

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
MATEMATINĖS INFORMATIKOS KATEDRA

Rimantė Žalytė
bioinformatikos studijų programa

Ciano bakterijų serino treonino baltymų kinazių analizė
Analysis of Cyanobacteria Serine/Threonine Protein Kinases

Bakalauro baigiamasis darbas

Darbo vadovė: lekt. **Kotryna Kvederavičiūtė**

Vilnius 2016

Turinys

Įvadas.....	3
1. Literatūros apžvalga.....	4
1.1 Ciano bakterijos	4
1.2 Baltymų kinazės.....	6
1.3 Serino treonino baltymų kinazė.....	7
2. Metodai	9
2.1 Perl programos.....	9
2.2 CLANS sekų grupavimas.....	10
2.3 Filogenetinių medžių sudarymas	10
2.4 Domenų identifikavimas.....	10
3. Rezultatai ir jų aptarimas.....	11
Išvados	18
Santrauka	19
Summary	20
Literatūros sąrašas	21
Priedai	24
1. Priedas.....	24
2. Priedas	25
3. Priedas.....	26
4. Priedas.....	27
5. Priedas	28
6. Priedas	37
7. Priedas	42

Ivadas

Organizmo genome yra užkoduoti visi gyvybiškai svarbūs mechanizmai, jų dalys, pavyzdžiui, būdingos baltymų kinazių rūšys. Jos sudaro daugumą tarp kinazių ir vykdydamos fosforilinimą gali reguliuoti baltymų funkcijas, inicijuoti ar sutrikdyti sąveikas su kitais baltymais, taip pat yra svarbios fermentų veiklai bei signalų perdavimui ląstelėje.

Šiame bakalauriniame darbe siekiama:

1. Išsiaiškinti serino treonino kinazių reikšmę ciano bakterijoms.
2. Iširti serino treonino baltymų kinazių pasiskirstymą tiriamose ciano bakterijose.
3. Tiriamas serino treonino kinazių sekas palyginti filogenetiškai.
4. Įvertinti kinazes, kaip galimą taksonominį markerį ciano bakterijų klasifikacijai.

Pirmojoje darbo dalyje apžvelgiama literatūra: aiškinamasi kokios yra ciano bakterijoms būdingos savybės, kinazių svarba organizmuose, kokia serino treonino baltymų kinazių reikšmė.

Antrojoje darbo dalyje aprašyti metodai, kuriais naudojantis analizuojami sekvenuoti ir anotuoti ciano bakterijų genomų duomenys ir juose užkoduotos kinazės.

Trečiojoje darbo dalyje pateikiami rezultatai, gauti panaudojus antroje dalyje aprašytus metodus, bei gautų rezultatų aptarimas.

1. Literatūros apžvalga

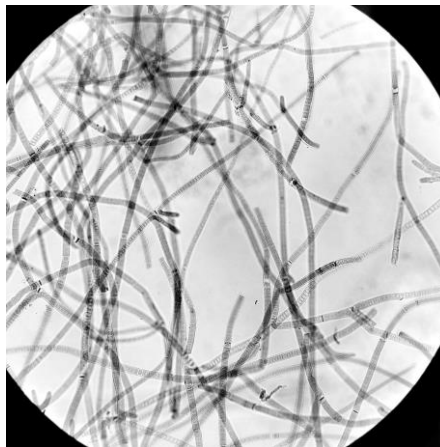
Didelės genų grupės koduoja baltymus, kurie yra membraniniai ir atlieka receptoriaus funkcijas ląstelės paviršiuje. Dalis šių baltymų yra kinazės receptoriai ir paprastos kinazės, kurios yra atsakingos už signalo perdavimą ir baltymų postransliacines modifikacijas.

Signalų perdavimas ląstelėse yra svarbus tuo, kad gali valdyti pagrindinę ląstelės veiklą ir gali koordinuoti ląstelės veiksmus. Tai padeda ląstelėms tinkamai reaguoti į mikro aplinkos plėtrą, audinių atstatymą, vykdyti homeostazę [1]. Šie procesai ir jų valdymai yra konservatyvūs daugelyje organizmų įskaitant ir ciano bakterijas.

1.1 Ciano bakterijos

Ciano bakterijos arba melsvadumbliai yra prokariotiniai organizmai (1 pav), galintys energijos gauti fotosintezės būdu. Ciano bakterijos taip pat geba vykdyti azoto fiksaciją.

Tai vieni sėkmingiausių organizmų, paplitimo atžvilgiu. Ciano bakterijų dėl itin gausios genetinės įvairovės galima aptikti gėlo vandens, jūrų ir sausumų ekosistemose bei kraštutinėse nišose, tokiose kaip karštosios versmės ar itin druskingi šaltiniai. Jos puikiai išsaugojo gebėjimus prisitaikyti prie labai įvairios aplinkos ir turėjo didelę įtaką Žemės atmosferos susidarymui. Vandens ciano bakterijos geriausiai žinomos dėl jų sukeliama vandens žydėjimo.



1 pav. Ciano bakterija *Tolypothrix* [2]

Ciano bakterijos yra skirstomos į 5 grupes (pagal R. Rippka). Pirmosios dvi, I ir II, iš jų yra vienaląsčiai organizmai, kurie randami kaip pavieniai organizmai arba ląstelių kolonijos susijungiančios papildomu išorinių dienelių sluoksniu. Šių grupių ciano bakterijoms būdingas dauginimasis dalijantis vienai ląstelei į dvi naujas. Likusių trijų grupių, III, IV ir V, ciano bakterijoms būdingos ląstelių grandinėlės [3]. Šiose grupėse identifikuoti organizmai skirstomi į smulkesnius pogrupius pagal gentis ir joms būdingus bruožus.

Ciano bakterijas sudaro daug skitingų šeimų, pavyzdžiui, *Aphanizomenon*, *Anabaena* ir kitos. Deja, bet tik nedidelė dalis šių ciano bakterijų yra daugiau tirtos ir tik nedidelės dalies atstovų genomai yra nusekvenuoti. Pavyzdžiui, lietuviai nusekvenavo *Aphanizomenon flos-aquae* [4], amerikiečiai *Anabaena sp. 39864*, *Anabaena sp. WA113* [4], *Nodularia spumigena* [6] ir t.t. Šiuo metu viešai prieinama tik iki trisdešimties ciano bakterijų.

Aphanizomenon flos-aquae - cianobakterijos, sukeliančios vandens žydėjimą Kuršių mariose. Šis kasmet intensyviai pasikartojantis reiškinys kelia pavojų biologiniams ištekliams, mažina pajūrio regiono turizmo patrauklumą [7]. Šios ciano bakterijos plinta gėlo vandens telkiniuose, tačiau šiuo atveju jos aptiktos ir pusiau sūriame vandenyje. *Aphanizomenon flos-aqua* populiacijos tyrinėtose ežeruose buvo įvairesnės nei eanti Baltijos jūroje, tad buvo nustatyta, kad vieno tipo ežeruose rasti genotipai natūralios atrankos būdu yra geriau prisitaikę išgyventi druskingame vandenyje [7].

Anabaena priklauso cianobakterijų genčiai, gaminančiai toksinus. Pavojingi šios populiacijos gaminami toksinai yra mikrocistinai, anatoksinais ir saksitoksinais [8]. *Anabaena*, pagal duomenis gautus *in vivo*, koduoja bent 38 II tipo toksinų antitoksinų sistemas, tačiau spėjama, kad individualus sistemų toksiškumas yra labai nežymus ir jų efektyvumas gali būti pastebimas tik veikiant kartu kelioms sistemoms [9].

