**Atlikimas:** 35 min.

**Laboratoriniai darbai, kuriuos reiktų atlikti prieš atliekant šį darbą:** BIOINF1.

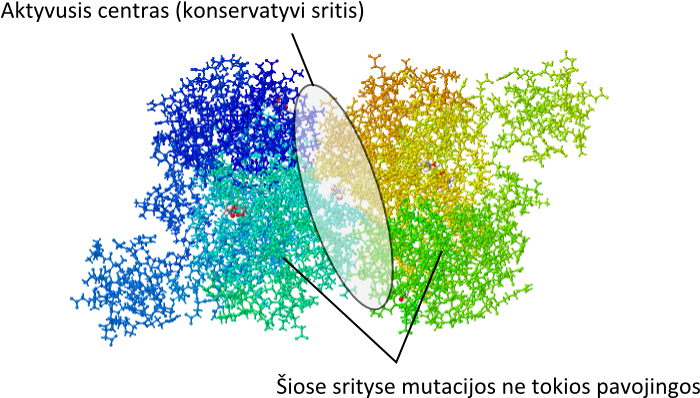
# Konservatyvių sričių nustatymas Ebola viruso baltymuose

**Teorinis įvadas**

*Ebolavirus* yra virusų gentis, kuriai priskiriamos 5 rūšys: *Reston ebolavirus*, *Zaire ebolavirus*, *Sudan ebolavirus*, *Tai Forest ebolavirus*, ir *Bundibugyo ebolavirus*. Iš jų tik *Reston ebolavirus*, manoma, neužkrečia žmonių (šis virusas aptiktas tik beždžionėse), o visi kiti sukelia mirtinai pavojingą Ebola karštligę. *Zaire ebolavirus* atsakingas už didžiausią 21-ojo amžiaus epidemiją, kuri 2013-2015 metais nusinešė beveik 10 000 žmonių gyvybes, taigi šis virusas šiuo metu aktyviai tyrinėjamas. Viruso išoriniame apvalkale yra baltymas, vadinamas Ebola viruso glikoproteinu. Manoma, kad šis baltymas būtinas virusui patekti į žmogaus ląsteles – jei galėtume tam sutrukdyti, galėtume sukurti vaistus nuo Ebola karštligės. O šiai užduočiai atlikti reikia išsiaiškinti daugiau apie patį glikoproteiną.

Norint nustatyti, kurios baltymo dalys svarbios atliekant tam tikrą funkciją, galima sulyginti to paties baltymo iš kelių giminingų organizmų aminorūgščių sekas. Pavyzdžiui, žmogaus ir beždžionės hemoglobinas tikrai atlieka tą pačią funkciją – beždžionei juk irgi reikia deguonies. Tačiau abiejų baltymų aminorūgščių sekos kiek skirsis dėl evoliucijos eigoje įvykusių mutacijų. Tai ypač būdinga toms baltymo sritims, kurios nėra labai svarbios jo funkcijai atlikti – aminorūgštys, kurios tiesiog palaiko bendrą baltymo formą, gali būti pakeistos kitomis be didesnių problemų (pav. 1). Iš kitos pusės, fermentų aktyvusis centras būna labai svarbus, ir net vienos aminorūgšties pakeitimas gali būti mirtinas. Tas pats būdinga ir kitoms sritims, kurios svarbios baltymo veikimui – pavyzdžiui, hemoglobino monomerams susijungti į ketvirtinę struktūrą būtina, kad tam tikros aminorūgštys baltymo išorėje būtų nepažeistos. Tokios sritys vadinamos konservatyviomis („išsaugotomis“). Palyginę daugelio organizmų baltymų struktūrą, konservatyviose srityse rastume žymiai mažiau mutacijų, nei kitur. Vadinasi, mutacijų kiekis parodo, ar tam tikra sritis svarbi baltymo funkcijai.

Šiame darbe turėsite palyginti glikoproteino sekas iš kelių giminingų virusų ir nustatyti, kurios sritys yra konservatyvios, taigi galėtų būti naujų vaistų taikiniu.



Pav. 1. Skirtingose fermento vietose aminorūgščių seka nevienodai svarbi.

