# Bioinformatikos 3 laboratornis darbas

Atliko: Kasparas Taminskas, Programų sistemos, 4 kursas, 5 grupė, 2019 rudens semestras

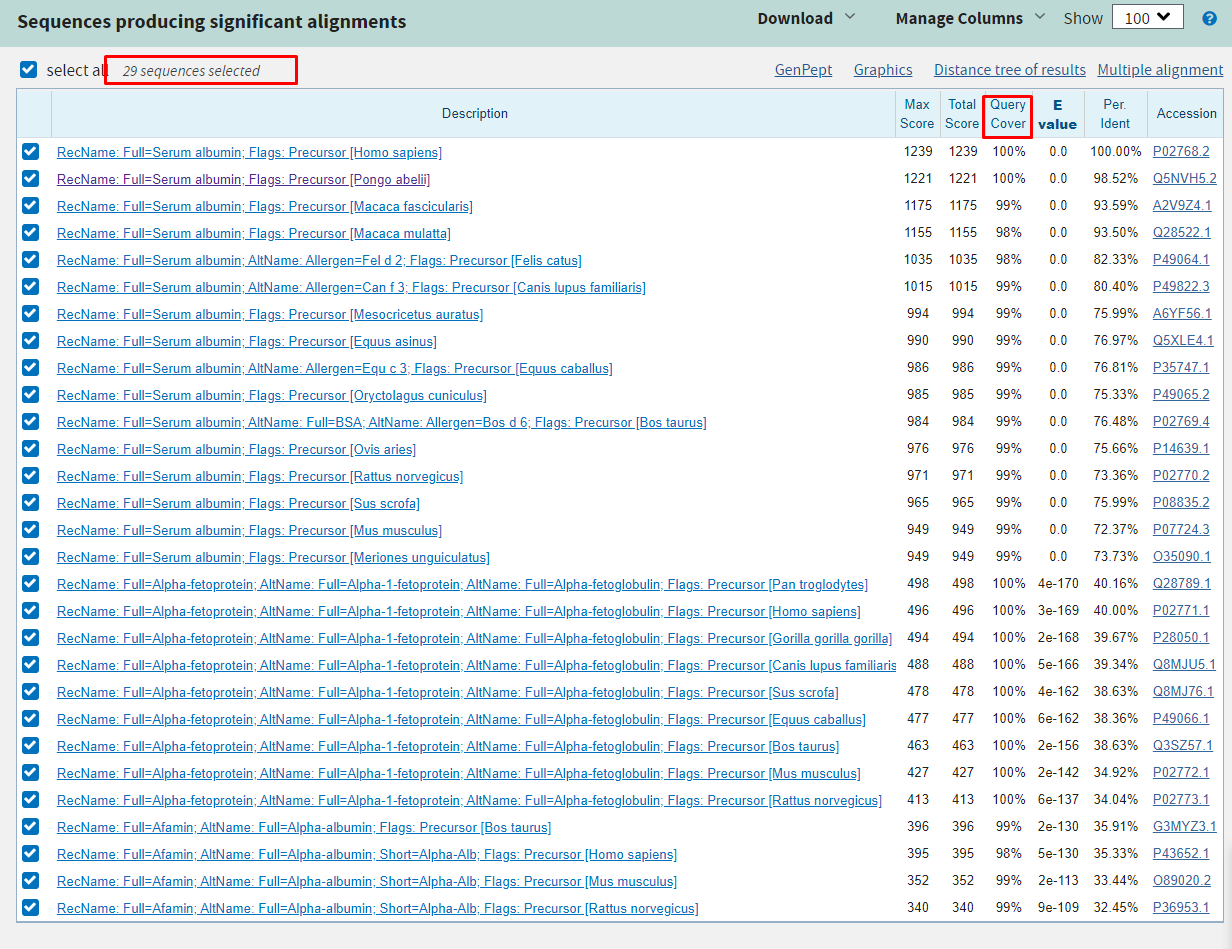
Tikslas: Parinkti baltymo sekos fragmentą, tinkamą gaminti antikūnus, kurie atpažintų žinduolių (kur vaikus maitina pienu ir turi plaukus – tame tarpe ir žmonės) pagrindinį kraujo ir plazmos baltymą - serumo  
albuminą Parinktas baltymo fragmentas būtų atskirai susintetintas ir naudojamas pelių imunizacijai. Išskirti antikūniai prieš žmogaus serumo albuminą (ŽSA) būtų naudojami kriminalistikoje - atpažinti  
kraujui ir nustatyti ar kraujas yra žmogaus.

