# Домашнее задание 3

## Егор Портнягин

#### 14.12.2022

#### Цели:

- 1. Проанализировать 9 геномов коронавируса SARS-CoV-2 из людей разных стран и геном коронавируса SARS-CoV-1. Сделать вывод о том, кого вирус заразил раньше, а кого самого последнего.
- 2. Взять 2 генома самого первого после SARS-CoV-1 по дереву и самого последнего и выписать координаты 5 любых мутаций. Посмотреть, куда они попадают в ген (какой?) или межгенное пространство (какое, между какими генами?).

Были скачаны следующие последовательности для SARS-CoV-2:

- OM283601.1 Austria 2021-12-23
- OX317378.1 Iceland 2021-12-28
- OP288474.1 Italy 2020-11-20
- OP963522.1 Kenya 2022-11-21
- OP720005.1 New Zealand 2022-02-17
- OP926910.1 Pakistan 2022-09-28
- ON032860.1 Russia 2021-10-08
- OP777367.1 United Kingdom 2020-04-09
- OP981546.1 USA 2022-11-12

Геном SARS-CoV-1: NC 004718.3

Выравниваем последовательности с помощью MUSCLE: alignment.meg. Строим деревья тремя методами: Maximum Likelihood(ML), Maximum Parsimony(MP), UPGMA:

### ML.mtsx



Ближе всего(в данной выборке заражены раньше) к SARS Russia, дальше всего(в данной выборке заражены позже) Pakistan

#### MP.mtsx



Ближе всего(в данной выборке заражены раньше) к SARS Russia, дальше всего(в данной выборке заражены позже) Iceland, Pakistan, Kenya, New Zealand, Austria

#### **UPGMA.mtsx**



Ближе всего(в данной выборке заражены раньше) к SARS Russia, дальше всего(в данной выборке заражены позже) Pakistan

Сравним **ON032860.1 Russia** и **OP926910.1 Pakistan**, проверим с помощью разметки, куда попадают мутации:

1. Мутация с координатой 20678 попадает на orf1ab polyprotein



| 2. | Мутация с<br><b>Sp * * * *</b>   | координатой                                 | 20690 попада               | ает      | на orf. | 1ab<br>* * | poly  | prote              | ein<br>* * * | * * | * * | * *  | * * | * * | * * | * *   | * * | *   | * *      | * 1 | * *        | *   |
|----|--|---|----------------------------|----------|---------|------------|-------|--------------------|--------------|-----|-----|------|-----|-----|-----|-------|-----|-----|----------|-----|------------|-----|
|    |  | CGTGGC                                      | A A C C G G G              | т        | GTT     | G C        | T A   | T G                | ст           | AA  | тс  | тт   | ΤА  | C A | A   | A A   | T G | c / | ΑА       | Α ( | 3 A        | A   |
|    |  |   | A A C C <mark>A</mark> G G |          |         |            |       |                    |              |     |     |      |     |     |     |       |     |     |          |     |            |     |
|    |  |   |                            |          |         |            |       |                    |              |     | -   |      |     |     |     |       |     |     |          |     |            |     |
|    | Site#  | 20690                                       |                            | *        |         | • \        | with  | 0                  | w/o          | gap | s S | elec | te  | d q | ene | tic ( | :00 | le: | Sta      | ınd | arc        | d   |
| 3. | Мутация с  | координатой                                 | 21617 попада               | ает      | на orf  | 1ab        | poly  | prote              | $_{ m ein}$  |     |     |      |     |     |     |       |     |     |          |     |            |     |
|    | Sp * * * *   | * * * * * *                                 | * * * * * * *              | *        | * * *   | * *        | * * : | * -                | * * *        | * * | * * | * *  | * * | * * | *   |       |     |     |          |     | *          | *   |
|    |  |   | CAGTGTG                    |          |         |            |       |                    |              |     |     |      |     |     |     | _     |     | -   |          |     | - C        | A   |
|    | 2. A G T C   | TCTAGT                                      | CAGTGTG                    | Т        | TAA     | ТС         | T T / | A G                | A C          | CA  | G A | A C  | ТС  | A A | T   | TA    | c c | C ( | : C      | T   | <b>G</b> C | Α   |
|    |  |   |                            |          |         |            |       |                    |              |     |     |      |     |     |     |       |     |     |          |     |            |     |
|    | Site#  | 21617                                       |                            | *        |         | • \        | with  | $\circ$            | w/o          | gap | s S | elec | te  | d q | ene | tic ( | 00  | le: | Sta      | ınd | аго        | d   |
| 4. | Мутация с  | координатой                                 | 19219 попада               | ает      | на orf  | 1ab        | poly  | prote              | ein          |     |     |      |     |     |     |       |     |     |          |     |            |     |
|    | Sp * * * *   | * * * * * *                                 | * * * * * * *              | *        | * * *   | * *        | * * : | * 1                | * * *        | * * | * * | * *  | * * | * * | * * | * *   | * * | * 1 | * *      | * * | * *        | *   |
|    | 1. G A A T   | $T \mathrel{G} C \mathrel{A} A \mathrel{T}$ | GTCGATA                    | G        | ATA     | ТС         | C T ( | G C                | ГАА          | ТТ  | СС  | ΑТ   | T G | Τī  | Т   | G T   | A G | Α.  | ГΤ       | Τ ( | 3 A        | C / |
|    | 2. G A A T   | TGCAAT                                      | GTCGATA                    | G        | ATA     | ТС         | C T ( | G T                | АА           | ТТ  | СС  | ΑТ   | T G | ТТ  | Т   | G T   | A G | Α   | ГΤ       | Τ ( | 3 A        | c / |
|    |  |   |                            |          |         |            |       |                    |              |     |     |      |     |     |     |       |     |     |          |     |            |     |
|    | Site#  | 19219                                       |                            | <b>+</b> |         | • \        | with  | 0                  | w/o          | gap | s S | elec | te  | d q | ene | tic ( | :00 | le: | Sta      | ınd | аго        | d   |
| 5  | Мутапия с  | коорлинатой                                 | 18162 попада               | ает      | на orf  | 1ab        | poly  | prote              | ein          |     |     |      |     |     |     |       |     |     |          |     |            |     |
| •  | Sp * * * *   | * * * * * *                                 | * * * * * * *              | * *      | * * *   | * *        | * * : | * 1                | * * *        | * * | * * | * *  | * * | * * | * * | * *   | * * | *   | * *      | * 1 | * *        | *   |
|    | 1. T C A A   | AACTGA                                      | AGGTTTA                    | Т        | G T G   | тт         | G A ( | c <mark>G</mark> 1 | A C          | СТ  | G G | C A  | ΤА  | c c | т.  | АА    | G G | А   | C A      | Т ( | 3 A        | C ( |
|    | 2. T C A A   | AACTGA                                      | AGGTTTA                    | Т        | GTG     | тт         | G A ( | C A                | ГАС          | СТ  | G G | C A  | ТА  | c c | Τ.  | АА    | G G | А   | <b>A</b> | Т ( | 3 A        | c ( |
|    |  |   |                            |          |         |            |       |                    |              |     |     |      |     |     |     |       |     |     |          |     |            |     |
|    | Site # 18162 • with \( \text{w/o gaps} \) Selected genetic code: Stand |   |                            |          |         |            |       |                    |              |     |     |      |     |     |     |       |     |     |          |     |            |     |
|    | Site#  | 18162                                       |                            | -        |         | • \        | with  |                    | w/o          | gap | s S | elec | te  | d q | ene | tic ( | :00 | le: | Sta      | ınd | arc        | d   |