylogenetic groupings. All trees are mid-point rooted.

2.6. Genetic variability and genetic relatedness

The genetic variability among the various viruses were assessed using the pairwise multiple sequencing alignment method implemented in the Geneious v. 9. The genetic distances, within and between groups, were determined based on groups designated as clades by phylogenetic analysis; the distances were expressed as percentage values. Additional analysis, using the boxplot method (Williamson et al., 1989), was applied when a given group was considered when its intra-group genetic distance was at least four-fold lower than the mean of the genetic distance in comparison to other groups (intergroup genetic distance). However, regardless the rate variation this approach has been largely used for several viruses and reinforces the phylogeny giving more confidence to our conclusions.

3. Results

3.1. Genome characterization

Nearly complete genome sequences were obtained the 91 virus isolates with a mean coverage of 84 times and quality values above 20. The genome sizes ranged from 7.039 nt to 10.054 nt in length (Table 2). Using the BlastN search tool (www. http://blast.ncbi.nlm.nih.gov/Blast.cgi), all of the generated sequences closely matched with previously published negeviruses genomes (Vasilakis et al., 2013; Kallies et al., 2014; Auguste et al., 2014; Nabeshima et al., 2014). Predictions of GC contents and open reading frames (ORFs) were made using the Geneious software. GC contents were estimated in 42.25% (40% to 43%) and one or three ORFs were, in general, observed in the recovered genomes. Potential genes were also detected (Table 3).

3.1.1. Prediction of proteins, domains and conserved motifs

The majority of viruses showed evidence of three proteins encoded by the viral mRNA in distinct ORFs. The largest ORF (ORF1) encodes the viral polymerase protein, while ORF2 (medium) and ORF3 (small) encode to the glycoprotein and membrane proteins, respectively. Conserved regions were observed in ORF 1 and were related to the high conserved domains of VMeTrfase (PFAM: accession number: PF01660) (Rozanov et al., 1992), Ribosomal RNA methyltransferase FtsJ domain [rRNA MeTrfase FtsJ dom], (PFAM: accession number: PF01728) (Bugl et al., 2000), VHEL (PFAM: accession number: PF01443) (Koonin and Dolja, 1993), and RdRp (PFAM: accession number: PF00978) (Lemm and Rice, 1993). In ORF2, the domain DiSB, a putative virion glycoprotein, N-terminal (accession number: IPR032433), and in ORF3, the SP24 (accession number: PF16504) domain, described and identified as a functional membrane protein domain, was found. Fig. 1 graphically represents the genome organization of the studied viruses according to their number of ORFs and predicted protein domains.

3.1.2. Cysteine (Cys residues and potential N-linked glycosylation [NGlyc]) sites in glycoproteins

Based on the Geneious software analysis, the presence of UGU/UGC codons that predict for Cys residues were assessed and a number of Cys residues for distinct arthropod-restricted viruses are presented as follows: Brejeira virus (BREV) and Loreto virus (LORV) (8 *Cys*); Dezidougou virus (DEZV), Cordoba virus (CDBV), Hibiscus Fort Crockett virus (HFCKV), Piura virus (PIUV), Wallerfield virus (WALV) (9 *Cys*); Gautanap virus (GANV), Ngewotan virus (NWTV), San Bernardo virus (SBDV), Tanai virus (TANV) (10 *Cys*); Santana virus (SANV) (11*Cys*); Biratnagar virus (BIRV) and Negev virus (NeGV) (12 *Cys*); and Big Cypress virus (BCPV) (13 *Cys*). For N*Glyc* sites, the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/) predicted different numbers of potential N*Glyc* sites were also observed for the studied viruses as follows: BCPV (N*Glyc*=1); DEZV, LORV, PIUV, SANV and SBDV (N*Glyc*=2); BIRV, BREV, GANV and TANV (N*Glyc*=3); NWTV and WALV (N*Glyc*=4); NEGV (N*Glyc*=5). No N*GLyc* sites were observed to HFCKV.

3.2. Genetic variability

Intra- and intergroup genetic differences were estimated, based on the nucleotide and amino acid sequence alignments; they demonstrated the existence of at least two major groups: group I including PIUV, BREV, NEGV, NWTV, SBDV, BCPV, LORV and CDBV, while group II was represented by WALV, GANV, DEZV, SANV, BIRV and TANV. The genetic similarity among these negeviruses ranged from 47.8% to 70.8% and from 14.8% to 78.6% at nucleotidic and amino acid sequences levels, respectively (Table 3). Additional analyses using the box plot method demonstrated the distinct graphical disposition for each virus group (here assumed to be virus species) with absence of overlapping groups between the inter-and intra-group areas (CI > 95%; $p < 0.005$). The box plot graphic representation for the two major groups and individual virus species is shown in Figs. S1a and b.

3.3. Phylogenetic analysis

The unrooted phylogenetic trees based on concatenated amino acid sequences (773 aa), obtained after Gblock for methyl transferase (VMTase), viral helicase (VHEL) and RdRp domains, divided the negeviruses into two major groups, hereafter named the *Nelorpivirus* and *Sandewavirus* clades, both supported by 100% bootstrap values (Fig. 2). The *Nelorpivirus* clade clustered eight well-supported genetic subclades (bootstrap value > 98%) which includes five previously described negeviruses (PIUV, BREV, NEGV, NWTV, and LORV) as well as three new viruses designated as SBDV, CDBV, and BCPV, which are more closely related genetically to two arthropod-borne plant viruses, HGSV (*Higrevirus* genus) (Pascon et al., 2006) and CiLV (*Cilevirus* genus) (Melzer et al., 2013). The *Sandewavirus* group also includes six well-supported subclades (100% bootstrap values) represented by WALV, GANV, DEZV, SANTV, TANV and the newly described BIRV, which are more closely related to the plant virus Blueberry necrotic ring blotch virus (BNRBV) (*Blunevirus* genus) (Quito-Avila et al., 2013). (Fig. 2).

Additional analyses based on nucleotide sequences of the entire ORFs of all included viruses revealed other branch ramifications for some groups of negeviruses. For example, viruses in the PIUV group can be subdivided into two subgroups: subgroup I, including isolates from Peru, Mexico, Colombia and USA; and subgroup II, including two isolates from the USA (bootstrap > 90%) (Fig. 3A).

The NWTV group also separated into two subgroups. Subgroup I includes three isolates from Nepal (NP/NWTV/Nepal_12/1984, NP/NWTV/Nepal_12_2/1984, NP/NWTV/Nepal_22/1984), and group II consists of a single isolate (ID/NWTV/JKT_9982/1981) from Indonesia (Fig. 3B).

The BREV group was composed by three distinct subgroups

Table 3
Genetic similarity of the negeviruses included in this study, based on concatenated nucleotide and amino acid sequences related to the predicted ORFs.

