

Bioinformatika
Pirmas laboratorini darbas
Ataskaita

Paruošė
Raimond Sobolevskij

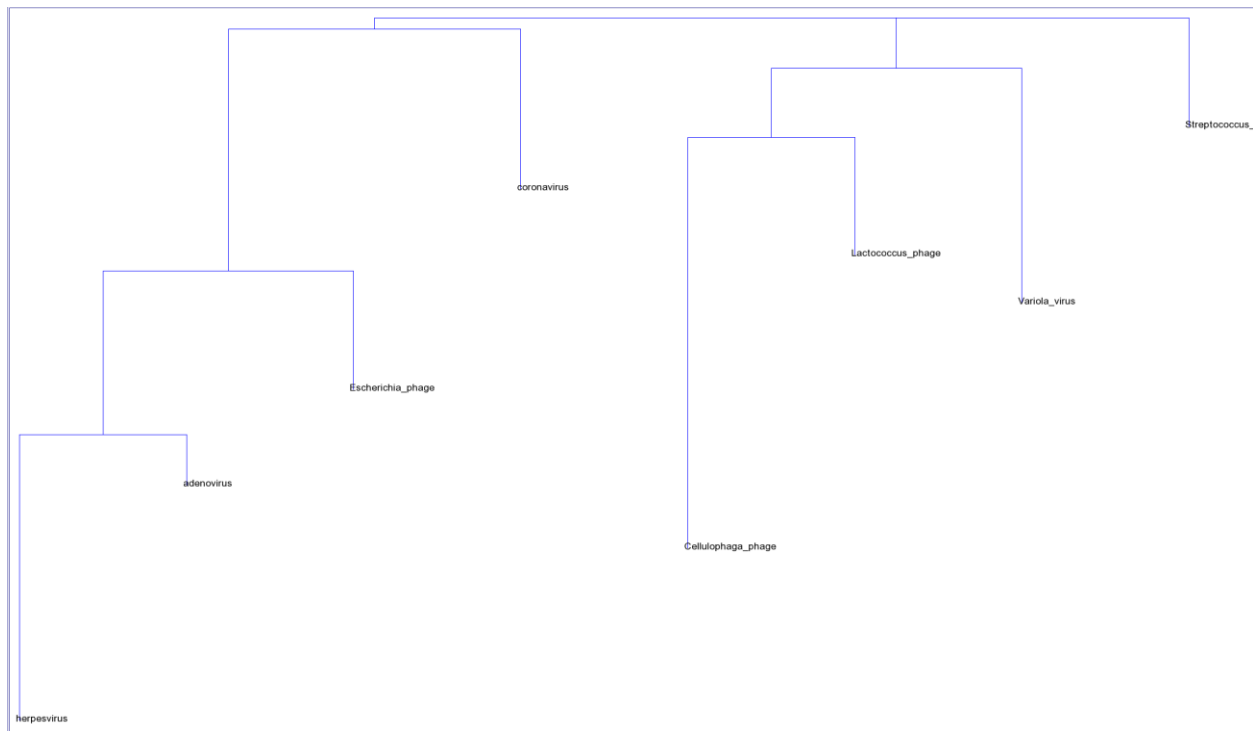
Vilnius
2021

Darbo eigos žingsniai:

1. Suradau koduojančias sritis duotuose sekose(suradau ORF ir toliausiais stovintį kodoną)
2. Gautuose sekose, skaičiavau kiekvieno įmanomo kodono ir dikodono dažnius
3. Pagal dažnius suradau atstumo matricą. Vienoje sekoje vieno kodono dažnį sudėjau su kitos sekos kodonu dažniu ir padalinau iš dviejų, gaunasi vidutinis atstumas(vėliau sugalvojau kitą formulę, iš vienos sekos kodono dažnio atimti kitą ir vėliau jas sudėti, bet rezultatas gavosi toks pat, todėl pasirinkau seną formulę). Analogiškai padariau ir su dikodonais. Vėliau gautus atstumus sudėjau ir gavosi, kad kuo arčiau nulio rezultatas, tuo daugiau yra panašesnės sekos.
4. Pagal gautas matricas gavau medžius:

