

1 **Complete genome sequences of Zika Virus strains isolated from the blood of patients in**  
2 **Thailand (2014) and Philippines (2012).**

3 Ellison,D.W.<sup>1</sup>, Ladner,J.T.<sup>3</sup>, Buathong,R.<sup>4</sup>, Alera,M.T.<sup>1</sup>, Wiley,M.R.<sup>3</sup>, Hermann,L.<sup>1</sup>,  
4 Rutvisuttinunt,W<sup>1</sup>., Klungthong,C.<sup>1</sup>, Chinnawirotpisan,P.<sup>1</sup>, Manasatienkij,W.<sup>1</sup>, Melendez,M.C.<sup>2</sup>,  
5 Maljkovicberry,I.<sup>2</sup>, Thaisomboonsuk,B.<sup>1</sup>, Ong-ajchaowlerd,P.<sup>1</sup>, Kaneechit,W.<sup>1</sup>, Velasco,J.M.<sup>1</sup>,  
6 Tac-An,I.A.<sup>5</sup>, Villa,D.<sup>5</sup>, Lago,C.B.<sup>1</sup>, Roque,V.G. Jr.<sup>6</sup>, Akrasewi,P.<sup>4</sup>, Plipat,T.<sup>4</sup>, Nisalak,A.<sup>1</sup>,  
7 Srikiatkachorn,A.<sup>7</sup>, Fernandez,S.<sup>8</sup>, Yoon,I.K.<sup>9</sup>, Haddow,A.D.<sup>3</sup>, Palacios,G.F.<sup>3</sup>, Jarman,R.G.<sup>2</sup>  
8 and Macareo,L.R.<sup>1</sup>

9

10 <sup>1</sup> Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand

11 <sup>2</sup> Walter Reed Army Institute of Research, Silver Spring, MD 20910, USA

12 <sup>3</sup> Center for Genome Sciences, United States Army Medical Research Institute of Infectious  
13 Diseases, 1425 Porter Street, Fort Detrick, Frederick, MD, 21702, USA

14 <sup>4</sup> Department of Disease Control, Bureau of Epidemiology, Ministry of Public Health,  
15 Nonthaburi, Thailand

16 <sup>5</sup> Cebu City Health Department, Cebu City, Philippines

17 <sup>6</sup> Department of Health, Manila, Philippines

18 <sup>7</sup> University of Massachusetts Medical School, Worcester, Massachusetts, USA

19 <sup>8</sup> United States Army Medical Materiel Development Activity, Frederick, MD, USA

20 <sup>9</sup> International Vaccine Institute, Seoul, Republic of Korea.

21

22 Running Head: Zika Virus Genomes

23

24 **ABSTRACT**

25 ZIKV is an arbovirus and member of the family *Flaviviridae* it is transmitted throughout Africa  
26 and Asia and is currently causing an outbreak in South America. Here we present the complete  
27 genome sequences of two Zika Virus (ZIKV) strains, Zika virus/H.sapiens-  
28 tc/THA/2014/SV0127-14 and Zika virus/H.sapiens-tc/PHL/2012/CPC-0740, isolated from the  
29 blood of patients collected in Thailand, 2014 and the Philippines, 2012. Sequencing and  
30 phylogenetic analysis showed that both strains belong to the Asian lineage.

31

32 Key words: Zika, *Flaviviridae*, *Flavivirus*, whole genome sequence

33

34 Zika Virus (ZIKV) has garnered worldwide attention as researches have linked the virus to an  
35 increase in microcephaly cases during the current outbreak in South America (1). Once thought  
36 to cause mild infections, ZIKV is now the subject of intensive worldwide research collaborations  
37 and efforts. ZIKV is a single-stranded positive-sense RNA arbovirus and a member of the  
38 *Flaviviridae* family, which includes dengue, yellow fever, St. Louis encephalitis, Japanese  
39 encephalitis and West Nile viruses (2). The ZIKV genome, approximately 11 kilobases in  
40 length, is similar in its arrangement to other members of *Flaviviridae* containing 5' and 3'  
41 untranslated regions flanking a single open reading frame (ORF). The 5' and 3' untranslated  
42 regions are thought to be important for host interaction, viral replication, and pathogenicity. The  
43 ORF codes for three structural proteins: the capsid (C), premembrane/membrane (prM) and  
44 envelope (E) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and  
45 NS5), which are responsible for viral replication and assembly (3). Previous phylogenetic  
46 analysis based on the nucleotide sequences of ZIKV indicated two major lineages: African and  
47 Asian (2).

