



Evolução molecular dos genótipos do vírus febre amarela através dos padrões de utilização de códons

Palavras-Chave: Vírus Febre Amarela, Genótipo, Evolução viral

Autores:

**Elisa Zamboin Oliveira, Laboratório de Bioinformática e Virologia, Departamento de Genética, Evolução, Microbiologia e Imunologia, Instituto de Biologia, Universidade de Campinas
Prof. Dr. Marielton dos Passos Cunha (orientador), Laboratório de Bioinformática e Virologia, Departamento de Genética, Evolução, Microbiologia e Imunologia, Instituto de Biologia, Universidade de Campinas**

INTRODUÇÃO:

A Febre Amarela é uma doença causada pelo *Orthoflavivirus flavi* (anteriormente chamado de vírus febre amarela, YFV), um vírus que é membro do gênero *Orthoflavivirus*. Esse grupo taxonômico agrupa diversas espécies virais que possuem um genoma constituído por uma molécula de RNA de fita simples com polaridade positiva, que codifica uma única poliproteína a qual após a tradução, é clivada para constituir as proteínas estruturais (C, prM e E) e não-estruturais (NS1, NS2A, NS2B, NS3, NS4A, NS4B e NS5). O YFV possui um único sorotipo, porém, diversifica-se em quatro genótipos: (i) Leste Africano, (ii) Oeste Africano, (iii) América do Sul I, e (iv) América do Sul II. Estes genótipos circulam em dois ciclos de manutenção na natureza, onde o ciclo urbano ocorre na África, e o ciclo silvestre ocorre nas Américas e também, na África. Apesar de ser clara a existência de uma associação temporal e espacial associada à diversidade genômica do vírus, existe ainda, uma ausência de

conhecimento biológico ligado a cada um dos genótipos do YFV, particularmente ligados à ecologia e evolução do vírus. Neste sentido, este projeto teve como objetivo caracterizar os padrões genômicos associados ao comportamento biológico do YFV. A caracterização molecular dos genótipos envolvidos nessa rede de relações biológicas, torna-se primordial na compreensão da sua história evolutiva passada e, também, na projeção de possíveis trajetórias ocultas neste contexto biológico, o que poderá auxiliar na inferência de possíveis novos hospedeiros ou na compreensão do seu ciclo em hospedeiros habituais.

METODOLOGIA:

Os dados genômicos virais utilizados foram obtidos do Centro Nacional de Informações Biotecnológicas (<https://www.ncbi.nlm.nih.gov/>) como um dataset contendo os dados de nucleotídeos e informações associadas ao

depósito. Para o desenvolvimento do projeto, recuperamos 439 sequências do YFV, e 9 sequências de vírus relacionados ao YFV (abrangendo os vírus *Edge Hill*, *Potiskum*, *Saboya*, *Sepik*, *Banzi*, *Uganda S virus*, *Bouboui*, *Jugra* e *Wesselsbron*) (aqui denominados *related-YFV* - rYFV), utilizando filtros específicos voltados à exclusão de sequências que poderiam ser problemas para as análises, como os replicons e genomas incompletos. A manipulação das sequências, envolvendo desde o processamento inicial do dataset, passando pelo cálculo dos valores de uso relativo de códons sinônimos (do inglês, *relative synonymous codon usage* - RSCU), cálculo dos valores do índice de adaptação de códons (do inglês, *Codon adaptation index* - CAI) e geração dos gráficos, foram realizados através de ferramentas utilizadas por meio da linguagem *Python3*.

RESULTADOS E DISCUSSÃO:

Os resultados de RSCU indicam a formação de grupos de sequências de acordo com o genótipo de YFV (Fig. 1A) e também de acordo com as espécies virais (YFV *versus* rYFV) (Fig. 1B), denotando padrões no uso de códons de acordo com essas *features*. Ao observamos o agrupamento das sequências pertencentes a um mesmo grupo viral (Fig. 1B) (considerando YFV e rYFV), os rYFV formam um grupo de *Orthoflavivirus* próximos de YFV, porém, diferente dos YFV, são organismos associados à circulação em mosquitos do gênero *Aedes spp.* com hospedeiros vertebrados diversos (dentre mamíferos e aves), e além disso, o

potencial de circulação dos rYFV em humanos não é ainda compreendido.

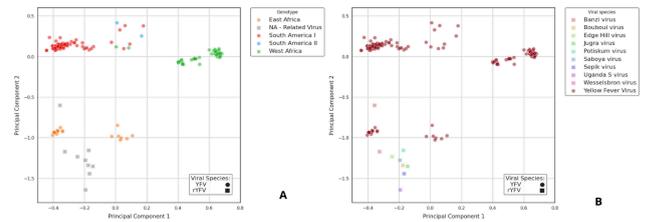


Figura 01. (A) Análise de Componentes Principais (PCA) entre os distintos genótipos distintos de YFV e **(B)** entre as diferentes espécies virais (YFV *versus* rYFV) através dos valores de RSCU. Na figura A, as cores indicam os genótipos virais, e em B, as cores indicam as espécies virais.

