Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Codon adaptation biases among sylvatic and urban genotypes of Dengue virus type 2

Marielton dos Passos Cunha\textsuperscript{a,1}, Ayda Susana Ortiz-Baeza\textsuperscript{a,b,1}, Caio César de Melo Freire\textsuperscript{c}, Paolo Marinho de Andrade Zanotto\textsuperscript{a,⁎}

\textsuperscript{a} Laboratory of Molecular Evolution and Bioinformatics, Department of Microbiology, Biomedical Sciences Institute, University of São Paulo, São Paulo, Brazil  
\textsuperscript{b} Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, Colombia  
\textsuperscript{c} Department of Genetics and Evolution, UFSCar - Federal University of São Carlos, São Paulo, Brazil

1. Introduction

Dengue virus (DENV) is the world’s most important mosquito-borne viral human pathogen. It is widespread throughout tropical and subtropical regions (Bhatt et al., 2013; Gubler, 1998). DENV has a natural viral human pathogen. It is widespread throughout tropical and sub-

DENV-2 genotypes. Moreover, CAI values were signi-

Aedes aegypti. Remarkably, we found no significant differences in codon adaptation to human between urban American/Asian and sylvatic DENV-2 genotypes. Moreover, CAI values were significantly different, when comparing all genotypes to Ae. aegypti codon preferences, with lower values for sylvatic than urban genotypes. In summary, our findings suggest the presence of a molecular signature among the genotypes that circulate in sylvatic and urban environments, and may help explain the trafficking of DENV-2 strains to an urban cycle.

1567-1348/ © 2018 Published by Elsevier B.V.

https://doi.org/10.1016/j.meegid.2018.05.017

Available online 21 May 2018

1567-1348/ © 2018 Published by Elsevier B.V.
amino acid can be coded by different codons. The redundancy in the
genetic code has an important role in controlling metabolic processes,
but not all species use its built-in codon redundancy in the same way.
The unequal preference of specific codons over other synonymous
codons during the translation process creates a bias in codon usage.
Codon usage biases are common throughout the Tree of Life, and for
viruses, the balance between mutation and natural selection allows for
changes in codon usage biases (Morton, 2003). Genetic plasticity and
the capacity to adapt to new hosts facilitate the emergence of RNA
viruses in novel, unexplored environments. This process entails both, (i)
adapting to different types of cellular machinery involved in viral re-
plication and (ii) the evasion from different types of cellular antiviral
responses. However, the mechanisms used for virus for trafficking
among hosts remain poorly understood (Bahir et al., 2009; Longdon
et al., 2014; Pal et al., 2014). Because DENV must infect successfully
and alternate between mosquito and human hosts, nucleotide analyses
support the notion that arboviruses use a restricted and balanced nu-
cleotide composition as a compromise to be able to infect both hosts,
thus successfully ensuring several processes such as translational effi-
ciency and replication (Shen et al., 2015). In this context, the Codon
adaptation index (CAI) correlates with gene expression levels and
adaptation of viral genes to their hosts (Gustafsson et al., 2012; Neame,
2009; Pal et al., 2014). In the present study, we performed compre-
hensive analyses of codon adaptation of DENV-2 from different habitat
settings and host systems.

2. Material and methods

2.1. Sequence datasets

The sequences used in this study were downloaded from the National Center for Biotechnology Information (NCBI) (https://www.
cbi.nlm.nih.gov/genbank/) website in GenBank format. The dataset
included 877 sequences from all six genotypes as follows: American
(16), American/Asian (435), Asian 1 (228), Asian 2 (20), Cosmopolitan
(163) and Sylvatic (15). Comparisons of codon usage preferences were
performed against reference sets from Homo sapiens (3803 CDSs)
(https://github.com/CaioFreire/CUB) and Aedes aegypti (585 CDSs)
(http://www.kazusa.or.jp/codon/), using the standard genetic code.
All the complete genomic sequences available for DENV-2 with in-
formation regarding the location and year of isolation were recovered,
and later converted into FASTA format.

