Cai et al., J Virol Antivir Res 2016, 5:3
DOI: 10.4172/2324-8955.1000158

Journal of Virology & Antiviral Research

Short Communication A SCITECHNOL JOURNAL

Diversity of and Implications from the Viral Genomes and Viral Proteins of Zika Virus

Hong Cai[1], Min-Hua Luo[2], Yufeng Wang[1*] and Qiyi Tang[3*]

Abstract

Zika virus (ZIKV) belongs to the Genus *Flavivirus* of Family *Flaviviridae* and is closely related to Dengue, West Nile, Japanese encephalitis and yellow fever viruses. Since 2007 ZIKV has caused a series of epidemics in Micronesia, the South Pacific, and, most recently, the Americas. Infection with ZIKV is likely linked to severe medical sequelae. Most recently, it has been observed that ZIKV infection is probably associated with microcephaly of neonates and this makes it urgent to investigate every aspect of the virus including its natural history, epidemiology, pathogenesis and interactions with hosts. In the present study, we analyzed the nucleotide acid (NA) and amino acid (Aa) sequences of the strains of Zika viruses responsible for the seven different epidemics. Sequence alignments and phylogenetic analysis revealed that the ZIKV can be divided into African and Asian types. Interestingly, we found that the Malaysia strain isolated in 1966 appears more divergent from other Asian strains and less divergent from the African type than other Asian strains. To understand why the recent ZIKV outbreak correlates with microcephaly in neonates, we analyzed the diversity of amino acid sequences between the French Polynesia (FP) strain and the Brazil strain. We found that they are more closely related to each other than to other strains being studied. This is consistent with the hypothesis of Brazil ZIKV being evolved from FP-type ancestors. Notably, there were 11 amino acid residues in the Brazil 2016 strain that were different from the consensus sequence. Further analysis of sequence divergence in individual proteins showed that the biggest difference between the FP strain and Brazil strain lies within the NS1 protein, which is related to neurovirulence as in the Dengue virus. Therefore, NS1 might be an important target for further investigation.

Keywords

Zika virus (ZIKV); Outbreak; Genomic diversity; Alignment; *Flaviviridae*

Background

The Zika virus (ZIKV), together with the West Nile virus, Yellow fever virus, Japanese encephalitis virus, Dengue fever virus, and other classified and unclassified viruses, forms the genus *Flavivirus* within the family *Flaviviridae*. The family *Flaviviridae* consists of many other viruses as summarized in a recent review [1]. This family of viruses has an enveloped icosahedral capsid containing a single strand, positive sense RNA genome (about 11,000 nucleotides) [2]. Therefore, the infected viral RNA can be directly translated to a large polyprotein precursor, which is co- and post-translationally processed by viral and cellular proteases into structural and non-structural proteins. The three structural proteins are critical for the formation of the envelope and capsid, and the seven non-structural (NS) proteins play important roles in virus replication. The three structural proteins are the enveloped, E; membrane precursor, PrM; and capsid, C. The seven non-structural (NS) proteins include NS1, NS2a, NS2b, NS3, NS4a, NS4b, and NS5. Both the structural and non-structural proteins are needed to determine the pathogenicity of ZIKV [3].

ZIKV is a re-emerging *Flavivirus* transmitted both via the mosquito Aedes aegypti and sexually [4,5]. The ZIKV was first isolated in 1947 from a monkey of the Zika forest [6], which represents the African strain, and was later isolated in Southern Asia [7]. Therefore, two types of ZIKV are considered to be the ancestors of the contemporary virus [8]. It was not considered as a severe medical concern until recent outbreaks of ZIKV disease in the Pacific Islands and in region of the Americas [9,10]. However, epidemiological data and clinical findings on laboratory-confirmed Zika virus disease remain limited. Here we report our recent genomic analysis of the Zika viruses and discuss their potential biological implications.

