VILNIAUS UNIVERSITETAS

MATEMATIKOS IR INFORMATIKOS FAKULTETAS

PROGRAMŲ SISTEMŲ BAKALAURO STUDIJŲ PROGRAMA

**PIRMOJO LABORATORINIO DARBO ATASKAITA**

Bioinformatika

I – asis laboratorinis darbas

Atliko: 4 kurso 3 grupės studentas Lukas Orliukas

Vilnius – 2023

Turinys

[Darbo kodas 3](#_Toc148553296)

[Atstumo funkcija 3](#_Toc148553297)

[Medžiai 3](#_Toc148553298)

[Skirtumai 4](#_Toc148553299)

# Darbo kodas

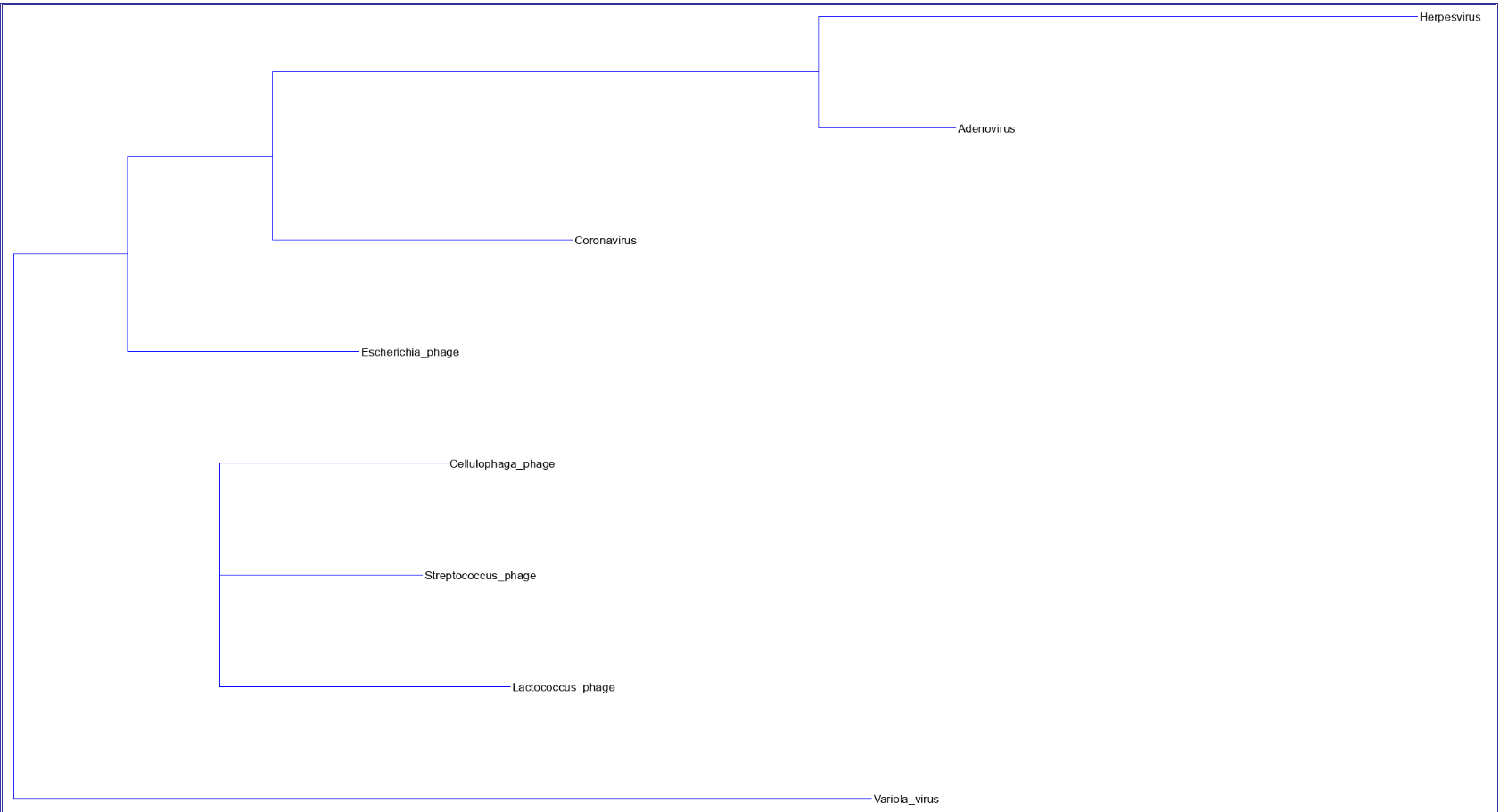
Atlikto darbo kodą galima rasti čia: https://colab.research.google.com/drive/1Kvo6e2QzkD8oGBK3Un5TId7Zo-oPE-OA?usp=sharing