	Aminoacid similarity													
	PIUV	**BREJV**	**NEGV**	**NWTV**	**SBDV**	**BCPV**	**LORV**	**CDBV**	**WALV**	**GANV**	**DEZV**	**SANV**	**BIRV**	**TANV**
PIUV		78.6	56.5	56.6	52.4	43.3	38	26	18.7	18.8	18.5	17.3	18.1	17.1
BREJV	70.8		55.7	55.7	52.1	42.3	37.4	26.2	18.9	18.5	18.4	17.7	18.5	17
NEGV	57	56.6		75.5	53.7	43.7	39	27.2	18.5	18.8	18.1	18.2	17.9	17.5
NWTV	56.9	56.3	69.6		53.3	43	38.7	26.8	18.2	18.6	18.2	18.2	17.7	17.1
SBDV	54.5	54.8	55.8	54.9		43.9	38.7	27.1	19.5	18.9	18	17.8	18.5	17.7
BCPV	48.9	48.3	50.8	49.4	49.6		38.7	26.5	18.3	18.5	18.4	18.2	18	18.4
LORV	45.7	46.9	46.6	46.2	45.4	45.1		26.4	18.9	18.4	19	18.6	17.7	18
CDBV	40.1	41.6	40.5	40.8	40.3	39.8	42.3		16.2	15.9	15.9	15.2	15.2	14.9
WALV	33.2	34.5	33.5	32.8	32.5	33.3	34.9	34.9		47.5	56.1	45.2	50.4	43.9
GANV	33	35.1	34	33.6	33.2	33.1	35.6	35.1	55.8		49.3	43.2	45.1	40.4
DEZV	33.3	34.4	33.5	33.5	32.8	32.8	34.5	34.1	60.1	55.5		47.7	53.3	44.4
SANV	32.5	33.5	32.8	32.8	32.5	32.8	33.7	32	52.1	50.6	52.9		50.9	40.1
BIRV	33.2	33.6	32.9	32.8	32.7	32.1	33.8	32.7	54.3	50.9	55.4	54.4		42.8
TANV	31.3	33	31.4	31.3	30.9	31.7	33.2	33.7	51.2	48.8	51.3	48.3	47.8	

Nucleotide similarity.

(Fig. 3C). Subgroup I is composed of eight isolates from Canaã dos Carajás, Brazil (BR/BREV/AR800208-B/2013, BR/BREV/AR800212-B/2013, BR/BREV/AR800213-B/2013, BR/BREV/AR800214/2013, BR/BREV/AR800224/2013, BR/ BREV/AR800227/2013, BR/BREV/AR800233-B/2013, BR/BREV/AR800239/2013). Subgroup II is composed of two isolates from Trairão, Brazil (BR/BREV/BEAR803410/2005, BR/BREV/BEAR803412/2005. Subgroup III includes three isolates from Colombia (CO/BREV/CoB_35/2013, CO/BREV/CoB_37(a)/2013, CO/BREV/CoB_37d(a)/2013).

The CDBV group clade also exhibited two phylogenetic subgroups (Fig. 3D). Subgroup I includes eight isolates from the USA and Colombia (US/CDBV/ EVG_9-226(b1)/2013, US/CDBV/ EVG_9-228(b)/2013, CO/CDBV/CoB_37(b)/2013, CO/CDBV/CoB_37d(c)/2013), (US/CDBV/ EVG_9- 196(b)/2013, US/CDBV/US/EVG_7-47(2)/2013, US/CDBV/EVG_7- 41(b)/2013 and US/CDBV/GMC_30/2013). Subgroup II is composed of two isolates from Nepal (NP/CDBV/Nepal_12-3/1984 and NP/CDBV/Nepal_24-3/1984).

In the case of the WALV group, six phylogenetic subgroups are well supported by bootstrap values over 95% and are represented as follows (Fig. 3E). Subgroup I is composed by nine isolates from Big Cypress Nature Preserve (BCNP-USA). Subgroup II consists of one isolate from Trinidad and Tobago (TT/WALV/ TR7904/2009). Subgroup III is represented by three isolates from Panama (PA/WALV/ GAM_058A/2013, PA/WALV/GAM_195/2013, and PA/WALV/ GAM_058A_2/2013). Subgroup IV includes two Colombian isolates CO/WALV/CoB_92/2013 and CO/WALV/ CoB_37d_b/2013). Subgroup V is represented by four Brazilian isolates BR/WALV/ AR800208_A/2013, BR/WALV/ AR800212_A/2013, BR/WALV/AR800213_A/2013, and BR/WALV/ AR800233_A/2013. Subgroup VI is composed of three other isolates also obtained in Brazil. The remaining viruses did not show potential subgroups.

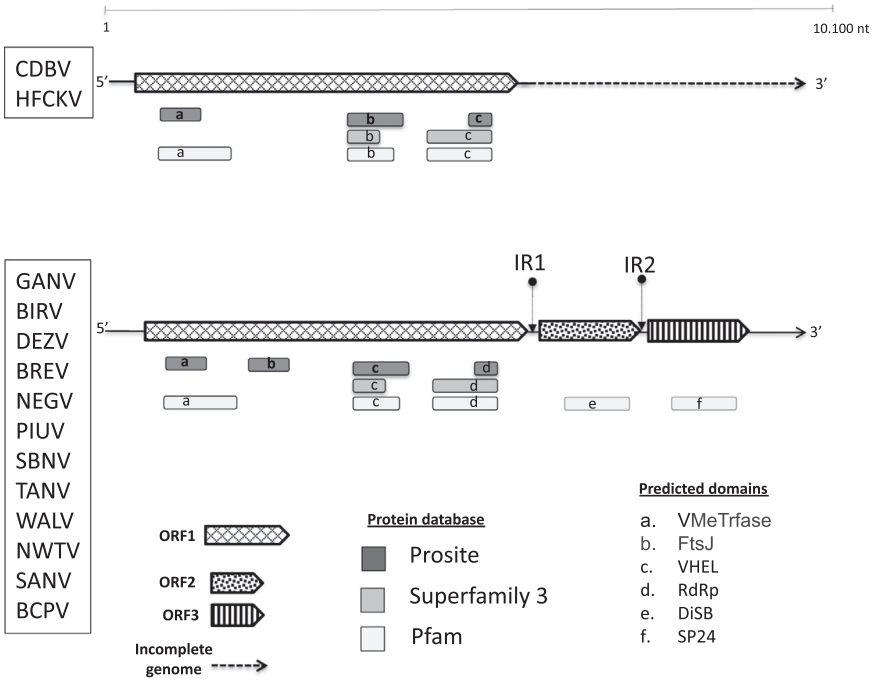


Fig. 1. Graphic representation of the genome organization of the open reading frame ORFs, proteins and conserved domains observed in the following groups of insect viruses. **(a)** CDBV (Cordoba virus) and HFCKV (Hibiscus Fort Crockett virus) **(b)** GANV (Gautanap virus), BIRV (Biratnagar virus), DEZV (Dezidougou virus), BREV (Brejeira virus), NEGV (Negev virus), PIUV (Piura virus), SBNV (San Bernardo virus), TANV (Tanay virus), WALV (Wallerfield virus), NWTV (Ngeowtan virus), SANV (Santana virus), BCPV (Big Cypress virus). IR1: Intergenic region 1, IR2: Intergenic region 2.

