VILNIAUS UNIVERSITETAS

MATEMATIKOS IR INFORMATIKOS FAKULTETAS

**Bioinformatikos pirmojo laboratorinio darbo ataskaita**

**Atstumo funkcijos skaičiavimas:**

Sukūriau compare\_freqs(arr1, arr2) funkciją, kurioje:

* Apibrėžtas rezultatų masyvas (resultArr)
* Iteruojama per pirmą dažnių masyvą (arr1)
* Kiekvienoje iteracijoje ieškomas arr[i] atitikmuo (vienodas dikodonas/kodonas).
* Jeigu jis randamas, atliekamas skaičiacimas: |arr1[i].daznis – atitikmuo.daznis|. Atsakymas keliauja į rezultatų masyvą.
* Ciklas pasibaigia, tikrinamas rezultatų masyvo ilgis.
* Jeigu ilgis lygus 0, funkcija grąžina -1.
* Jeigu ilgis didesnis už 0, atsumo funkcijos rezultatas apskaičiuojamas taip: (sum(resultArr) / len(resultArr)) \* 1000

**Gauti medžiai:**

<http://www.trex.uqam.ca/index.php?action=trex&menuD=1&method=2>

Graphical user interface

Description automatically generated

pav. 1 - Genomų sekų kodonų atstumų matricos medis

Graphical user interface

Description automatically generated with medium confidence

pav. 2 - Genomų sekų dikodonų atstumų matricos medis

**Pastebėjimai**

Atstumų matricoje virusas U18337.1 yra įvertintas -1 atstumu, nes jo kodonų/dikodonų nepavyko rasti. Kodonų medyje labiausiai nutolę buvo KC821626.1 ir herpesvirus genomai. Dikodonų medyje labiausiai nutolęs Lactococcus\_phage virsusas. Kodonų ir dikodonų dažnis tarp žinduolių ir bakterijų virusų gerokai skiriasi. Labiausiai išsiskyręs virusas yra U18337.1, kadangi nepavyko rasti jam kodonų/dikodonų porų.

Labiausiai varijuojantys kodonai:

* TCC
* TTG
* TTT
* CCG

Labiausiai varijuojantys dikodonai:

* TTCCTC
* CTCCCC
* GGGAGG
* TCCCCCC
* TTTTCT
* ACCCCA
* GCCCCC

**Atstumų matricos**

Kodonų:

8

Lactococcus\_phage 0.0 5.7146 7.6465 8.8265 5.9784 8.8761 -1.0 11.0277

KM389305.1 5.7146 0.0 6.3972 8.3431 4.6095 6.7081 -1.0 7.9787

NC\_028697.1 7.6465 6.3972 0.0 8.4531 7.3422 7.1699 -1.0 8.1296

KC821626.1 8.8265 8.3431 8.4531 0.0 6.5945 9.9228 -1.0 11.4814

coronavirus 5.9784 4.6095 7.3422 6.5945 0.0 6.429 -1.0 7.9431

adenovirus 8.8761 6.7081 7.1699 9.9228 6.429 0.0 -1.0 3.9756

U18337.1 -1 -1 -1 -1 -1 -1 -1 -1

herpesvirus 11.0277 7.9787 8.1296 11.4814 7.9431 3.9756 -1.0 0.0

Dikodonų:

8

Lactococcus\_phage 0.0 2.6261 1.1936 2.1741 2.5014 3.3669 -1.0 3.2458

KM389305.1 2.6261 0.0 1.8879 1.3049 0.8512 0.8848 -1.0 1.0853

NC\_028697.1 1.1936 1.8879 0.0 1.9221 1.6569 2.8988 -1.0 2.4624

KC821626.1 2.1741 1.3049 1.9221 0.0 0.9831 1.3012 -1.0 1.065

coronavirus 2.5014 0.8512 1.6569 0.9831 0.0 0.7495 -1.0 0.8204

adenovirus 3.3669 0.8848 2.8988 1.3012 0.7495 0.0 -1.0 0.4891

U18337.1 -1 -1 -1 -1 -1 -1 -1 -1

herpesvirus 3.2458 1.0853 2.4624 1.065 0.8204 0.4891 -1.0 0.0

Laura Mučinytė