2.2. Molecular phylogenetic approach

FASTA sequences were aligned using Clustal Omega (Larkin et al.,
2007) and recombinant sequences were screened using all algorithms
implemented in RDP4 program (RDP, GENECONV, BootScan, MaxChi,
Chimaera, Siscan and 3Seq) using the standard settings (Martin et al.,
2015). The alignment of recombinant free sequences was manually
inspected and edited using the program AliView v.1.18 (Larsson, 2014),
resulting in a final dataset with 877 coding sequences (Table S1). Viral
phylogenies based on full-length coding sequences were estimated
using Maximum Likelihood (ML) implemented in FastTree 2 (Price
et al., 2010). We first evaluated the best transition model to be GTR + I
using JModelTest v. 0.1.1. The final tree was then visualized and plotted
using FigTree v.1.4.3 (http://tree.bio.ed.ac.uk). All sequences
used in this work are presented in the format: genotype/accession
number/country/year of isolation.

2.3. CAI of DENV to humans and Aedes aegypti mosquitoes

CAI is a measure of silent, synonymous codon usage bias based on
the codon preference of a viral strain and a codon usage table for a
given host (Sharp et al., 1986; Sharp and Li, 1987). We applied the CAI
using a frequency table for housekeeping of human genes (Eisenberg
and Levanon, 2013) (available at https://github.com/CaioFreire/CUB)
and for Ae. aegypti using the table available in the Codon Usage Data-
base (Nakamura, 2000). The CAI values were calculated to measure the
synonymous codon usage bias using the CAICal program (Puigbò et al.,
2008a). To evaluate the statistical support of the CAI values, we define
a threshold value or expected CAI (e-CAI) (Puigbò et al., 2008b) by
generating random sequences with GC content, amino acid composition
and sequence length similar to the DENV-2 query sequences. CAI values
above the e-CAI are interpreted as statistically significant, meaning that
codon similarity arises from codon preferences rather than from in-
ternal biases (Puigbò et al., 2008b).

2.4. Statistical analysis

Statistical significance was determined by the Mann-Whitney U test
to compare the difference between CAI values for humans and the
mosquito, assuming a significance level of 0.05. Multiple comparisons
among genotypes were performed using a Kruskal-Wallis test im-
plemented in the pgirmess v1.6.5 package for the R statistical en-
vironment. To evaluate levels of independence, we further carried out a
Dunn’s post hoc test. Further, to avoid Type I error inflation, we applied
a Bonferroni adjustment of p-values, using the PMCMR package, at a
significance level of 0.05. Perl and R scripts were used to analyze most
of the data in this study and are available from the authors upon re-
quest.

3. Results

As has been well documented, we recovered six phylogenetically
structured groups or genotypes for DENV-2, based on the open reading
frame of 877 non-recombinant sequences. The most basal group en-
compassed lineages from West Africa and Asia of Sylvatic genotype,
which reinforced the notion that the urban genotypes emerged from the
sylvatic cycle, as previously described (Vasilakis et al., 2011; Vasilakis
et al., 2007a). Urban genotypes were also identified as: (i) American,
(ii) American/Asian, (iii) Asian 1, (iv) Asian 2 and (v) Cosmopolitan
genotypes (Fig. 1A, Fig. S1). Comparison of the CAI values obtained
from human and Ae. aegypti codon usage tables, shown in Fig. 1B,
revealed significant differences between hosts (p-value < 0.05), demon-
strating that DENV-2 codon usage is more similar to that of humans
than to Ae. aegypti.

