



Fundação Universidade Federal do ABC

Pró reitoria de pesquisa

Av. dos Estados, 5001, Santa Terezinha, Santo André/SP,
CEP 09210-580

Bloco L, 3ºAndar, Fone (11) 3356-7617

iniciacao@ufabc.edu.br

Projeto de Iniciação Científica
submetido para avaliação no
Edital: 04/2022

Título do projeto: Análise de seleção positiva em hemaglutinina (HA) de Influenza A e B, hemaglutinina-esterase de fusão (HEF) de Influenza C e D e hemaglutinina-esterase (HE) em β -coronavírus e torovírus.

Palavras-chave do projeto: Hemaglutinina-esterase (HE), hemaglutinina-

esterase de função(HEF), hemaglutinina (HA), evolução molecular, seleção positiva, Influenza, β -coronavírus e torovírus.

Área do conhecimento do projeto:

Sumário

1 Resumo	2
2 Introdução e Justificativa	2
2.1 Ácidos siálicos	3
2.2 Hemaglutinina (HA)	4
2.3 Hemaglutinina-esterase de fusão (HEF)	5
2.4 Hemaglutinina (HE)	5
2.5 Vírus Influenza	7
2.6 <i>Betacoronavirus</i>	9
2.7 <i>Torovirus</i> (ToV)	10
2.7 Justificativa	10
3 Objetivos	11
3.1 Objetivo geral	11
3.2 Objetivos específicos	11
4 Metodologia	11
5 Cronograma de atividades	13
Referências	13

1 Resumo

Vírus constantemente sofrem mutações e até recombinação genômica por transmissão horizontal, como é possível verificar no caso das glicoproteínas hemaglutinina (em Influenza A e B), hemaglutinina-esterase de fusão (em influenza C e D) e hemaglutinina-esterase (em *Betacoronavirus* e *Torovirus*), homólogas entre si. Por serem proteínas de membrana do vírus e que atuam se ligando aos receptores do hospedeiro, possuem regiões que constantemente podem sofrer com a ação da seleção positiva, o que permite aumentar a afinidade do vírus com o hospedeiro, ou permitir afinidade do vírus com novos hospedeiros e até um escape do sistema imunológico. O objetivo do projeto é entender essas mutações ao longo da história evolutiva do vírus, através de análise filogenética, e correlacionar com informações biológicas importantes para verificar um possível potencial pandêmico destes.

2 Introdução e Justificativa

Doenças virais estão continuamente emergindo, causando epidemias e até pandemias pelo mundo. Em vírus envelopados, a interação entre os receptores das células do hospedeiro e das glicoproteínas presentes na

membrana do vírion é a primeira etapa para iniciar alterações conformacionais necessárias para a fusão entre as membranas virais e celulares e, então, permitir a entrada do material genético do vírus para futura replicação (Mahy, 2010).

A Hemaglutinina (HA) é uma glicoproteína de fusão encontrada na superfície do envelope do gênero *Alphainfluenzavirus* e *Betainfluenzavirus* (ICTV, 2011 [1]), sendo esta homóloga à hemaglutinina-esterase de fusão (HEF) (Zeng, 2008) de *Gammainfluenzavirus* e *Deltainfluenzavirus* (Viral zone, 2022) e à hemaglutinina-esterase (HE) encontrada em 4 espécies de *Betacoronavirus* (BCoV) e no gênero *Torovirus* (Viral zone, 2022). A visão geral do envelope viral e suas respectivas glicoproteínas são mostradas na figura 1.

