Phylogenetic Analysis of Chikungunya Virus Strains Circulating in the Western Hemisphere

Robert S. Lanciotti and Amy J. Lambert*

Division of Vector-Borne Disease, Centers for Disease Control and Prevention, Fort Collins, Colorado

Abstract. In December 2013, chikungunya virus (CHIKV) was isolated for the first time in the Western Hemisphere (WH) during an epidemic on the island of St. Martin. Subsequently, the virus has spread to 42 countries or territories in the Caribbean, Central, South, and North America. In this study, we have determined the full genomic sequences of 29 temporally and geographically diverse CHIKV strains from 16 countries of the WH. Phylogenetic analyses revealed minimal evolution among compared emergent CHIKV strains of the New World.

Chikungunya virus (CHIKV) is a single-stranded RNA virus (family *Togaviridae*, genus *Alphavirus*) that is transmitted by *Aedes* species mosquitoes. The virus was first isolated in Tanzania in 1953; however, descriptions of probable CHIKV epidemics occurring in the nineteenth century are found in the literature.^{1–3} Of particular interest, it appears likely that CHIKV caused epidemics in the Caribbean (St. Thomas) and the southeast coastal United States during the early nineteenth century.^{1–3} CHIKV-related illness is characterized by fever, rash, and acute polyarthralgia. Transmission in urban epidemics is between humans and *Aedes aegypti* and *Aedes albopictus* mosquitoes with no intervening amplifying host. CHIKV epidemics have been described in Africa, the Middle East, Europe, India, and southeast Asia. Epidemics can be explosive, marked by large numbers of human cases and rapid virus dissemination.⁴

Genetic studies have revealed that the virus has evolved into three distinct genotypes, west African, East/Central/south African (ECSA), and Asian. Comprehensive phylogenetic analyses of historic strains suggest that CHIKV originated in Africa and has spread episodically into Asia in approximated 50-year intervals.⁵ More recently, however, the movement of virus genotypes has increased dramatically, likely as a direct result of increased human travel and commercial trade. Beginning in 2005, the ECSA genotype virus was responsible for an explosive epidemic in which the virus moved from coastal Kenya to islands adjacent to southeastern Africa and on to India, where well over 1 million cases were recorded and are believed to be dramatically underestimated.⁶ During this time, imported cases were reported worldwide, and in some instances, autochthonous transmission was detected in distal locations.⁷ An E1 mutation shared among some ECSA strains that facilitates replication in *Ae. albopictus* mosquitoes has been associated with this epidemic.⁸ In December 2013, CHIKV was isolated for the first time in the Western Hemisphere (WH) during an outbreak on the island of St. Martin. Subsequently, the epidemic has spread to 42 countries and territories of the Caribbean, Central, South, and North America; collectively, over 1 million suspected cases have been reported throughout this region.⁹

Previously, we demonstrated that CHIKV strains circulating in the Caribbean are of the Asian genotype and, based upon phylogenetic analysis of a limited number of strains,

were likely introduced from southeast Asia.¹⁰ This conclusion is supported by phylogenetic trees generated from complete genome sequences, which demonstrate that CHIKV strains from the British Virgin Islands are most closely related to strains from the Philippines, China, and Yap (Federated States of Micronesia).¹⁰ Here, we extend and confirm these studies by analyzing the complete genomes of 29 newly sequenced CHIKV strains from 16 countries, including multiple strains isolated at different time points from the same geographic location (GenBank accession nos. KR559470–98). The complete nucleotide sequences of these viruses were derived by next-generation sequencing technologies using the Ion Torrent Personal Genome Machine platform (Life Technologies, Grand Island, NY), as previously described.¹⁰ DNA libraries for sequencing were constructed directly from viremic serum previously shown to contain at least 10⁶ CHIK viral genome copies/mL. The sequences were assembled and mapped to a reference CHIKV genome using SeqMan Pro (DNASTAR, Madison, WI). The complete genome sequences were aligned with one another and with additional reference CHIKV strain sequences representing a diversity of genotypes from GenBank. Phylogenetic trees were constructed with the MEGA software package (version 6) using both the complete genome or only the coding region of the genome (nucleotide positions 77–11,301) by neighbor joining, minimum evolution, and maximum likelihood algorithms. Phylogenetic trees constructed by all of these methods, using either the complete genome or the coding region, are nearly identical and a maximum likelihood tree generated from the coding region is displayed in Figure 1.

The phylogenetic tree (Figure 1) demonstrates that Asian CHIK viruses are subdivided into two well-supported clades that show primarily temporal grouping and are not reciprocally monophyletic. These two clades include CHIKV strains isolated in the 1950s–1990s and those isolated after 2005 that are nested within the 1950s–1990s clade. The > 2005 clade is further subdivided into two, well-supported clades; one containing viruses isolated from southeast Asia (highlighted in yellow) and another clade containing CHIK viruses isolated from the Philippines, the Western and South Pacific, and all of the 2013–2014 WH viruses (highlighted in blue). A previous study noted that the southeast Asian clade showed a spatial and temporal pattern in which it appeared that CHIKV had moved from Thailand to Indonesia and then to Malaysia.⁵ The phylogenetic trees generated in this study support this general view, with the added new observation that CHIK viruses from the > 2005 southeast Asian clade appear to have moved into the Western and South Pacific (American Samoa and Yap) and, subsequently, into the WH.