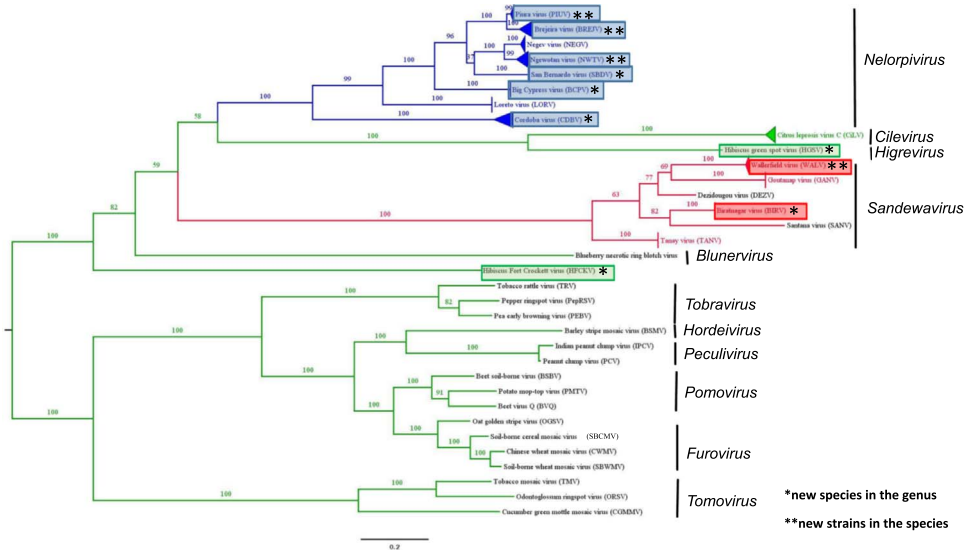


Fig. 2. Phylogenetic relationship of negeviruses and plant viruses based on the Metyl transferase, Helycase, and RdRp amino acid concatenated sequences (~ 780 aa in length). Branches in red and blue represent the two major insect-specific virus groups (*Nelorpivirus* and *Sandewavirus*). Branches in green correspond to plant viruses. Bootstrap values are placed over each node of the tree.

3.4. Virus stability

Because of the genetic similarity of the negeviruses to some plant viruses, an experiment was performed to compare the stability (infectivity) of a representative negevirus (Piura) held at 28 °C in a 10% sucrose solution with the stability of a representative alphavirus (Sindbis) and flavivirus (West Nile) held under the same conditions. The results are shown in Fig. 4. The initial infectivity titer of the WNV-sucrose solution was $10^{9.0}$ $TCID_{50}$/ml; however, it rapidly declined and no virus was detected by day 4. The Sindbis-sucrose solution showed a similar but slightly longer survival pattern; the initial virus titer was $10^{7.8}$ $TCID_{50}$/ml, but no infectious virus was detected after day 6. In contrast, the initial titer of Piura-sucrose solution was $10^{10.0}$ $TCID_{50}$/ml, but it declined very slowly over the next 76 days. On day 76, the titer of infectious Piura virus was still $10^{5.0}$. $TCID/_{50}$.

4. DISCUSSION

At present, our knowledge of the ecology of viruses in the negevirus taxon is still quite limited (Vasilakis et al., 2013; Kallies et al., 2014; Auguste et al., 2014; Nabeshima et al., 2014; Carapeta etal., 2015; Kawakami et al., 2015), but several observations can be made. First, negeviruses are very widely distributed geographically and occur in both tropical and temperate regions of the world. Isolates of these viruses have been cultured from field-collected mosquitoes (*Culicidae*) and phlebotomine sandflies (*Phlebotominae*) orginating from North and South America, Europe, the Middle East, Asia and Pacific islands. Based on their diversity and wide gegraphic distribution, it seems likely that additional negevirus types exist and will be found.

Second, negeviruses have a wide host range among Diptera. To date, the reported negeviruses have only been associated with mosquitoes and sandflies (Vasilakis et al., 2013; Kallies et al., 2014; Auguste et al., 2014; Nabeshima et al., 2014; Carapeta etal., 2015; Kawakami et al., 2015); however, other types of insects may also be naturally infected with these viruses. Most of the currently known negeviruses were detected during arbovirus surveillance studies; but arbovirologists generally collect and test only hematophagous insects. Consequently, the total number of negevirus types and their host associations may be much greater. For example, just within the family Culicidae, nine different genera of mosquitoes (*Aedes, Culex, Anopheles, Armigeres, Psorophora, Uranotaenia, Deinocerites, Wyeomyia* and *Trichoprosopon*) have been found naturally infected with negeviruses. Given their world-wide distribution and their broad host range among mosquitoes, it seems likely that negeviruses infect other types of insects as well.

Third, negeviruses replicate to high titers (up to 10^{10} PFU/ml) in some mosquito and sandfly cell lines (Vasilakis et al., 2013 and Fig. 4). Their ease of replication in these insect cell lines indicates that negeviruses infect and propagate in their insect hosts and are not just surface or gut contaminants picked up during culture or metagenomics studies. In contrast, negeviruses appear unable to infect laboratory animals (mice and hamsters) or a variety of vertebrate cell lines. These observations imply that mosquitoes and sandflies do not acquire these viruses by feeding on viremic vertebrates, as is the case with true arboviruses.

Based on their genomic organization and phylogeny, the negeviruses are most closely related to plant viruses in the genera *Cilevirus, Higrevirus* and *Blunervirus* (Pascon et al., 2006; Melzer et al., 2013; Quito-Avila et al., 2013; Carapeta el al, 2015). Viruses in these genera and the related family *Virgaviridae* (Adams et al., 2009) are transmitted to plants by a variety of modes, including mites, soil nematodes, seeds and protozoa (*Polymyxa*) (King et al., 2012). How negeviruses are acquired by mosquitoes and sandflies in nature is unknown. The recent finding of a negevirus in larval *Aedes* mosquitoes in Japan (Kawakami et al., 2015) is suggestive of pre-adult acquisition and many insect-specific flaviviruses are known to be transovarially (vertically) transmitted (Bolling et al., 2015; Tesh et al., 2016). However, the genetic similarity of negeviruses with some plant viruses suggests that plants may be involved in the natural transmission cycle. This could occur through horizontal transmission involving plant floral or extra-floral nectaries, as mosquitoes and sandflies are known to obtain essential carbohydrates from plants (Foster, 1995; Schlein and Jacabson, 1999). Alternatively, larval forms of these insects could opportunistically acquire infection when they feed upon the microbial fauna and associated plant detritus which many species are dependent upon (Fish and Carpenter, 1982; Merritt et al., 1992; Guzman and Tesh, 2000). The observed stability of Piura virus in sucrose solution (Fig. 4) is analogous to that of tobacco mosaic virus *(Tobamovirus: Virgaviridae*), which remains viable for years in the dried leaves and stems of infected plants (Lazarowitz, 2001). Another possible route of horizontal transmission could be facilitated by aquatic mites, which parasitize adult mosquitoes when they emerge from pupae (Mullen, 1975). Future laboratory and field experiments should reveal the mode of transmission for negeviruses and their maintenance in nature.