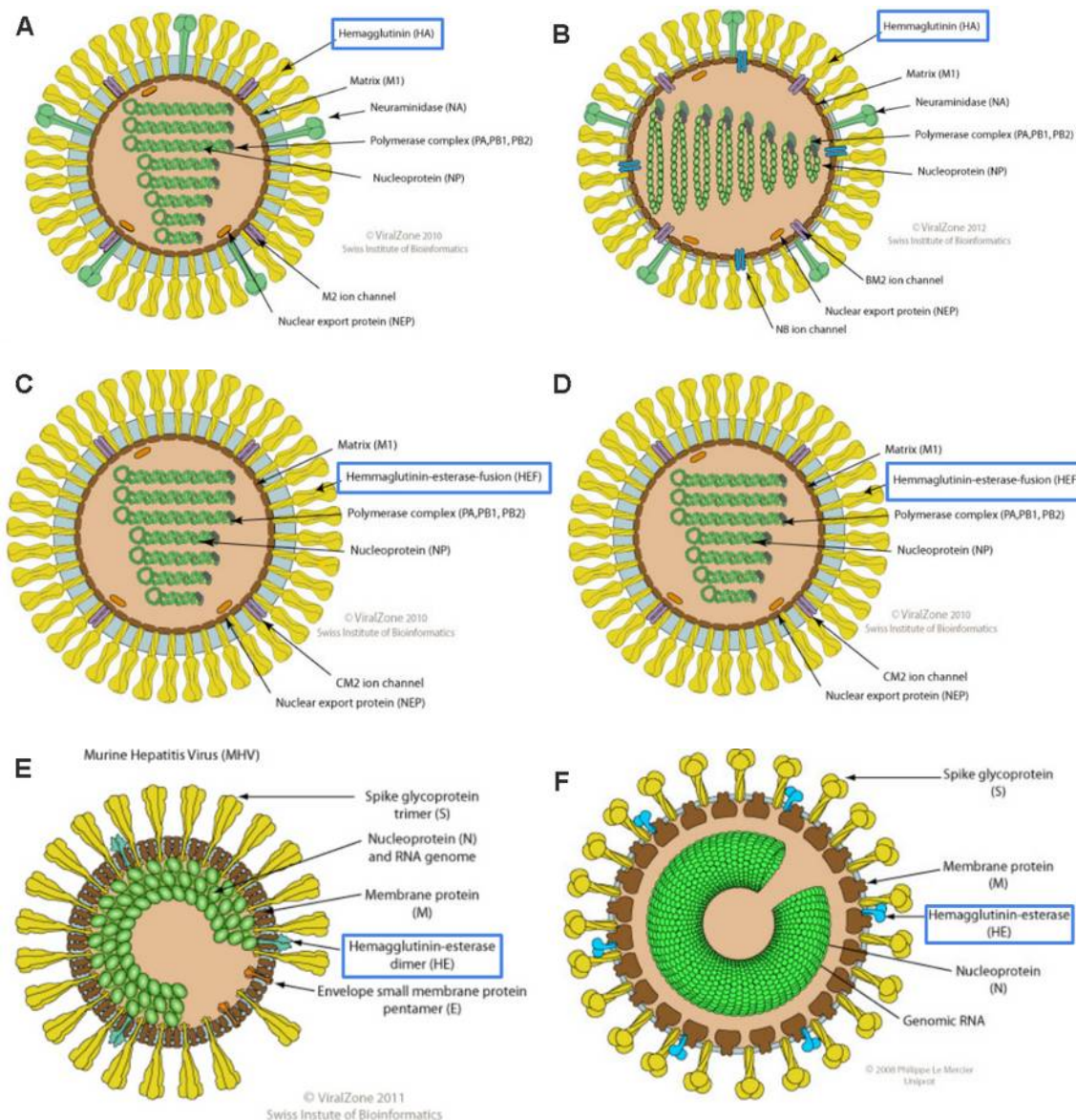


Figura 1. Destaque para HA presente no *Alphainfluenzavirus* (A) e *Betainfluenzavirus* (B). Destaque para HEF de *Gammainfluenzavirus* (C) e *Deltainfluenzavirus* (D). Destaque para HE do vírus MHV, um *Betacoronavirus* (E) e *Torovirus* (F) (Adaptada do Viral Zone, 2022).

2.1 Ácidos siálicos

Os ácidos siálicos (SIA) são uma variedade de monossacarídeos de nove carbonos, derivados do ácido 5-N-acetil neuramínico (Zeng, 2008) e que são carregados negativamente (Mahy, 2010). É a via de ligação de muitos vírus entéricos e respiratórios, que possuem uma variedade de receptores para se ligar à variedade de SIA, porém, também pode ser encontrado em várias moléculas no ambiente extracelular, como também em células não-alvo dos hospedeiros (Zeng, 2008) e no próprio vírus (deGroot, 2006).

2.2 Hemaglutinina (HA)

Sendo uma glicoproteína de fusão trimérica, a HA apresenta uma cabeça globular e uma haste, como mostrado na figura 2 (Mahy, 2010). Esta proteína se liga aos receptores celulares que contém ácido siálico em seus terminais. A eficiência dessa proteína em realizar esse processo está relacionada com a patogenicidade do vírus (Mahy, 2010). Os hospedeiros possuem uma resposta imunológica para essa proteína estrutural, porém, devido à deriva antigênica, o vírus sofre mutações rapidamente e consegue escapar da imunidade.

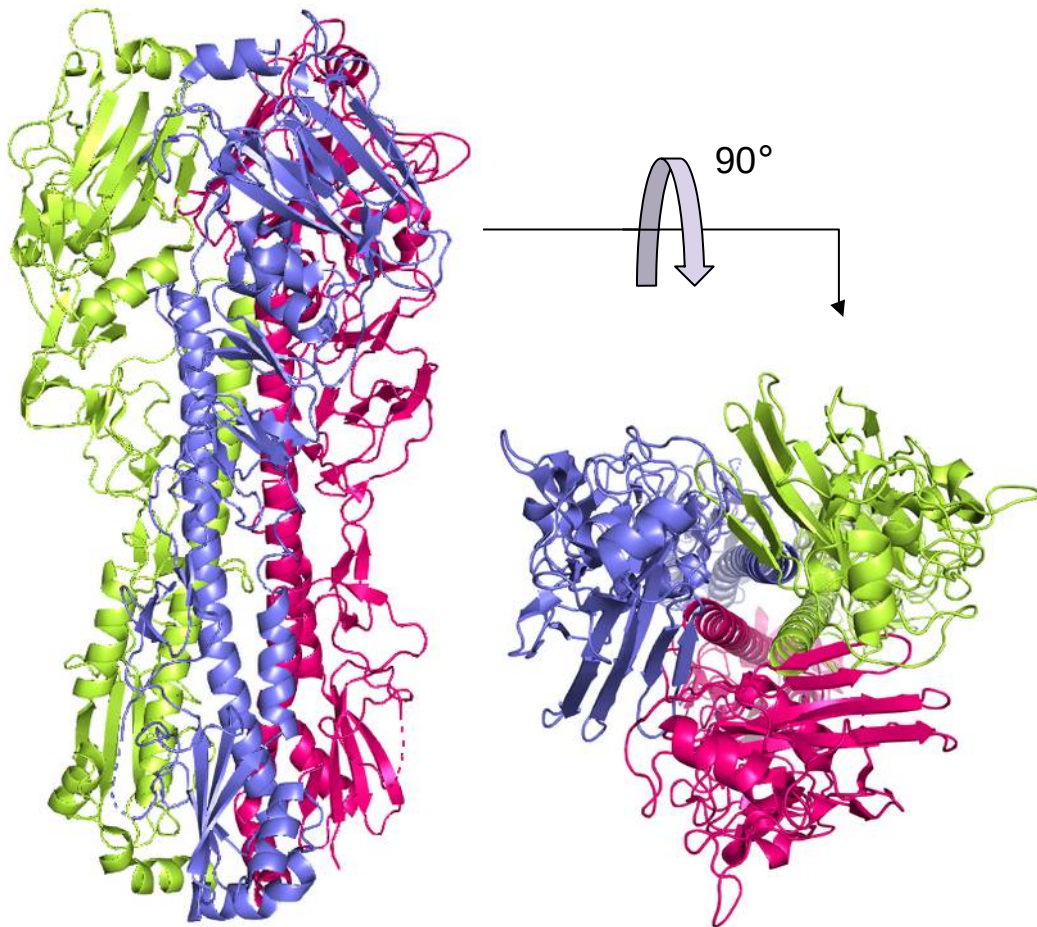


Figura 2. Estrutura de trímero de HA do vírus Influenza A. Vista lateral, à esquerda, e a 90°

para frente, à direita, mostrando a região dos sítios de ligação com o receptor (Mahy, 2010). Imagem obtida através do PDB 7VDF (Fan, 2021)

A HA também medeia a fusão entre as membranas, pois, após a endocitose do vírus, o pH baixo dentro do ambiente celular causa mudança conformacional por alterar a protonação dos resíduos do peptídeo de fusão, localizado na haste da glicoproteína HA, o que facilita a fusão da membrana. Existem 16 tipos de HA (Mahy, 2010).