*Address correspondence to Amy J. Lambert, Arbovirus Diseases Branch, Division of Vector Borne Diseases, Centers for Disease Control and Prevention, 3150 Rampart Road, Foothills Research Campus, Fort Collins, CO 80521. E-mail: ahk7@cdc.gov

FIGURE 1. Phylogeny of chikungunya virus strains. The evolutionary history was inferred by using the maximum likelihood method. The tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

In support of this hypothesis, the tree demonstrates that all of the CHIK viruses from the WH (along with one strain isolated from French Polynesia, see below) form a single clade that is most closely related to a clade containing viruses from

Yap and American Samoa. This clade, containing the WH viruses along with viruses from Yap and American Samoa, has as its closest neighbor a clade of CHIK viruses from China and the Philippines isolated in 2012. However, as mentioned

TABLE 1
Amino acid differences among WH and related CHIKV strains

Amino acid position non-structural poly; protein/position	WH consensus difference*	Strains
84; NSP1/84	Met/Val	China 2012
101; NSP1/101	Ala/Thr	American Samoa July 2014
147; NSP1/147	Val/Ala	El Salvador September 2014
224; NSP1/224	Lys/Asn	Honduras July 2014
224; NSP1/224	Lys/Arg	Guatemala September 2014
329; NSP1/329	Ser/Pro	Philippines 2012
476; NSP1/476	Pro/Thr	U.S. Virgin Islands August 2014
592; NSP2/57	Ala/Val	Philippines 2012
688; NSP2/153	Val/Ala	Panama November 2014
1,205; NSP2/670	Leu/Met	Philippines 2012, China 2012
1,381; NSP3/48	Ser/Cys	U.S. Virgin Islands August 2014, U.S. Virgin Islands September 2014
1,430; NSP3/97	Thr/Ile	Yap 3462 October 2013
1,618; NSP3/285	Ile/Val	Puerto Rico July 2014
1,711; NSP3/378	XXXX/LPTT	All non-WH Asian genotype strains with the exception of Yap 2013, American Samoa, China 2012, Philippines 2012, French Polynesia 2014
1,958; NSP4/99	Arg/Gln	Guyana August 2014, Guyana July 2014
2,318; NSP4/459	Ala/Val	Haiti July 2014
Amino acid position structural poly; protein/position	WH consensus difference	Strains
24 C/24	Thr/Ile	American Samoa July 2014
28 C/28	Ile/Thr	American Samoa July 2014
78 C/78	Arg/Gln	St. Barts June 2014; Columbia August 2014
280 E3/19	Arg/Gln	Philippines 2012; China 2012
336 E3/75	Ala/Val	Honduras July 2014
362 E3/101	Ile/Val	Puerto Rico November 2014
438 E2/113	Val/Ala	British Virgin Islands December 2013
604 E2/279	Gly/Glu	Dominican Republic July 2014
638 E2/313	His/Leu	Yap 3462 October 2013; Yap 3807 October 2013
693 E2/368	Ala/Val	Yap 3462 October 2013; Yap 3807 October 2013; Philippines 2012; China 2012; American Samoa July 2014
731 E2/406	Ala/Thr	Jamaica October 2014
768 6K/20	Met/Leu	Yap 3462 October 2013; Yap 3807 October 2013; Philippines 2012; China 2012; American Samoa July 2014
864 E1/55	Ile/Thr	U.S. Virgin Islands September 2014

CHIKV = chikungunya virus; WH = Western Hemisphere.

*WH consensus and related CHIKV strain amino acids appear before/and after, respectively.

in a previous manuscript,¹⁰ the exact origin of the China 2012 CHIKV is unknown. The China 2012 virus was isolated from a sailor who had traveled throughout southeast Asia.¹⁰ Collectively, the data suggest a scenario in which a CHIKV strain with recent ancestry in southeast Asia seeded the WH epidemic, possibly via the Western and South Pacific.