Nodularia spumigena, priklausanti *Nodularia* genčiai, pirmoji nustatyta toksiška kitiems gyvūnams ciano bakterijų rūšis, įprastai aptinkama sūriame vandenyje. Įtariama, kad dėl gyvūnų apsinuodijimo vandens žydėjimo periodu šio Baltijos regiono pakrantėse kaltas būtent šios ciano bakterijų rūšies dominavimas [10]. *Nodularia spumigena* gamina antrinį metabolitą nodulariną ir ciklinį baltymą hepatoksiną. Taip pat nuolatinis nodularino išskyrimas sukelia įvairius pakitimus žinduoliams, todėl daugelis žmonių ir gyvūnų apsinuodijimų žydinčiu vandeniu priskiriama būtent šiai ciano bakterijai [11].

Cylindrospermum yra siūlinių melsvadumblių rūšis. Jų galima rasti ir sausumoje (dirvožemyje), ir vandens ekosistemoje (dažniausiai kartu su kitais organizmais sudaro dumblo sluoksnį ant vandens augalų), yra labiausiai paplitę tropiniuose ir subtropiniuose regionuose. Analizuojami šiai genčiai priklausančių *Cylindrospermum stagnale* ir *Cylindrospermopsis raciborskii* ciano bakterijų genomai, kurie priskiriami toksiškoms ciano bakterijoms, sukeliančioms vandens žydėjimą [12].

Dolichospermum genčiai priklausanči *Dolichospermum circinale* ciano bakterija dažniausiai aptinkama gėlame vandenyje. Kaip ir *Aphanizomenon flos-aquae* gali gaminti toksiną priskiriamas medžiagas, bet be reikiamų sąveiku nėra toksiška. Šių ciano bakterijų genomai nemažai analizuotas ir lyginti dėl šio siejančio požymio [13].

Nostoc (analizuojamos *Nostoc azollae*, *Nostoc punctiforme*) cianobakterijos, kaip ir *Anabaena*, *Aphanizomenon* ar *Cylindrospemum*, priklauso *Nostocales* šeimai, kurios gali būti randamos dirvožemyje, ežerų ar šaltinių dugnuose, retais atvejais jūrose, gali gyventi kaip simbiontai su augalais. Dauguma šių melsvabakterių yra toksiškos, bet ne visos, pavyzdžiui, *Nostoc punctiforme* negamina jokių toksinų [14].

Raphidiopsis gentis dažniausiai apibūdinama kaip PSP toksinų gamintoja. Dėl šios priežasties gan didelė *Raphidiopsis brookii* genomo dalis sutampa su *Cospermopsis raciborskii*, kuri gamina tuos pačius toksinus. Šios rūšies ciano bakterijos dažniausiai aptinkamos gėlo vandens telkiniuose [15].

Richelia genčiai priklausanči *Richelia intracellularis* aptinkame vandens ekosistemose, dažnai tropinio ir subtropinio klimato regionų jūrose. Tyrinėti galimi šios ciano bakterijos simbiotiniai sąryšiai su kitais tokiais pačiomis sąlygomis išgyvenančiais organizmais, siekianti išsiaiškinti ar tai daro įtaką sukeliama vandens žydėjimui ir azoto fiksacijai [16].

1.2 Baltymų kinazės

Baltymų kinazės, tai fermentų šeima katalizuojanti baltymo ir ATP (angl. Adenosine triphosphate, vienas pagrindinių ląstelės makroenergetinių junginių) konversiją į ADP (angl. Adenosine diphosphate) ir fosfoproteiną. Tai reguliuoja įvairius biologinius procesus.

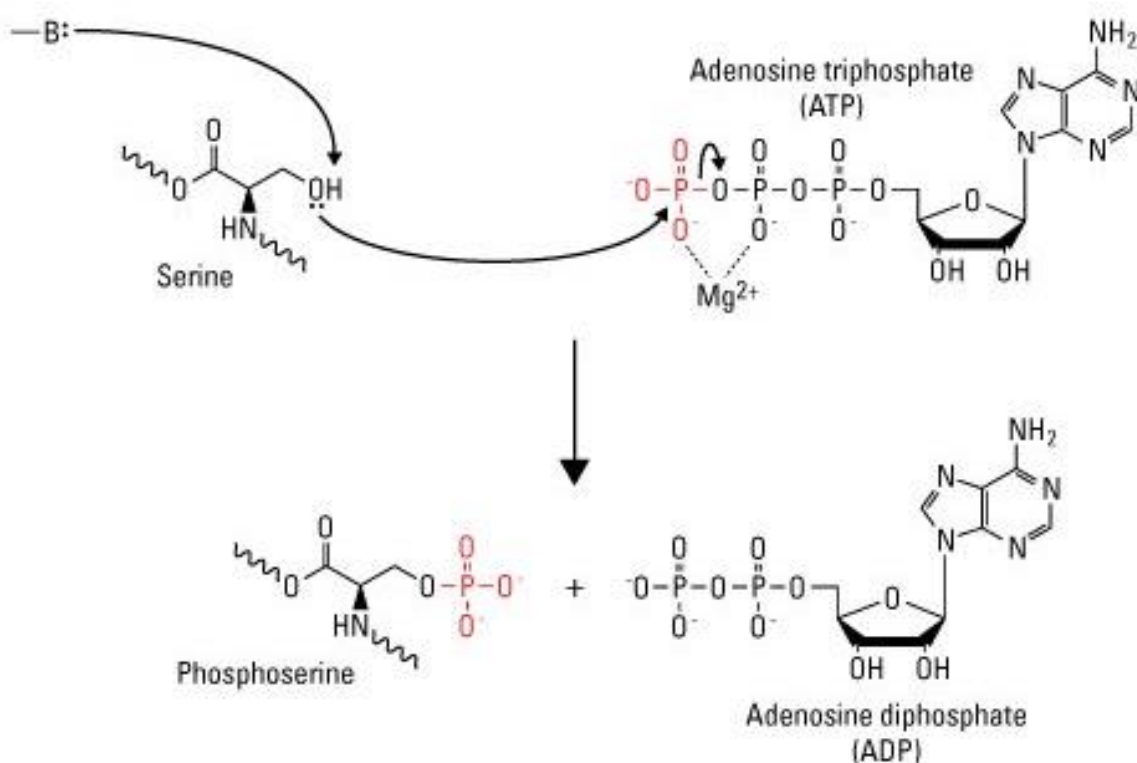
Bakterijoms, taip pat, būdingos eukariotinio tipo baltymų kinazės, kurios atlieka svarbų vaidmenį ląstelių diferenciacijai, patogeniškumo ir antrinio metabolismo reguliacijai. Šio tipo kinazės dažnai paaiškina bakterijų gebėjimus, tokius kaip signalinė reakcija į šviesą ir kitus, kurie yra panašūs į augalų. Ciano bakterijoms, pavyzdžiui, *Anabaena*, tai garantuoja pakankamą atsaką ir reakciją į aplinką. Pagal turimus genome užkoduotus signalinės sistemos komponentus, serino treonino kinazės ir atsako reguliavimo domeno, jos signalinis potencialas yra vienas didžiausių [17].

Baltymų kinazių klasifikavimo schema pagrindžiama kinazių domenų filogenetika, pagal S. K. Hanksą ir T. Hunterį. Remiantis šia teorija visa baltymų kinazių šeima persiskiria į dvi dalis: serino treonino ir tirozino kinazes. Tačiau ši teorija galioja nagrinėjant tik eukariotų baltymų kinazes [18]. Prokariotų organizmuose randamos baltymų kinazės gali būti suskirstytos į 72 subšeimas, pavyzdžiui, iš kurių 38 būdinga itin savita taksonomija [19].

1.3 Serino treonino baltymų kinazė

Serino treonino baltymų kinazės yra fermentai, atliekantys serino treonino OH grupės fosforilinimą. Jų receptoriai yra atsakingi už proliferacinį reguliavimą, programuotą ląstelės mirtį (apoptazę), ląstelių diferenciaciją ir embrioninį vystymąsi [20].

Kaip ir daugumos kinazių, pagrindinė serino treonino baltymų kinazių funkcija yra fosforilinimas (2 pav.). Tai postransliacinė baltymų modifikacija, kurios metu aminorūgščių liekana yra fosforinama baltymų kinazės, kovalentiškai prijungiant fosfato grupę [21].