48 These two ZIKV isolates were obtained by intrathoracic inoculation of ZIKV real-time RT-PCR  
49 positive patient's serum samples into *Toxorhynchites splendens* mosquitoes followed by  
50 inoculation of mosquito derived C6/36 cells (4, 5). Viral RNA was extracted using the Direct-  
51 zol RNA extraction kit (Zymo Research), converted to cDNA using SuperScript III (Invitrogen)  
52 and amplified using sequence-independent single-primer amplification (6) combined with  
53 primers for rapid amplification of cDNA ends (7). Sequencing libraries were constructed using  
54 PrepX ILM 32 DNA Library Kit (Wafergen) and sequenced by using Illumina NextSeq platform  
55 (2 x 151 bp). Adaptors and primers were clipped from the sequence reads using Cutadapt version  
56 1.21 (8) and low-quality reads/bases were filtered using Prinseq-lite version 0.20.4 (1) ZIKV

57 consensus genomes were assembled using Ray Meta (9), Bowtie2 v. 2.0.6 (10) and Samtools v.  
58 0.1.18 (11).

59 The complete genome sequence with a total length of 10,807 nucleotides (nt) containing the 5'  
60 (107 nt) and 3' (428 nt) UTRs and one ORF (10,272 nt) were obtained from both isolates (12).  
61 Our sequences and other ZIKV sequences from GenBank were used in maximum-likelihood  
62 phylogenetic analysis by using PhyML 3.1 (13) with GTR+G model (-lnL 33630.580); the tree  
63 revealed that these two isolates belong to the Asian lineage and are closely related to 2015  
64 Brazilian isolates. This confirms our previous analyses based of NS5 genes of these two isolates  
65 (4, 5). Data suggests that Asian lineage isolates are 95-98% identical on the nt level and 98-99%  
66 identical on the amino acid level. Nevertheless, some substitutions were found within  
67 primer/probe binding sites in the genomes of these two isolates sequences. The two mismatches  
68 found on Zika virus/H.sapiens-tc/PHL/2012/CPC-0740 genome for primer 835 (residue 1/23),  
69 and for probe 1107 (residue 1/31) binding sites (14). There are two mismatches found on Zika  
70 virus/H.sapiens-tc/THA/2014/SV0127-14 genome for primer 911c (residue 21/22), and for probe  
71 1107 (residue 19/31) binding site (14). Further ZIKV genome studies and comparisons will not  
72 only elucidate factors involved in the virulence and pathogenicity of ZIKV, but will also lend  
73 insight into the evolution of this virus and help with vaccine design.

74 **Nucleotide sequences accession numbers.** The assembled complete genome sequences of the  
75 Zika virus/H.sapiens-tc/THA/2014/SV0127-14 and Zika virus/H.sapiens-tc/PHL/2012/CPC-  
76 0740 isolates were submitted to GenBank under the accession numbers KU681081 and  
77 KU681082, respectively. The versions described in this paper are the third versions.

78

79 **ACKNOWLEDGEMENTS**

80 All military service members or employees of the U.S. Government and this work was prepared  
81 as part of their official duties. Title 17 U.S.C. §105 provides that ‘Copyright protection under  
82 this title is not available for any work of the United States Government.’ Title 17 U.S.C. §101  
83 defines a U.S. Government work as a work prepared by a military service member or employee  
84 of the U.S. Government as part of that person’s official duties. This work is supported by the  
85 Global Emerging Infections Surveillance and Response System, a Division of the Armed Forces  
86 Health Surveillance Center. Work at USAMRIID was funded by DTRA/JSTO under project  
87 CB10246. The views expressed in this article are those of the author and do not necessarily  
88 reflect the official policy or position of the Department of the Army, the Department of Defense,  
89 nor the U.S. Government.

