**Projeto Metodologia Científica: Identificação de modificações nos genomas e genes entre coronavírus.**

Este projeto deverá ser feito pelos alunos (individualmente ou em duplas) com apoio de ferramentas de Bioinformática (*in silico*) e dos docentes. O objetivo é comparar genomas e genes de vários coronavírus humanos e de animais. Este trabalho permitirá encontrar regiões do genoma e genes específicos que sejam conservadas ou variáveis, na busca de informações sobre a origem do vírus SARSCov-2, responsável pela pandemia COVID-19. Comparações como essa permitirão também discutir o aparecimento de variantes do SARS-Cov2 e que genes apresentam mutações nessas variantes. Discutiremos as tecnologias das vacinas atuais para esse patógeno, e as respostas imunológicas induzidas por elas.

Cada Grupo deverá entregar um relatório científico referente a esse projeto, no qual deve apresentar os principais resultados obtidos, assim como uma discussão de seu significado.

Metodologia a ser usada:

Busca de sequências de genomas e genes no banco de dados do NCBI.

Comparar sequências de nucleotídeos e de proteínas com a ferramenta BLAST Two Sequences.

Plano de trabalho:

1. Comparar 5 genomas inteiros identificando os genomas mais próximos ou distantes entre si, além das regiões do genoma mais próximas ou distantes, em relação ao SARS-Cov2.
2. Dividir os genomas em 10 segmentos de 3.000 bases (o genoma de coronavírus apresenta cerca de 30.000 bases no total) e comparar os genomas, em relação ao SARS-Cov2. Apresentar os resultados em gráficos.
3. Estudar 5 proteínas desses 5 genomas e realizar a comparação ao nível de nucleotídeo e de proteína, em relação ao SARS-Cov2. Especial atenção deve ser dada às proteínas SPIKE e Replicase.
4. Identificar INDELS nos 5 genes analisados.
5. Buscar os variantes e verificar quantas mutações os principais variantes considerados problemáticos (alfa, beta, gama e delta) apresentam no genoma e também nos genes analisados. Responder porque eles são considerados preocupantes.

Discussão:

Várias questões podem ser respondidas com esses dados incluindo a distância evolutiva entre os coronavírus estudados, proximidade e relação entre os coronavírus mais próximos ao SARS-COV2. Quais genes são mais passíveis de mutações e maior?

Discutir também qual o impacto essas informações têm em relação à origem do vírus SARS-Cov2 e nas vacinas atuais.

Referências interessantes a serem lidas (e citadas em seus relatórios):

Lv L, Li G, Chen J, Liang X, Li Y. Comparative Genomic Analyses Reveal a Specific Mutation Pattern Between Human Coronavirus SARS-CoV-2 and Bat-CoV RaTG13. Front Microbiol. 2020 Nov 30;11:584717.

Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nat Med. 2020 Apr;26(4):450-452

***Genomas a serem estudados:***

*Clade 1*

SARS-Cov2- human coronavirus- MN996528.1

Bat coronavirus RaTG13- MN996532.2

Bat coronavirus CoVZC45- MG772933.1

Pangolin Coronavirus MP789- MT121216.1

*Clade 2*

Bat coronavirus LYRa11 - KF569996

Civet coronavirus civet007- AY572034

Human SARS-Cov- Tor2- AY274119

*MERS- related:*

Human MERS- EMC2012- JX869059

Camel coronavirus 363- KJ713298

Bat coronavirus RSA2011- KC869678

*Other beta coronavirus:*

Human coronavirus HKU1- NC\_006577

Mouse coronavirus MHV-A59- NC\_001846

Bovine coronavirus BCoV-ENT - NC\_003045

*Alpha coronavirus:*

Pig coronavirus PEDV- NC\_003436

Human coronavirus 229E - NC\_002645

Bat coronavirus HKU2- NC\_009988

***Genes a serem estudados:***

***Ligados à ORF1ab- [SARS-Cov2]:***

NSP3- papain-like protease YP\_009742610.1

3C-like proteinase YP\_009742612.1

3’-5’exonuclease- YP\_009725309.1

helicase - YP\_009725308.1

RNA-dependent RNA polymerase - YP\_009725307.1

***Outras proteínas a serem estudadas:***

S- Surface glycosylase- Spike protein

E- envelope protein

M- Membrane protein

N- Nucleocapside protein