2 pav. Serino fosforilinimo schema. Fermentas katalizuoja protonų perdavimą iš serino sužadintos (OH) grupės, kad γ -fosfato grupė būtų pernešama serinui. Rezultatas: fosfoserinas ir ADP [22].

Dažniausiai serino treonino baltymų kinazė yra transmembraninė. Tai naudinga perduodant receptoriams signalus, vykdant fermentų aktyvaciją. Yra bandymais įrodyta, kad tokio tipo kinazės baltymų fosforilinimą gali vykdyti ir *in vitro*. Tai naudinga norint pritaikyti baltymų gavimą dirbtinėmis sąlygomis, ypač aktulų medicinoje [23].

Ciano bakterijoms serino treonino baltymų kinazių signalo perdavimo sistemos itin svarbios dėl galimybės prisitaikyti prie labai įvairios aplinkos. Šis dviejų komponentų mechanizmas yra plačiai paplitęs kaip bakterijų reguliacinė struktūra reaguojant į stresą,

patogeniškumą ir galimybę prisitaikyti prie kinatamumo. Melsvadumblių mutacijų tyrimai parodė, kad serino treonino baltymų kinazės gali dalyvauti ne tik azoto apykaitos reguliavime, bet net ir ląstelių judrumo valdyme [24]. Toks platus šių baltymų kinazių suteikiamų funkcijų spektras įrodo, kad tai eukariotinio tipo struktūros. Būtent serino treonino baltymų kinazės ir yra vienos labiausiai paplitusios, nes būdingos tiek prokariotams, tiek eukariotams [25].

2. Metodai

2.1 Perl programos

Išanalizuoti trisdešimties cianobakterijų genomuose esančias kinazes buvo parašyta keletas programų Perl kalba. Šios nedidelės programos buvo panaudojamos skirtinguose analizės etapuose, apibendrinti turimus duomenis ar gautus rezultatus.

Pirmoji programa (1 priedas) suskaičiuoja statistiką kiek ir kokios rūšies kinazių buvo aprašytą visų šių tiriamų ciano bakterijų genomų anotacijose. Pagal Genbank pateiktus duomenis panaudojant šią programą gaunama statistika, kurioje pateiktas kiekvienos rastos skirtingos kinazės pavadinimas ir pasikartojimo dažnį šiuose organizmuose nurodantis skaičius. Rezultatų duomenys gaunami programa nuskaitant visus genomų anotacijų failus vienu kartu.

Antroji Perl programa (2 priedas) skirta pasirinktos kinazių grupės sekų išrinkimui iš turimų genomų anotacijų. Panaudojant reguliariąsias išraiškas ieškoma gausiausiai besikartojančios serino/treonino kinazių (angl. Serine/threonine protein kinase) rūšies. Radus reikiamą pavadinimą išgaunama ir ją aprašanti seka. Programos išvestyje gaunami FASTA formatą aptitinkantys sekų aprašai: pirmoje eilutėje pradedant simboliu „>“ prasideda sekos aprašas, kuriame nurodomas sekos numeris (automatiškai priskiriamas unikalus identifikatorius), organizmo pavadinimas (pagal anotacijos failų pavadinimus) ir identifikatorius pagal sekos locus tag; antroje (ir tolimesnėse) eilutėje išrašoma aprašyme minimą kinazę koduojanti seka. Rezultatų duomenys, taip pat, gaunami programos įvesčiai nurodant pilną genomų anotacijų failų sąrašą.

Trečioji programa (3 priedas) skirta gautų serino treonino kinazių grupių statistikai suskaičiuoti. Kiekvienai grupei skaičiuojamas organizmų pasikartojimo dažnis joje. Programos įvesčiai pateikiami failai, kuriuose yra išvardintos kiekvienai grupei priklausančios serino/treonino kinazių sekos, o išvestyje pateikiamas abstraktus (vartotojo sukurtas) grupės pavadinimas ir organizmų sąrašas su jų dažniais.

Ketvirtoji Perl programa (4 priedas) skirta iškirpti baltymų kinazių domenams iš sekų, kurios buvo paimitos iš ciano bakterijų anotacijų. Pagal po palyginimų gautas serino treonino baltymų kinazių domenų koordinatas iškerpama reikiama dalis. Vykdam sekų paiešką pagal suteiktus pavadinimus iš duomenų failo apskaičiuojamas intervalas aprašytai baltymų kinazei ir jis iškerpamas. Programos išvestyje pirmoje eilutėje pateikiamas sekos, kurioje buvo identifikuotas atitinkantis domenas, pavadinimas (suformuotas ir priskirtas su antrąją Perl programa) ir iškirpta seka antroje eilutėje. Tai kartojama su visomis sekomis, tad tokių tipų

išvesties eilutėmis užpildoma pagal kiekvieną palyginimo būdu identifikuotą serino treonino baltymų kinazės domeną.

2.2 CLANS sekų grupavimas

CLANS yra duomenų vizualizacijos Java įrankis naudojamas sekų klasifikavimui pagal jų panašumus arba filogenetinius medžius. Panaudojant su antrąja perl programa atsiskirtas serino treonino baltymų kinazių sekas šiuo įrankiu gaunama grafinė vizualizacija. Inicijuojant sekų klasterizaciją *pi* reikšmė (angl. *P-values*) buvo nurodyta 1.0E-70.

Pagal gautus sekų klasterius išrenkamos kiekvienai grupei būdingos sekos. Taip turėtos šių kinazių rūšies sekos suskirstomos į atskiras grupių, iš kurių kiekvienai priklauso bent 10 sekų. Pagal kiekvieną pasirinktą grupę sudaromi jai priklausančių sekų sąrašai.

2.3 Filogenetinių medžių sudarymas

Filogenetinių medžių sudarymui buvo panaudota keli skirtingi programiniai įrankiai. Pirmiausia gauti grupių sekų rinkiniai buvo išlyginti. Tam panaudota programa Muscle [26]. Tai laisvai prieinamas programinis įrankis skirtas daugybiniam baltymų ir nukleotidų sekų palyginimui. Gavus sekų grupių palyginius duomenys dar kartą peržiūrėti ir esant galimybei pakoreguojamas išlyginimas rankiniu būdu.

Tuomet iš išlygintų sekų rinkinių suformuojami filogenetiniai medžiai. Tam naudojamas Iq-tree [27] programa. Pagal kiekvienos su CLANS pagalba atskirtos grupės sekų rinkinį suformuotas atskiras filogenetinis medis. Jų tikslumui įvertinti taip pat ir suformuotas filogenetinis medis iš visų analizuotų serino treonino kinazių sekų.

2.4 Domenų identifikavimas

Tiksliam baltymų kinazių domentų identifikavimui naudotas Pfam įrankis. Lokaliai kompiuteryje įdiegu Pfamscan ir Hmmer įrankius atliekamas pateikiamų sekų palyginimas su esančiomis duomenų bazėje pagal paslėptą Markovo modelį (angl. *hidden Markov model*). Radus atitikmenis identifikuojamas ir įvardijamas kokios struktūros tai domenai, pateikiamos jo koordinatės, pavadinimas, tipas, šeima, ilgis ir panašūs parametrai [28].