2.3 Hemaglutinina-esterase de fusão (HEF)

Assim como a HA, possui atividade de ligação com ácido-9-O-acetil siálico, mediando a fusão entre as membranas. Além disso, a HEF (figura 3) também quebra a ligação do ácido siálico com o vírus tal quanto das proteínas celulares para que este seja facilmente liberado, prosseguindo no ciclo viral, função que em influenzas A e B é realizada pela proteína Neuraminidase (NA) ausentes em Influenza C (*Gammainfluenzavirus*) e D (*Deltainfluenzavirus*).

Apesar de HEF se ligar com o ácido-9-O-acetil siálico e não com os mesmos SIAs que Influenza A e B, seu domínio de ligação é semelhante ao HA, assim como a região da haste, diferenciando-se apenas no feixe α -helicoidal na extremidade e o peptídeo de fusão fica exposto parcialmente no solvente.

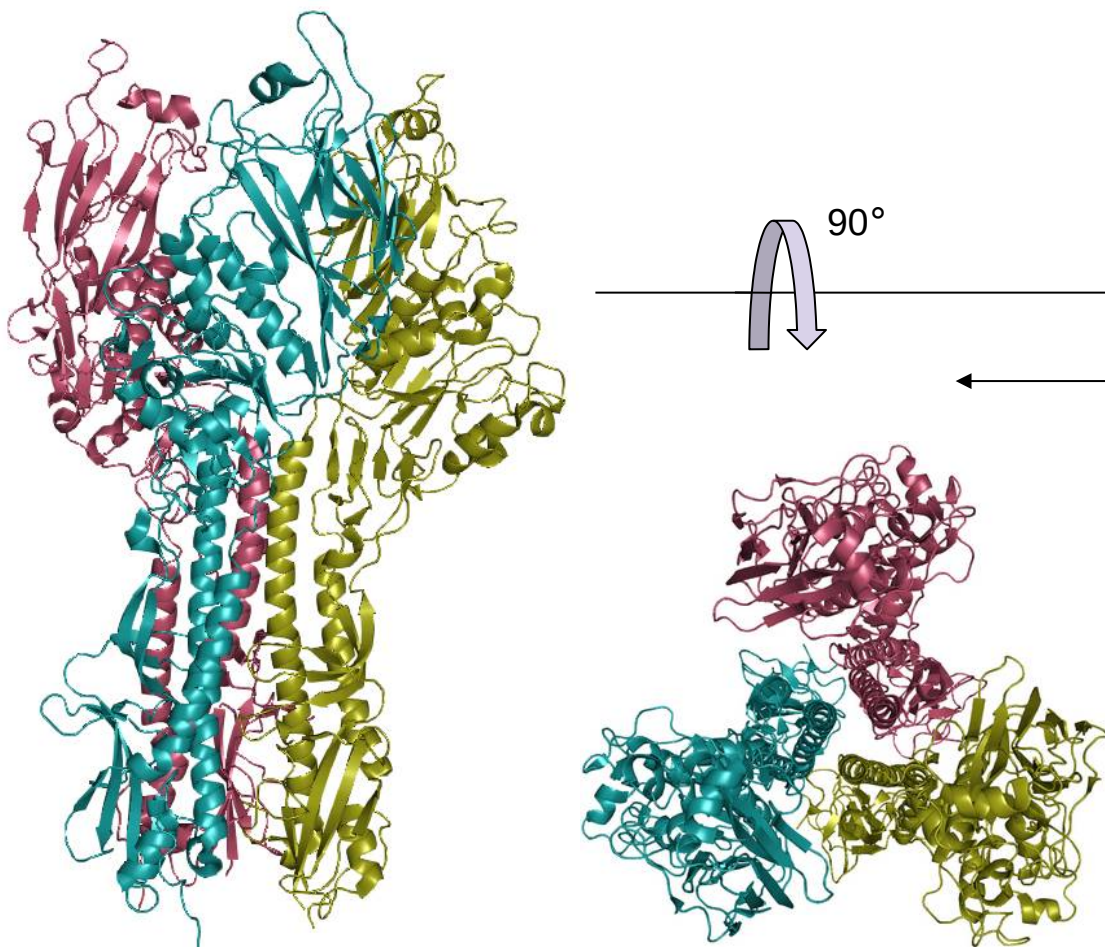


Figura 3. Estrutura de trímero de HEF do vírus Influenza C. Vista lateral, à esquerda, e a 90° para frente, à direita, mostrando a região dos sítios de ligação com o receptor (Mahy, 2010). Imagem obtida através do PDB 6YI5 (Halldorsson, 2021)

2.4 Hemagglutinina (HE)

Presente no gênero *Torovirus* e em quatro *Betacoronavirus* (HCoV-HKU1, HCoV-OC43, CoV-Parker e o CoV-HKU24) possivelmente como resultado de transferência horizontal, as glicoproteína HE (figura 4) são dímeros que tem a ligação estabilizada pelo que era a haste da HEF, região do domínio de fusão em influenza C e D. HEs agem como lectinas e também como uma enzima destruidora de receptores (EDR) pela presença do sialato 9(4)-O-Acetil esterase (Zeng, 2010).