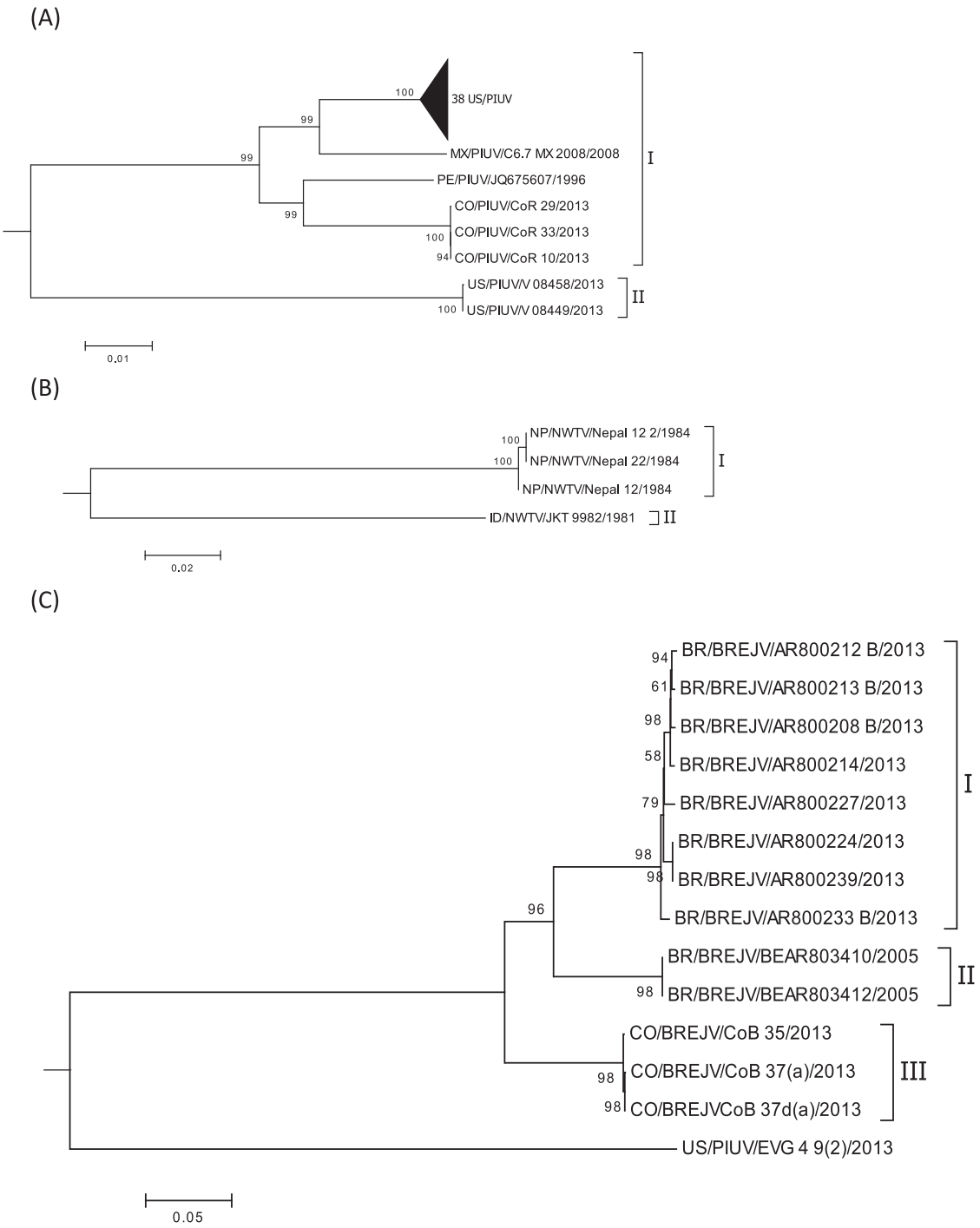
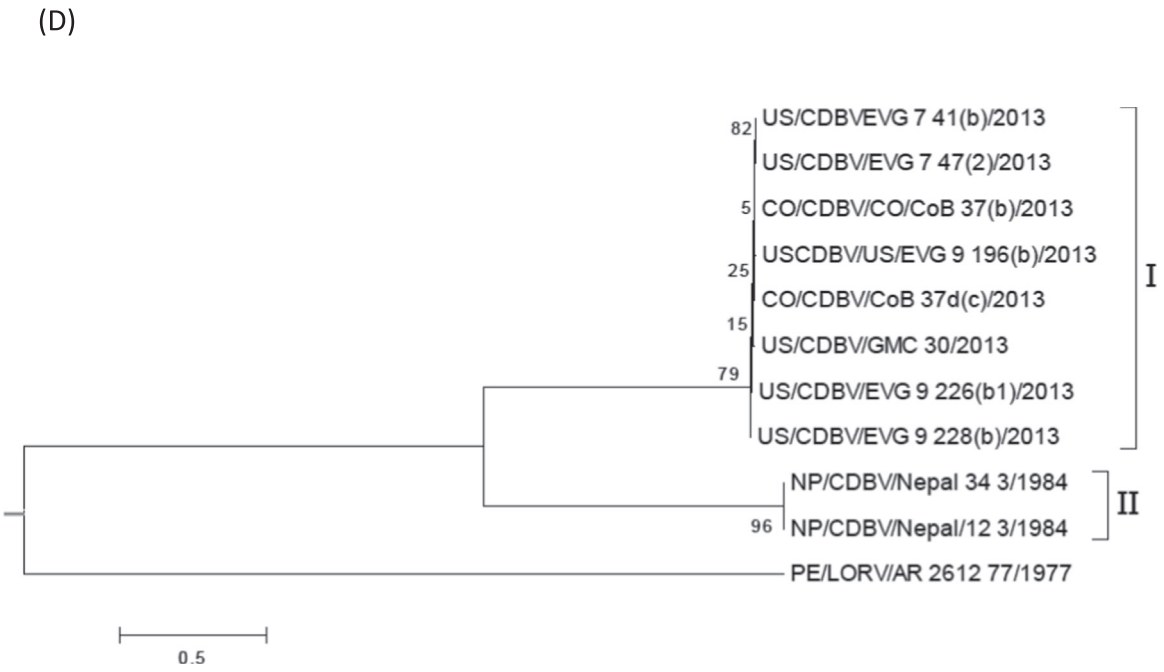


Fig. 3. Nucleotide sequence analyses of the entire ORFs of selected negevirus groups showing branch ramifications and subgroups. Fig. 3**A.** Piura virus (PIUV); Fig. 3**B.** Ngewoton virus (NWTV); Fig. 3**C.** Brejeira virus (BREJV); Fig. 3**D.** Cordova virus (CDBV); Fig. 3**E.** Wallerfield virus (WALF) groups. Roman numbers (I to VI) represent the genetic subgroups for each virus group. Numbers over each node represent the bootstrap value supports. Scale bars over each tree represent the number of nucleotide substitutions/site.