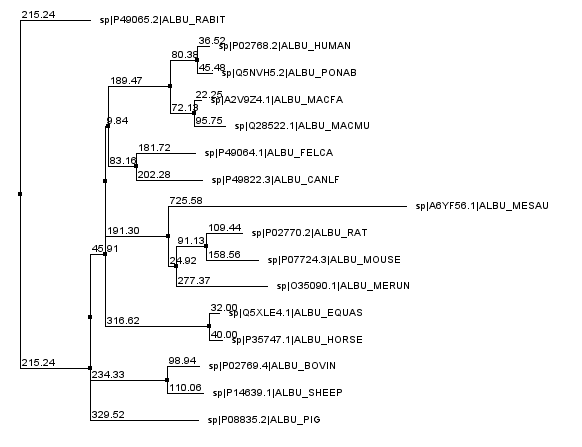
Darbo eiga:

1. Išnagrinėjama prieinama atitinkamo ŽSA geno informacija NCBI puslapyje:
   1. Atlikta paieška
   2. Baltymo prieigos kodas: NP\_000468.1
   3. NCBI pagal paiešką "[Protein name] AND mammals[Organism]" informacija yra šioms

žinduolių rūšims:

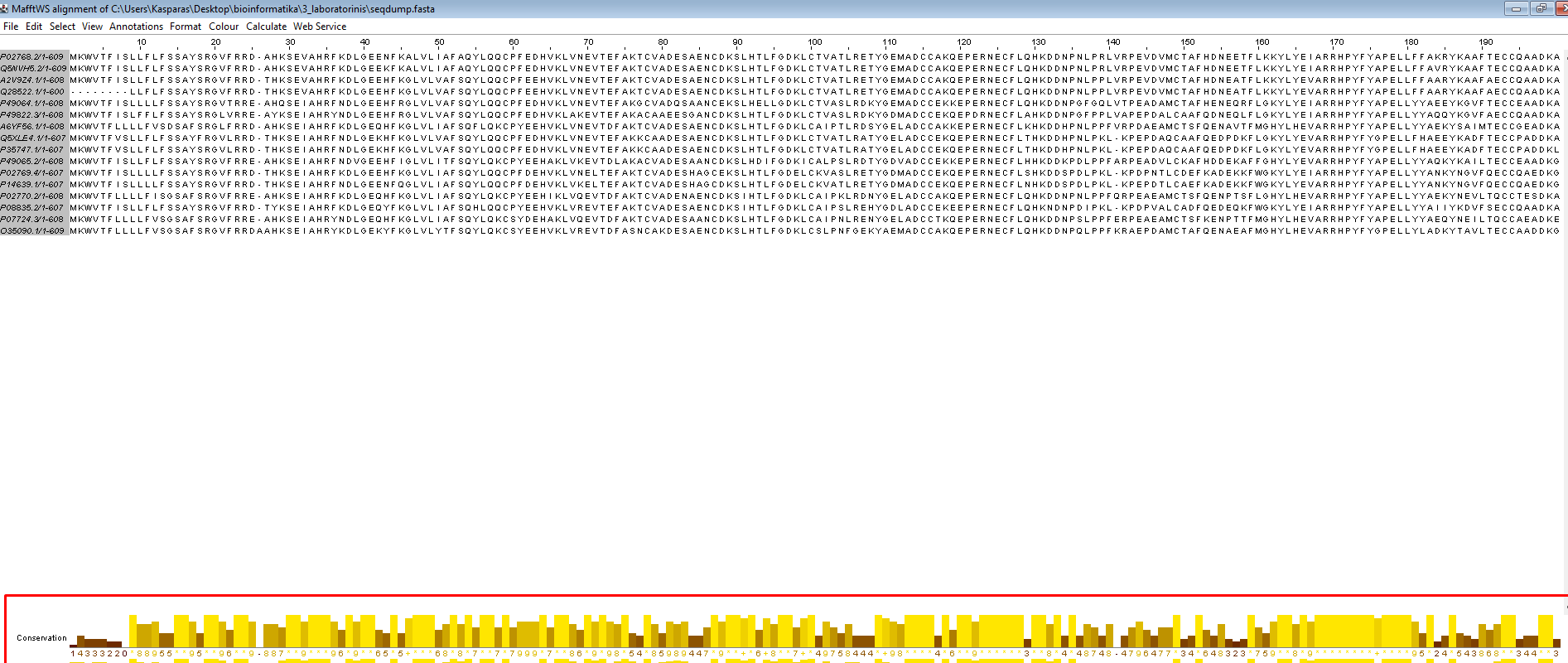
* + 1. Homo sapiens (66) - žmogus
    2. Mus musculus (13) - naminė pelė
    3. Rattus norvegicus (8) - pilkoji žiurkė
    4. Bos taurus (5) - naminis jautis
    5. Pongo abelii (4) - sumatrinis orangutanas
    6. Macaca mulatta (4) - rezusas
    7. Canis lupus familiaris (4) - šuo
    8. Equus caballus (4) - arklys
    9. Pan troglodytes (3) - paprastoji šimpanzė
    10. Felis catus (3) - Felis - rainuotoji katė
    11. Callithrix jacchus (2) - paprastoji marmozetė
    12. Nomascus leucogenys (1) - baltaskruostis gibonas
    13. Papio anubis (1) - žaliasis pavianas
    14. Ailuropoda melanoleuca (1) - didžioji panda