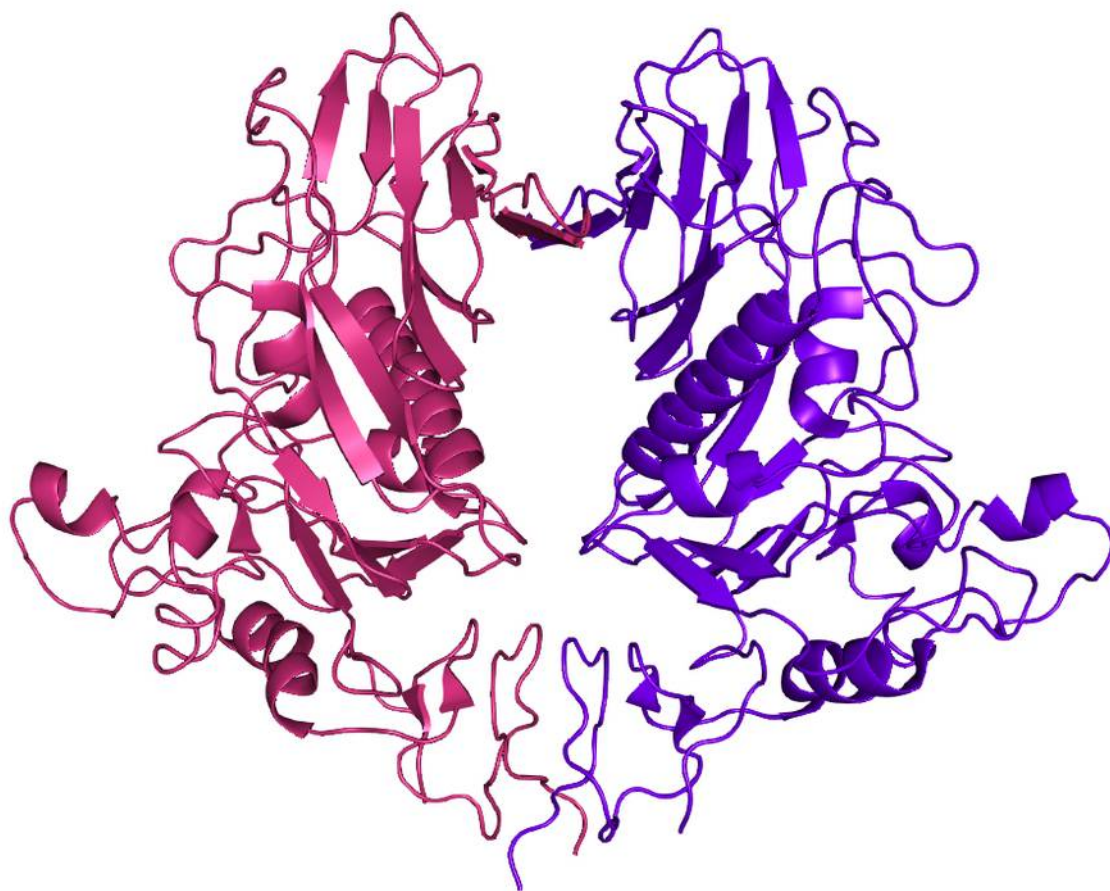


Figura 4. Estrutura de dímero HE do vírus HKU1. Imagem obtida através do PDB 6Y3Y (Hurdiss, 2020)

As HEs presentes em BCoV são acessórias à proteína Spike (S), que é responsável pela ligação de fusão com o hospedeiro. Cada monômero é composto por 3 módulos: domínio de ligante de receptor (R), um domínio de acetilesterase (E) e um domínio de membrana proximal (MP), todos eles equivalentes a uma subunidade de HEF, conforme figura 5.

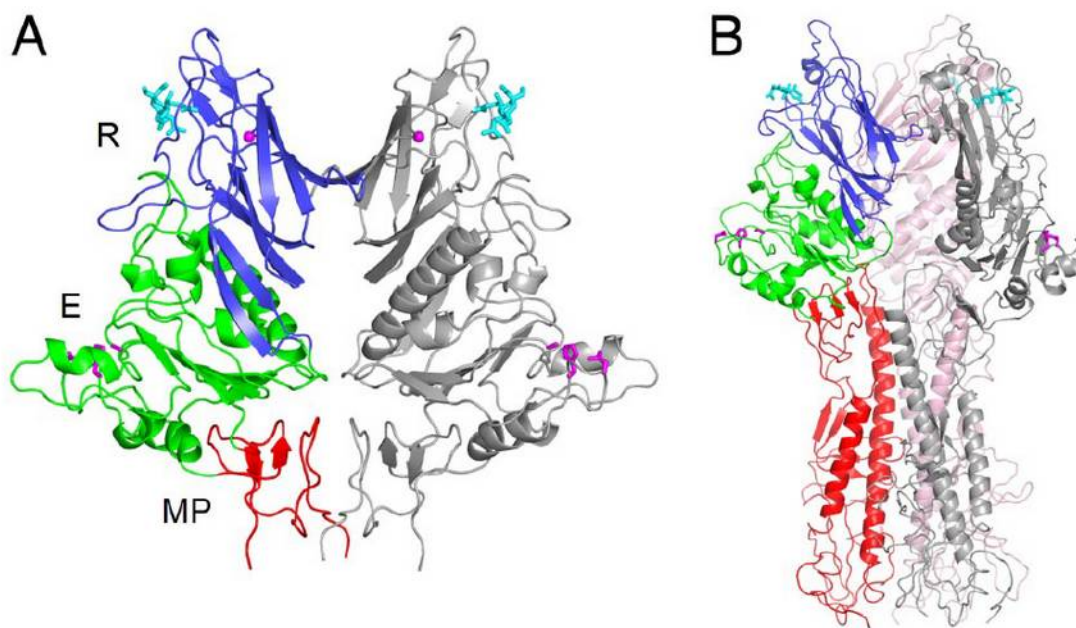


Figura 5. Estrutura de HE (A) comparada com a estrutura de HEF (B) (Adaptada de Zeng, 2010).

2.5 Vírus Influenza

Os vírus da gripe, classificados como Influenza A, B, C e D, são de RNA fita simples e negativa (ssRNA(-)), apresentando um genoma segmentado, sendo que A e B possuem 8 segmentos, e os C e D possuem 7 segmentos, como é possível observar na figura 6 . Cada segmento é um complexo de ribonucleoproteína, pois o RNA é revestido com proteína de nucleocapsídeo, mantendo-se em formato helicoidal (Mahy, 2010).

São vírus envelopados (derivado da membrana celular) (ICTV, 2011 [2]) e esféricos (alguns sendo filamentosos), pertencentes à família Orthomyxoviridae e foram responsáveis por pandemias em 1918, 1957 e 1968, além da de 2009 que surgiu de uma combinação entre o H3N2 e influenza B.