Quando observamos a clusterização entre os genótipos (Fig. 1A), percebemos que os genótipos Oeste Africano, América do Sul I e América do Sul II apresentam-se mais próximos, o que pode indicar a proximidade evolutiva, e também semelhança entre para o padrão no utilização de códons. Os genótipos americanos foram introduzidos no continente Americano durante a diáspora africana, e após a sua introdução, diversificaram-se a partir do genótipo do oeste africano, o que pode ser compreendido comparando a distância entre os grupos de sequências de cada um dos genótipos. Entretanto, quando observamos o genótipo do Leste Africano, denotamos uma maior distância de seu cluster para os outros genótipos de YFV.

Essa distância do genótipo do Leste Africano aos demais também é observada na reconstrução filogenética de YFV, mostrando que o grupo possui um ancestral comum com os outros genótipos mais antigo que o último ancestral comum aos outros genótipos (Oeste Africano, América do Sul I, e América do Sul II)

de YFV. Além disso, o genótipo encontra-se próximo aos rYFV, o que sugere uma possível proximidade na utilização de códons entre o genótipo e o grupo o que, por sua vez, pode também assemelhá-los em características ecológicas e epidemiológicas.

Os resultados obtidos através da análise de *Codon Adaptation Index* (CAI), comparando as sequências virais à tabelas de utilização de códons por possíveis hospedeiros, reforçam as observações prévias baseadas em RSCU.

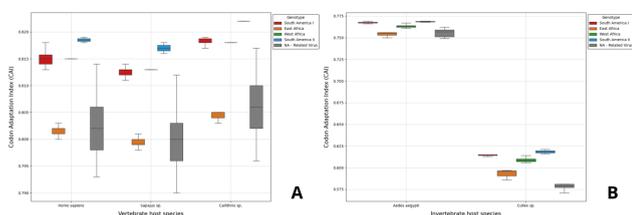


Figura 02. Comparação dos valores de CAI para os principais hospedeiros considerados, **(A)** vertebrados e **(B)** invertebrados, considerando todos os genótipos de YFV, e também os rYFV. As cores representam os possíveis hospedeiros considerados na análise.

Os valores de CAI tanto para hospedeiros vertebrados como invertebrados (Fig. 2 A e B), denotam um padrão de distribuição semelhante entre si. Os valores de CAI para os genótipos Americanos e do Oeste Africano apresentam uma distribuição de valores similar, indicando uma possível correspondência no modo que se relacionam com os seus respectivos hospedeiros. Essa característica pode impactar diretamente no ciclo de circulação destes vírus. Mesmo com registros acerca de um ciclo urbano para o genótipo Oeste Africano, ainda observamos a predominância do estado selvagem para o genótipo, sendo similar ao

observado nos genótipos que circulam no continente Americano.

Seguindo o mesmo padrão, o genótipo do Leste Africano conserva sua similaridade com os vírus relacionados a YFV, indicando também semelhança na forma que os grupos utilizam o maquinário celular para tradução nas células hospedeiras. Esse aspecto torna-se ainda mais relevante quando olhamos para os valores de CAI para humanos (*Homo sapiens*). A proximidade entre eles denota uma baixa adaptação para esse hospedeiro. A escassez de informações sobre a ocorrência de rYFV em humanos e a ausência de surtos de febre amarela no leste africano são informações epidemiológicas que reforçam os padrões observados nos nossos resultados.

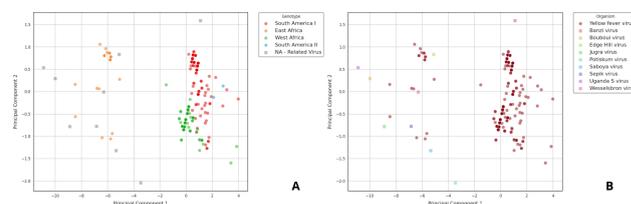


Figura 03. **(A)** Clusterização entre os distintos genótipos de YFV e **(B)** entre as diferentes espécies virais (YFV versus rYFV) através dos valores de CAI. As cores distintas representam os possíveis hospedeiros vertebrados e invertebrados considerados na análise.

Ao observarmos os gráficos a partir dos resultados de CAI (Fig. 03), denotamos o mesmo padrão de agrupamento denotado para os valores de RSCU. Isso indica que as métricas distintas sugerem um mesmo resultado Além disso, ao observarmos dentre as espécies de rYFV, observamos que a espécie *Banzi virus* apresenta proximidade com os genótipos de YFV (Oeste Africano, América

do Sul I e América do Sul II), se distinguindo do padrão observado para os demais rYFV, o que sugere que essa espécie viral pode representar um possível evento de emergência à circulação entre humanos.