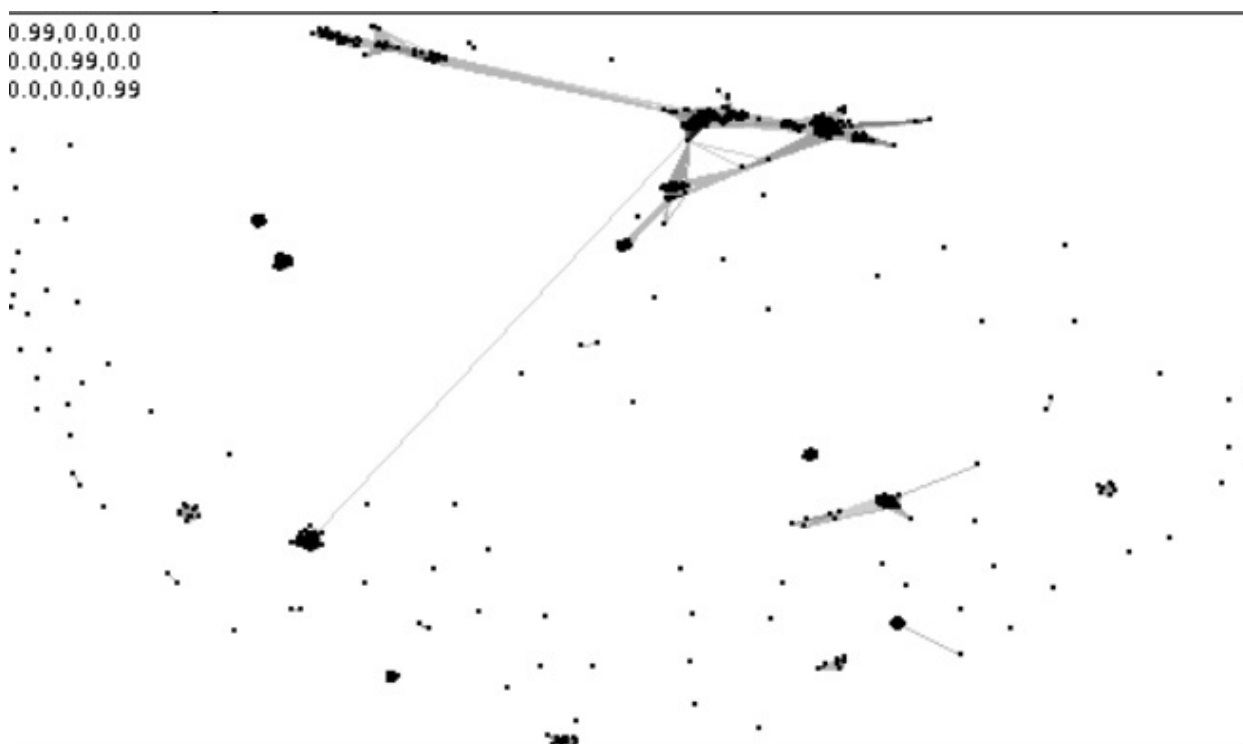
Palyginamos visos turimos 811 pagal anotacijas atrinktas serino treonino baltymų kinazių sekos ir iš gautų domenų atrenkami tik priklausantys baltymų kinazių (angl. *Pkinase*) tipai. Tolimesniam gautų duomenų panaudojimui rezultatai apdorojami, kad būtų turima informacija tik apie sekas ir domenus, kuriuose yra serino treonino baltymų kinazės domenai.

3. Rezultatai ir jų aptarimas

Panaudojus pasirašytą Perl programą iš 30 organizmų anotacijų buvo gauti duomenys, kiek ir kokių kinazių buvo aptikta tirtų ciano bakterijų genomuose. Iš viso aptikta virš 3000 kinazių sekų, iš kurių dažniausios serino treonino kinazių, pasikartojusios net 811 kartų. Toks didelis vienos rūšies baltymų kinazių dažnis paskatino daryti hipotezę, kad jų gausus paplitimas skirtingose ciano bakterijose gali turėti įtakos rūšių giminiškumui ir klasifikacijai. Tam išsiaiškinti buvo nuspręsta smulkiau nagrinėti būtent serino treonino baltymų kinazių sekas.

Tikslesniam minimų ciano bakterijų analizavimui buvo atsiskirtos tik serino treonino baltymų kinazes koduojančios sekos. Tai padaryta su antrosios Perl programos pagalba. Gautas sekų rinkinys tapo svarbiausiais duomenimis, atspindinčiais ir kinazes koduojančias sekas, ir jai būdingą organizmą.

Pagal gautas serino treonino sekas bandoma išsiaiškinti smulkesnį jų klasifikavimą turimose ciano bakterijų genomuose. Pagal įvairias sekas esančias prieinamose duomenų bazėse nepavyko rasti atitikmenų, parodančių galimą šios rūšies kinazių atsiskyrimą į grupes, tad nuspręsta panaudoti sekų klasterizavimo ir vizualizavimo įrankį CLANS. Iš įvesto sekų rinkinio gauta grafinė vizualizacija parodė tam tikrų sekų susigrupavimą į „debesėlius“ (3 pav.).



3 pav. Serino treonino kinazių sekų grafinė vizualizacija naudojant CLANS

Remiantis tokių serino treonino baltymų kinazių vizualių susiklasterizavimų buvo sudarytos sekų grupės, kurioms priklausė ne mažiau nei 10 sekų. Iš viso tolimesniai analizei

tokiu principu atrinkta 15 grupių, kurioms dėl paprastesnio atskyrimo suteikti atsitiktiniai pavadinimai raidėmis.

Išnagrinėti gautoms kinazių grupėms panaudota dar viena Perl programa. Taip gauti duomenys apie tai kiek ir kokio organizmo serino treonino baltymų kinazių sekų patenka į kiekvieną grupę (4 pav.).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
afakm1d3.gbk	3	6	0	4	0	0	3	0	0	0	0	0	2	1	1
ana29413.gbk	4	15	0	4	0	0	2	11	1	0	1	1	2	1	1
ana7122.gbk	3	11	0	4	0	1	1	1	1	1	4	1	2	1	0
anawa102.gbk	3	11	0	3	0	0	2	1	1	2	1	0	2	1	0
dsp131c.gbk	3	5	0	1	0	1	1	0	1	0	0	0	0	1	1
nos7120.gbk	3	13	0	2	0	0	2	12	1	1	1	1	2	1	0
riche.gbk	1	4	0	0	0	0	1	0	0	0	0	0	1	1	0
afamdt13cul.gbk	3	8	0	2	2	1	1	0	1	1	0	1	0	1	0
ana39858.gbk	4	13	0	1	3	0	2	10	1	0	1	1	2	1	1
ana90.gbk	2	13	0	1	5	0	4	0	1	1	1	0	0	1	1
anawa113.gbk	3	9	0	1	2	1	1	0	1	1	0	1	0	1	0
dsp310f.gbk	3	4	0	0	1	1	1	0	1	1	1	0	0	1	0
nos73102.gbk	3	15	0	1	2	1	2	8	1	0	2	1	2	1	0
richehh01.gbk	1	3	0	0	0	0	0	0	0	0	0	0	0	1	0
afamdt14.gbk	3	7	0	1	2	1	1	0	1	1	0	1	0	1	0
ana39864.gbk	2	8	0	1	0	1	2	0	1	0	1	1	0	1	0
anaaml93.gbk	3	10	0	1	2	0	1	1	1	2	1	0	2	1	0
cylind7417.gbk	3	7	0	1	1	1	0	1	1	0	3	0	2	0	0
nodul9414.gbk	2	7	2	1	1	1	3	1	1	0	3	2	0	1	1
nos7524.gbk	3	6	8	1	1	0	1	0	1	1	2	0	3	0	0
richehm01.gbk	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
afanies81.gbk	3	6	4	1	2	0	1	1	1	1	1	0	2	1	1
ana39865.gbk	3	5	4	1	1	0	2	0	1	1	1	1	0	1	0
anacrks33.gbk	3	6	3	0	3	1	1	0	1	0	1	2	0	1	1
cylindcs505.gbk	2	3	2	0	1	0	1	1	0	0	0	0	0	0	1
nos0708.gbk	1	3	3	1	0	0	0	1	1	0	1	0	1	1	0
raphd9.gbk	2	2	2	0	0	0	1	0	0	0	0	0	0	0	1
afawa102.gbk	4	8	0	2	0	1	1	0	1	1	0	1	0	1	0
ana7108.gbk	1	4	0	0	0	0	0	0	1	1	1	0	1	1	0
anamdt14.gbk	3	8	0	2	0	0	2	0	1	0	0	1	0	1	0
nos7107.gbk	3	15	0	2	0	2	1	4	1	0	2	1	3	1	0

4 pav. Serino treonino sekų pasiskirstymas grupėse pagal skirtingas ciano bakterijas

Pastebima, kad kaikuriose grupėse (A, B, G, N) yra serino treonino kinazių iš beveik visų tiriamų ciano bakterijų, o kai kurios grupės (C, E, O) ribojasi tik nedidele organizmų dalimi. Taip pat, keletas organizmų, pavyzdžiui, *richehh01* ir *richehm01* (*Richelia intracellularis*) atsiskyrė akivaizdžiu sekų negausumu. Tam įtakos galėjo turėti netikslus organizmų sekų sekvenavimas (surinktas netikslus ir per mažas genomai) arba kitokie faktoriai nulėmė kinazių sekų nesigrupavimą prie filogenetiškai artimiausių.