Possui sequências conservadas em todos os segmentos: AGUAGAAACAAGG (extremidade 5') e UCG(U/C)UUUCGUCC (extremidade 3') que funcionam como promotor viral para a replicação e transcrição (Mahy, 2010). As principais variações encontradas no segmento linear do genoma dos gêneros de influenza é o segmento 4 do Influenza A e B codifica HA e o NA é codificado no segmento 6, enquanto no Influenza C e D há codificação de HEF e não há codificação de NA (figura X).

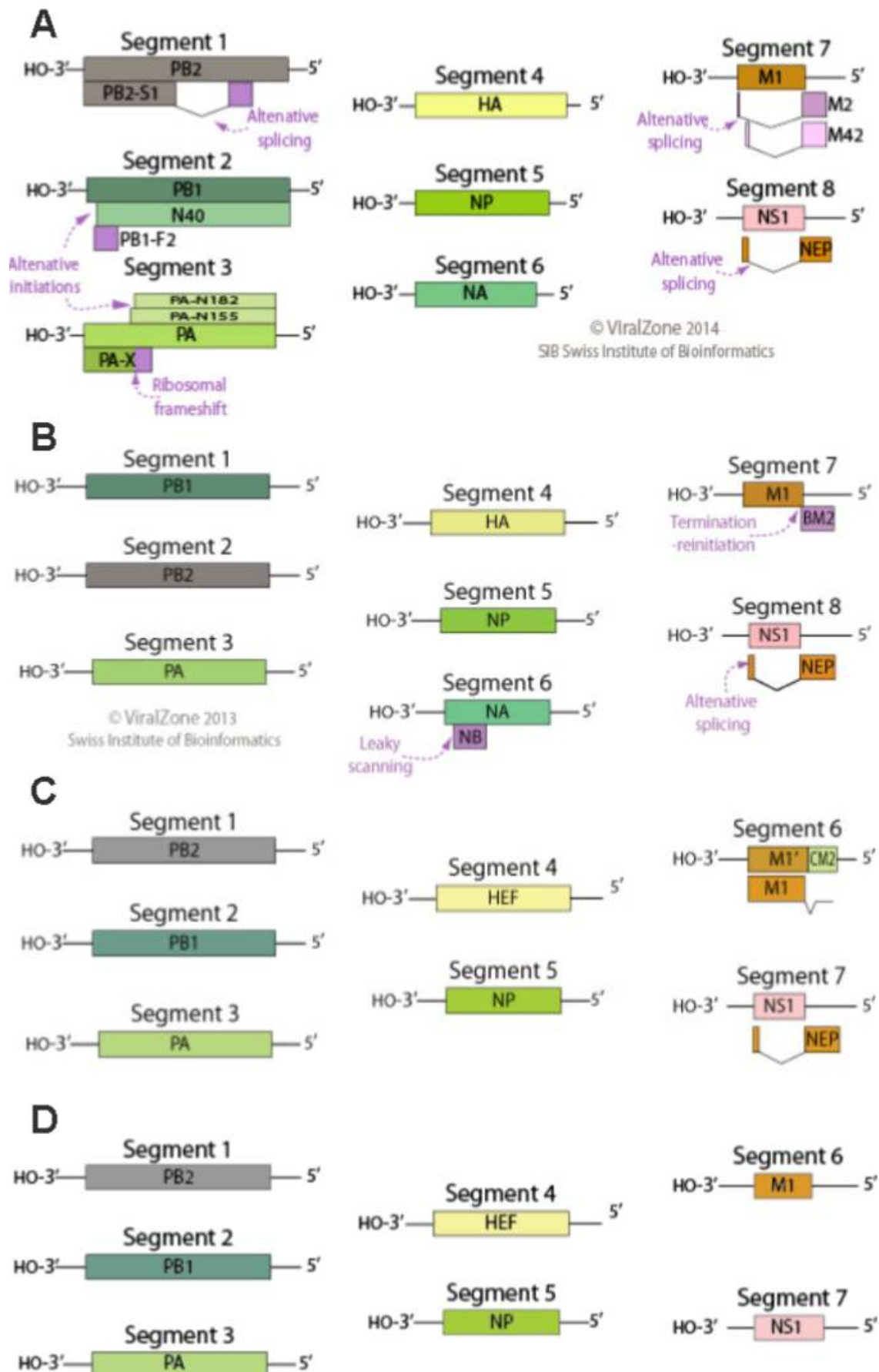


Figura 6. Segmentos de genomas dos vírus Influenza. (A) segmento de *Alphainfluenza*, que codifica de 12-14 proteínas dependendo do segmento, que variam de 0.89 à 2.341 kb, sendo que o tamanho total do genoma é de 13.5 kb. (B) segmento de *Betainfluenza*, que codifica de 11 proteínas e possui segmentos que variam de 0.89 à 2.341 kb, sendo que o tamanho total do genoma é de 13.5 kb. (C) segmento de *Gammainfluenza*, que codifica 9 proteínas e possui segmentos que variam de 0.9 à 2.3 kb, sendo que o tamanho total do genoma é de 10 kb. (D) segmento de *Deltainfluenza*, que codifica 9 proteínas e possui segmentos que variam de 0.76 à 2.32 kb, sendo que o tamanho total do genoma é de 10 kb. (Viral Zone, 2022 [1], [2], [3], [4])

2.6 Betacoronavirus

O gênero *Betacoronavirus* é um dos quatro gêneros de coronavírus: alfa, beta, gama e deltacoronavírus, que foram classificados de acordo com as duas regiões *open reading frames* (ORF1a e ORF1b) conservadas. Todos eles são vírus envelopados de RNA fita simples e positiva (ssRNA(+)). Pertencentes à subfamília Orthocoronavirinae, eles estão envolvidos com diversas doenças que emergiram e que tem potencial de emergir entre humanos (Shang, 2020) e animais (Zeng, 2008).

A sua capacidade de infecção está principalmente pela habilidade de fazer revisões genômicas, minimizando as chances de diminuir seu tropismo, além de recombinar-se facilmente com outros vírus, o que pode levar a capacidade de infectar nova células, facilitando a contaminação entre espécies (Cyranoski, 2020), conforme ilustrado na figura 7.

