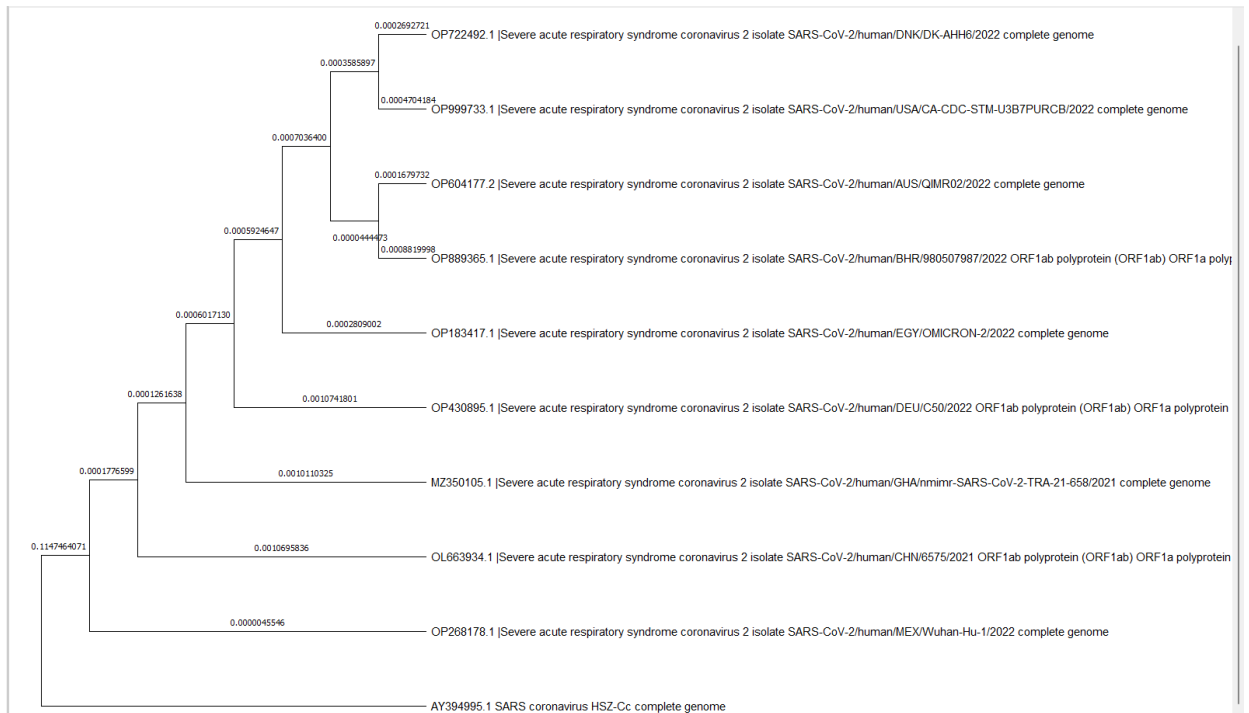


Кондратюк Александр Николаевич

Группа 3.

1. Выравниваем последовательности алгоритмом Muscle.
2. Построим деревья с учетом генетических расстояний различными методами – Maximum Likelihood (ML), UPGMA, Maximum Parsimony.

