



Short communication

Zika virus complete genome from Salvador, Bahia, Brazil



Marta Giovanetti ^{a,b,*}, Nuno Rodrigues Faria ^c, Marcio Roberto Teixeira Nunes ^d, Janaina Mota de Vasconcelos ^d, José Lourenço ^c, Sueli Guerreiro Rodrigues ^d, Joao Lidio Vianez Jr ^d, Sandro Patroca da Silva ^d, Poliana Silva Lemos ^d, Fernando Neto Tavares ^d, Darren Patrick Martin ^e, Mateus Santana do Rosário ^{a,f}, Isadora Cristina Siqueira ^a, Massimo Ciccozzi ^g, Oliver G. Pybus ^c, Tulio de Oliveira ^h, Luiz Carlos Alcantara Junior ^a

^a Fundação Oswaldo Cruz, Salvador, Bahia, Brazil

^b University of Rome "Tor Vergata", Italy

^c Department of Zoology, University of Oxford, UK

^d Instituto Evandro Chagas, Pará, Brazil

^e Division of Computational Biology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

^f Hospital Santa Isabel, Salvador, Bahia, Brazil

^g National Institute of Health, Rome, Italy

^h Africa Center, University of KwaZulu-Natal, Durban, South Africa

ARTICLE INFO

Article history:

Received 2 March 2016

Received in revised form 28 March 2016

Accepted 29 March 2016

Available online 9 April 2016

Keywords:

Zika virus

Full-length sequence

Phylogeny

ABSTRACT

In May 2015 the first autochthonous Zika virus infection was reported in Brazil. Rapid and urgent measures are needed to contain the ongoing outbreak. Here we report the full-length ZIKV coding sequence from Bahia. Genetic analysis of outbreak sequences will be essential for characterizing the diversity of circulating strains, identifying hotspots of virus transmission and guiding public health control. Rapid and urgent measures are needed to contain the ongoing outbreak.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Zika virus (ZIKV) is an emerging arthropod born Flavivirus with a positive-sense, single-stranded RNA genome about 10,794-nt in length (Kuno et al., 2007). Its natural transmission cycle involves primarily *Aedes* mosquito species including *Aedes aegypti* and *Aedes albopictus*, which are commonly found in tropical and sub-tropical areas of the globe (Mourão et al., 2012). Since its first isolation in Uganda 1947, sporadic isolations have occurred from humans and a variety of mosquito species in both Africa and Asia (Duffy et al., 2009). However in 2007 a Zika fever epidemic took place in the Yap Island, Micronesia and a pediatric case of ZIKV infection was also reported in Cambodia demonstrating not only that the virus could cause outbreaks in human populations but also that its geographic distribution was expanding (Grad et al., 2014; Faye et al., 2014). In May 2015 the first autochthonous ZIKV

infection was reported in Brazil. Worryingly, by January 2016, more than 17,000 suspected cases had been notified in Salvador, the capital city of the state of Bahia (Zanluca et al., 2015). To date Bahia is the most affected region with the greatest number of ZIKV notifications in America. Rapid and urgent measures are needed to contain the ongoing outbreak. Here we report the full-length ZIKV coding sequence from Bahia. Genetic analysis of outbreak sequences will be essential for characterizing the diversity of circulating strains, identifying hotspots of virus transmission and guiding public health control. Rapid and urgent measures are needed to contain the ongoing outbreak.

2. Materials and methods

On 30th June 2015, a 52 year-old previously healthy woman presented with moderate arthralgia and conjunctival hyperemia to Santa Isabel Hospital in Salvador. The following day (day 2 of symptoms) the patient developed fever (38 °C), pruritus and bilateral swelling in knees, ankles, wrists and hands. During an outpatient visit on the 1st of July 2015, the medical doctor identified a maculopapular rash on the patient's face, abdomen and thighs and blood samples were collected. Painkillers and dexchlorpheniramine were prescribed to control

* Corresponding author at: Laboratório de Hematologia, Genética e Biologia Computacional, Centro de Pesquisas Gonçalo Moniz, Fundação Oswaldo Cruz, Rua Cícero Simões 225, Candeal, 42296710 Salvador, Bahia, Brazil.