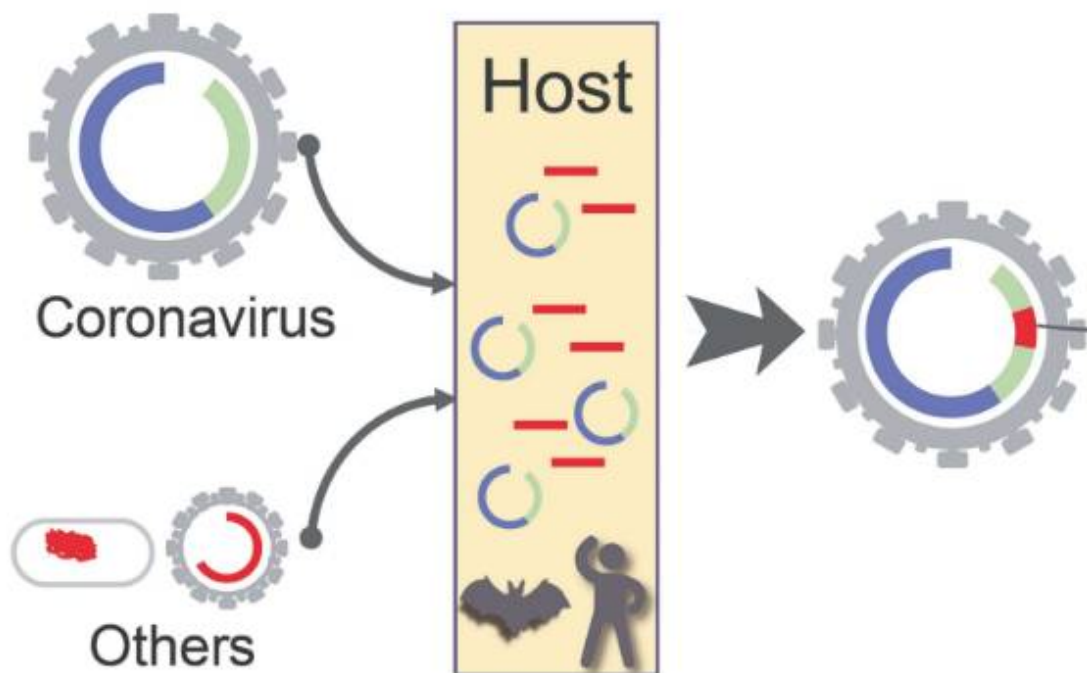


Figura 7. Representação da recombinação do material genético do coronavírus com o de outros organismos (adaptado de Shang, 2020)

Devido às características citadas e também graves doenças em

2002/2003 pelo SARS-CoV, em 2012 pelo MERS-CoC e em 2020 pelo SARS-CoV-2 (Cyranoski, 2020), CoVs passaram a ser visto como vírus com grande potencial pandêmico . Antes destas datas, HCoV-OC43 e HCoV-HKU1 foram vírus que entraram na população humana a partir de diferentes zoonoses e permanecem até hoje (Bakkers, 2017). Estes últimos são conhecidos por causar resfriados comuns, mas de acordo com estudos recentes, o OC43, que passou de vacas para humanos, causou uma pandemia e levou mais de um milhão de pessoas à morte entre 1889 e 1890 (Cyranoski, 2020).

A organização do genoma dos CoVs com HE está na figura 8.

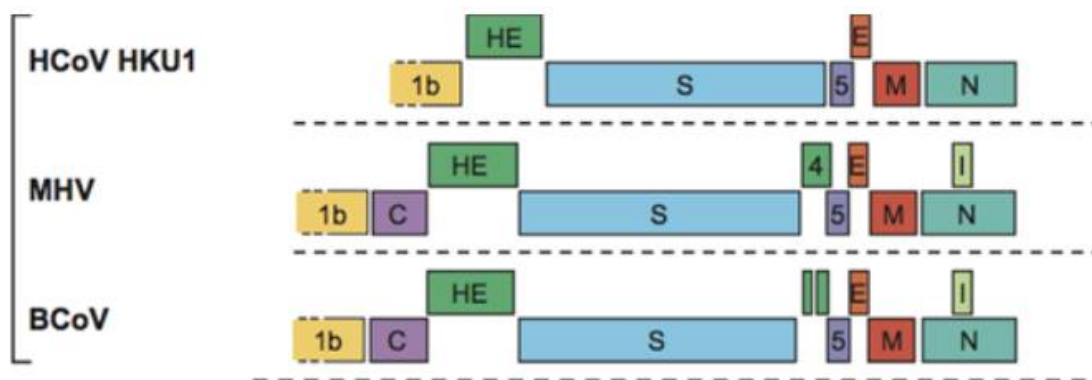


Figura 8. Organização do genoma de CoVs com HE (adaptado de ICTV, 2011 [3])

2.7 Torovirus (ToV)

Os ToV pertencem a ordem Nidovirales, assim como os Covs. Também são vírus ss-RNA(+) e polimorfos. ToVs foram encontrados em equinos (Berne virus), bovinos e porcos. Em 1984, foi encontrado nas fezes de humanos com gastroenterite, mas ainda requer evidências conclusivas (Ujike, 2021). A organização do genoma está na figura 9.

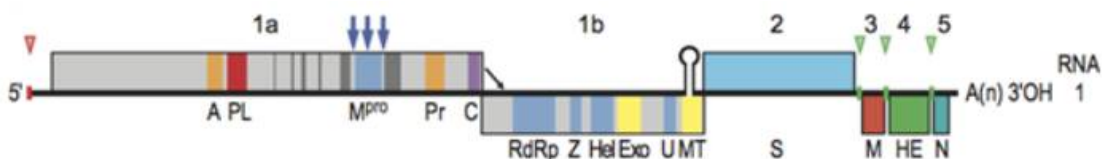


Figura 9. Organização do genoma de *Torovirus* (adaptado de ICTV, 2011 [3])

2.7 Justificativa

As glicoproteínas de superfície em vírus, como a HA, HEF e HE, sofrem mutações como qualquer outra parte do genoma viral grande parte dessas mutações pode não ter nenhum efeito no fitness viral. Entretanto justamente por serem proteínas de superfície e estarem na superfície do

virions elas um tem potencial gigante para influenciar no fitness viral. Isso pode se dever principalmente a 3 possibilidades: 1) por estarem mas expostas estão suscetíveis a interação com proteínas do sistema imune dessa forma certas mutações poderiam ajudar na evasão do sistema imune; 2) essas proteínas por estarem na superfície tem o papel de ligar com receptores na superfície da célula hospedeiro, dessa forma mutações nessas proteínas poderiam ter capacidade de aumentar afinidade com receptores qual a qual já interagem ou poderiam adquirir a capacidade de se ligar a novos tipos de receptores; 3) proteínas na superfície não tem só de ligar com a ligação a receptores mas também fazer a etapa seguinte de fusão de membranas e determinadas mutações podem melhorar essa capacidade.