In addition, we assessed the values of CAI for each genotype in
correlation to human and Ae. aegypti hosts. Because CAI values for the
polyprotein sequences were not normally distributed, as assessed using
Shapiro-Wilk and Kolmogorov-Smirnov tests, we determined statistical
significance with the Mann-Whitney U test to compare among CAI es-
timates. The distribution of CAI values for different genotypes showed
that all the genotypes fell above the e-CAI estimated for each host.
Furthermore, statistical analyses showed significant differences (p-
value < 0.05) among urban genotypes for both hosts (Table 1).

Significant differences between urban and sylvatic genotypes were
identified in Ae. aegypti (Table 1). Likewise, the genotypes with the
highest CAI were Asian 1, Asian 2 and the American genotypes. We also
observed that the CAI values for the sylvatic genotype were the lowest
when comparing with urban genotypes in A. aegypti (p-value < 0.05)
(Fig. 1B and Table 1). This suggested that genotypes from urban setting
could have experienced some fine-tuning process in codon optimization
to translation in A. aegypti.

For humans, the only pairwise comparison that was not significantly
different between urban and sylvatic genotypes was that between the
American/Asian and the Sylvatic genotype (Fig. 1B and Table 1), sig-
ifying that they share similar silent selective pressures. The American,
Asian 2 and Cosmopolitan genotypes had the highest CAI values for
humans (Fig. 1B). Finally, we observed that these values were similar
and roughly constant over time (Fig. 2). Crucially, values were higher
than the e-CAI for all the genotypes in human and mosquito cells,
indicating that codon usage preferences were not random in DENV-2 genotypes (Fig. 2).

4. Discussion

The potential emergence of sylvatic strains into urban cycles has become an increasing focus of research and public concern (Moncayo et al., 2004; Vasilakis et al., 2011; Vasilakis et al., 2007b). Recently, this concern become more tangible after some studies have suggested human infection cases with presumably sylvatic DENV-2 strains. Phylogenetic analyses showed that the highly divergent strains (QML22/2015, D2Sab2015) of DENV-2 was basal and strikingly different from all other previously isolated strains (Liu et al., 2016; Pyke et al., 2017). In both cases, we excluded these sequences from our analyses, because including highly divergent sequences without knowing which genotype belong, may impact on the estimates for the whole dataset due to sampling bias and differences in the evolutionary rates.

Previous data suggest that the nature of selective pressures is broadly equivalent in urban and sylvatic strains (Vasilakis et al., 2011; Vasilakis et al., 2007b). This could mean that the successful emergence of a novel sylvatic DENV strains in an urban cycle is unlikely to require a major adaptive challenge (Vasilakis et al., 2007b).

Insects and mammals are separated by about 1 billion years of evolutionary history, and they differ at the organismal level and in many biochemical processes (Lobo et al., 2009; Shen et al., 2015). Since DENV needs to negotiate with different hosts (i.e. human and Aedes mosquitoes), it is fair to assume that the efficiency in the viral translation process is a key factor sustaining the urban cycle.

Our results suggested that codon adaptation was higher for human cells than for Ae. aegypti (Fig. 1, Fig. 2). This was also reported for other flavivirus (Butt et al., 2016; Freire et al., 2015; Moratorio et al., 2013; Pal et al., 2014). In addition, similar CAI values recovered for urban and sylvatic genotypes suggests their ability to possibly make similar use of human cellular machinery. This notion is consistent with kinetic experiments comparing the replication of sylvatic and urban DENV-2 strains, which could explain that the emergence of urban DENV-2 from sylvatic progenitors may not have required adaptation to replicate more efficiently in humans (Vasilakis et al., 2007b). Notwithstanding, when using the CAI tables from Ae. aegypti, we observed a conspicuous difference between urban and sylvatic genotypes with lower CAI values for the sylvatic genotype (Fig. 2). This finding suggested that the emergence of urban genotypes into the urban cycle may have required viral adaptation towards increased competence in Ae. aegypti, which should then entail higher efficiency in viral replication. This is in good agreement with susceptibility experiments reported by Moncayo et al. (2004). Furthermore, competence studies with Ae. aegypti from regions where sylvatic genotype circulation occurs indicate its low vector capacity for Dengue virus (Amarasinghe et al., 2011; Diallo et al., 2008). Despite this, there was no previous attempt to establish an association with codon usage biases.