The extent of evolution among CHIK viruses as the epidemic has spread throughout the Caribbean and Central/South America in 2014–2015 appears to be minimal. Results from a distance calculation of all of the WH strains collected from December 2013 to February 2015 reveals an average number of 7.7 nucleotide differences over the entire genome during this time (14 months), with a range of two to 16 changes between compared sequences (data not shown). This degree of evolution is in general agreement with an evolutionary rate model that predicts that CHIK viruses during an epidemic would display a mutation rate of 8.41E-4 mutations/site/year⁵ (or approximately 10 nucleotide mutations over the entire genome during the course of a 1-year epidemic). Nearly all of nucleotide changes present in the WH strains are silent and amino acid differences are rare with an overall average 1.5 amino acid changes between compared strains (range of 0–4). The presented Table 1 displays the location and nature of these amino acid changes and includes additional viruses for comparison. Interestingly, there is a 12-nucleotide/four-amino acid deletion in the NSP3 protein (polyprotein positions 1,712–1,715) found among all of the

WH CHIK viruses. Of note, there is one CHIK strain from French Polynesia, which possesses this exact deletion. However, the French Polynesia CHIK virus was reported to have come from an infected person from the WH (Guadeloupe) in May 2014. This same 12-nucleotide deletion also occurs among the phylogenetically closely related CHIK viruses isolated from Yap, American Samoa, the Philippines 2012, and Indonesia 2007 (see Figure 1; Table 1). The remaining viruses of the > 2005 genotype have a 21-nucleotide/seven-amino acid deletion in this same location; the 12-nucleotide deletion is adjacent to an additional 9-nucleotide deletion (data not shown). There exist discordant data in GenBank concerning the Malaysian Bagan Panchor 2006 viruses; one institution reports no deletion in this location; however, a second institution verified this 21-nucleotide deletion by sequencing five viruses directly from human viremic serum.¹¹ It appears, therefore, that a deletion at this location, of either 12 or 21 nucleotides, is a common feature of all CHIK viruses of the > 2005 genotype. Finally, there are several unique signature sequences found only among WH CHIK viruses. All of the CHIK viruses from the WH possess two unique amino acid differences in the structural protein region: an Ala at E2/368 and a Met at 6K/20). In addition, all WH CHIK viruses possess a three-base addition in the 3' non-coding region (11,408–11,410) of the genome. The epidemiological and ecological significance of these mutations is not known.

Taken together, the data support the hypothesis that a strain of CHIKV that had been recently circulating throughout southeast Asia moved into the Western and South Pacific and more recently into the WH. The signature 12-nucleotide deletion and the two unique amino acids found among all these closely related WH viruses will be particularly informative in tracking the movement of this strain.

Received May 19, 2015. Accepted for publication December 5, 2015.

Published online February 8, 2016.

Authors' addresses: Robert S. Lanciotti, Centers for Disease Control and Prevention, Fort Collins, CO, E-mail: rsl2@cdc.gov. Amy J. Lambert, Division of Vector-Borne Infectious Diseases, National Center for Zoonotic, Vector-Borne, and Enteric Diseases, Fort Collins, CO, E-mail: ahk7@cdc.gov.

REFERENCES

- Carey DE, 1971. Chikungunya and dengue: a case of mistaken identity. *J Hist Med Allied Sci* 26: 243–262.
- Dickson SH, 1839. *On Dengue: Its History, Pathology and Treatment*. Philadelphia, PA: Haswell, Barrington and Haswell, 5–23.
- Dickson SH, 1859. *Elements of Medicine: A Compendious View of Pathology and Therapeutics; or the History and Treatment of Diseases*. Philadelphia, PA: Blanchard and Lea, 744–749.
- Weaver SC, Lecuit M, 2015. Chikungunya virus and the global spread of a mosquito-borne disease. *N Engl J Med* 372: 1231–1239.
- Volk SM, Chen R, Tssetsarkin KA, Adams AP, Garcia TI, Sall AA, Nasar F, Schuh AJ, Holmes EC, Higgs S, Maharaj PD, Brault AC, Weaver SC, 2010. Genome-scale phylogenetic analyses of chikungunya virus reveal independent emergences of recent epidemics and various evolutionary rates. *J Virol* 84: 6497–6504.
- Mavalankar D, Shastri P, Raman P, 2007. Chikungunya epidemic in India: a major public-health disaster. *Lancet Infect Dis* 7: 306–307.
- Lanciotti RS, Kosoy OL, Laven JJ, Panella AJ, Velez JO, Lambert AJ, Campbell GL, 2007. Chikungunya virus in US travelers returning from India, 2006. *Emerg Infect Dis* 13: 764–767.
- Tssetsarkin KA, Vanlandingham DL, McGee CE, Higgs S, 2007. A single mutation in chikungunya virus affects vector specificity and epidemic potential. *PLoS Pathog* 3: e201.
- Halstead SB, 2015. Reappearance of chikungunya, formerly called dengue, in the Americas. *Emerg Infect Dis* 21: 557–561.
- Lanciotti RS, Valadere AM, 2014. Transcontinental movement of Asian genotype chikungunya virus. *Emerg Infect Dis* 20: 1400–1402.
- Sam IC, Loong SK, Michael JC, Chua CL, Wan Sulaiman WY, Vythilingam I, Chan SY, Chiam CW, Yeong YS, AbuBakar S, Chan YF, 2012. Genotypic and phenotypic characterization of chikungunya virus of different genotypes from Malaysia. *PLoS One* 7: e50476.