If these Diptera acquire negeviruses by feeding or vertical transmission, then what is the insects' role in the ecology of these viruses? Do mosquitoes and sandflies serve as vectors and transmit negeviruses to plants during sugar feeding, or are they simply dead-end hosts for the viruses? Do negeviruses also infect plants and produce some yet unrecognized plant disease, or are they simply maintained in a silent cycle in their insects hosts? Since negeviruses are part of the mosquito and sandfly microbiome, does infection have deleterious effects on the insect host or alter its vector competence for certain protozoal or arboviral pathogens of humans? Much remains to be learned about this ubiquitous and potentially important group of insect viruses.

Current knowledge of the molecular biology and evolution of the negeviruses is also still quite limited. Here, we have performed a comprehensive analysis of the genome organization and phylogeny of this arthropod-restrictive virus taxa, including 91 novel negevirus isolates obtained from distinct geographic areas around the world. Genome organization of the new viruses was similar to that observed in previous studies for negeviruses (Vasilakis et al., 2013; Kallies et al., 2014; Auguste et al., 2014; Nabeshima et al., 2014; Carapeta et al., 2015; Nunes et al., 2015a, 2015b), demonstrating that these agents are single-stranded RNA viruses with at least three ORFs related to the main virus protein: polymerase, glycoprotein and membrane protein.



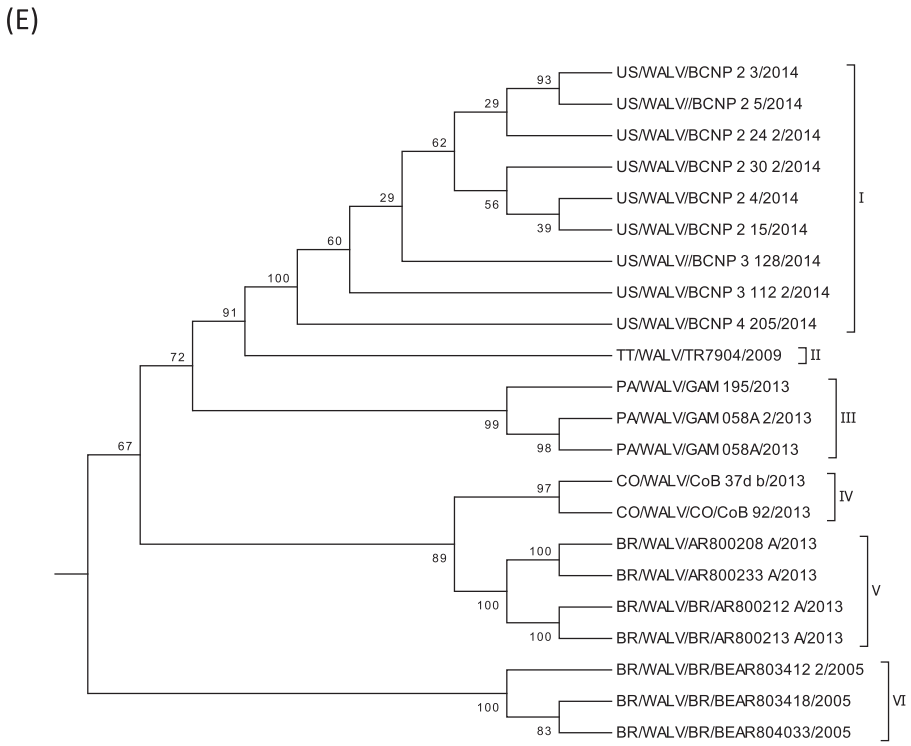


Fig. 3. (*continued*)

In the present study, we isolated and characterized four additional new negeviruses detected in mosquitoes and phlebotomine sand flies from Brazil, Colombia, USA, and Nepal.

In light of the observed size heterogeneity among the 91 studied viruses, one possible explanation for the differences could be due to the lack of complete sequences in the 5′ and 3′ ends, or it could be related to specific differences in size usually observed in untranslated terminal regions, such as have been reported for yellow fever virus (YFV) *(Flaviviridae, Flavivirus*). These differences in genome size of YFV were clearly associated with the geographic area where the isolates were obtained (e.g. West Africa, Central Africa, and Americas) (Bryant et al., 2005). However, for *Negevirus*, this kind of genome fingerprint was not observed.

Cysteine (*Cys*) residues and potential glycosylation sites (*Glyc*) have been associated with protein secondary structure maintenance and virus pathogenicity or ability to infect a given host, respectively (Lipton et al., 2002; Murrell et al., 2004). In the case of negeviruses, distinct numbers of *Cys* residues were found in glycoproteins. It suggests that virus species could possess different conformational three-dimensional (3D) structures for glycoproteins, possibly related to a given receptor in the host cell. In the case of potential glycosylation sites, differences in number were also found. Although, *Cys* residues and N-linked *Glyc* sites have a direct influence on virus attachment and pathogenicity properties, further studies on these mechanisms are necessary to

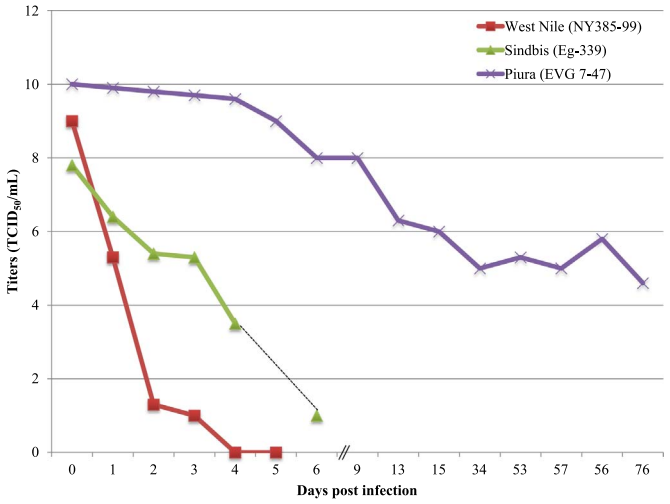


Fig. 4. Graphic representation of the stability of Piura, Sindbis and West Nile viruses in a 10% sucrose solution and maintained a 28 °C. Viruses are represented by different colored lines. Traced line indicates no sample tested. Titers are expressed as $TCID_{50}$/ml.

understand how the insect-specific viruses infect their host cells.

The genetic variability in RNA viruses, as well as genetic relationships, and evolutionary aspects have been extensively studied for several arboviruses, such as dengue, yellow fever, West Nile, and Chikungunya viruses (Allicock et al., 2012; Weaver and Vasilakis, 2009; Nunes et al., 2012, 2015a, 2015b; Volk et al., 2010) 46–50), and more recently for Zika virus (Rodrigues Faria et al., 2016; Lanciotti et al., 2016). For the *Negevirus* taxon, few analyses have been performed. In this work, the assessment of the nearly complete genomes of 91 different viruses confirms the presence of at least two distinct phylogenetic groups (*Nelorpivirus* and *Sandewavirus*), as previously described (Kallies et al., 2014).

