

Atstumo funkcija: Atimame kodono/dikodono dažnį tarpusavyje tarp sekų, skirtumui panaudojame absoliutaus dydžio (modulio) $\text{abs}()$ formulę, visus dažnius susumuojame ir iš sumos ištraukiame kvadratinę šaknį. Pavyzdžiui:

Seka1: („AAA“, 120), („GGT“, 115)

Seka: („AAA“, 151), („GGT“, 179)

$$\text{Atstumas} = \sqrt{\text{abs}(120 - 151) + \text{abs}(115 - 179)}$$

Kodonai atstumų matrica

8

bacterial1	0.0	8496.5	22897.0	6635.5	125038.0	15560.5	13006.5	12057.0
bacterial2	8496.5	0.0	21751.5	11048.0	122746.5	11420.0	15394.0	11189.5
bacterial3	22897.0	21751.5	0.0	23196.5	103135.0	18701.5	35522.5	31709.0
bacterial4	6635.5	11048.0	23196.5	0.0	125256.5	18882.0	12859.0	12590.5
mamalian1	125038.0	122746.5	103135.0	125256.5	0.0	113461.5	137963.5	133727.0
mamalian2	15560.5	11420.0	18701.5	18882.0	113461.5	0.0	26335.0	21867.5
mamalian3	13006.5	15394.0	35522.5	12859.0	137963.5	26335.0	0.0	5161.5
mamalian4	12057.0	11189.5	31709.0	12590.5	133727.0	21867.5	5161.5	0.0

Kodonų medis

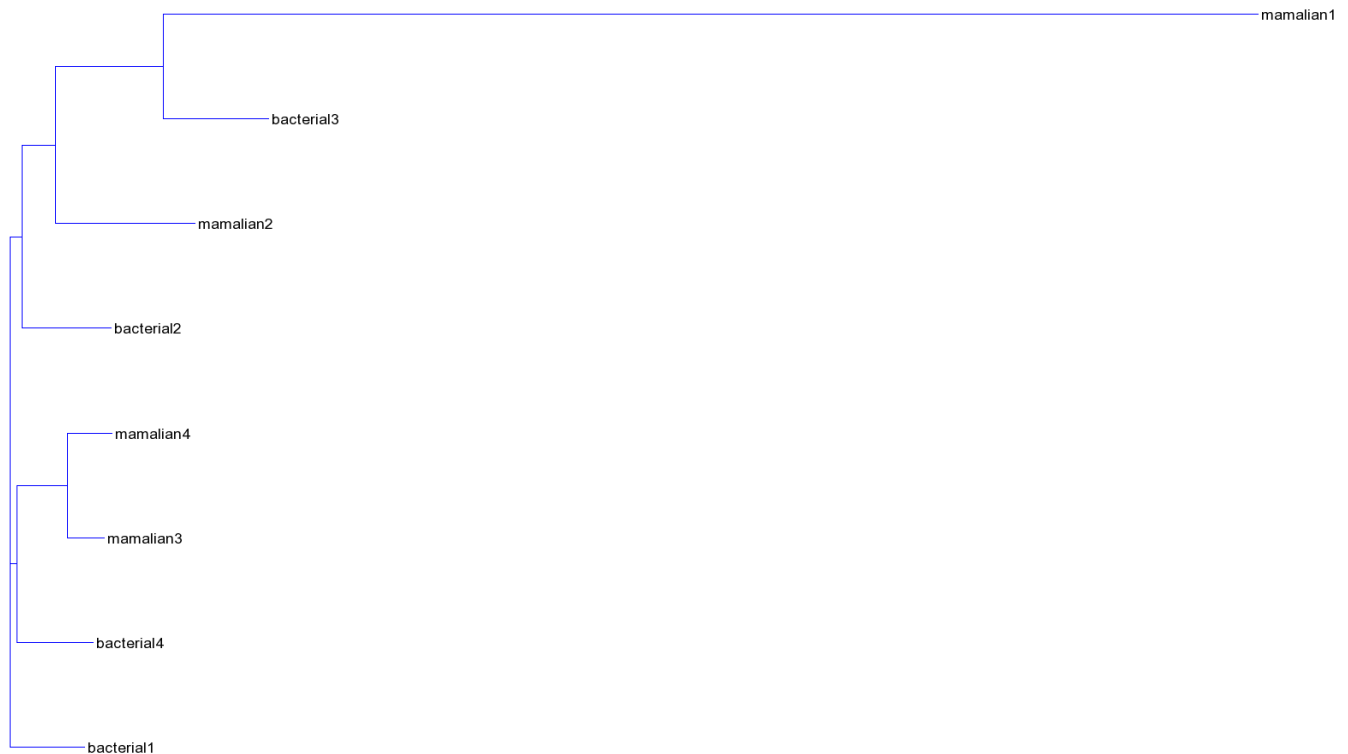


Dikodonai atstumų matrica

8

bacterial1	0.0	10025.0	15660.5	8578.5	65183.5	14034.5	7609.5	8965.5
bacterial2	10025.0	0.0	17031.5	10601.5	66883.5	12520.5	8640.5	8646.5
bacterial3	15660.5	17031.5	0.0	15662.0	60274.0	18914.0	18386.0	18552.0
bacterial4	8578.5	10601.5	15662.0	0.0	64638.0	15011.0	7539.0	9129.0
mamalian1	65183.5	66883.5	60274.0	64638.0	0.0	66138.0	69462.0	69885.0
mamalian2	14034.5	12520.5	18914.0	15011.0	66138.0	0.0	13795.0	12587.0
mamalian3	7609.5	8640.5	18386.0	7539.0	69462.0	13795.0	0.0	4080.0
mamalian4	8965.5	8646.5	18552.0	9129.0	69885.0	12587.0	4080.0	0.0

Dikodonų medis



Dažniausi kodonai tarp visų sekų buvo : ('AAA', 20294) ('GCT', 19328) ('AAT', 18444)

Dažniausi dikodonai tarp visų sekų buvo: ('GGTGTT', 640) ('GCTGGT', 600) ('TAAATG', 586)

Labiausiai paplitę:

bacterial1: 'AAA', 'TTCTTT'

bacterial2: 'GAT', 'TAAATG'

bacterial3: 'GCT', 'GCTGGT'

bacterial4: 'AAA', 'GATGAT'

mamalian1: 'GTT', 'GGTGTT'

mamalian2: 'ATG', 'TAAATG'

mamalian3: 'ATA', 'TAAATG'

mamalian4: 'ATG', 'TGAATG'

- Labiausiai išsiskyrė mamalian1 virusas, kuris turėjo žymiai daugiau kodonų nei kiti.
- Lyginant b1,b2,b3,b4 su m2 kodonų kiekis panašus
- Lyginant b1, b2, b3, b4 su m3, m4, bakteriniai virusai turėjo daugiau kodonų

Klasterizacija:

- Šakninis klasteris:
 - bacterial1
 - (bacterial4, mamalian4, mamalian3)
 - (bacterial2, mamalian2, bacterial3, mamalian1)
- (bacterial4, mamalian4, mamalian3) skyla į:
 - bacterial4
 - (mamalian4, mamalian3)
- (bacterial2, mamalian2, bacterial3, mamalian1) skyla į:
 - bacterial2
 - (mamalian2, bacterial3, mamalian1)
- (mamalian2, bacterial3, mamalian1) skyla į:
 - mamalian2
 - bacterial3, mamalian1