**Tikslas**: nustatyti konservatyvias sritis Ebola viruso glikoproteino sekoje.

**Reikalingi failai**: virusu\_baltymai.fasta, virusu\_baltymai\_TEST.fasta

**Eiga**

Jums pateiktame faile **virusu\_baltymai.fasta** yra glikoproteinų iš penkių *Ebolavirus* rūšių sekos. Dar viena artimai gimininga rūšis yra Marburgo virusas – pirmiausia suraskime ir jo glikoproteino seką.

1. Naršyklėje atsidarykite NCBI svetainę adresu <http://www.ncbi.nlm.nih.gov>.
2. Čia sutelkta informacija iš daugelio duomenų bazių. Lango viršuje matysite paieškos laukelį, o šalia jo – išskleidžiamą meniu, kur pasirinkta „All Databases“. Išskleiskite šį meniu ir pasirinkite „Protein“, kadangi ieškosime baltymo sekos.
3. Paieškos laukelyje įrašykite „glycoprotein marburgvirus“ (tai baltymo ir organizmo pavadinimai) ir spauskite „Search“ mygtuką. Paieška pateiks nemažai rezultatų, kurie atrodo panašiai:

glycoprotein [Marburg marburgvirus] **baltymo pavadinimas [organizmas]**

681 aa protein **aminorūgščių skaičius**

Accession: ACT79243.1 GI: 254688087 **sekos numeris**

Pav. 2. „Protein“ duomenų bazės paieškos rezultatų formatas.

1. Kad apribotumėte rezultatų skaičių, kairėje, po antrašte „Source databases“, spustelkite užrašą „RefSeq“. Taip paliekamos tik būdingiausios organizmo sekos, o įvairios retų atmainų, neaiškios kilmės, nepilnos sekos yra pašalinamos.
2. Iš likusių rezultatų išsirinkite bet kurį patinkantį ir spustelkite ant pavadinimo eilutės.
3. Atsidariusiame vaizde matysite įvairios informacijos – kur buvo rastas šis virusas, kokie autoriai nustatė baltymo seką ir t.t. Jums reikalinga pati seka FASTA formatu, taigi viršuje, po sekos pavadinimu, spustelkite nuorodą „FASTA“.
4. Matysite tipišką FASTA sekos vaizdą (kaip pav. 3). Nukopijuokite visas eilutes su baltymo seka, **o taip pat ir pavadinimo eilutę**.

>gi|3253216|gb|AAC24346.1| virion spike glycoprotein

**Pavadinimas**

precursor [Reston ebolavirus]

MKKTKIVCTIGPKTESVEKLTELVNAGMNVMRLNFSHGDYVEHGTR

ITNFRKVMEVTGKQLAILLDTKGPEIRTIKLENGDDVDLVAGQEFT

**Seka**

FTTDTKVVGNKERVAVTYSGFTKDLNVGNRILVDDGLIEMEVLATT

DTEVKCKVLNNGALGENKGVNLPGVSVNLPALSEKDKNDLKFGCEQ

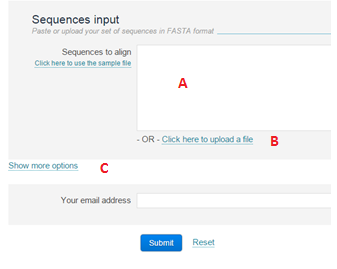
Pav. 3. FASTA formato sekos vaizdas.

1. Nukopijuotą seką įterpkite savo failo virusu\_baltymai.fasta pradžioje. (Failą galite taisyti su „Notepad programa, kaip aprašyta darbe BIOINF1.)
2. Papildytą failą išsaugokite.

Dabar galime atlikti baltymų sulyginimą ir ieškoti konservatyvių sričių.

1. Naršyklėje atidarykite svetainę T-COFFEE adresu <http://tcoffee.crg.cat>.
2. Užveskite pelę ant „Proteins“ lauko ir pasirinkite „Structural alignments (Expresso)“. Tai paprasta baltymų sulyginimo (angl. *alignment*) programa.