Our analyses identified the presence of four new negevirus species: three belonging to the *Nelorpivirus* genus, namely SBDV, BCPV and CDBV isolated in Nepal, USA, and Colombia, respectively. A fourth new virus species in the genus *Sandewavirus*, designated Biratnagar virus (BIRV), isolated in Nepal, was described (Fig. 3). Distinct and highly supported phylogenetic subgroups were also observed for BREV (two subgroups), CDBV (three subgroups) and WALV (six subgroups). Interestingly for the species BREV and WALV, most subgroups separated according to the geographic location from where the isolates were made, suggesting that geography may play an important role in virus sub-speciation.

Evolutionary aspects of this taxon of insect-specific viruses are still unclear. In this work, we observed the presence of two insect-restricted negevirus groups (*Nelorpivirus* and *Sandewavirus* genera) that are genetically related to plant viruses from the genera *Blunevirus, Higrevirus* and *Cilevirus* (Pascon et al., 2006; Melzer et al., 2013; Quito-Avila et al., 2013; Carapeta et al., 2015) (Fig. 3). Narrow temporal window and the existence of few previous negevirus sequences composing the dataset must be considered as a limitation for producing reliable chronological dating, and for this reason no further molecular clock analysis was performed. However, the existence of two distinct well supported branches (bootstrap values of 100%) suggested a different origin for the two groups of negeviruses, since the sandewaviruses are more related to BNRBV and HFCV and the nelorpiviruses more closely related to higreviruses and cileviruses (Fig. 3). Nonetheless, it is noteworthy that several of the nelorpiviruses and sandewaviruses share the same host and geographic location. But as noted before, some of the negeviruses are extremely diverse and are worldwide in their geographic distribution.

In addition to new negeviruses, we also described Hibiscus Fort Crocket virus (HFCKV), which was isolated from a pool of 12 mealy bugs (*Paracoccus marginatus*), a hemipteran pseudococcid that feeds on *Hibiscus* sp. plants (Miller and Miller, 2002). These insects were collected in Galveston, TX from a hibiscus plant in one of the author's garden. A homogenate of the bugs was prepared, filtered and inoculated into a culture of C6/36 cells. No CPE was observed in the mosquito cells and no virions were observed by transmission electron microscopy. RNA was also extracted from a sample of the bug homogenate and sent for next generation sequencing, Only part of the HFCKV genome was recovered. Consequently, the polymerase ORF was found and used for further genetic characterization, phylogenetic reconstructions, and evolutionary analysis. Based on the polymerase genetic traits and phylogenetic data, HFCKV is more genetically related to plant viruses (Fig. 3). HFCKV is also genetically distance from other plant viruses, as well as from the nelorpiviruses and sandewaviruses. This suggests that HFCKV is a new viral species as well as a possible new genus within the plant virus group.

In conclusion, the information described here indicates that the nelorpiviruses, sandewaviruses and HFCKV, have a close evolutionary relationship to certain plant viruses (*Vigaviridae* and members of the genera *Higrevirus, Blunevirus* and *Cilevirus*. The common factor to most of them is an insect or arthropod host. This finding provides further evidence to the hypothesis that these viruses evolved with arthropods (Maramorosch, 1955; Andrewes, 1957; Dudas and Obbard, 2015; Li et al., 2015; Fort et al., 2012; Ballinger et al., 2013; Markelwitz et al., 2015; Reidenbach et al., 2009).

Acknowledgements

Work at UTMB was supported by NIH grant U54 AI057156, R24 AI120942 (HG, RBT, NV) and by NIH contract HHSN272202000040I/HHSN200004/D04 (HG, RBT, NV). MAC was supported by Programa de Doctorados Nacionales – Colciencias" (Convocatoria 567) (MAC) PhD fellowship. The field collection in Colombia was supported in part by Colciencias C.T-695-2014 and by Colciencias grant 111549326198. This work was supported in part by CNPq (Brazilian National Council for Research and Development) grant no. 302032/2011-8 (MRTN). We also acknowledge support from the Yale Institute for Biospheric Studies Center for EcoEpidemiology and from the U.S. National Park Service (Collecting Permit #EVER-2013-SCI-0032). We thank to Dr. Juan Carlos Gallego-Gómez and Richard Hoyos from Universidad de Antioquia for assistance in field collections and entomological processing of the samples in Colombia. Joao L. Vianez for contributing the box plot analyses.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2017.01.022.

References

Adams, M.J., Antoniw, P.F., Kreuze, J., 2009. Virgaviridae: new family of rod-shaped plant viruses. Arch. Virol. 154, 1967-–1972.
Allicock, O.M., Lemey, P., Tatem, A.J., Pybus, O.G., Bennett, S.N., Mueller, B.A., Suchard, M.A., Foster, J.E., Rambaut, A., Carrington, C.V.F., 2012. Phylogeography and population dynamics of dengue viruses in the Americas. Mol. Biol. Evol. 29, 1533-–1543.
Andrewes, C., 1957. Factors in virus evolution. Adv. Virus Res. 4, 1-–24.
Auguste, A.J., Carrington, C.V.F., Forrester, N.L., Popov, V.L., Guzman, H., Widen, S.G., Wood, T.G., Weaver, S.C., Tesh, R.B., 2014. Characterization of a novel Negevirus and a novel Bunyavirus isolated from Culex (Culex) declarator mosquitoes in Trinidad. J. Gen. Virol. 95, 481-–485.
Ballinger, M.J., Breuenn, J., Kotov, A., Taylor, D.J., 2013. Selectively maintained paleoviruses in Holarctic water fleas reveal an ancient origin for phleboviruses. Virology 446, 276-–282.
Bolling, B.G., Weaver, S.C., Tesh, R.B., Vasilakis, N., 2015. Insect-specific virus discovery: significance for the arbovirus community. Viruses 7, 4911-–14928.
Bryant, J.E., Vasconcelos, P.F.C., Rijnbrand, R.C.A., Mutebi, J.P., Higgs, S., Barrett, A.D.T., 2005. Size heterogeneity in the 3′ noncoding region of South American isolates of yellow fever virus. J. Virol. 79, 3807-–3821.
Bugl, H., Fauman, E.B., Staker, B.L., Zheng, F., Kushner, S.R., Saper, M.A., Bardwell,