Results and Discussion

ZIKV has caused several epidemics at different scales since it was first isolated. The geographic locations of viral isolation and epidemics include: South Africa [6], Malaysia [7], Cambodia [11], Yap island [12], French Polynesia [13] and Brazil [14]. Strains of the virus were isolated from each of the epidemics. RNA viruses keep varying to avoid host immunity, leading to generation of new strains. We wondered whether the Zika virus outbreaks are related to the continued mutations of viruses. To examine this possibility, we chose seven strains of ZIKV from different geographic isolations: an African strain (MR477) isolated in 1947 [6], an Asian strain [11], an Malaysian strain [7], a Yap island strain [12,15], a more recent African strain [16], an FP strain [17,18], and a Brazilian strain [19]. The complete nucleotide (Nt) and Aa sequences of all these strains are available in the PubMed database, their database accession numbers are shown in the figures. First, as shown in Figure 1S, we performed a multiple sequence alignment of the whole Aa sequence of the precursor of ZIKV protein that contains 3,421 amino acids in total. Conserved amino acid residues are shown in red and other colors (gray or blue) indicate where sequences differ from the consensus sequence. The mutations in the Brazil strain are shown in gray and highlighted in yellow. The consensus sequence, shown in the end of each block, was derived using the BoxShade program based on the majority rule (>=50% agreement). Overall, the seven sequences were highly conserved.

The large polyprotein precursor must be cleaved to generate active functional proteins. The distribution of mutations varies in 10 proteins. As summarized in Figure 1, PrM protein presents the highest number of the mutations (7.2%), and NS2b has the smallest number of mutations (1.5%). Interestingly, the region spanning

***Corresponding authors:** Qiyi Tang, Biology, Department of Microbiology, Howard University College of Medicine, Seeley Mudd Building, Room 315, 520 W Street, NW, Washington, DC 20059, USA, Tel: (202) 806 3915; Fax: (202) 238-8518; E-mail: qiyi.tang@howard.edu

Yufeng Wang, South Texas Center for Emerging Infectious Diseases, University of Texas at San Antonio, One UTSA Circle, San Antonio, Texas 78249, USA, Tel: (210)458-6492; Fax: (210)458-5658; E-mail: Yufeng.Wang@utsa.edu

Received: July 11, 2016 **Accepted:** August 02, 2016 **Published:** August 09, 2016

NS2b, NS3, and NS4a seems to be more conserved than the other areas. The functional consequence of these mutations awaits further investigation.

We then performed phylogenetic analysis for both the Aa and Nt sequences (Figures 2 and 3) [20-23]. The neighbor-joining (NJ) trees inferred from Aa and Nt sequences yielded a consistent topology. Brazil 2016 and FP 2014 were most closely related, forming a clade in the trees with strong bootstrapping support. Both strains are most closely related to an Asian strain, the Cambodia strain, and the Yap island strain. The African strains (1947 and 2014) appear to be highly homologous. However, Malaysia 1966, a previously identified Asian strain, appears to be divergent from the other Asian strains. Notably,

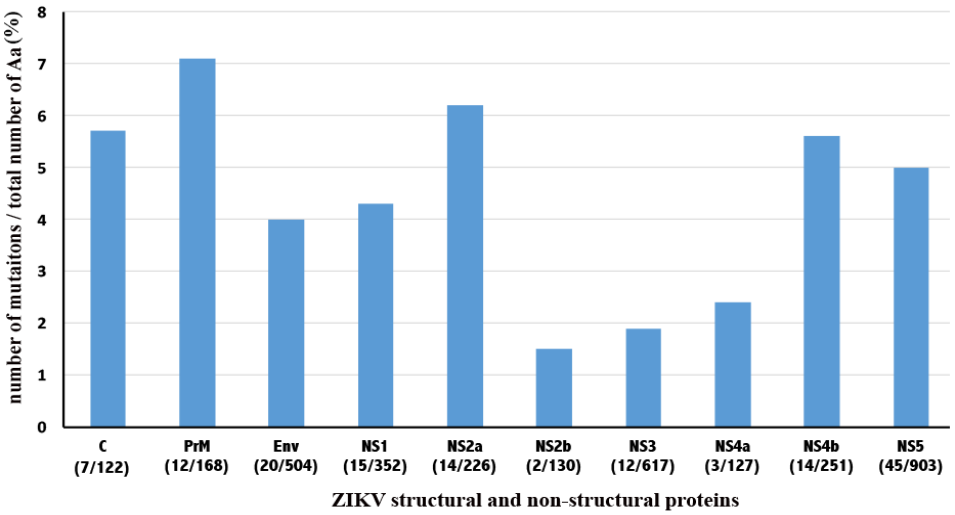

Figure 1: **Summary of the mutations of ZIKV genes**: There are a total of 144 mutations in the ZIKV proteins, most of which are substitutions except one deletion in the NS5 protein of the African 2014 strain. The percentages were shown by dividing the mutation number to the total number of amino acid residues of each protein. The numbers were shown under the name of each protein. Abbreviations: C (capsid), PrM (precursor of M), Env (envelope), NS (non-structure protein).