Atsivers langas su minimaliais programos nustatymais. Jie atrodys taip:



Pav. 4. Programos „Expresso“ darbo langas.

1. Pirma išbandykime programą tik su dviem baltymų sekomis. Faile **virusu\_baltymai\_TEST.fasta** yra tik dvi sekos. Šio failo turinį nukopijuokite į langelį A („Sequences to align“), arba pasirinkite failą paspaudę nuorodą B („Click here to upload a file“).
2. Paspauskite nuorodą C („Show more options“), kad atsivertų papildomi nustatymai. Apačioje, skiltyje „Output options“, prie „Residue number“ vietoj „off“ nustatykite „on“ – taip aminorūgštys bus sunumeruotos, ir bus lengviau aptarti rezultatus.
3. Spauskite mygtuką „Submit“. Netrukus turėtų pasirodyti ir rezultatai, kur jūsų baltymų sekos bus pateiktos dviem eilutėmis, maždaug taip:

BAD AVG GOOD

Zaire\_virus 1 LLQLNETI-YTSGKRSNTTG 19

Sudan\_virus 1 LFQLNDTIHLHQ-QLSNTTG 19

cons 1 \*:\*\*\*:\*\* : \*\*\*\*\* 20

Pav. 5. Sulyginimo rezultatas. Pirmos dvi eilutės – tai dviejų baltymų aminorūgščių sekos. Skaičiai reiškia, kad šios eilutės rodo nuo 1-os iki 19-os kiekvieno baltymo aminorūgšties. Sekų panašumą rodo trečioji eilutė (žvaigždutė rodo, kad aminorūgštys sutampa abiejose sekose, taškai – kad aminorūgštys panašios, o tarpai – kad gana skirtingos). Tą patį koduoja ir spalvos (mėlyna ir žalia – sekos nesutampa, geltona – panašios, raudona – sutampa). Brūkšnys eilutėje reiškia, kad toje vietoje viena aminorūgštis buvo prarasta (delecija) arba kitoje sekoje atsirado papildoma (insercija).

K1. Trumpai aprašykite, kiek panašios šios dvi sekos. Ar to ir reikėtų tikėtis?

|  |
| --- |
|  |

1. Dabar naršyklės mygtuku „Back“ grįžkite į pradinį paieškos langą. Šįkart naudosime visas sekas, taigi mygtuku B pasirinkite failą **virusu\_baltymai.fasta** arba nukopijuokite šio failo turinį į langelį A.
2. Paleiskite programą mygtuku „Submit“.

Atsidarius rezultatams (tai gali užtrukti), matysite visų šešių baltymų sulyginimą.

K2. Kuris iš baltymų labiausiai skiriasi nuo kitų? Ar to ir reikėjo tikėtis?

|  |
| --- |
|  |

K3. Kurios baltymo sritys visuose virusuose panašios, o kurios – skiriasi? Lentelėje žemiau nurodykite, ar tokiais numeriais pažymėtos aminorūgštys sutampa gerai (žymėkite \*), ar vidutiniškai (žymėkite : ), ar nesutampa visai (palikite tuščią laukelį).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Baltymo sritis (aminorūgščių nr.)** | 1-200 | 200-340 | 340-510 | 510-670 |
| **Panašumas** |  |  |  |  |

K4. Kurios iš šių sričių labiausiai konservatyvios?

|  |
| --- |
|  |

K5. Kurią iš šių sričių naudotumėte kaip taikinį naujiems vaistams prieš šį virusą?

|  |
| --- |
|  |

K6. Tarkime, norite sukurti vaistus, kurie apsaugotų nuo visų šių virusų. Jums reikia parinkti taikinį – kuo ilgesnę nepertraukiamą visiškai sutampančių aminorūgščių seką. Suraskite geriausią taikinį, nurodykite jo seką ir aminorūgščių numerius (apytiksliai).

|  |
| --- |
|  |

K7. O kurią baltymo sritį pasirinktumėte, jei norite sukurti diagnostinį testą nustatyti pacientą užkrėtusio viruso rūšiai? Pagrįskite savo pasirinkimą.

|  |
| --- |
|  |