

Asian lineage of Zika virus RNA pseudoknot may induce ribosomal frameshift and produce a new neuroinvasive protein ZIKV-NS1'

Tiago Tambonis^{1*}, Vinícius G. Contessoto^{2*}, Cíntia Bittar³, Marília F. Calmon³, Maurício L. Nogueira⁴, Paula Rahal³ & Vitor B. P. Leite¹

¹*São Paulo State University (UNESP), Department of Physics, São José do Rio Preto - SP, 15054-000, Brazil*

²*Brazilian Bioethanol Science and Technology Laboratory (CTBE), Campinas - SP, 13083-100, Brazil*

³*São Paulo State University (UNESP), Department of Biology - Genomic Studies Laboratory, São José do Rio Preto - SP, 15054-000, Brazil*

⁴*São José do Rio Preto School of Medicine (FAMERP), São José do Rio Preto - SP, 15090-000, Brazil*

**These authors contributed equally to this work*

Zika virus (ZIKV) is a threat to humanity, and understanding its neuroinvasiveness is a major challenge. Microcephaly observed in neonates in Brazil is associated with ZIKV that belongs to the Asian lineage. What distinguishes the neuroinvasiveness between the RNA lineages from Asia and Africa is still unknown. Here we identify an aspect that may explain the different behavior between the two lineages. The distinction between the two groups is the occurrence of an alternative protein NS1' (ZIKV-NS1'), which happens through a pseu-

doknot in the virus RNA that induces a ribosomal frameshift. Presence of NS1' protein is also observed in other *Flavivirus* that are neuroinvasive, and when NS1' production is suppressed, neuroinvasiveness is reduced.¹ This evidence gives grounds to suggest that the ZIKV-NS1' occurring in the Asian lineage is responsible for neuro-tropism, which causes the neuro-pathologies associated with ZIKV infection, of which microcephaly is the most devastating. The existence of ZIKV-NS1', which only exists in the Asian lineage, was inferred through bioinformatic methods, and it has yet to be experimentally observed. If its occurrence is confirmed, it will be a potential target in fighting the neuro-diseases associated with ZIKV.

Zika virus (ZIKV) is an *Arbovirus* belonging to the *Flaviviridae* family and *Flavivirus* genus like Dengue virus (DENV), West of Nile virus (WNV), Yellow Fever virus (YFV) and Japanese encephalitis virus (JEV). The ZIKV genome is similar to other viruses in the *Flavivirus* genus and encodes three structural proteins and seven non-structural (NS) proteins.² ZIKV was first isolated in Africa,³ and subsequently an Asian lineage was also reported.^{4,5} Neurological complications such as Guillain-Barré syndrome in adults in French Polynesia and microcephaly in newborns in Brazil were associated with the Asian ZIKV lineage.⁶⁻⁸ There is well-known evidence of neurological problems associated with *Flavivirus* genus.⁹ In particular, for JEV and WNV, the production of an eleventh protein related to the viral neuroinvasiveness was observed, the NS1'.^{1,10} This protein was first observed in the 80s in JEV-infected cells,¹¹ and its production mechanism was recently discovered.¹⁰ NS1' is produced as the result of a -1 ribosomal frameshift at the 5' end of the NS2A coding sequence, which consists of the NS1 protein with the addition of 52 more residues.¹²

The -1 ribosomal frameshift occurs through a pseudoknot formation preceded by a ‘slippery’ heptanucleotide, which is followed by a ‘spacer’ region.¹³ This mechanism is also observed in other viruses.¹⁴

Here, using bioinformatic approaches, we investigated the likelihood of a ribosomal frameshift occurrence in ZIKV, which may account for the translation of a NS1’ protein. 29 different Zika virus gene sequences from GenBank¹⁵ were analyzed, of which 15 were from Asian lineage and 14 from African lineage. Probing for the conditions for the occurrence of a ribosomal frameshift, a ‘slippery’ heptanucleotide *CCCAAU* and a pseudoknot formation propensity in a position near the extremity of the NS1 gene of ZIKV was observed. These features allow the ribosomal frameshift to happen in ZIKV, in similar positions to those occurring in JEV and WNV. This frameshift occurrence in ZIKV was observed in all Asian lineage sequences but not in African lineage sequences. The Brazil strain in which microcephaly cases were observed comes from the Asian lineage.^{16,17} On the other hand, no cases of microcephaly associated with the ZIKV African lineage were reported.¹⁸