E-mail address: giovanetti.marta@gmail.com (M. Giovanetti).

symptoms. The fever lasted only one day. Arthralgia worsened on days 3 and 4 and the patient developed an inability to walk. Improvement of symptoms occurred on day 5, with total recovery on day 6. Because the course of ZIKV infection is frequently asymptomatic (Zanluca et al., 2015) and the clinical presentation may be mistaken for dengue virus (DENV) or chikungunya virus (CHIKV) infection, both of which co-circulate in Bahia (Rodrigues Faria et al., 2016), molecular detection was undertaken using RT-PCR and the patient tested positive for ZIKV and negative for both CHIKV and DENV.

2.1. Virus isolation and RNA isolation and genome sequencing

A serum sample of the patient was subjected to virus isolation attempt in *A. albopictus* C6/36 cell cultures as previously described (Zamree et al., 2005) however no positive result was observed for ZIKV. Total RNA was isolated direct from the clinical sample (serum) by using the Qiaamp Viral RNA Minikit (Qiagen), and first subject to RTqPCR for genome detection (Lanciotti et al., 2008). After confirmation of presence of ZIKV genome in the RNA, viral genomes were recovered using both Next Generation Sequencing platforms, Ion Torrent PGM (Life Technologies) and GSFLX 454 (Roche, Life Sciences) and the ion semiconducting and pyrosequencing methods, respectively.

Raw reads were combined and assembled using Mira v4.0 software and sequences were inspected for quality, coverage and confidence by the Geneious v.7 software (Kearse et al., 2012). The ZIKV genome (10,648 bp) from Salvador Bahia collected on 01 July 2015 was deposited in the GenBank, accession number KU707826.

2.2. Phylogenetic analysis

In order to identify the origin of the ZIKV genome from Salvador, we performed a maximum likelihood (ML) phylogenetic analysis using ZIKV whole genome sequences (19 sequences) published in peer-reviewed journals for which sampling year and geographic location is available. Sequences were codon aligned using Clustal X and manually edited by Bioedit as already described (Hall, 1999) and sub-genomic regions were identified.

The ML tree reconstructed from the complete coding region, generated with the HKY + I + G substitution nucleotide model by using Phym1 v 3.0 (Guindon et al., 2003). The evolutionary model was chosen as the best-fitting nucleotide substitution model in accordance with the results of the hierarchical likelihood ratio test (HLRT) implemented in Modeltest software version 3.7 (Posada et al., 2004). The statistical robustness and reliability of the branching order within the phylogenetic trees was confirmed by the bootstrap analysis and considering as significant statistical support a bootstrap value >90%.

2.3. Evolutionary analysis

In order to identify if specific amino-acids were positively selected in the ZIKV viral genome, we performed a site-specific positive selection analysis.

The non-synonymous (dN) and synonymous (dS) rates (ω) per codon site was estimated by the ML approach implemented in the program HyPhy assuming a significance level of 1% ($\alpha = 0.01$) to infer the selection pressures acting on the Envelope, NS3 and NS5 gene (Pond et al., 2005) of ZIKV. In the HyPhy output, values of ω are