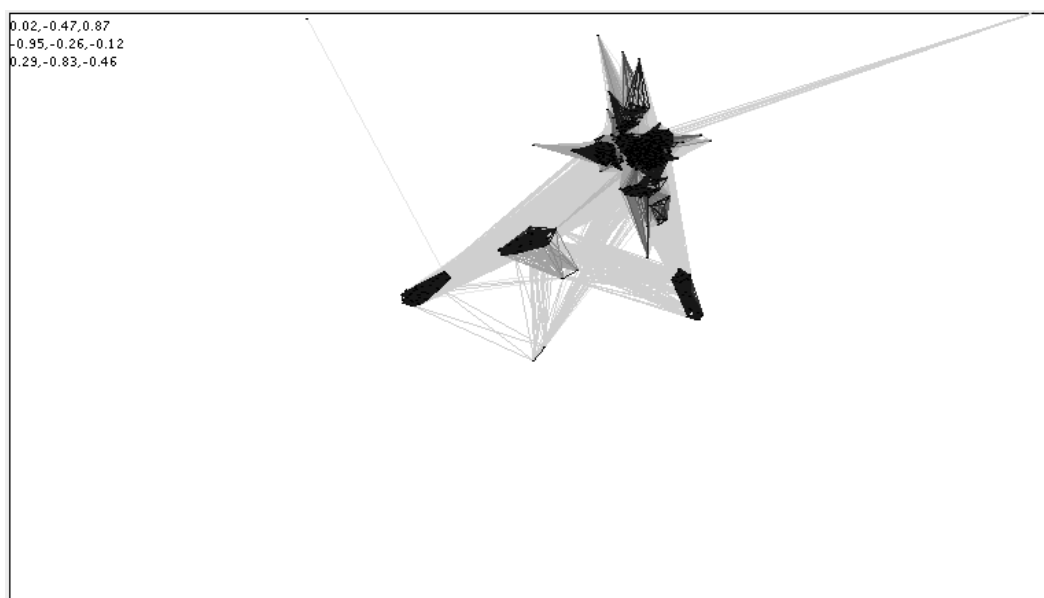
Išsiaiškinti kiek filogenetiškai artimos yra serino treonino kinazių sekos kiekvienoje grupėje buvo sudaryti filogenetiniai medžiai pagal CLANS gautas grupes. Atlikti sekų išlyginimą, reikalingą medžių sudarymui.

Sudarius visus 15 (pagal atsiskirtų grupių skaičių) filogenetinių medžių (5 priedas), itin aiškaus sekų atsiskyrimo dėl filogenetikos nepastebėta. Galima išskirti tik keletą organizmų, kurie šiek tiek, bet tik dalimi savo sekų, atsiskyrė panašumu, tai: *Aphanizomenon flos-aquae* ir *Anabaena* (*afanies81*, *anawa102*, *ana90* ir *anamdt14*). Pagal užšifruotus ciano bakterijų pavadinimus, tai žinomi kaip labai artimi organizmai.

Patikrinus visas turimas iš anotacijų serino treonino baltymų kinazių sekas su Pfam sekų palyginimo įrankiu gauti rezultatai, kuriuose galima pastebėti, kad dauguma sekų yra multidomeninės. Pašalinus sekas, kuriuose nebuvo atpažinti baltymų kinazi domenai analizuojami serino treonino baltymų kinazių aibė susiaurėjo iki 701 vietoj buvusių 811.

Naudojant trečiąją Perl programą iš ankščiau analizėje naudotų sekų iškerpami tik identifikuoti konservatyvūs baltymų kinazių domenai. Gautas sekų rinkinys naudojamas visai toliau aprašytai analizei kaip pagrindiniai duomenys.

Su turimomis tiksliau identifikuotomis baltymų kinazių sekomis vėl bandoma išsiaiškinti kaip jie gali susigrupuoti. Tam panaudojamas CLANS įrankis. Tokiomis pačiomis sąlygomis kaip ir su kinazių sekomis, gautomis iš anotacijų pirmuoju atveju, gaunama serino treonino baltymų kinazių domenų klasterių vizualizacija (5 pav.).



5 pav. Serino treonino kinazių sekose identifikuotų domenų grafinė vizualizacija naudojant CLANS

Šiuo atveju analizuojamas sekų rinkinys susigrupavo į 8 grupes, atitinkančias ankščiau vykdyto grupavimo reikalavimus. Taip pat pastebima, kad šiuo atveju į grupes (klasterius) patenka beveik visos sekos, o ne tik dalis kaip buvo gauta pradinės analizės metu. Tolimesnei analizei panaudota trečioji Perl programa ir gauta statistika, kiek kurios ciano bakterijos serino treonino baltymų kinazių sekų patenka į kiekvieną grupę (6 pav.).

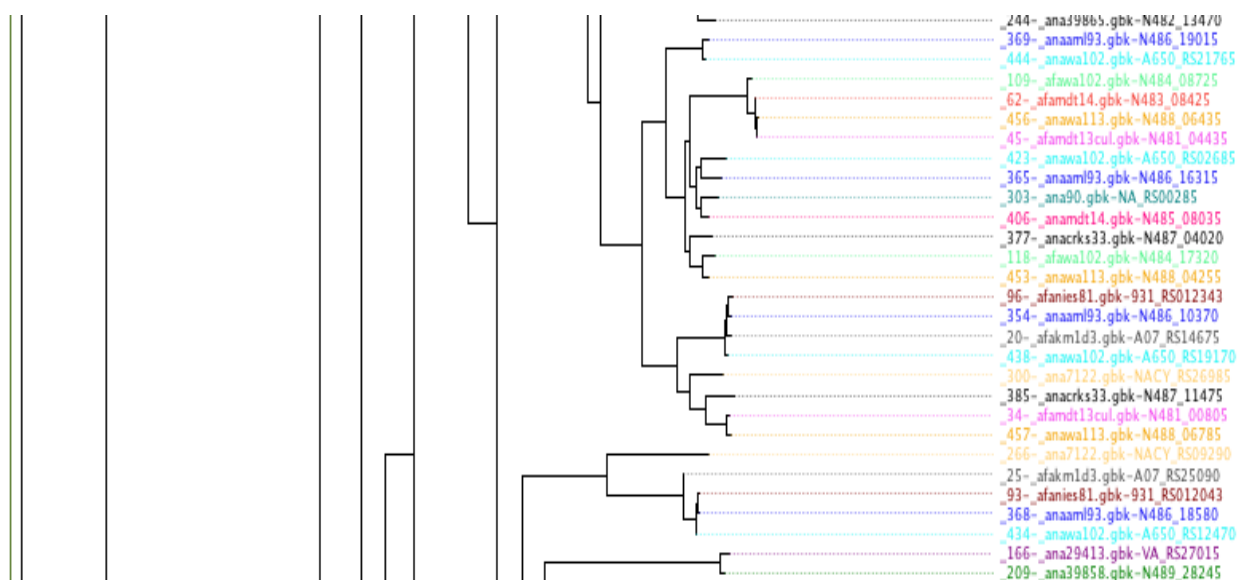
	A	B	C	D	E	F	G	H
afakm1d3.gbk	0	0	2	11	3	0	0	2
ana29413.gbk	11	0	2	23	3	1	1	3
ana7122.gbk	1	1	2	19	4	1	6	2
anawa102.gbk	1	2	2	18	1	0	1	2
dsp131c.gbk	0	0	0	8	0	0	0	0
nos7120.gbk	0	0	0	0	0	0	0	36
riche.gbk	0	0	1	6	0	0	0	1
afamdt13cul.gbk	0	1	0	13	4	1	0	2
ana39858.gbk	10	0	2	22	3	1	1	3
ana90.gbk	0	1	0	18	7	0	1	5
anawa113.gbk	0	1	0	14	2	1	0	2
dsp310f.gbk	0	15	0	0	0	0	0	0
nos73102.gbk	0	0	2	22	2	1	2	2
richehh01.gbk	0	0	0	4	0	0	0	1
afamdt14.gbk	0	1	0	12	2	1	0	2
ana39864.gbk	0	0	0	13	0	1	1	2
anaam193.gbk	1	2	2	16	1	0	1	2
cylind7417.gbk	1	0	2	13	2	0	3	0
nodul9414.gbk	1	0	0	16	2	2	5	1
nos7524.gbk	0	0	3	21	1	0	2	1
richehm01.gbk	0	0	0	1	0	0	0	1
afanies81.gbk	1	1	2	16	2	0	0	2
ana39865.gbk	0	1	0	13	1	1	1	3
anacrks33.gbk	0	0	0	14	3	0	1	2
cylindcs505.gbk	1	0	0	9	2	0	0	0
nos0708.gbk	1	0	1	10	0	0	0	1
raphd9.gbk	0	0	0	7	1	0	0	0
afawa102.gbk	0	1	0	13	1	1	0	2
ana7108.gbk	0	0	2	5	0	0	4	0
anamdt14.gbk	0	0	0	11	1	1	0	2
nos7107.gbk	4	0	0	24	1	1	2	3