# Atstumo funkcija

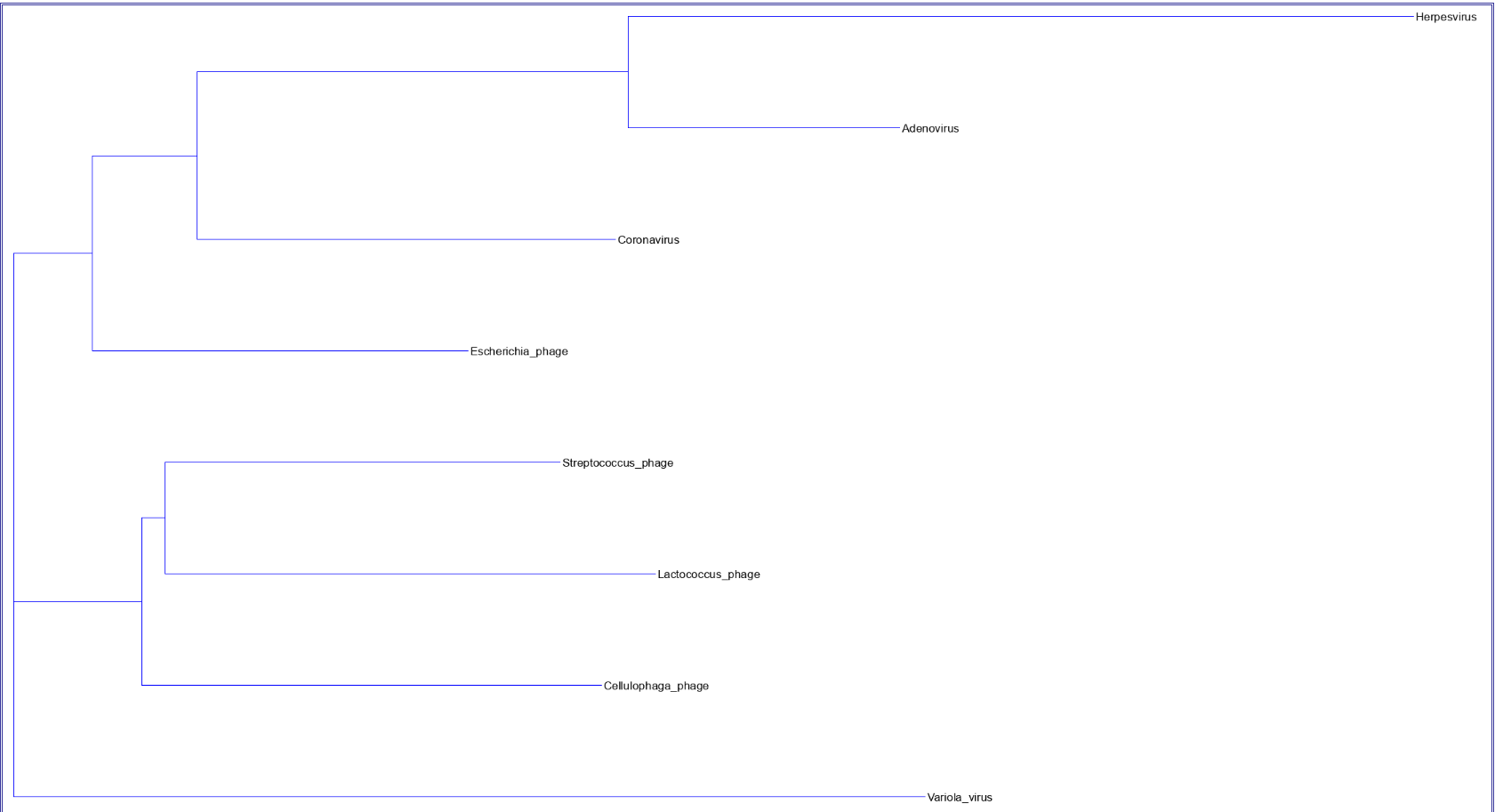
Atstumo skaičiavimui buvo naudojamas Euklidinis atstumas. Jam suskaičiuoti į pagalbą buvo pasitelkta „Python“ programavimo kalbos biblioteka „SciPy“ ir jos metodas „distance.euclidean“. Tai buvo taikyta kodonų ir dikodonų dažnių pasiskirstymo skaičiavimui kiekvienai failų (virusų sekų) porai. Didesnis atstumas rodo didesnį virusų nepanašumą.

# Medžiai

Gauti du medžiai: vienas kodonams, kitas dikodonams.

Medį gautą naudojantis kodonų atstumo matrica galite matyti pirmame paveiksle (žr. 1 pav.)

1 pav. Kodonų medis

Medį gautą naudojantis dikodonų atstumo matrica galite matyti antrame paveiksle (žr. 2 pav.)

2 pav. Dikodonų medis

# Skirtumai

Skirumą tarp žinduolių ir bakterijų virusų tikrai galima įžvelgti iš gautų rezultatų. Vieni turi didesnį panašumą, kiti mažesnį, bet bendra tendencija rodo, kad bakterijų virusai labiau skiriasi nuo žinduolių virusų. Tai geriau matosi stebint kodonų atstumo skirtumą ir prasčiau dikodonų. Klasterizacija tai ir parodo (kad sudaromi atskiri klasteriai) didžiąja dalimi, tačiau yra išimčių.

Labai įdomu, kad rezultatai rodo, kad Variola\_virus, kuris yra žinduolių, yra panašesnis į bakterijų virusus, nei į žinduolių virusus. Skirtumai nėra dideli, tačiau vis tiek panašumas didesnis į bakterijų virusus (didesnis skirtumas matosi kodonų matricoje). Taip pat bakterijų virusas Escherichia\_phage turi ne tokį pastebimą skirtumą nuo žinduolių ar bakterijų virusus. Labai artimai panašus ir į tuos ir į tuos (ypač iš dikodonų matricos).

Pagrindinis skirumas kurį pastebėjau, tai bakterijų virusuose būna mažiau C kodono (maždaug iki 1,4%), o žinduolių daugiau (maždaug nuo 2,2% iki 3%).