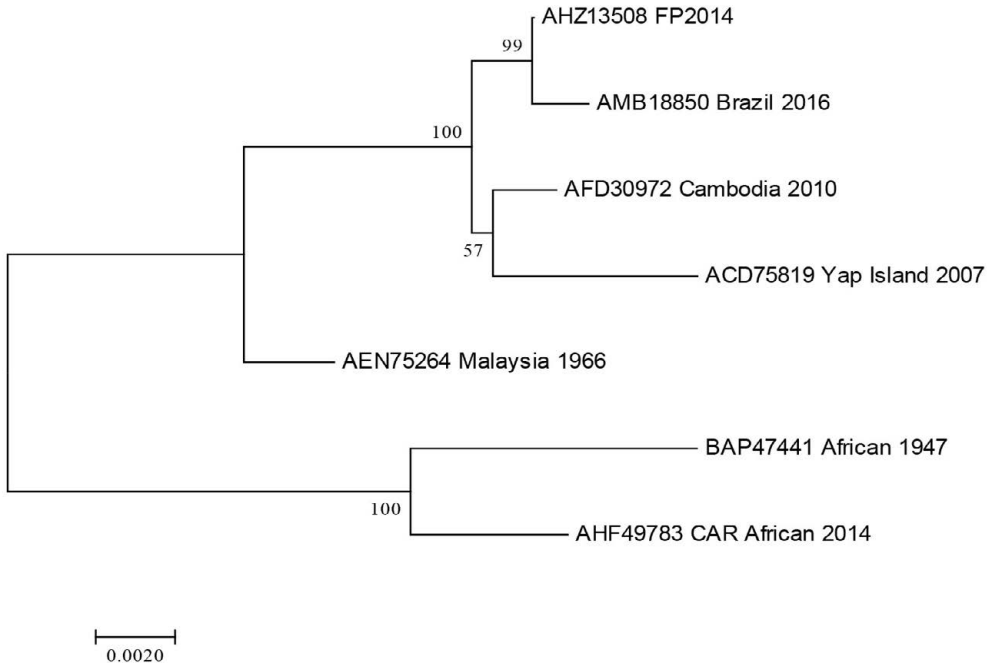

Figure 2: **Phylogenetic tree of seven strains of the Zika virus based on amino acid sequences:** The evolutionary history of amino acid sequences was inferred using the Neighbor- Joining method [20]. The optimal tree with the sum of branch length = 0.04561554 is shown.The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) are shown next to the branches [21]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [22]. and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA7 [23].