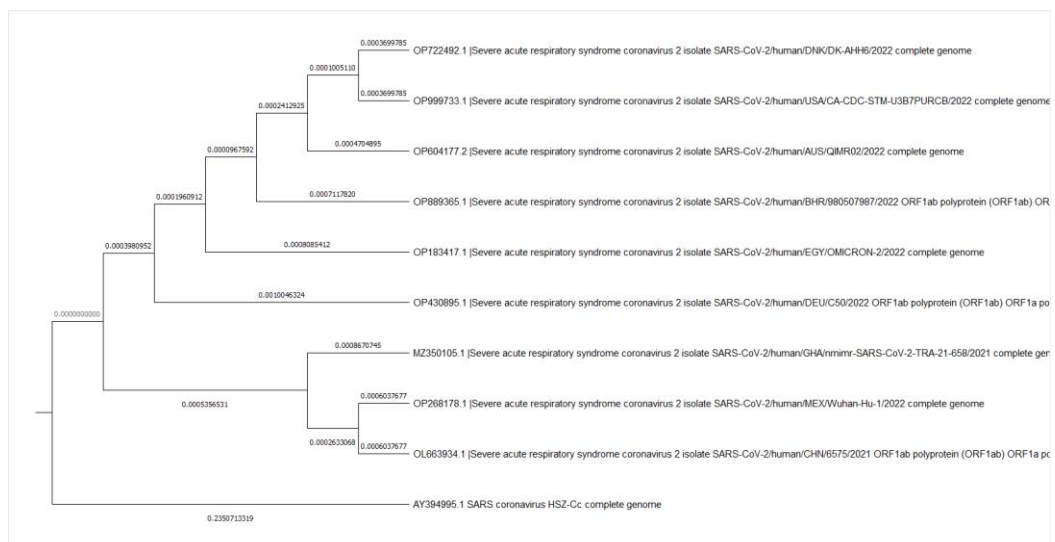
Maximum Likelihood (ML)



Рис_1 - ML

Заметим, что ближайшим к SARS-CoV-1 является геном SARS-CoV-2 из Ганы. Следовательно, из данной подборки он заразился первым. С учетом генетических расстояний, самым дальним от SARS-CoV-1 является геном из США, следовательно из перечисленных геномов он появился позже всех.

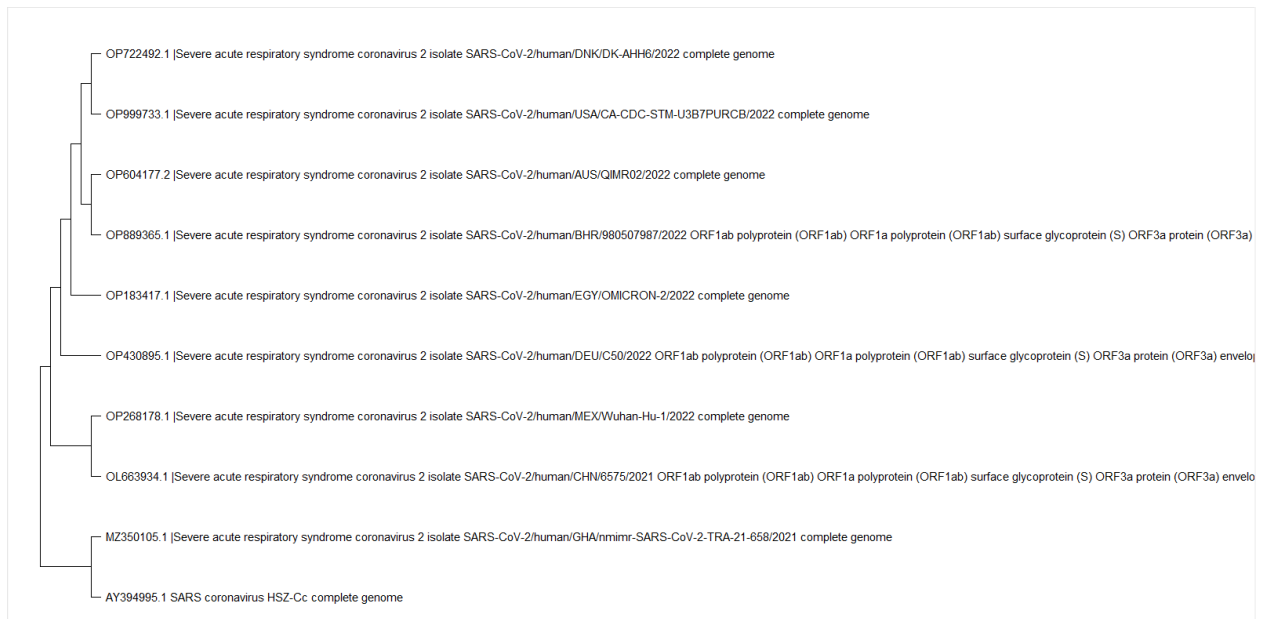
UPGMA.



