

# 2 Homework Assignment

Armintas Pakenis

2022-11-18

## Abstract

Phylogenetic trees of Betacoronaviruses genomes related to NC\_045512

## Introduction

Phylogenetic trees were generated following homework instructions. Outputs of some inner steps can be found here<sup>1</sup>. The results consist of two trees: one without an outgroup and the other with MN514967.1 as an outgroup.

## Methodology

The following sequences were collected from NCBI: - Betacoronaviruses relating to NC\_045512 with coverage better than 50%. - NC\_045512 sequence itself. - MN514967.1 (camel coronavirus).

Genomes of organisms belonging to SARS-CoV-2 were explicitly excluded to isolate the subset of interest and avoid unnecessary clustering caused by other human coronavirus variants (as NC\_045512 will be more closer to them).

## Results

Phylogenetic trees are scaled to emphasize differences between different sequences.

Phylogenetic tree without an outgroup can be found here<sup>2</sup>.

Phylogenetic tree with an outgroup can be found here<sup>3</sup>.

---

<sup>1</sup><https://github.com/ArmintasP/Bioinf/tree/main/Lab2>

<sup>2</sup>[https://raw.githubusercontent.com/ArmintasP/Bioinf/main/Lab2/5/tree\\_outgrouped.png](https://raw.githubusercontent.com/ArmintasP/Bioinf/main/Lab2/5/tree_outgrouped.png)

<sup>3</sup>[https://raw.githubusercontent.com/ArmintasP/Bioinf/main/Lab2/5/tree\\_outgrouped.png](https://raw.githubusercontent.com/ArmintasP/Bioinf/main/Lab2/5/tree_outgrouped.png)

## **Conclusion**

### **Evolution of Covid-19**

Judging by the results of both trees, NC\_045512 closely clustered with sarbecoviruses, bat coronaviruses. Pangolin coronaviruses are not far away from that cluster.

### **Outgrouping MN514967.1**

Camel coronavirus had the largest distance and was not close to any cluster in the phylogenetic tree. Hence, it became sound to make a tree with camel coronavirus outgrouped to get more accuracy. The result of outgrouping is evident - there are 3 major clusters, while before there were more than 6.

Nonetheless, the interpretation of NC\_045512 hosts' chain would not differ a lot even, except for the case that camel coronavirus would have a common ancestor as NC\_045512.

### **Evolution of Urbani SARS**

Urbani SARS viruses have same came from the same parent as bat coronavirus (KP886808.1). In addition, their ancestor is close to the family of bat coronaviruses. That family of bat coronaviruses is more distant from bat coronaviruses family clustered next to NC\_045512. The relationship between NC\_045512 and Urbani viruses is relatively miniscule.

### **The Origin of Palm Civet viruses**

The origin of Palm Civet viruses remains unknown as viruses were not included after applying various filtering criteria.