Fluctuations in CAI distribution across genotypes (Fig. 1B) indicated possible differences in codon adaptation for both hosts (Table 1).

Table 1

| Genotype     | American | American/Asian | Asian 1 | Asian 2 | Cosmopolitan | Sylvatic |
|--------------|----------|----------------|---------|---------|--------------|----------|
| American     | -        | *              | ns      | *       | ns           | *        |
| American/Asian| *       | -              | *       | ns      | *            | *        |
| Asian 1      | ns       | *              | -       | *       | ns           | *        |
| Asian 2      | ns       | *              | ns      | -       | ns           | *        |
| Cosmopolitan | *        | ns             | *       | *       | -            | ns       |
| Sylvatic     | *        | *              | *       | *       | *            | -        |

ns – Non statistical significance.
Differential selection pressures exerted by the complex interplay of distinct virus factors, such as infection efficiency, population density and herd immunity (Plowright et al., 2017; Salazar et al., 2010). For example, a recent study reported by Salje et al. (2017) demonstrated the association between population density and the establishment of transmission chains. The increase in mosquito populations has also been identified as a potential determinant of the emergence and increased transmission of arbovirus, particularly in immunologically naïve human populations (Dudas et al., 2018; Pettersson et al., 2016).

Our findings bring into question whether sylvatic strains could reach similar levels of transmission as the urban genotypes in Asian and American populations if the vector competence of \textit{Ae.aegypti} is increased. In fact, recent outbreaks caused by other emerging viruses \textit{e.g.} Ebola virus, Influenza A virus (H1N1), Middle East respiratory syndrome coronavirus (MERS-CoV)], have occurred due to the zoonotic spillover and adaptation of the virus to replication in human cells (Dudas et al., 2018; Mänz et al., 2016; Plowright et al., 2017; Urbanowicz et al., 2016). In this context, as previous analyses have suggested (Moncayo et al., 2004), our findings support the notion that for the sylvatic strains to effectively colonize the urban environment, the virus needs a number of silent, adaptive nucleotide substitutions to optimize the codon usage to invertebrate host cells, while maintaining a compositional base balance suitable for efficient alternate spread among both human and insect hosts. Apparently, even when humans are susceptible to infection by DENV vectored by \textit{Ae. aegypti}, its reduced vectorial competence ultimately constitutes a hurdle rather than an enabler of virus transmission.

5. Conclusions

In conclusion, our findings provide a comprehensive assessment of the codon adaptation of DENV-2 in different habitats \textit{i.e.} urban and sylvatic settings and host systems \textit{i.e. Homo sapiens and the mosquito vector \textit{Aedes aegypti}}. In this context, although the virus replicates in both human and mosquitoes, our analysis suggested that DENV-2 codon usage is better adapted to humans than to the main cosmopolitan vector \textit{(Ae. aegypti)}. This would imply that there is still room for adaptation and improved transmission among new DENV-2 strains for causing future pandemics.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2018.05.017.

Acknowledgments and funding

This work was supported by the Brazilian National Council of Scientific and Technological Development (CNPq) process No. 441105/2016-5. MPC and ASOB received FAPESP grants: No. 2016/08204-2 and No. 2013/25434-3, respectively. The funders had no role in the data collection or analysis, decision to publish, or preparation of the manuscript. We thank Felipe G. Naveca and Valdinete A. Nascimento for the support.

References

Amarasinghe, A., Kuritsky, J.N., William Letson, G., Margolis, H.S., 2011. Dengue virus infection in Africa. Emerg. Infect. Dis. 17, 1349–1354. http://dx.doi.org/10.3201/eid1708.101515.
