Encontrando sequencias de proteinas formadas a partir da ORF1a/b

1. procurar o genoma de SARS-Cov2- MN996528.1 no All data base;
2. No site do genoma clique na ORF1a/b (/protein\_id="[QHR63259.1](https://www.ncbi.nlm.nih.gov/protein/1802633809)");
3. Clique em Graphics, onde vão aparecer as proteinas relacionadas a essa ORF (formadas após clivagem da poliproteina com protease).
4. Para encontrar a proteina passe o cursor em cima da proteína. Você verá o nome e também verá informação sobre a proteína, incluindo uma breve descrição da sua função, e tamanho da proteína. Ao clicar no FASTA RECORD você já tem a sequencia da proteína.
5. Para encontrar a sequencia de nucleotídeos: use a sequencia de sua proteína para encontrar os nucleotídeos fazendo um Query com sua proteína contra a sequencia de nucleotídeos do seu genoma inteiro ou com a sequencia do gene da ORF1ab. No alinhamento anote o início e o final da sequencia de nucleotídeos (para RNA polimerase RNA dependente- por exemplo: 13462-16236).

Para achar a sequencia de outro coronavírus refaça a análise para seu outro genoma. Entretanto você pode encontrar uma grande variação de genoma para genoma e anotações diferentes.

Uma outra forma de encontrar as sequencias de proteínas e nucleotídeos em outros genoma é usar a própria sequencia da proteína dos SARS-Cov2:

Use a sequencia da proteína ou a sequencia de nucleotídeo de SARS-COV2 como isca (Query) contra o genoma do outro vírus. Atenção para garantir que a cover do seu query é total, ou não (certamente em algumas proteínas e vírus mais distantes a cobertura do query não será total).