6 pav. Serino treonino baltymų kinazių sekose identifikuotų domenų pasiskirstymas grupėse pagal skirtingas ciano bakterijas

Pastebima, kad D grupė ypatingai išsiskiria iš visumos, nes joje yra sekų beveik iš visų ciano bakterijų, bet ir sekų skaičius itin gausu (nuo 4 iki 24). Kiekvienos ciano bakterijos serino treonino baltymų kinazių sekos kurioje nors vienoje grupėje pasižymi atsiskiriančiu gausumu. Tai leidžia daryti interpretaciją, kad vienoje grupėje pagrindą sudaro vienos ar kelių ciano bakterijų baltymų kinazių sekos ir įsimašo vos kelios artimesnių organizmų sekos.

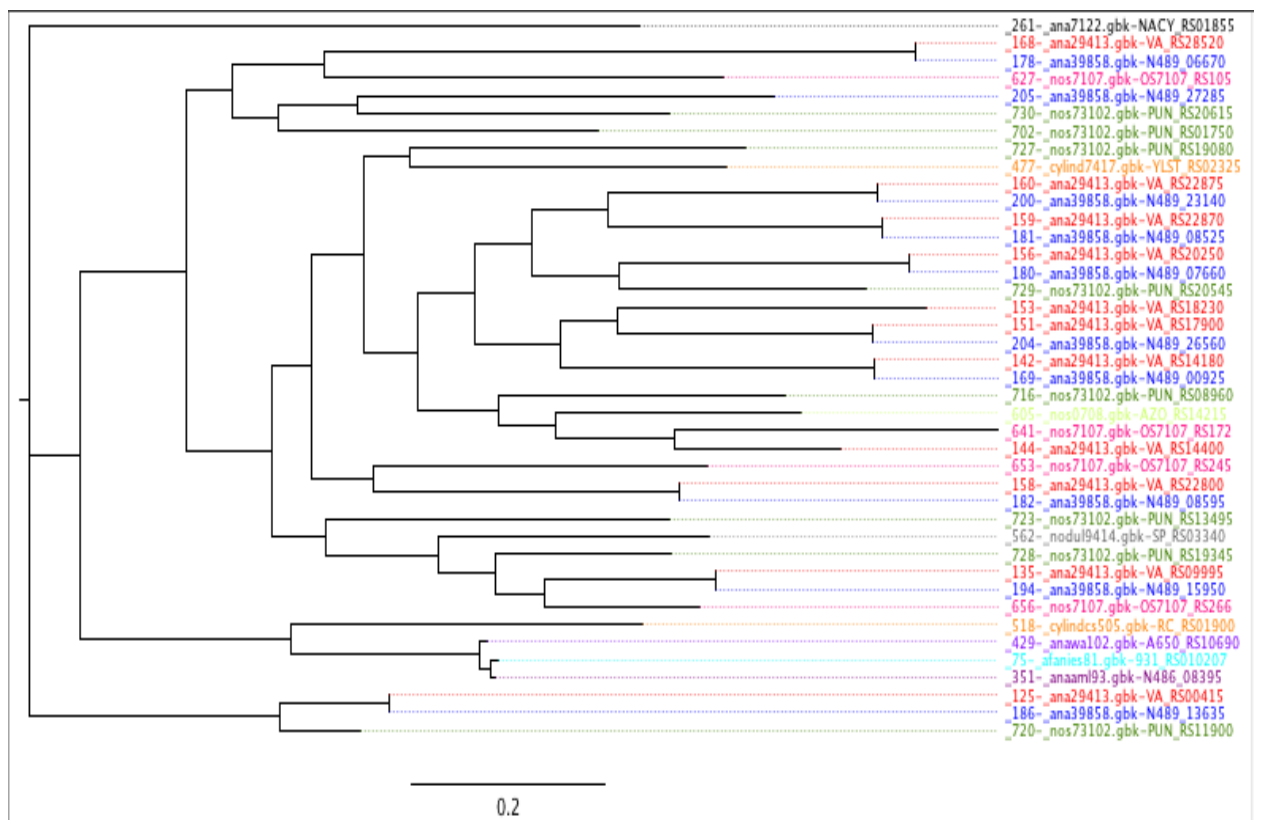
Patikrinti kiekvienos grupės serino treonino baltymų kinazių filogenetiškumą sudaryti filogenetiniai medžiai visoms aštuonioms grupėms naudojant CLANS įrankį (6 priedas). Pastebimi sekų susigrupavimai kiekvienoje grupėje.

Gausiausia D grupė parodo, kad toliausiai nuo medžio šaknies nutolusios šakos paskutiniame lygmenyje susigrupuoja beveik visada pagal tą pačią ciano bakteriją. Gretimose šakose išsidėsto giminiškiausių organizmų sekų rinkinių grupelės. Kai kuriose platesnėse šakose išsidėsto ne tik tos pačios ciano bakterijos, bet tos pačios genties organizmų serino treonino baltymų kinazių domenai, pavyzdžiui, iš *Anabaena sp. MDT14b* ir *Anabaena sp. wa102 strain WA102* (7 pav.). Šioje grupėje šakos, esančios netoli medžio šaknies, taip pat atskiria baltymų kinazių grupes, kurios evoliuciškai yra itin artimos.



