

Ataskaita

1 laboratorinis darbas

Darbą atliko: Aleksandra Kondratjeva

Ivados

Laboratorinio darbo tikslas buvo apskaičiuoti ir palyginti kodonų ir dikodonų dažnius tarp žinduolių ir bakterijų virusų, sukonstruoti atstumų matricą.

Kodonų ir dikodonų dažniai buvo apskaičiuoti naudojant Python programavimo kalbą ir NumPy, collections ir Biopython bibliotekas. Programa apdoroja aštuonis skirtingus FASTA failus.

Algoritmas

Iš pradžių gauname 6 frame'us – pirmus tris iš pradinės sekos, kitus iš reverse compliment sekos. Juose ieškome sekas kurios prasideda start-kodonu (ATG) ir baigiasi bet kuriuo stop-kodonu (TAA, TAG, TGA). Toliau, paverčiant kodonus į amino rūgštis, suskaičiuojami kodonų ir dikodonų dažniai, kurie normalizuojami pagal bendrą kodonų ir dikodonų skaičių.

Naudojant Euklido atstumo formulę, sugeneruojamos kodonų ir dikodonų atstumo matricos. Rezultatai įrašomi „*codon_distance_matrix.txt*“ ir „*dicodon_distance_matrix.txt*“ failuose.

Rezultatai

Apskaičiuotos atstumo matricos atskleidžia kodonų ir dikodonų dažnių skirtumus tarp žinduolių ir bakterijų virusų. Kuo mažesnė reikšmė atstumo matricoje, tuo panašesnę kodonų ar dikodonų naudojimą turi skirtingi organizmai. Kodonų atstumas tarp žinduolių ir bakterijų virusų svyruoja nuo 0.0117 iki 0.0207, o dikodonų atstumas nuo 0.0003 iki 0.0008.

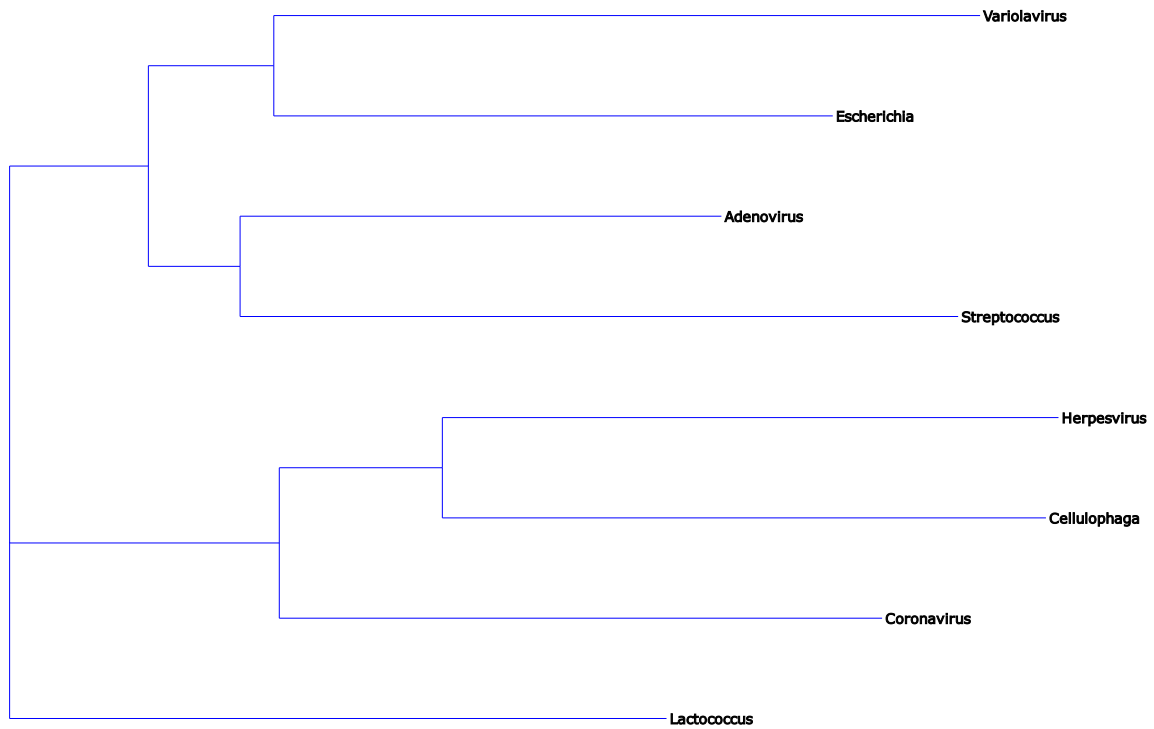


Diagram 1. Codon tree

- Virusai Cellulophaga ir Herpesvirus (0.01075) turi mažiausį kodonų atstumą.
- Virusai Streptococcus ir Herpesvirus (0.02013) turi didžiausį kodonų atstumą.
- Virusai Cellulophaga ir Herpesvirus panašūs pagal kodonų naudojimą, tačiau skirtingi pagal dikodonų.