doi: 10.4172/2324-8955.1000158

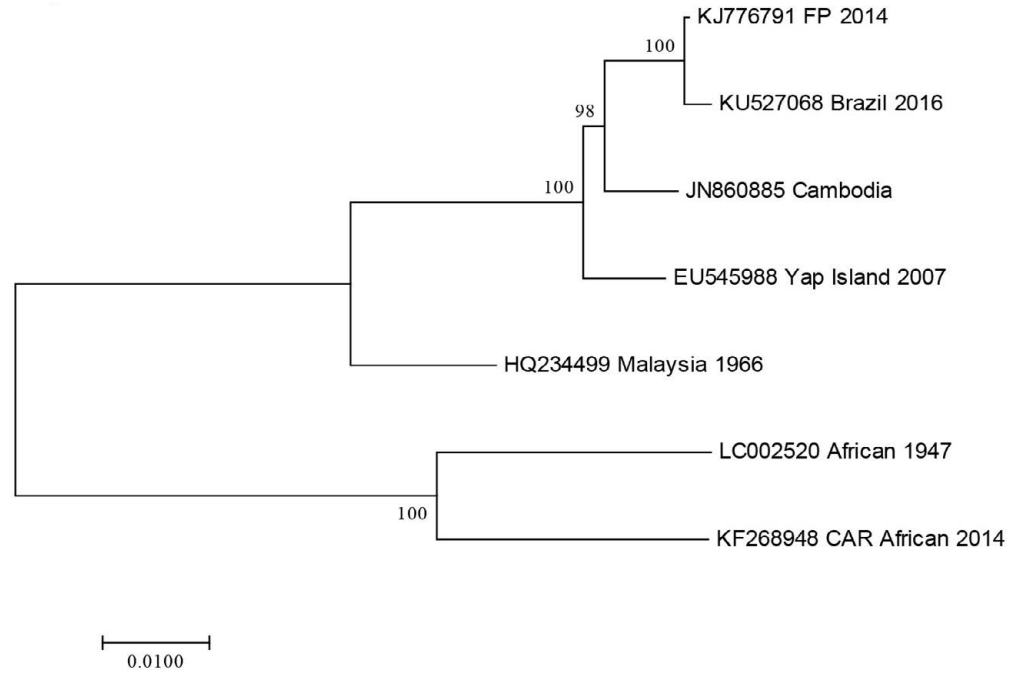

Figure 3: Phylogenetic trees of seven strains of the Zika virus based on nucleotide sequences. The evolutionary history of nucleotide sequences was inferred using the Neighbor-Joining method [20]. The optimal tree with the sum of branch length = 0.18463975 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) is shown next to the branches [21]. The evolutionary distances were computed using the Maximum Composite Likelihood method [22]. and are in the units of the number of base substitutions per site. Evolutionary analyses were conducted in MEGA7 [23].

there were 11 amino acid residues in the Brazil 2016 strain that were different from the consensus sequence, as shaded in yellow color in Figure 1S. The possibility that some of these mutations are related to the pathogenicity of the ZIKV is intriguing.

Finally, we examined the types of amino acid mutations. The mutations are diverse, most of which are substitutions from Isoleucine to Valine, and Valine to Alanine. The change from Lysine to Arginine is also observed. Most mutations are not continuously distributed in the genome, except that six amino acid mutations were observed in the capsid gene of the Yap island strain, from kksggf to EEIRRI. All the mutations are substitutions except one deletion mutation in the NS5 protein of African strain 2014.

It would be highly interesting to know whether the observed mutations affect the biological functions of the proteins. The cleavage of the polyprotein precursor is a sophisticated process and is completed collaboratively by cellular proteases of the PACE (Paired basic Amino acid Cleaving Enzyme)-type or other Golgi-localized proteases and the viral serine protease embedded in the N-terminal domain of non-structural protein 3 (NS3Pro), which requires NS2b for its activity [1]. A distinct feature of genus *Flavivirus* from other genera of *Flaviviridae* is that the 5′-end of the (+) ssRNA genome of genus *Flavivirus* is decorated with an RNA cap structure (N7meGpppA2′Ome-RNA). The 5'end capping of the viral RNA is as important as that for eukaryotic mRNAs, not only to initiate the process of translation but also to protect the viral RNA from degradation by endogenous RNA exonucleases. The protein translation happens immediately after the uncoating of viral particle in the cytoplasm. The (+) ssRNA genome is used as a template not only for gene expression but also for viral genome replication. Both viral RNA replication and gene translation occur in the cytoplasm. For RNA replication, viral non- structural (NS) proteins and cellular proteins interact to form a replication compartment (RC).

During the period of viral RNA replication in the cytoplasm, the RC consists of morphologically distinct, membrane-bound compartments that also differ with respect to both function and NS proteins composition [24]. The NS3 and NS5 proteins are central to the viral RC, as together, they harbor most, if not all, of the catalytic activities required to both cap and replicate the viral RNA. Following replication, the protected genomic RNA is packaged by the C protein to form a capsid in a host-derived lipid bilayer in which the E protein is embedded and later integrated into viral envelope. The mature particles subsequently exit from the host cell by exocytosis.

Recent events in Brazil have earned the Zika virus the world's attention. It has become another member of Genus *Flavivirus* to move to the center of virological research. There is an urgent need to solve this problem, but time is needed to achieve a better understanding of its pathogenesis, prevention, and treatment. The genomic analyses presented here have identified mutations that may have driven the conversion of Zika virus from an infectious agent of little concern to a virulent pathogen, and it is our hope that these results will help guide the research community to a full understanding of ZIKV.