The predicted -1 frameshift in the ZIKV RNA code is located at nucleotide position 794 in the NS1 region, allowing a translation of 40 different residues until the stop codon *UAA* appears at nucleotide position 913 after the occurrence of the frameshift. The translated ZIKV-NS1’ protein has 303 residues with ~35 kDa of molecular weight. ZIKV-NS1’ is 49 residues smaller than the NS1 protein, and both proteins are similar in residue sequence until position 263, but after this position the residue sequences are completely different. The ZIKV-NS1’ amino acid sequence

based on the KU321639 Genbank entry is: GCSVDFSKKETRCGTGVFVYNDVEAWRDRY
KYHPDSPRRLAAAVKQAWEDGICGISSVSRMENIMWRSVEGELNAILEENGVLTVVV
GSVKNPMWRGPQRLPVPVNELPHGWKAWGKSHFVRAAKTNSFVVDGDTLKECPLK
HRAWNSFLVEDHGFGVFHTSVWLKVREDYSLECDPAVIGTAVKGKEAVHSDLGYWIE
SEKNDTWRLKRAHLIEMKTCEWPKSHTLWTDGIEESDLIIPKSLAGPLSHHNTREGYRT
QMKARGGNMWNKRTISEINHCKRKGDRGMVLQGVHNAPTVPVPG, in which the bold
part of the sequence is associated with the 40 residues that are translated after the pseudoknot,
and they are nearly identical (two different AA at the most) for all ZIKV Asian sequences. The
results, along with experimental evidence from JEV and WNV studies,^{1,10} suggest that the Asian
lineage ZIKV RNA has the essential characteristics to produce the pseudoknot, which allows for a
ribosomal frameshift to produce the unexpected ZIKV-NS1' protein. Moreover, the appearance of
ZIKV-NS1' protein may be associated with the neuroinvasiveness of ZIKV, and may be related to
microcephaly in neonates.

Studies in JEV and WNV show that the addition of mutations in nucleotide which makes it
difficult for the pseudoknot in viral RNA to form, correlates with the reduction of neuroinvasive-
ness in mice.^{1,19,20} It is plausible to assume that ZIKV neuroinvasiveness may be related to the
occurrence of ZIKV-NS1', just like to the NS1' observed in JEV and WNV. ZIKV-NS1' is ex-
pected to be rarely expressed, similarly to NS1' for JEV and WNV, but should be experimentally
detectable in a straightforward manner.

A recent study of ZIKV with x-ray crystallography showed pseudoknot formation in its RNA,

and these structures were associated with exonuclease Xrn1 resistance, which also points to the importance of fully understanding pseudoknot formation and function.²¹ Experimental groups with the necessary expertise are urged to validate this conjecture. The corroboration of the existence of ZIKV-NS1' protein may open a window of opportunity in the understanding of its relationship with microcephaly and ZIKV neuroinvasiveness, as well as in the development of drugs and vaccines with this specific target.

Methods

The frameshift mechanism consists of a 'slippery' heptanucleotide followed by a 'spacer' region of 5-9 nucleotides and then a stable RNA secondary structure such as a pseudoknot. The 'slippery' heptanucleotide consensus sequence is NNNWWWH, where NNN are any three identical nucleotides, WWW represents AAA or UUU and H represents A, C or U.¹⁴ The 29 studied gene sequences of Zika virus were obtained from GenBank,¹⁵ being 15 from Asian lineage: KU497555, KU729218, KU321639, KU729217, KU365778, KU365777, KU365780, KU365779, KU707826, KJ776791, KU922923, KU922960, KU820899, KU744693, KU963796 and 14 from the African lineage: KF268948, KF268950, KF268949, AY632535, LC002520, HQ234501, HQ234500, KF383117, NC_012532, KF383118, KF383115, KF383119, DQ859059, KF383121. All possible 'slippery' heptanucleotide combinations were exhausted constructed and performed in R environment²² *version 3.3.2*. The formation of pseudoknot structures was analyzed using *pKiss* tool²³ from nucleotide sequence: GGCCAUGGCACAGUGAAGAGCUUGAAAUUCGGUUUGAGGAAUG CCCAGGCACUAAGGUC and structure [[[[.....(((((((.....)))))).....)]]. All se-

quence alignments were executed using the MUSCLE algorithm²⁴ from *Bioinformatic toolkit* web server.²⁵