1. Gauti ŽSA ir panašių sekų žinduoliuose sekų palyginį:
   1. 
   2. Atlikus BLAST paiešką pagal ŽSA seką, gavau 35 hit‘us, iš jų liko 29 atmetus visas sekas, kurios turi mažiau nei 80% “Query Cover“. O iš pastarųjų atmetus sekas, kurios nėra serumo albuminai (buvo alfa-fetoproteinai ir afaminai), taip pat sekas, kurios yra žmogaus, liko 15 hit‘ų.
   3. Su JalView programa atlikau išsaugotų 15 sekų MAFFT (L-INS-i preset) palyginio radimą ir Neighbour joining metodu nubražiau filogenetinį medį:



Iš jo matome, jog panašiausias serumo albumino baltymas į žmogaus yra ALBU\_PONAB – tai **Pongo abelii (Sumatrinis orangutanas).**

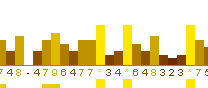
1. Surasti amino sekas:
   1. 20 AR fragmentas iš ŽSA sekos, kuris labiausiai panašus į kitų žinduolių ŽSA:
      1. JalView programoje atlikau MAFFT palyginimą sekoms ir apačioje gavau konservatyvumo grafiką, kuriame pozicionaliai matosi, kurie regionai skiriasi mažiausiai ir daugiausiai:

Pagal šį grafiką ir ieškojau regiono, kuris turėtų didžiausią ištęstą konservatyvumą ir radau tokį: [430-450] sekos pozicijos, o pats fragmentas:

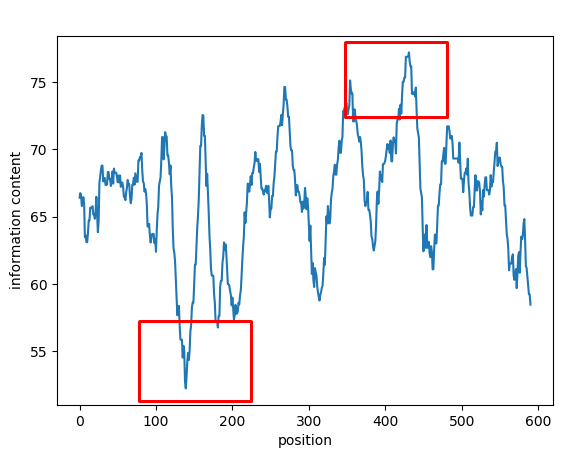
**NALLVRYTKKVPQVSTPTLVE –** jis būtų naudojamas sintetinant ir gaminant antikūnus žinduolių kraujo atpažinimui. Sekos vietos konservatyvumas buvo tikrai didelis: 

* 1. ŽSA 20 AR fragmentas, kuris labiausiai skirtųsi:

Ta pati logika, tiesiog ieškojau mažiausio konservatyvumo regiono, radau tokį [137-157] pozicijoje, seka: **LPRLVRPEVDVMCTAFHDNEE,** regiono konservatyvumo vaizdas, matosi žymiai mažesnis**:**



1. Atliktoje programavimo dalyje MAFFT palyginiui bandžiau skaičiuoti information content: <http://biopython.org/DIST/docs/tutorial/Tutorial.html#sec:getting_info_content> reikšmę kiekvienam 20 AR fragmentui iš MAFFT palyginio. Gavau tokį information content grafiką, kur x ašyje yra pozicija palyginyje, o y ašyje information content įvertis:



Šis grafikas, kurį mano programa sugeneruoja, patvirtina mano 3 dalies fragmentų pasirinkimus, nes žemiausias IC įvertis atitinka žemiausius konservatyvumo įverčius JalView programoje ir atvirkščiai.

Programos duomenys:

* Išsaugota Homo sapiens ŽSA seka: **serum\_albumin\_homo\_sapiens.fasta** faile
* Blast.xml faile – Blast paieškos rezultatai
* Similar\_albumins.fasta – Blast paieškoje rastos sekos
* Mafft\_msa.fasta – MAFFT palyginys

Programa randa sekų fragmentus:

'Similar sequence: ', Seq('**LVRYTKKVPQVSTPTLVEVS**', SingleLetterAlphabet()))

'Differing sequence: ', Seq('**VRPEVDVMCTAFHDNEETFL**', SingleLetterAlphabet()))

Matome, jog tai sekos iš tų pačių regionų, kuriuos identifikavau 3 dalyje.