Acknowledgement

This work was supported by grants from National Institutes of Health, SC1AI112785 to QT, and GM100806 to YW. This study was supported by an American Cancer Society grant (RSG-090289-01-MPC) to QT.

doi: 10.4172/2324-8955.1000158

References

1. Bollati M Alvarez K, Assenberg R, Baronti C, Canard B, et al. (2010) Structure and functionality in flavivirus NS-proteins: perspectives for drug design. Antiviral Res 87: 125-148.
2. Faye O, Freire CC, Iamarino A, Faye O, de Oliveira JV, et al. (2014) Molecular evolution of Zika virus during its emergence in the 20(th) century. PLoS Negl Trop Dis 8: e2636.
3. Kuno G, Chang GJ (2007) Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. Arch Virol 152: 687-696.
4. Musso D, Roche C, Robin E, Nhan T, Teissier A, et al. (2015) Potential sexual transmission of Zika virus. Emerg Infect Dis 21: 359-361.
5. Faria NR, Azevedo Rdo S, Kraemer MU, Souza R, Cunha MS, et al. 2016. Zika virus in the Americas: Early epidemiological and genetic findings. Science 352: 345-349.
6. Dick GW, Kitchen SF, Haddow AJ (1952) Zika virus. I. Isolations and serological specificity. Trans R Soc Trop Med Hyg 46: 509-520.
7. Marchette NJ, Garcia R, Rudnick A (1969) Isolation of Zika virus from Aedes aegypti mosquitoes in Malaysia. Am J Trop Med Hyg 18: 411-415.
8. Haddow AD, Schuh AJ, Yasuda CY, Kasper MR, Heang V, et al. (2012) Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage. PLoS Negl Trop Dis 6:e1477.
9. Ventura CV, Maia M, Bravo-Filho V, Góis AL, Belfort R Jr (2016) Zika virus in Brazil and macular atrophy in a child with microcephaly. Lancet 387: 228.
10. Mlakar J, Korva M, Tul N, Popovia M, Poljaak-Prijatelj M, et al. (2016) Zika Virus Associated with Microcephaly. N Engl J Med 374: 951-958.
11. Heang V, Yasuda CY, Sovann L, Haddow AD, Travassos da Rosa AP, et al. (2012) Zika virus infection, Cambodia, 2010. Emerg Infect Dis 18: 349-351.
12. Duffy MR, Chen TH, Hancock WT, Powers AM, Kool JL, et al. (2009) Zika virus outbreak on Yap Island, Federated States of Micronesia. N Engl J Med 360: 2536-2543.
13. Cao-Lormeau VM, Roche C, Teissier A, Robin E, Berry AL, et al. (2014) Zika virus, French polynesia, South pacific, 2013. Emerg Infect Dis 20: 1085-1086.
14. Campos GS, Bandeira AC, Sardi SI (2015) Zika Virus Outbreak, Bahia, Brazil. Emerg Infect Dis 21: 1885-1886.
15. Lanciotti RS, Kosoy OL, Laven JJ, Velez JO, Lambert AJ, et al. (2008) Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. Emerg Infect Dis 14: 1232-1239.
16. Berthet N, Nakoune E, Kamgang B, Selekon B, Descorps-Declere S, et al. (2014) Molecular characterization of three Zika flaviviruses obtained from sylvatic mosquitoes in the Central African Republic. Vector Borne Zoonotic Dis 14: 862-865.
17. Hancock WT, Marfel M, Bel M (2014) Zika virus, French Polynesia, South Pacific, 2013. Emerg Infect Dis 20: 1960.
18. Baronti C, Piorkowski G, Charrel RN, Boubis L, Leparc-Goffart I, et al. (2014) Complete coding sequence of zika virus from a French polynesia outbreak in 2013. Genome Announc 2: 500-414.
19. de MCR, Cirne-Santos C, Meira GL, Santos LL, de Meneses MD, et al.(2016) Prolonged detection of Zika virus RNA in urine samples during the ongoing Zika virus epidemic in Brazil. J Clin Virol 77: 69-70.
20. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406-425.
21. Felsenstein J (1988) Phylogenies from molecular sequences: inference and reliability. Annu Rev Genet 22: 521-565.
22. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci USA 101: 11030-11035.
23. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013)MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30: 2725-2729.
24. Mackenzie J (2005) Wrapping things up about virus RNA replication. Traffic 6: 967-977.

Author Affiliations

Top

[1]*Department of Biology, South Texas Center for. Emerging Infectious Diseases, University of Texas at San Antonio, San Antonio, Texas, USA*

[2]*State Key Laboratory of Virology, CAS Center for Excellence in Brain Science and Intelligence Technology (CEBSIT), Wuhan Institute of Virology, China*

[3]*Department of Microbiology, Howard University College of Medicine, USA*