1. Melian, E. B. *et al.* NS1' of flaviviruses in the Japanese encephalitis virus serogroup is a product of ribosomal frameshifting and plays a role in viral neuroinvasiveness. *Journal of Virology* **84**, 1641–1647 (2010).
2. Mukhopadhyay, S., Kuhn, R. J. & Rossmann, M. G. A structural perspective of the flavivirus life cycle. *Nature Reviews. Microbiology* **3**, 13–22 (2005).
3. Dick, G. W. A., Kitchen, S. F. & Haddock, A. J. Zika Virus (I). Isolations and serological specificity. *Transactions of The Royal Society of Tropical Medicine and Hygiene* **46**, 509–520 (1952).
4. Hayes, E. B. Zika Virus Outside Africa. *Emerging Infectious Diseases* **15**, 1347–1350 (2009).
5. Cao-Lormeau, V.-M. *et al.* Zika virus, French polynesia, South pacific, 2013. *Emerging Infectious Diseases* **20**, 1085–1086 (2014).
6. Oehler, E. *et al.* Zika virus infection complicated by Guillain-Barre syndrome—case report, French Polynesia, December 2013. *Euro Surveillace: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin* **19** (2014).
7. Schuler-Faccini, L. *et al.* Possible Association Between Zika Virus Infection and Microcephaly - Brazil, 2015. *MMWR. Morbidity and mortality weekly report* **65**, 59–62 (2016).

8. Cugola, F. R. *et al.* The Brazilian Zika virus strain causes birth defects in experimental models. *Nature* **534**, 267–271 (2016).
9. Sips, G. J., Wilschut, J. & Smit, J. M. Neuroinvasive flavivirus infections. *Reviews in Medical Virology* **22**, 69–87 (2012).
10. Firth, A. E. & Atkins, J. F. A conserved predicted pseudoknot in the NS2a-encoding sequence of West Nile and Japanese encephalitis flaviviruses suggests NS1' may derive from ribosomal frameshifting. *Virology Journal* **6**, 14 (2009).
11. Mason, P. W. Maturation of Japanese encephalitis virus glycoproteins produced by infected mammalian and mosquito cells. *Virology* **169**, 354–364 (1989).
12. Melian, E. B. *et al.* Programmed Ribosomal Frameshift Alters Expression of West Nile Virus Genes and Facilitates Virus Replication in Birds and Mosquitoes. *PLoS Pathogens* **10** (2014).
13. Staple, D. W. & Butcher, S. E. Pseudoknots: RNA Structures with Diverse Functions. *PLOS Biol* **3**, e213 (2005).
14. Atkins, J. F., Loughran, G., Bhatt, P. R., Firth, A. E. & Baranov, P. V. Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic Acids Research* gkw530 (2016).
15. Benson, D. A. *et al.* GenBank. *Nucleic Acids Research* **43**, D30–D35 (2015).
16. Campos, G. S., Bandeira, A. C. & Sardi, S. I. Zika Virus Outbreak, Bahia, Brazil. *Emerg. Infect. Dis.* **21**, 1885–1886 (2015).

17. Faria, N. R. *et al.* Zika virus in the Americas: Early epidemiological and genetic findings. *Science* aaf5036 (2016).
18. Mlakar, J. *et al.* Zika Virus Associated with Microcephaly. *New England Journal of Medicine* **374**, 951–958 (2016).
19. Liu, W. J. *et al.* A single amino acid substitution in the West Nile virus nonstructural protein NS2a disables its ability to inhibit alpha/beta interferon induction and attenuates virus virulence in mice. *Journal of Virology* **80**, 2396–2404 (2006).
20. Ye, Q. *et al.* A single nucleotide mutation in NS2a of Japanese encephalitis-live vaccine virus (SA14-14-2) ablates NS1' formation and contributes to attenuation. *The Journal of General Virology* **93**, 1959–1964 (2012).
21. Akiyama, B. M. *et al.* Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. *Science* aah3963 (2016).
22. R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2008). ISBN 3-900051-07-0.
23. Janssen, S. & Giegerich, R. The RNA shapes studio. *Bioinformatics* **31**, 423–425 (2015).
24. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797 (2004).

25. Alva, V., Nam, S.-Z., Sding, J. & Lupas, A. N. The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Research* **44**, W410–W415 (2016).

Acknowledgements TT was supported by Coordination for the Improvement of Higher Education Personnel (CAPES), Brazil. VGC was funded by grant 2016/13998-8, São Paulo Research Foundation (FAPESP) and CAPES. VBPL was supported by the National Council for Scientific and Technological Development (CNPq) and FAPESP grant 2014/06862-7. CB was supported by the grant CNPq 165802/2015-4. MFC, MLN and PH supported by FAPESP grant 2014/22198-0 and by CNPq grant 440723/2016-7.

Author Contributions C.B., P.R., T.T., V.G.C. and V.B.P.L. conceived the experiment, T.T. and V.G.C. conducted the experiment and analysed the results. All authors reviewed the manuscript.

Author Information Correspondence and requests for materials should be addressed to V.B.P.L. (vleite@sjrp.unesp.br).