Dessa forma identificar mutações de seleção positiva eliminaria o ruído das mutações neutras e permite nos concentrar nas mutações que tem chance de ter um papel no aumento no fitness viral. Estudar esse processo evolutivo nessas proteínas pode nos ajudar a identificar padrões evolutivos e regiões chave na evolução dessas proteínas isso pode ajudar a prever mutações novas com potencial de risco pandêmico em novos vírus de Influenza e *Coronavirus*.

3 Objetivos

3.1 Objetivo geral

Entender a evolução das HE nos gêneros *Torovirus* e *Betacoronavirus* e relacionar com a evolução da HEF e HA, dos gêneros de Influenza da família Orthomyxoviridae, após fazer análise filogenéticas (a partir da sequência das proteínas) e correlacioná-la com dados estruturais e informações biológicas importantes.

3.2 Objetivos específicos

Coletar sequências das proteínas HE, HEF e HA de membro dos gêneros *Torovirus*, *Betacoronavirus* e Influenza da família Orthomyxoviridae para análise e construção da filogenia molecular com o objetivo de separar os membros em grupos monofiléticos, seguido de avaliação de seleção positiva e coevolução. As informações obtidas serão relacionadas com as análises a nível estrutural, com foco na interface com receptores de hospedeiros, para que então seja feita correlação com as informações biológicas disponíveis em banco de dados (RCSB pdb, NCBI e uniprot) que indiquem melhorias no tropismo viral.

4 Metodologia

1. Revisão bibliográfica

Pesquisa e leitura de fontes bibliográficas para fazer uma revisão constante dos assuntos envolvidos no projeto.

2. Seleção de sequências e estruturas

Busca, em bancos de dados como o NCBI e Uniprot, por sequências de nucleotídeos e aminoácidos da glicoproteína hemaglutinina em membros de Influenza A e B, hemaglutinina-esterase de fusão em membros de Influenza C e D e hemaglutinina-esterase em membros de Coronavírus e Torovírus. Além de busca e coleta de suas estruturas no RCSB pdb.

É esperado que, ao final desta etapa, a aluna consiga organizar e estruturar os dados de forma eficaz.

3. Alinhamento de sequências

As sequências de nucleotídeos obtidas da etapa anterior serão alinhadas usando o servidor revtrans (<http://www.cbs.dtu.dk/services/RevTrans-2.0/web/>) e, para alinhamento de aminoácidos, será utilizado o software MUSCLE (Edgar, 2004) e também o T-Coffee (Notredame, 2000).

Após, será utilizado o software JmodelTest (Posada,2008) para criar uma matriz de probabilidade que irá calcular o modelo de substituição de códons, necessário para a construção da árvores filogenéticos (item 5 da metodologia).

Ao final desta etapa, além de aprender a trabalhar com dados de aminoácidos e nucleotídeos, espera-se que a aluna se familiarize com regiões importantes e conservadas de cada hemaglutinina e para identificação destas na estrutura 3D.

4. Alinhamento estrutural e construção de modelos

A partir da comparação dos pdbs obtidos, serão calculados os valores de RMSD obtidos pelo software MUSTANG (Konagurthu et al., 2006) com o objetivo de verificar a conservação estrutural das proteínas e entender a sua diversidade.

Caso algum pdb não seja encontrado, será possível a criação de modelos teóricos da sua estrutura através do software modeller (Webb, 2016).

5. Construção das árvores filogenéticas

Com base no alinhamento de códons, será feita a filogenia molecular através de modelos probabilísticos obtidos por Máxima Verossimilhança (ML) com o programa PhyML (Guindon et al., 2010) e por inferência Bayesiana com o programa BEAST (Drummond et. al., 2012). A visualização dos cladogramas obtidos será através do software FigTree.

6. Análise de seleção positiva e coevolução

A análise de seleção positiva será feita pelo software mrBayes (Ronquist, 2003) e sítios com alta probabilidade serão identificados nas estruturas analisadas no item 4 da metodologia.

7. Correlação dos dados seleção positiva/estruturais com biológicos

8. Avaliação dos resultados e redação

5 Cronograma de atividades

1. Revisão Bibliográfica
2. Seleção de sequências e estruturas
3. Alinhamento de sequências
4. Alinhamento estrutural e construção de modelos
5. Construção das árvores filogenéticas
6. Análise de seleção positiva e coevolução
7. Correlação dos dados seleção positiva/estruturais com biológicos
8. Avaliação dos resultados e redação

Tabela 1 – Exemplo de cronograma de atividades previstas

[illegible]

Referências

Bakkers, Mark J. G., et al. "Betacoronavirus Adaptation to Humans Involved Progressive Loss of Hemagglutinin-Esterase Lectin Activity". *Cell Host & Microbe*, vol. 21, no 3, março de 2017, p. 356-66. DOI.org (Crossref), <https://doi.org/10.1016/j.chom.2017.02.008>.

CYRANOSKI, David. "Profile of a Killer: The Complex Biology Powering the Coronavirus Pandemic". *Nature*, vol. 581, nº 7806, maio de 2020, p. 22-26. DOI.org (Crossref), <https://doi.org/10.1038/d41586-020-01315-7>.

de GROOT, Raoul J. "Structure, Function and Evolution of the Hemagglutinin-Esterase Proteins of Corona- and Toroviruses". *Glycoconjugate Journal*, vol. 23, no 1-2, fevereiro de 2006, p. 59-72. DOI.org (Crossref), <<https://doi.org/10.1007/s10719-006-5438-8>>.

