

Laboratorinis #1

Atstumo funkcija buvo pasirinkta skirtumo kvadratų vidurkio kvadratinė šaknis:

$$RMSD = \sqrt{\frac{\sum_{i=1}^N (p_i - \hat{p}_i)^2}{N}}$$

Atitinkamai \hat{p} vienos sekos kodonų/dekodonų dažnis atimtas iš p kitos sekos kodonų/dekodonų dažnio.

Kitame puslapyje pateikiami kokie medžiai gavosi, atitinkamai, su kodonais ir dikodonais.

Atsakant į klausimus: vertinant kodonų/dikodonų dažnių aspektu, bakteriniai ir žinduolių virusai nesudaro visiškai atskirų klasterių: matoma, jog žinduolių variola (U18337.1) virusas yra panašesnis į bakterinius virusus (taip sakant yra vienoje grupėje – klasteryje); bakterinis escherichia faga (KM389305.1) virusas atvirkščiai – arčiau žinduolių virusų, o coronavirus jau, šiuo atveju, panašus tiek į bakterinius, tiek ir į žinduolių virusus. Bendrai žvelgiant, labiausiai nuo visų kitų virusų išsiskyrė cellulophaga faga (KC821626.1) virusas. Aukščiausias kodonų/dekodonų dažnio išsiskyrimas atitinkamai matomas tarp cellulophaga faga (KC821626.1) ir herpes virusų, o mažiausias – tarp streptococcus faga (NC_028697.1) ir corona virusų.