J.C., Jakob, U., 2000. RNA methylation under heat shock control. Mol. Cell 6, 349-–360.
Carapeta, S., do Bem, B., McGuinness, J., Esteves, A., Abecasis, A., Lopes, Â., de Matos, A.P., Piedade, J., de Almeida, A.P.G., Parreira, R., 2015. Negeviruses found in multiple species of mosquitoes from southern Portugal: isolation, genetic diversity, and replication in insect cell culture. Virology 483, 318-–328.
Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Muller, W.E.G., Wetter, T., Suhai, S., 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res. 14, 1147–1159.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2011. ProtTest-HPC: Fast selection of best-fit models of protein evolution. European Conference on Parallel Processin., 177-184.
Dudas, G., Obbard, D.J. 2015. Are arthropods at the heart of virus evolution? DOI:10.7554/eLife.06837.
Fish, D., Carpenter, S.R., 1982. Leaf litter and larval mosquito dynamics in tree-hole ecosystems. Ecology 63, 283–288.
Fort, P., Albertini, A., Van-Hua, A., Berthomieu, A., Roche, S., Delsuc, F., Pasteur, N., Capy, P., Gaudin, Y., Weill, M., 2012. Fossil rhabdoviral sequences integrated into arthropod genomes: Ontogeny, evolution, and potential functionality. Mol. Biol. Evol. 29, 381-–390.
Foster, W.A., 1995. Mosquito sugar feeding and reproductive energetics. Ann. Rev. Entomol. 40, 443–474.
Guidon, S., Dufayard, J.F., Lefort, V., Anisimova, Mm, Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.
Guzman, H., Tesh, R.B., 2000. Effects of temperature and diet on the growth and longevity of phlebotomine sandflies (Diptera: Psychodidae). Biomedica 20, 190-–199.
Hsiung, G.D., 1994. Virus isolation and identification methods. (In)In: Hsiuig, G.D., Fong, C.K.Y., Landry, M.L. (Eds.), Hsiung's Diagnostic Virology4th ed.. Yale University Press, New Haven, 37–45.
Kallies, R., Kopp, A., Zirkel, F., Estrada, A., Gillespie, T.R., Drosten, C., Junglen, S., 2014. Genetic characterization of Goutanap virus, a novel virus related to negeviruses, cileviruses and higreviruses. Viruses 6, 4346–4357. http://dx.doi.org/10.3390/v6114346.
Kawakami, K., Kurnia, Y.W., Fujita, R., Ito, T., Isawa, H., Asano, S.J., Binh, N.D., Bando, H., 2015. Characterization of a novel negevirus isolated from Aedes larvae collected in a subartic region of Japan. Arch. Virol.. http://dx.doi.org/10.1007/s/00705-015-2711-9.
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. BMC 28, 1647–1649.
King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J., 2012. Virus Taxonomy. Classification and Nomenclature of Viruses. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego.
Koonin, E.V., Dolja, V.V., 1993. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Crit. Rev. Biochem. Mol. Biol. 28 (5), 375-–430.
Lanciotti, R.S., Lambert, A.J., Holodniy, M., Saavedra, S., Carmen Castillo, del, Signor, L., 2016. Phylogeny of Zika virus in Western hemisphere, 2015 [letter]. Emerg. Infect. Dis.. http://dx.doi.org/10.3201/eid2005.160065.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.
Lazarowitz, S.G., 2001. Plant viruses. In: Knipe, D.M., Howley, P.M. (Eds.), 4th editionFields Virology 1. Lippincott, Williams & Wilkins, Philadelphia, 599–628.
Lemm, J.A., Rice, C.M., 1993. Roles of nonstructural polyproteins and cleavage products in regulating Sindbis virus RNA replication and transcription. J. Virol. 67, 1916–1926.
Li, C., Shi, M., Tian, J., Lin, X., Kang, Y., Chen, L., Qin, X., Xu, J., Holmes, E.C., Zhang, Y.Z., 2015. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. eLife. http://dx.doi.org/10.7554/eLife.05378.
Lipton, S.A., Choi, Y.B., Takahashi, H., Zhang, D., Li, W., Godzik, A., Bankston, L.A., 2002. Cysteine regulation of protein function-as exemplified by NMDA-receptor modulation. Trends Neurosci. 25, 474-–480.
Lohse, M., Bolger, A.M., Nagel, A., Fernie, A.R., Lunn, J.E., Stitt, M., Usadel, B., 2012. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res. 40, W622-–627. http://dx.doi.org/10.1093/nar/gks540.
Maramorosch, K., 1955. Multiplication of plant viruses in insect vectors. Adv. Virus Res. 3, 221–249.
Markelwitz, M., Zirkel, F., Kurth, A., Drosten, C., Junglen, S., 2015. Evolutionary and phenotypic analysis of live virus isolates suggests arthropod origin of a pathogenic RNA virus family. PNAS 112, 7536-–7541.
Melzer, M., Simbjon, N., Borth, W.B., Freitas-Astua, J., Kitajima, E.W., Neupane, K.R., 2013. A cilevirus infects ornamental hibiscus in Hawaii. Arch. Virol. 15, 2421-–2424.
Merritt, R.W., Dadd, R.H., Walker, E.D., 1992. Feeding behavior, natural food and nutritional relationships of larval mosquitoes. Ann. Rev. Entomol. 37, 349-–376.
Miller, D.R., Miller, G.L., 2002. Redescription of Paracoccus marginatus. Williams and Granara de Willink (Hemiptera: Coccoidea: Pseudococcidae), including descriptions of the immature stages and adult male. Proc. Entomol. Soc. Wash. 104, 1–23.
Mullen, G.R., 1975. Acarine Parasites of osquitoes: 1. A critical review of all known records of mosquitoes parasitized by mites. J. Med. Entomol. 12, 27-–36.
Murrell, M.P., Yarema, K.J., Levchenko, A., 2004. The systems biology of glycosylation. Chembiochem 5, 1334-–1347.
Nabeshima, T., Inoue, S., Okamoto, K., Posadas-Herrera, G., Yu, F., Uchida, L., Ichinose, A., Sakaguchi, M., Sunahara, T., Buerano, C., Tadena, F., Orbita, I., Natividad, F., Morita, K., 2014. Tanay virus, a new species of virus isolated from mosquitoes in the Philippines. J. Gen. Virol. 95, 1390-–1395.
Nunes, M.R.T., Palacios, G., Cardoso, J.F., Martins, L.C., Sousa, Jr.E.C., de Lima, C.P.S., Medeiros, D.B.A., Savji, N., Desai, A., Rodrigues, S.G., Carvalho, V.L., Lipkin, W.I., Vasconcelos, P.F.C., 2012. Genomic and phylogenetic characterization of Brazilian yellow fever virus strains. J. Virol. 86, 13263-–13271.