90

91

92

93

94

95

96

97

98

99 **REFERENCES**

- 100 1. **Calvet G, Aguiar RS, Melo AS, Sampaio SA, de Filippis I, Fabri A, Araujo ES, de**  
101 **Sequeira PC, de Mendonça MC, de Oliveira L, Tschoeke DA, Schrago CG, Thompson FL,**  
102 **Brasil P, Dos Santos FB, Nogueira RM, Tanuri A, de Filippis AM.** 2016. Detection and  
103 sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case  
104 study. *Lancet Infect Dis.* [http://dx.doi.org/10.1016/S1473-3099\(16\)00095-5](http://dx.doi.org/10.1016/S1473-3099(16)00095-5).  
105
- 106 2. **Haddow AD, Schuh AJ, Yasuda CY, Kasper MR, Heang V, Huy R, Guzman H, Tesh**  
107 **RB, Weaver SC.** 2012. Genetic characterization of Zika virus strains: geographic expansion of  
108 the Asian lineage. *PLoS Negl Trop Dis.* 6(2):e1477.  
109 <http://dx.doi.org/10.1371/journal.pntd.0001477>  
110
- 111 3. **Kuno G and Chang GJ.** 2007. Full-length sequencing and genomic characterization of  
112 Bagaza, Kedougou, and Zika viruses. *Arch Virol.* 152(4):687-96.  
113 <http://dx.doi.org/10.1007/s00705-006-0903-z>  
114
- 115 4. **Buathong R, Hermann L, Thaisomboonsuk B, Rutvisuttinunt W, Klungthong C,**  
116 **Chinnawirotpisan P, Manasatienkij W, Nisalak A, Fernandez S, Yoon IK, Akrasewi P,**  
117 **Plipat T.** 2015. Detection of Zika Virus Infection in Thailand, 2012-2014. *Am J Trop Med*  
118 *Hyg.* <http://dx.doi.org/10.4269/ajtmh.15-0022>.  
119
- 120 5. **Alera MT, Hermann L, Tac-An IA, Klungthong C, Rutvisuttinunt W, Manasatienkij W,**  
121 **Villa D, Thaisomboonsuk B, Velasco JM, Chinnawirotpisan P, Lago CB, Roque VG Jr,**

122 **Macareo LR, Srikiatkachorn A, Fernandez S, Yoon IK.** 2015. Zika virus infection,  
123 Philippines, 2012. *Emerg Infect Dis.* (4):722-4. <http://dx.doi.org/10.3201/eid2104.141707>.  
124

125 6. **Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, Afonso C,**  
126 **Zhang X, Anderson NG, Ghedin E, Spiro DJ.** 2008. Viral genome sequencing by random  
127 priming methods. *BMC genomics* 9:5. <http://dx.doi.org/10.1186/1471-2164-9-5>.  
128

129 7. **Leguia M, Loyola S, Rios J, Juarez D, Guevara C, Silva M, Prieto K, Wiley M, Kasper**  
130 **MR, Palacios G.** 2015. Full Genomic Characterization of a Saffold Virus Isolated in Peru.  
131 *Pathogens* 4:816-825. <http://dx.doi.org/10.3390/pathogens4040816>.  
132

133 8. **Martin M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing  
134 reads. *EMBnet.journal* 17:10-12. <http://dx.doi.org/10.14806/ej.17.1.200>.  
135

136 9. **Boisvert, S., Raymond, F., Godzaridis, E., Laviolette, F., and Corbeil, J.** 2012. Ray Meta:  
137 scalable de novo metagenome assembly and profiling. *Genome biology* 13, R122.  
138 <http://dx.doi.org/10.1186/gb-2012-13-12-r122>.  
139

140 10. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. *Nature*  
141 *methods* 9:357-359. <http://dx.doi.org/10.1038/nmeth.1923>.  
142



- 143 11. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G,**  
144 **Durbin R.** 2009. The sequence alignment/map format and SAMtools. *Bioinformatics.*  
145 25(16):2078-9. <http://dx.doi.org/10.1093/bioinformatics/btp352>.  
146
- 147 12. **Ladner JT, Beitzel B, Chain PS, Davenport MG, Donaldson E, Frieman M, Kugelman**  
148 **J, Kuhn JH, O’Rear J, Sabeti PC.** 2014. Standards for Sequencing Viral Genomes in the Era  
149 of High-Throughput Sequencing. *mBio* 5:e01360-01314. <http://dx.doi.org/10.1128/mBio.01360->  
150
- 151 13. **Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O.** 2010. New  
152 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the  
153 performance of PhyML 3.0. *Systematic biology.* 59(3):307-21.  
154 <http://dx.doi.org/10.1093/sysbio/syq010>.  
155
- 156 14. **Lanciotti RS, Kosoy OL, Laven JJ, Velez JO, Lambert AJ, Johnson AJ, Stanfield SM,**  
157 **Duffy MR.** 2007. Genetic and serologic properties of Zika virus associated with an epidemic,  
158 Yap State, Micronesia. *Emerg Infect Dis.* (8):1232-9. <http://dx.doi.org/10.3201/eid1408.080287>.  
159  
160  
161  
162  
163  
164  
165