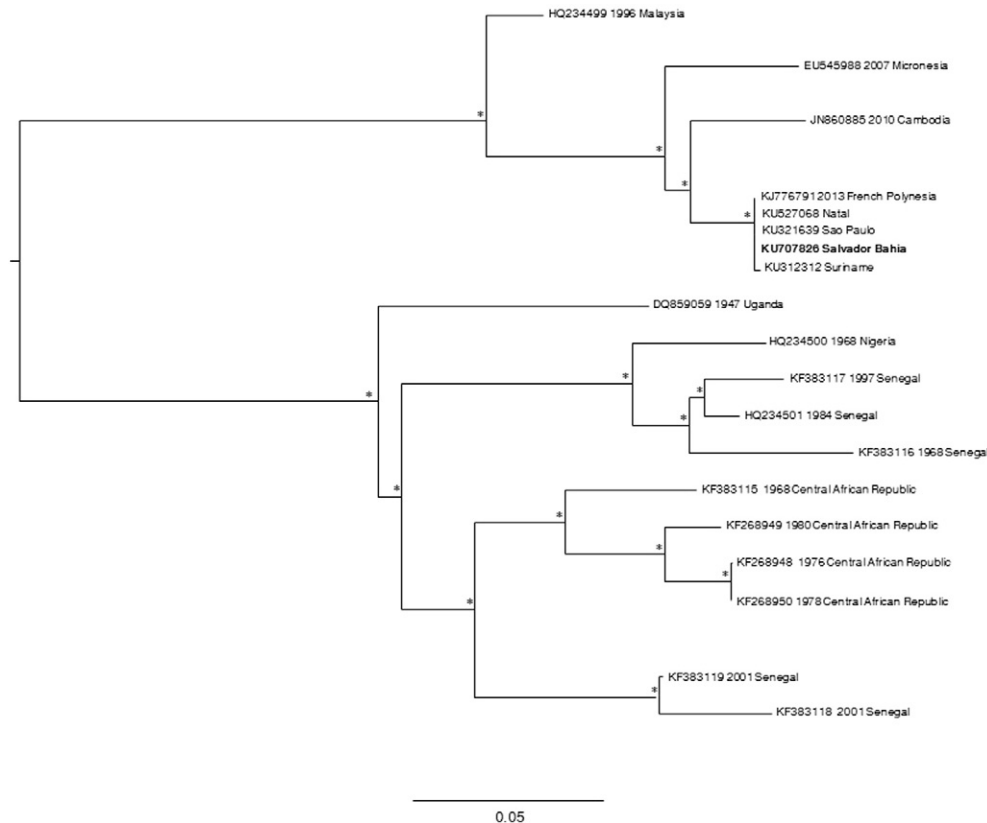


Fig. 1. Phylogenetic tree of ZIKV full-length coding sequences. Phylogenetic analysis of the complete genome sequences of Zika Virus. The tree was mid-point rooted. Scale bar is in units of nucleotide substitutions per site. Asterisks represent bootstrap values >90%. The GenBank accession number, year of isolation, and country of origin of each isolate are indicated on the tips of the tree for all strains except for those identified in 2015 and 2016. ZIKV strain from Salvador Bahia, Brazil (KU707826), was obtained in this study is highlighted in bold.

Table 1

Selection analysis for the Envelope, NS3 and NS5 genes of ZIKV showing the effect of molecular sequence variation on amino-acid substitution.