DRUMMOND, A. J. et al. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, Oxford Academic, v. 29, n. 8, p. 1969-1973, aug 2012. ISSN 07374038. Disponível em: <<http://beast-mcmc.googlecode.com>>.

EDGAR, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, Oxford Academic, v. 32, n. 5, p. 1792-1797, mar 2004. ISSN 03051048. Disponível em: <<http://www.drive5.com/muscle>>.

FAN, Hongcheng, et al. "A Cryo-Electron Microscopy Support Film Formed by 2D Crystals of Hydrophobin HFBI". *Nature Communications*, vol. 12, nº 1, dezembro de 2021, p. 7257. DOI.org (Crossref), <https://doi.org/10.1038/s41467-021-27596-8>.

GUINDON, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, Oxford Academic, v. 59, n. 3, p. 307-321, may 2010. ISSN 10635157. Disponível em: <<http://evolution.genetics.washington.edu/phymlip>>.

HALLDORSSON, Steinar, et al. "In Situ Structure and Organization of the Influenza C Virus Surface Glycoprotein". *Nature Communications*, vol. 12, nº 1, março de 2021, p. 1694. DOI.org (Crossref), <https://doi.org/10.1038/s41467-021-21818-9>.

HURDISS, Daniel L., et al. "Cryo-EM Structure of Coronavirus-HKU1 Haemagglutinin Esterase Reveals Architectural Changes Arising from Prolonged Circulation in Humans". *Nature Communications*, vol. 11, nº 1, setembro de 2020, p. 4646. DOI.org (Crossref), <https://doi.org/10.1038/s41467-020-18440-6>.

ICTV - Orthomyxoviridae - Figures - Negative Sense RNA Viruses - Negative Sense RNA Viruses (2011) - ICTV. Disponível em:

<https://talk.ictvonline.org/ictv-reports/ictv_9th_report/negative-sense-rna-viruses-2011/w/negrna_viruses/210/orthomyxoviridae-figures>. Acessado 4 de junho de 2022. [1]

ICTV - Orthomyxoviridae - Negative Sense RNA Viruses - Negative Sense RNA Viruses (2011). Disponível em: <https://talk.ictvonline.org/ictv-reports/ictv_9th_report/negative-sense-rna-viruses-2011/w/negrna_viruses/209/orthomyxoviridae>. Acessado 3 de junho de 2022.[2]

ICTV - Coronaviridae - Figures - Positive Sense RNA Viruses - Positive Sense RNA Viruses (2011). Disponível em: <https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/223/coronaviridae-figures>. Acessado 5 de junho de 2022.[3]

KONAGURTHU, A. S. et al. MUSTANG: A multiple structural alignment algorithm. *Proteins: Structure, Function and Genetics*, John Wiley Sons, Ltd, v. 64, n. 3, p. 559-574, aug 2006. ISSN 08873585. Disponível em: <<https://onlinelibrary.wiley.com/doi/full/10.1002/prot.20921>>.

MAHY, Brian W. J., e M. H. V. Van Regenmortel, organizadores. *Desk Encyclopedia of General Virology*. Elsevier Academic, 2010.

NOTREDAME, C.; HIGGINS, D. G.; HERINGA, J. T-coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, Academic Press, v. 302, n. 1, p. 205-217, sep 2000. ISSN 00222836.

POSADA, D. jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, Oxford Academic, v. 25, n. 7, p. 1253-1256, jul 2008. ISSN 07374038. Disponível em: <<https://academic.oup.com/mbe/article/25/7/1253/1045159>>.

RONQUIST, F.; HUELSENBECK, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, Oxford Academic, v. 19, n. 12, p. 1572-1574, aug 2003. ISSN 13674803. Disponível em: <<https://academic.oup.com/bioinformatics/article/19/12/1572/257621>>.

SHANG, Jingzhe, et al. "Compositional Diversity and Evolutionary Pattern of Coronavirus Accessory Proteins". *Briefings in Bioinformatics*, vol. 22, nº 2, março de 2021, p. 1267-78. DOI.org (Crossref), <https://doi.org/10.1093/bib/bbaa262>.

TORTORICI, M. Alejandra, et al. "Structural Basis for Human Coronavirus Attachment to Sialic Acid Receptors". *Nature Structural & Molecular Biology*, vol. 26, nº 6, junho de 2019, p. 481-89. DOI.org (Crossref), <https://doi.org/10.1038/s41594-019-0233-y>.

UJIKE, Makoto, e Fumihiro Taguchi. "Recent Progress in Torovirus Molecular Biology". *Viruses*, vol. 13, no 3, março de 2021, p. 435. DOI.org (Crossref), <https://doi.org/10.3390/v13030435>.

VIRAL ZONE - *Alphainfluenzavirus*. Disponível em: <<https://viralzone.expasy.org/6>>. Acessado 3 de junho de 2022.

VIRAL ZONE - *Beta*influenzavirus. Disponível em: <https://viralzone.expasy.org/80?outline=all_by_species>. Acessado 3 de junho de 2022.

VIRAL ZONE - *Gama*influenzavirus. Disponível em: <https://viralzone.expasy.org/81?outline=all_by_species>. Acessado 3 de junho de 2022.

VIRAL ZONE - *Delta*influenzavirus. Disponível em: <https://viralzone.expasy.org/6077?outline=all_by_species>. Acessado 3 de junho de 2022.

VIRAL ZONE - Torovirus. Disponível em: <https://viralzone.expasy.org/127?outline=all_by_species>. Acessado 3 de junho de 2022.

WEBB, Benjamin, e Andrej Sali. "Comparative Protein Structure Modeling Using MODELLER". Current Protocols in Bioinformatics, vol. 54, n° 1, junho de 2016. DOI.org (Crossref), <https://doi.org/10.1002/cpbi.3>.

ZENG, Qinghong, et al. "Structure of Coronavirus Hemagglutinin-Esterase Offers Insight into Corona and Influenza Virus Evolution". Proceedings of the National Academy of Sciences, vol. 105, n° 26, julho de 2008, p. 9065-69. DOI.org (Crossref), <https://doi.org/10.1073/pnas.0800502105>.