Nunes, M.R.T., Silva, S.P., Carvalho, V.L., Vasconcelos, J.M., da Silva, D.E.A., Oliveira, L.F., Nunes Neto, J.P., Rodrigues, S.G., Azevedo, R.S.S., Monteiro, H.A.O., Cardosa, J.F., Guzman, H., Tesh, R.B., Vasconcelos, P.F.C., Vianez-Junior, J.L.S.G., Martins, L.C., 2015b. Emergence of new insect-restrictive viruses in the Amazon Region. Genome Announc. 3 (2). http://dx.doi.org/10.1128/genomeA.00131-15.
Nunes, M.R.T., Faria, N.R., de Vasconcelos, J.M., Golding, N., Kraemer, M.U.G., de Oliveira, L.F., da Silva Azevedo, R.S., da Silva, D.E.A., da Silva, E.V.P., da Silva, S.P., Carvalho, V.L., Coelho, G.E., Cruz, A.C.R., Rodrigues, S.G., Vianes, J.L.S.G., Nunes, B.T.D., Cardoso, J.F., Tesh, R.B., Hay, S.I., Pybus, O.G., da Costa Vasconcelos, P.F., 2015a. Emergence and potential for spread of Chikungunya virus Brazil. BMC Med. 13, 102.
Pascon, R.C., Kitajima, J.P., Breton, M.C., Assumpcao, L., Greggio, C., Zanca, A.S., Okura, V.K., Alegria, M.C., Camargo, M.L., Silva, G.G.C., Cardozo, J.C., Vallim, M.A., Franco, S.F., Silva, V.H., Jordao Junior, H., Oliveira, F., Giachetto, P.F., Ferrari, F., Aguilar-Vildoso, C.I., Franchiscini, F.J.B., Silva, J.M.F., Arruda, P., Ferro, J.A., Reinach, F., Rasera da Silva, A.C., 2006. The complete nucleotide sequence and genomic organization of Citrus leprosis associated virus, cytoplasmatic type (CiLV-C). Virus Genes 32, 289–298.
Pease, J., Sooknanan, R., 2012. A rapid, directional RNA-seq library preparation workflow for Illumina® sequencing. Nat. Methods 9, 1–2.
Pei, J., Kim, B.H., Grishin, N.V., 2008. PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res. 36, 2295–2300.
Quito-Avila, D.F., Brannen, P.M., Cline, W.O., Harmon, P.F., Martin, R.R., 2013. Genetic characterization of Blueberry necrotic ring blotch virus, a novel RNA virus with unique genetic features. J. Gen. Virol. 94, 1426-–1434.
Reidenbach, K.R., Cook, S., Bertone, M.A., Harbach, R.E., Wiegmann, B.M., Besansky, N.J., 2009. Phylogenetic analysis and temporal diversification of mosquitoes (Diptera: culicidae) based on nuclear genes and morphology. BMC Evol. Biol. 9 (298), 1-–14.
Rio, D.C., Ares, M., Hannon, G.J., Nilsen, T.W., 2010. Purification of RNA using TRIzol (TRI Reagent). In: Rio, D.C., Ares, M., Hannon, G.J., Nilsen, T.W. (Eds.), RNA: A Laboratory Manual. CSHL Press, Cold Spring Harbor, NY, USA, 1276–1288.
Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. Nat. Biotechnol. 29, 24-–26.
Rodrigues Faria, N., do Socorro da Silva Azevedo, R., Kraemer, M.U.G., Souza, R., Sequetin Cunha, M., Hill, S.C., Theze, J., Bonsall, M.B., Bowden, T.A., Rissanen, I., Rocco, I.M., Silva Nogueira, J., Maeda, A.Y., da Silva Vasami, F.G., de Lima Macedo, F.L., Suzuki, A., Guerreiro Rodrigues, S., Ribeiro Cruz, A.C., Tardeli Nunes, B., Barbosa de Almeidda Medeiros, D., Guerreiro Rodrigues, D.S., Nunes Queiroz, A.L., Pinto da Silva, E.V., Freitas Henriques, D., Salve Travassos da Rosa, E., Silva de Oliveira, C., Caricio Martins, L., Baldez Vasconcelos, H., Medeiros Neves Casseb, L., de Brito Simith, D., Messina, J.P., Abade, L., Lourenco, J., Alcantara, Jr.L.C., de Lima, M.M., Giovanetti, M., Hay, S.I., de Oliveira, R.S., da Silva Lemons, P., Freitas de Oliveira, L., Pereira Silva de Lima, C., Patroca da Silva, S., Mota de Vasconcelos, J., Franco, L., Ferreira Cardoso, J., da Silva Goncalves Vianez, Jr.J.L., Mir, D., Bello, G., Delatorre, E., Khan, K., Creatore, M., Coelho, G.E., de Oliveira, W.K., Tesh, R.B., Pybus, O.G., Nunes, M.R.T., Vasconcelos, P.F.C., 2016. Zika virus in the Americas: early epidemiological and genetic findings. Science 352, 345–349.
Rothberg, J.M., Hinz, W., Rearick, T.M., Schultz, J., Mileski, W., Davey, M., Leamon, J.H., Johnson, K., Milgrew, M.J., Edwards, M., Hoon, J., Simons, J.F., Marran, D., Myers, J.W., Davidson, J.F., Branting, A., Nobile, J.R., Puc, B.P., Light, D., Clark, T.A., Huber, M., Branciforte, J.T., Stoner, I.B., Cawley, S.E., Lyons, M., Fu, Y., Homer, N., Sedova, M., Miao, X., Reed, B., Sabina, J., Feierstein, E., Schorn, M., Alanjary, M., Dimalanta, E., Dressman, D., Kasinskas, R., Sokolsky, T., Fidanza, J.A., Namsaraev, E., McKernan, K.J., Williams, A., Roth, G.T., Bustillo, J., 2011. An integrated semiconductor device enabling non-optical genome sequencing. Nature 475, 348-–352.
Rozanov, M.N., Koonin, E.V., Gorbalenya, A.E., 1992. Conservation of the putative methyltransferase domain: a hallmark of the 'Sindbis-like' supergroup of positive-strand RNA viruses. J. Gen. Virol. 73 (Pt 8), 2129-–2134.
Schlein, Y., Jacabson, R.L., 1999. Sugar meals and longevity of the sandfly Phlebotomus papataci in an arid focus of Leishmania major in the Jordon Valley. Med. Vet. Entom. 13, 65–71.
Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J., Birol, I., 2009. ABySS: a parallel assembler for short read sequence data. Genome Res. 19, 1117–1123.
Talavera, G., Gastresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56, 564–577.
Tesh, R.B., Bolling, B.G., Beaty, B.J., 2016. Role of vertical transmission in arbovirus maintenance and evolution. In: Gubler, D.J., Vasilakis, N. (Eds.), Arboviruses: Molecular Biology, Evolution and Control. Caister Academic Press, Norwalk, U.K., 191–218.
Vasilakis, N., Forrester, N.L., Palacios, G., Nasar, F., Savji, N., Rossi, S.L., Guzman, H., Wood, T.G., Popov, V., Gorchakov, R., Gonzalez, A.V., Haddow, A.D., Watts, D.M., da Rosa, A.P.T., Weaver, S.C., Lipkin, W.I., Tesh, R.B., 2013. Negevirus: a proposed new taxon of insect-specific viruses with wide geographic distribution. J. Virol. 87, 2475-–2488.

Volk, S.M., Chen, R., Tsetsarkin, K.A., Adams, A.P., Garcia, T.I., Sall, A.A., Nasar, F., Schuh, A.J., Holmes, E.C., Higgs, S., Maharaj, Brault, A.C., Weaver, S.C., 2010. Genome-scale phylogenetic analyses of chikungunya virus reveal independent emergence of recent epidemics and various evolutionary rates. J. Virol. 84, 6497-–6504.

Weaver, S.C., Vasilakis, N., 2009. Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis. 9, 523–540.

Williamson, D.F., Parker, R.A., Kendrick, J.S., 1989. The box plot: a simple visual method to interpret data. Ann. Intern. Med. 110, 916–921.