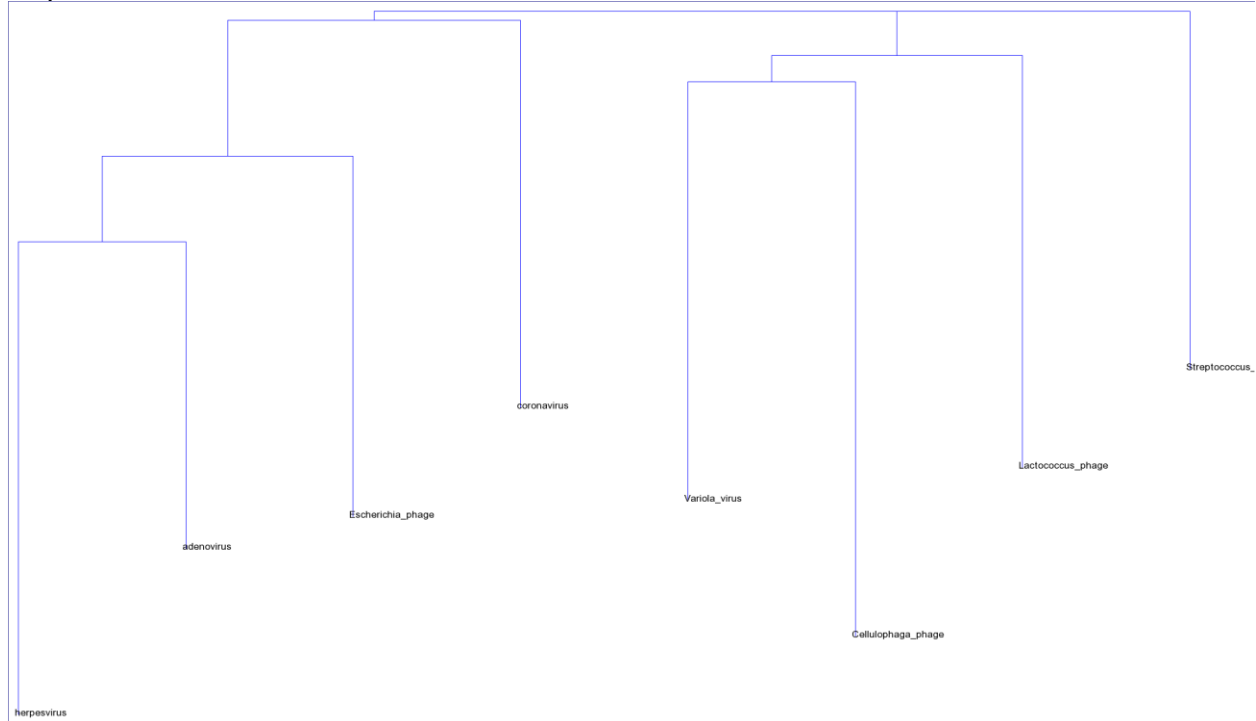
Kodonų klasterizavimas:

Lactococcus_phage	0.000000	22.479071	12.746277	19.488247	15.057989	26.012020	15.577873	34.702412
Escherichia_phage	22.479071	0.000000	17.729023	33.322479	19.212234	12.267434	24.260656	20.957825
Streptococcus_phage	12.746277	17.729023	0.000000	23.589684	10.307942	21.261972	14.527863	29.952364
Cellulophaga_phage	19.488247	33.322479	23.589684	0.000000	25.901396	36.855427	26.421280	45.545818
Coronavirus	15.057989	19.212234	10.307942	25.901396	0.000000	22.745184	16.839575	31.435575
Adenovirus	26.012020	12.267434	21.261972	36.855427	22.745184	0.000000	27.793606	12.392833
Variola_virus	15.577873	24.260656	14.527863	26.421280	16.839575	27.793606	0.000000	36.483997
Herpesvirus	34.702412	20.957825	29.952364	45.545818	31.435575	12.392833	36.483997	0.000000



Dikodonų klasterizavimas:

Lactococcus_phage	0.000000	50.980408	43.154545	52.610386	45.201710	52.667023	45.410488	61.494442
Escherichia_phage	50.980408	0.000000	45.768349	59.938667	46.868801	39.892769	52.738770	48.720188
Streptococcus_phage	43.154545	45.768349	0.000000	52.112804	39.989651	47.454964	44.912907	56.282383
Cellulophaga_phage	52.610386	59.938667	52.112804	0.000000	54.159969	61.625282	51.624565	70.452705
Coronavirus	45.201710	46.868801	39.989651	54.159969	0.000000	48.555416	46.960075	57.382835
Adenovirus	52.667023	39.892769	47.454964	61.625282	48.555416	0.000000	54.425385	41.324326
Variola_virus	45.410488	52.738770	44.912907	51.624565	46.960075	54.425385	0.000000	63.252804
Herpesvirus	61.494442	48.720188	56.282383	70.452705	57.382835	41.324326	63.252804	0.000000



Iš gautų medžių galima pastebėti, jog „Streptococcus phage“ daugiausiai išsiskiria iš kitų virusų, kadangi yra atsiskyręs nuo kitų bakterijų ir žinduolių virusų. Taip pat „Escherichia phage“ yra daugiau panašesnis į žinduolių virusus, kaip tarp kodonų taip pat ir tarp dikodonų. Didžiausias skirtumas yra tarp „herpesvirus“ ir „Cellulophaga“ virusų, kadangi jie priklauso skirtingoms virusų grupėms, nes kaip žinduoliai taip ir bakterijos yra linkę klasterizuotis kartu, nors ir yra išimčių, kaip pavyzdžiui „Variola virus“, kuris klasterizavosi kartu su bakterijų virusais. Tarp kodonų ir dikodonų skirtumas nedidelis, tik apsikeitė bakterijų klasteris.