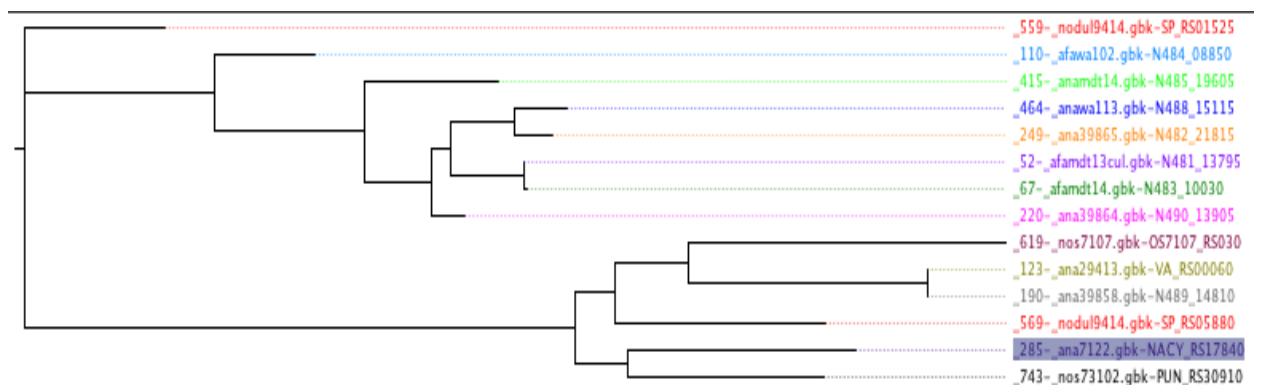
7 pav. D grupės baltymų kinazių domenų filogenetiškumą atvaizduojančio medžio (6 priedo fragmentas) fragmentas (skirtingomis spalvomis pažymėti skirtingoms ciano bakterijoms priklausančios sekos)

Ne tokioje didelėje A grupėje beveik visi baltymų kinazių domenų sekų rinkiniai išsidėsto poromis. Daugumą porų sudaro po serino treonino baltymų kinazę iš tos pačios genties, bet skirtingų ciano bakterijų, pavyzdžiui, *Anabaena variabilis* ATCC 29413 ir *Anabaena sp. 39858* yra viena dažniausiai pasikartojančių porų (8 pav.). Beveik visame filogenetiniame medyje genties ciano bakterijų susigrupavimas išsilaiko bent dvejuose medžio lygiuose. Šioje grupėje vienas ryškiausių ir tam pačiam organizmui būdingų sekų išsidėstymas (gretimose šakose) vientisomis grupėmis.



8 pav. A grupės baltymų kinazių domenų filogenetiškumą atvaizduojantis medis (6 priedo fragmentas) (skirtingomis spalvomis pažymėti skirtingoms ciano bakterijoms priklausančios sekos)

Tos pačios ciano bakterijos arba itin artimos serino treonino baltymų kinazės yra filogenetiškai artimos, tai matomas daugelyje grupių. Mažiau tikslus išsidėstymas pastebimas esant mažesnei grupei, pavyzdžiui, F, kuri sudaryta tik iš 15 sekų arba J, kurioje matoma tik viena itin artimų ciano bakterijų pora (9 pav.).



9 pav. F grupės baltymų kinazių domenų filogenetiškumą atvaizduojantis medis (6 priedo fragmentas) (skirtingomis spalvomis pažymėti skirtingoms ciano bakterijoms priklausančios sekos)

Įsitikinti, ar tikrai tiksliai atskirti serino treonino baltymų kinazių domenai taip gan tvarkingai išsidėsto filogenetiniuose medžiuose sudarytas bendras visų šių sekų filogenetinis medis (7 priedas). Šiuo atveju taip pat tų pačių ar labai artimų ciano bakterijų sekos išsidėsto artimose šakose ir patvirtina galimybę pagal tai identifikuoti kiek artimi organizmai.

Išvados

Tiriamų ciano bakterijų genomuose dėl gausos pasirinktos analizuoti serino treonino baltymų kinazės. Taikant pasirinktus metodus išsiaiškinta, kad šios kinazių grupės sekos, paimtos iš trisdešimties tiriamų ciano bakterijų genomų, gali dar klasterizuotis į mažesnes grupes. Atsirinkus reprezentatyvias grupes ir pagal jų sekų rinkinius sudarius filogenetinius medžius, gauti rezultatai parodė, kad serino treonino baltymų kinazių sekos filogenetiškai negali parodyti akivaizdžių galimybių kaip tiksliau suklasifikuoti tiriamus organizmus. Turimuose sekų rinkiniuose identifikavus ir tiksliau iškirpus serino treonino baltymų kinazių domenų bei pakartojus su jais prieš tai vykdytą analizę rezultatuose gauti nauji rezultatai. Šiuo atveju baltymų kinazių sekos išsidėstė atitinkamai pagal ciano bakterijų filogenetiškumą. Tuo remiantis galime teigti, kad atliekant ciano bakterijų serino treonino baltymų kinazių analizę galime patvirtinti šių organizmų taksonomiją. Turint tikslesnių duomenų apie tai, kokių organizmų genomus tiriam, galime įžvelgti, kad tų pačių ar labai artimų ciano bakterijų serino treonino baltymų kinazių sekos filogenetiniame medyje susigrupuoja atitinkamai artimai, jei sekų rinkinių imtis yra pakankamo dydžio pagal siekiamą tikslumą.

Santrauka

Rimantė Žalytė

Šiame bakalauro baigiamajame darbe buvo siekiama išanalizuoti kinazes ciano bakterijose ir nustatyti ar remiantis jų sekomis galima daryti prielaidas organizmų tikslesniam klasifikavimui. Siekiant šio tikslo buvo pasirinkta viena didžiausia serino treonino baltymų kinazių grupė, būdinga tiriamoms ciano bakterijoms. Šios kinazės su CLANS įrankiu suskaidytos į mažesnes, filogenetiškai artimesnes grupes ir atvaizduotos filogenetinis medžiais. Tai pakartota ir su patikslintais baltymų kinazių domenais. Pagal gautus rezultatus įvertinta galimybė, ar kinazių filogenetika gali atspindėti ciano bakterijų taksonomiją.

Raktiniai žodžiai: ciano bakterijos, kinazės, serino/treonino baltymų kinazės.

Summary

Rimantė Žalytė

The main purpose of this research is phylogenetic analyses of serine/threonine protein kinases in cyanobacteria and better cyanobacteria taxonomy. Analysis of those organisms genome's annotations was carried out using Perl programs, CLANS, Muscle, Iq-tree, and Pfam software. Results were formed like phylogenetic trees. It shows, that serine/threonine protein kinases can be used for cyanobacteria taxonomy classification determination.

Keywords: cyanobacteria, Pkinases, serine/threonine protein kinases.

Literatūros sąrašas

- [1] Joseph Schlessinger. 2000. Cell Signaling by Receptor Tyrosine Kinases.
- [2][https://en.wikipedia.org/wiki/Cyanobacteria#/media/File:Tolypothrix_\(Cyanobacteria\).JPG](https://en.wikipedia.org/wiki/Cyanobacteria#/media/File:Tolypothrix_(Cyanobacteria).JPG)
- [3] Rippka, Rosmarie, Josette Deruelles, John B. Waterbury, Michael Herdman & Roger Y. Stanier (1979). Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria.
- [4] Šulčius S., Alzbutas G., Kvederavičiūtė K., Koreivienė J., Zakrys L., Lubys A., Paškauskas R. 2015. Draft genome sequence of the cyanobacterium *Aphanizomenon flos-aquae* strain 2012/KM1/D3, isolated from the Curonian Lagoon (Baltic Sea).
- [5] Hao Wang, Kaarina Sivonen Email author, Leo Rouhiainen, David P Fewer, Christina Lyra, Anne Rantala-Ylinen, Johanna Vestola, Jouni Jokela, Kaisa Rantasärkkä, Zhijie Li and Bin Liu. 2012. Genome-derived insights into the biology of the hepatotoxic bloom-forming cyanobacterium *Anabaena* sp. strain 90.
- [6] Voss B, Bolhuis H, Fewer DP, Kopf M, Möke F, Haas F, El-Shehawey R, Hayes P, Bergman B, Sivonen K, Dittmann E, Scanlan DJ, Hagemann M, Stal LJ, Hess WR.. 2013. Insights into the physiology and ecology of the brackish-water-adapted Cyanobacterium *Nodularia spumigena* CCY9414 based on a genome-transcriptome analysis.
- [7] Maria J. Laamanen, Laura Forsström, Kaarina Sivonen. 2002. Diversity of *Aphanizomenon flos-aquae* (Cyanobacterium) Populations along a Baltic Sea Salinity Gradient
- [8] Hao Wang, Kaarina Sivonen, Leo Rouhiainen, David P Fewer, Christina Lyra, Anne Rantala-Ylinen, Johanna Vestola, Jouni Jokela, Kaisa Rantasärkkä, Zhijie Li, Bin Liu. 2012. Genome-derived insights into the biology of the hepatotoxic bloom-forming cyanobacterium *Anabaena* sp. strain 90
- [9] Degang Ning, Yan Jiang, Zhaoying Liu, Qinggang Xu. 2013. Characterization of a Chromosomal Type II Toxin–Antitoxin System *mazEaFain* in the Cyanobacterium *Anabaena* sp. PCC 7120
- [10] K. Sivonen, K. Kononen, W. W. Carmichael, A. M. Dahlem, K. L. Rinehart, J. Kiviranta, S. Niemela. 1989. Occurrence of the Hepatotoxic Cyanobacterium *Nodularia spumigena* in the Baltic Sea and Structure of the Toxin