CONCLUSÕES:

Desse modo, a partir do reconhecimento de distintos padrões moleculares característicos de genótipos virais, os quais possuem características biológicas específicas, torna-se possível entender os processos históricos da relação vírus-hospedeiro. Assim, o avanço do estudo poderá elucidar as bases que sustentam o surgimento, permanência, ou, até mesmo, a reemergência viral em diversas cadeias de transmissão, com um destaque ao estudo da análise dos padrões moleculares do YFV.

BIBLIOGRAFIA

01. Bryant, Juliet E; Holmes, Edward C; Barrett, Alan. Out of Africa: A Molecular Perspective on the Introduction of Yellow Fever Virus into the Americas. *PLoS Pathog*, v. 3, p. e75, May, 2007.
02. Koonin, Eugene V.; Dolja, Valerian V.; Krupovic, Mart. The logic of virus evolution. *Cell Host & Microbe*, v.30, p. 917-929, July, 2022.
03. Hanley, Kathryn A.; Monath, Thomas P.; Weaver, Scott C.; Rossi, Shannan L.; Richman, Rebecca L.; Vasilakis, Nikos. Fever versus fever: The role of host and vector susceptibility and interspecific competition in shaping the current and future distributions of the sylvatic cycles of dengue virus and yellow fever virus. Publisher: Elsevier BV, Oxford, v. 19, p. 292-311, October 2013.
04. Figueiredo, Luiz Tadeu Moraes. Human Urban Arboviruses Can Infect Wild Animals and Jump to Sylvatic Maintenance Cycles in South America. *Frontiers in Cellular and Infection Microbiology*, v. 9,p. 259, July, 2017.
05. Mares-Guia, Maria Angélica Monteiro De Mello et al. Yellow fever epizootics in non-human primates, Southeast and Northeast Brazil (2017 and 2018). *Parasites & Vectors*, v.13, 90 p., December, 2020.
06. Cunha, Marielton Dos Passos et al. Phylogeographic patterns of the yellow fever virus around the metropolitan region of São Paulo, Brazil, 2016–2019. *PLoS Negl Trop Dis*, v.16, p. e0010705, September, 2022.
07. Maramorosch, Karl; Shatkin, J. Aaron. Advances in Virus Research. In: Gritsun, T.S.; Gould, E.A. Origin and Evolution of 3'Utr of Flaviviruses: Long Direct Repeats as A Basis for the Formation of Secondary Structures and Their Significance for Virus Transmission, 2006, v. 69, p. 203-248.
08. Mutebi, John-Paul; Barrett, Alan D.T. The epidemiology of yellow fever in Africa. *Microbes and Infection*, v.4, p. 1459-1468, November, 2002.
09. Galler, Ricardo.; Chambers, Thomas J.; Hahn, Chang S.; Rice, Charles M. FLAVIVIRUS GENOME ORGANIZATION, EXPRESSION, AND REPLICATION. *Annual Review of Microbiology, California*, v. 44, p. 649-688, 1990.
10. Ellis, Brett R.; Barrett, Alan D. T. The enigma of yellow fever in East Africa. *Reviews in Medical Virology*, v.18, p. 331-346, September, 2008.
11. Moureau, Gregory et al. New Insights into Flavivirus Evolution, Taxonomy and Biogeographic History, Extended by Analysis of Canonical and Alternative Coding Sequences. *PLoS ONE*, v. 10, p. e0117849, February, 2015.
12. Macdonald, Joanne, et al. Molecular Phylogeny of Edge Hill Virus Supports its Position in the Yellow Fever Virus Group and Identifies a New Genetic Variant. *Evol Bioinform Online*, v. 6, p. 91-96, January, 2010.
13. Mahy; Brian W. J. (org.). Desk encyclopedia of human and medical virology. Oxford: Elsevier Academic Press, 2010.
14. Wang, Eryu; Weaver, Scott C.; Shope, Robert E.; Tesh, Robert B.; Watts, Douglas M.; Barrett, Alan D.T. Genetic Variation in Yellow Fever Virus: Duplication in the 3' Noncoding Region of Strains from Africa. *Virology*, v. 225, p. 274-281, November, 1996.
15. Fulop, L. D.; Barrett, A. D. T.; Titball, R. W. Nucleotide sequence of the NS5 gene of Banzi virus: comparison with other flaviviruses. *Journal of General Virology*, v. 76, p. 2317-2321, September, 1995.
16. Hansen, Clairissa A.; Barrett, Alan D. T. The Present and Future of Yellow Fever Vaccines. *Pharmaceuticals*, v.14, 891 p., September, 2021.

17. Scott, Kitchener. Viscerotropic and neurotropic disease following vaccination with the 17D yellow fever vaccine, ARILVAX®. Vaccine, v.22, p. 2103-2105, June, 2004.