| Env | | | |
|---------------------------------|--|--|---|
| Negatively selected sites | 527(A), 528(H), 529(A), 530(K), 531(R), 532(Q), 533(T), 534(V), 535(V), 566(G), 568(L), 571(R), 572(L), 576(K), 577(L), 579(L), 582(V,A), 585(I), 604(T), 606(T), 608(A), 614(T,R), 616(G), 618(C), 619(K). | 536(V), 537(L), 538(G), 540(Q), 541(E), 542(G), 544(V), 545(H), 546(T), 548(L), 586(S), 587(C), 589(A), 590(A,V), 591(F,C), 592(T), 594(T,A), 595(K), | 549(A), 550(G), 553(L), 555(E,A), 559(A), 560(K,N), 561(G), 564(F,S), 596(V,I), 597(P), 598(A), 601(L), 602(H), 603(G), |
| ($\omega < 1$) HYPHY software | | | |
| NS3 | | | |
| Negatively selected sites | 1644(G), 1645(V), 1647(I), 1649(N), 1650(G), 1651(S), 1653(V), 1654(S), 1655(A), 1656(I), | 1657(T), 1658(Q), 1663(E), 1666(P), 1668(E), 1669(C), 1670(F), 1671(E), 1672(P), | 1673(S), 1683(L), 1684(D), 1685(L), 1686(H), 1687(P), 1689(A), 1690(G), |
| ($\omega < 1$) HYPHY software | 1691(K), 1692(T), 1693(R), 1694(R) 1700(V), 1703(A), 1705(K), 1708(A), 1709(R), 1742(T,A), 1746(T), 1747(E), 1748(I), 1749(V), 1750(D), 1751(L), 1752(M), 1778(A), 1779(H), 1782(D), 1783(P), 1785(S), 1786(I), 1787(A), 1789(R), 1816(A), 1817(F), 1818(P), 1819(D), 1820(S), 1824(I), 1834(R), 1840(F), 1855(P), | 1710(T), 1711(V), 1712(I), 1713(L), 1716(T), 1718(V), 1719(V,E), 1723(M,S), 1756(T), 1758(T), 1760(R), 1761(L), 1762(L), 763(Q), 1764(P), 1765(I), 1768(P), 1790(G), 1791(Y), 1792(I), 1794(T), 1796(V), 1797(E), 1799(G), 1801(A), 1802(A), 1803(A), 1845(D), 1846(H), 1848(G), 1853(F), 1856(S), 1858(R), 1859(N,S), 1862(E), 1866(C), 1 | 1725(E,W), 1726(A,S), 1738(A) 1739(V,I), 1740(K,N), 1741(V), 11769(N), 1770(Y), 1771(N), 1772(L), 1774(I), 1776(D), 1777(D), 1804(I), 1805(F), 1806(M), 1807(T), 1808(A), 1812(G), 1814(R), 1815(D), 869(K), 1874(V), 1878(S), 1880(K), 1886(F), 1888 (K), 1896(F). |
| NS5 | | | |
| Negatively selected sites | 3051(R), 3053(S), 3057(L), 3061(A), 3062(L), 3063(I), 3064(T), 3071(H), 3123(F), 3125(N), 3126(L), 3130(L), 3131(I), 3132(R), 3135(E), 3138(E), 3198(V), 3200(K), 3201(D), 3203(Q), 3204(E), 3206(K), 3207(P), 3208(S), | 3075(A), 3077(A), 3079(I), 3080(K), 3088(V), 3090(V), 3091(L), 3097(G), 3100(V), 3103(I), 3140(L), 3144(D), 3145(L), 3148(L), 3149(R), 3160(S), 3161(S,N), 3168(R), 3171(V), 3175(D), 3213(N), 3215(E), 3217(V), 3218(P), 3220(C), 3221(S), 3222(H), 3223(H), 3224(F). | 3104(I), 3109(Q), 3110(R), 3112(S), 3116(V), 3117(T), 3119(A), 3121(N), 3176(C), 3178(V), 3179(K), 3181(I), 3191(F), 3192(L), 3194(D), 3196(G), |
| ($\omega < 1$) HYPHY software | | | |

expressed as $\omega = dN - dS$. Therefore, ω smaller than zero ($\omega < 0$) indicate purifying, negative selection.

3. Results and discussion

The phylogenetic analysis showed that the isolate from Salvador (**KU707826**) belongs to the Asiatic genotype and clusters closely with strong bootstrap support (>90%) with the recent sequences isolated in Sao Paulo, Natal and Suriname (Fig. 1) (Enfissi, 2016; Mlakar et al., 2016). These full-length isolates from America form a single monophyletic clade that clusters with strong with an isolate from French Polynesia. Since recombination has been reported for ZIKV (Zanluca et al., 2015), we also confirmed these relationships by generating individual trees Envelope, NS3 and NS5 regions (Data not show).

Selection analyses uncovered several sites under strong negative selection indicated by $\omega < 0$, suggesting frequent purging of deleterious polymorphisms in functionally important genes (Table 1). Likewise, the lack of positively selected sites, indicated by $\omega > 0$, is typical of highly adapted phenotypes and shows no detectable directional change on the available data.