Рис_2 - UPGMA

Посчитав расстояния, заметим, что ближайшим к SARS-CoV-1 является Ghana, а самым дальним – USA.

Maximum Parsimony



Рис_3 – MP

Заметим, что ближайшим к SARS-CoV-1 является Ghana, а самым дальним – USA.

Посмотрим даты фиксации геномов на сайте:

Germany 14.09.2022

China 30.11.2021

Ghana 06.06.2022

Bahrain 24.11.2022

Egypt 09.08.2022

Australia 11.10.2022

Mexico 19.08.2022

USA 13.12.2022

Denmark 26.10.2022

Как мы можем заметить, топология у всех деревьев получилась разная. Однако все три дерева SARS-CoV-1 сближают сильнее всего с Ghana, из чего можно сделать вывод, что он был заражён первым. Место последней заражённой страны, проанализировав все деревья, это USA.

Запишем 5 любых мутаций для Ghana и USA.

1) Мутация попадает в 44 сайт; так как самые первые гены начинаются с 266 сайта (кодирующие «orf1ab polyprotein» и «orf1a polyprotein»), то мутация попадает в межгенное пространство до этих генов.

DNA Sequences Translated Protein Sequences

Species/Abbrv

1. USA | OP999733.1 | Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/CA-CDC-STM-U3B7PURCB/2022 complete gen

2. GHANA | MZ350105.1 | Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/GHA/nmimr-SARS-CoV-2-TRA-21-658/2021 cor

Site # 44

with w/o gaps

Selected genetic code: Standard

2) Данная мутация попадает в 9565 сайт. Согласно таблице, она попадает в ген, кодирующий «orf1ab polyprotein».

DNA Sequences Translated Protein Sequences

Species/Abbrv

1. USA | OP999733.1 | Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/CA-CDC-STM-U3B7PURCB/2022 complete gen

2. GHANA | MZ350105.1 | Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/GHA/nmimr-SARS-CoV-2-TRA-21-658/2021 cor

Site # 9565

with w/o gaps

Selected genetic code: Standard

3) Данная мутация попадает в 10060 сайт. Согласно таблице, она попадает в ген, кодирующий «orf1ab polyprotein».

DNA Sequences Translated Protein Sequences

Species/Abbrv

1. USA | OP999733.1 | Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/CA-CDC-STM-U3B7PURCB/2022 complete gen

2. GHANA | MZ350105.1 | Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/GHA/nmimr-SARS-CoV-2-TRA-21-658/2021 cor

Site # 10060

with w/o gaps

Selected genetic code: Standard

4) Данная мутация попадает в 23114 сайт. Согласно таблице, данная мутация попадает в ген, кодирующий «surface glycoprotein».

DNA Sequences Translated Protein Sequences

Species/Abbrv

1. USA | OP999733.1 | Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/CA-CDC-STM-U3B7PURCB/2022 complete gen

2. GHANA | MZ350105.1 | Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/GHA/nmimr-SARS-CoV-2-TRA-21-658/2021 cor

Site # 23114

with w/o gaps

Selected genetic code: Standard

5) Данная мутация попадает в 26778 сайт. Согласно таблице, данная мутация попадает в ген, кодирующий «membrane glycoprotein».

DNA Sequences Translated Protein Sequences

Species/Abbrv

1. USA | OP999733.1 | Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/CA-CDC-STM-U3B7PURCB/2022 complete gen

2. GHANA | MZ350105.1 | Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/GHA/nmimr-SARS-CoV-2-TRA-21-658/2021 cor

Site # 26778

with w/o gaps

Selected genetic code: Standard