Although the ZIKV infection was initially considered a benign and self-limited disease, recent concerns about severe infection with neurological presentations and Guillain-Barré syndrome have been reported (Oehler et al., 2014). This coincided with the suggestion that ZIKV infection is correlated with birth malformations and microcephaly (Ventura et al., 2016), a hypothesis that is currently under investigation. Finally, given the rapid spread of ZIKV throughout the Americas, its potential for neurological complications, and a lack of effective diagnosis, vaccines and therapies, ZIKV is a public health issue of the greatest importance throughout the Americas.

Acknowledgments

The authors are grateful to Dalila Luciola Zanette for the collection of the sample and to the patient RCJS who agreed to participate in this study.

References

- Duffy, M.R., Chen, T.H., Hancock, W.T., Powers, A.M., Kool, J.L., Lanciotti, R.S., 2009. Zika virus outbreak on Yap Island, Federated States of Micronesia. *N. Engl. J. Med.* 360, 2536–2543.
- Enfissi, A., 2016. Zika virus genome from the Americas. *Lancet* 10015, 227–228.
- Faye, O., Freire, C.C., Iamarino, A., Faye, O., de Oliveira, J.V., Diallo, M., 2014. Molecular evolution of Zika virus during its emergence in the 20(th) century. *PLoS Negl. Trop. Dis.* 9, 8–2636.
- Grard, G., 2014. Zika virus in Gabon (Central Africa) – 2007: a new threat from *Aedes albopictus*? *PLoS Negl. Trop. Dis.* 8–2681.
- Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.
- Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.
- 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. *Bioinformatics* 28, 1647–1649.
- Kuno, G., Chang, G.J.J., 2007. Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. *Arch. Virol.* 152, 687–696.
- Lanciotti, R.S., Kosoy, O.L., Laven, J.J., Velez, J.O., Lambert, A.J., Johnson, A.J., Stanfield, S.M., Duffy, M.R., 2008. Genetic and serologic properties of Zika virus associated with an epidemic, Yap state, Micronesia, 2007. *Emerg. Infect. Dis.* 14, 1232–1239.
- Mlakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., Kolenc, M., Resman, R., Vesnaver, Vipotnik T., Fabjan, Vodušek V., Vizjak, A., Pižem, J., Petrovec, M., Avšič, Županc T., 2016. Zika virus associated with microcephaly. *N. Engl. J. Med.*
- Mourão, M.P., Souza, Bastos M., Figueiredo, R.P., Gimaque, J.B., Galusso, Edos S., Kramer, V.M., 2012. Mayaro fever in the city of Manaus, Brazil, 2007–2008. *Vector Borne Zoonotic Dis.* 2012, 12–42.
- Oehler, E., Watrin, L., Larre, P., Leparc-Goffart, I., Lasterre, S., Valour, F., Baudouin, L., Mallet, H.F., Musso, D., Ghawche, F., 2014. Zika virus infection complicated by Guillain-Barre

- syndrome — case report, French Polynesia, December 2013. *Euro Surveill.* 19, 9–20720.
- Pond, S.L., Frost, S.D., Muse, S.V., 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21, 676–679.
- Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Syst. Biol.* 53, 793–808.
- Rodrigues, Faria N., Lourenço, J., Marques de Cerqueira, E., Maia de Lima, M., Pybus, O., Alcantara Junior, L. Carlos, 2016. Epidemiology of chikungunya virus in Bahia, Brazil, 2014–2015. *PLoS Curr. Outbreaks*.
- Ventura, C.V., Maia, M., Bravo-Filho, V., Gois, A.L., Belfort, R.J., 2016. Zika virus in Brazil and macular atrophy in a child with microcephaly. *Lancet* 16 (387–228).
- Zamree, I., Drakes, N., Rohani, A., Lee, H.L., 2005. Sensitivity of *Aedes albopictus* C6/36 cells line for the detection and infectivity titration of dengue virus. *Trop. Biomed.* 22, 217–219.
- Zanluca, C., de Melo, V.C., Mosimann, A.L., Dos Santos, G.I., Dos Santos, C.N., Luz, K., 2015. First report of autochthonous transmission of Zika virus in Brazil. *Mem. Inst. Oswaldo Cruz* 110, 569–572.