- [11] Hanna Mazur-Marzec, Lidia Zeglin'ska & Marcin Plin'ski. 2005. The effect of salinity on the growth, toxin production, and morphology of *Nodularia spumigena* isolated from the Gulf of Gdańsk, southern Baltic Sea
- [12] Kim M. Wilson, Mark A. Schembri, Peter D. Baker, and Christopher P. Saint. 1999. Molecular Characterization of the Toxic Cyanobacterium *Cylindrospermopsis raciborskii* and Design of a Species-Specific PCR
- [13] Pawan K. Dadheech, Géza B. Selmeczy, Gábor Vasas, Judit Padisák, Wolfgang Arp, Kálmán Tapolczai, Peter Casper, Lothar Krienitz. 2014. Presence of Potential Toxin-Producing Cyanobacteria in an Oligo-Mesotrophic Lake in Baltic Lake District, Germany: An Ecological, Genetic and Toxicological Survey
- [14] Muriel Gugger, Christina Lyra, Irmgard Suominen, Irina Tsitko, Jean-Francois Humbert, Mirja S. Salkinoja-Salonen, Kaarina Sivonen. 2002. Cellular fatty acids as chemotaxonomic markers of the genera *Anabaena*, *Aphanizomenon*, *Microcystis*, *Nostoc* and *Planktothrix* (cyanobacteria).
- [15] Katia Soto-Liebea, Marco A. Méndezb, Loreto Fuenzalidac, Bernd Krockd, Allan Cembellad, Mónica Vásquez. 2012. PSP toxin release from the cyanobacterium *Raphidiopsis brookii* D9 (Nostocales) can be induced by sodium and potassium ions.
- [16] Zeev EB1, Yogev T, Man-Aharonovich D, Kress N, Herut B, Béjà O, Berman-Frank I. 2008. Seasonal dynamics of the endosymbiotic, nitrogen-fixing cyanobacterium *Richelia intracellularis* in the eastern Mediterranean Sea.
- [17] Cheng-Cai Zhang. 1996. Bacterial signalling involving eukaryotic-type protein kinases.
- [18] Hanks SK, Hunter T. 1995. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification.
- [19] Nidhi Tyagi, Krishanpal Anamika, Narayanaswamy Srinivasan. 2010. A Framework for Classification of Prokaryotic Protein Kinases
- [20] Nathalie Josso, Nathalie di Clemene. 1977. Serine/threonine kinase receptors and ligands.
- [21] Tony Hunter. 1995. Protein kinases and phosphatases: The Yin and Yang of protein phosphorylation and signaling

[22] <https://www.thermofisher.com/lt/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/phosphorylation.html>

[23] Herbert Y. Lin, Xiao-Fan Wang, Elinor Ng-Eaton, Robert A. Weinberg, Harvey F. Lodish. 1992. Expression cloning of the TGF- β type II receptor, a functional transmembrane serine/threonine kinase

[24] Xiaowen Zhang, Fangqing Zhao, Xiangyu Guan, Yu Yang, Chengwei Liang, Song Qin. 2007. Genome-wide survey of putative Serine/Threonine protein kinases in cyanobacteria.

[25] Cheng – Cai Zhang. 1996. Bacterial signalling involving eukaryotic-type protein kinases.

[26] Robert C. Edgar. 2004. MUSCLE: multiple sequences alignment with high accuracy and high throughput.

[27] Lam-Tung Nguyer, Heiko A. Schmidt, Arndt von Heaseler, Bui Quang Minh. 2014. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies.

[28] Erik L. L. Sonnhammer, Sean R. Eddy, Ewan Birney, Alex Bateman and Richard Durbin. 1998. Pfam: Multiple sequence alignments and HMM-profiles of protein domains.

Priedai

1. Priedas

Perl programa skirta skirtingo tipo kinazių dažniams skaičiuoti.

```
#!/usr/bin/perl
use warnings;
my $file="dazniai";
open( FILE, ">", $file ) or
    die "Can't open output file '$file' -- " . lcfirst($!);
my $i=0;
my $line;
my @name;
my @rate;
while (<>){
    $line=$_;
    chomp $line;
    if ($line=~ /product/){
        $line =~ s/\ /product=|d|\"|> //g;
        my $type = substr $line, 21;
        my $l = $#name+1;
        print "$type $l\n";
        my $j=0;
        my $k=0;
        if ($l == 0){
            $name[0] = $type;
            $rate[0] = 1;
        } else {
            while ($j < $l){
                if ($name[$j] eq $type){
                    $rate[$j] = $rate[$j] + 1;
                    $j++;
                    $k = 1;
                } elsif ($type ne $name[$j]){
                    $j++;
                }
            }
            if ($k == 0 && ($j == $l)){
                $name[$l] = $type;
                $rate[$l] = 1;
            } else {
                $k = 0;
            }
        }
    }
}

my $n = @name;
```

```

print "$n\n";
while ($i < $n){
    print FILE "$name[$i]  $rate[$i]\n";
    $i++;
}

```

```
close (FILE);
```

2. Priedas

Perl programa, skirta pasirinktos kinazių grupės sekų išrinkimui.

```

#!/usr/bin/perl
use warnings;

my $file="sekos";
open( FILE, ">", $file ) or
    die "Can't open output file '$file' -- " . lcfirst($!);
my $i=1;
my $line;
my $seq;
my $locus;
unshift(@ARGV, '-') unless @ARGV;
while ($ARGV = shift) {
    open(ARGV, $ARGV);
    while (<ARGV){
        $line=$_;
        chomp $line;
        if ($line=~\/locus_tag/){
            $locus=substr $line, 34, 11;
            $locus =~ s/\s//g;
        }
        if ($line=~\/product/ && $line=~\/serine\/threonine protein kinase/){
            print FILE "> \\\$i-\\$ARGV-\\$locus\\n";
            #print FILE "> \\\$i \\$ARGV  \\$line \\$locus\\n";
            do {$line=<>;
                chomp $line;
            }until ($line=~\/translation=/);
            chomp $line;
            $line =~ s/\/translation=|\\s|\\\"//g;
            do{
                $seq=$seq.$line;
                $line=<>;
                chomp $line;
                $line=~s/\s//g;
            } until ($line=~\/gene/ || $line=~\/CONTIG/ || $line=~\/BASECOUNT/ ||
$line=~\/repeat_region/ || $line=~\/assembly_gap/);
            print FILE "$seq\\n\\n";
            $i++;
            $seq="";

```


4. Priedas

Perl programa skirta serino treonino baltymų kinazių domenam iškirpti iš mišrių sekų.

```
#!/usr/bin/perl
```

```
use warnings;
```

```
open( FILE2, "<", "kinazes" ) or  
die "Can't open input file 'kinazes.txt' -- " . lcfirst($!);
```

```
my $file = "domenai.txt";  
open( FILE3, ">", $file ) or  
die "Can't open output file '$file' -- " . lcfirst($!);
```

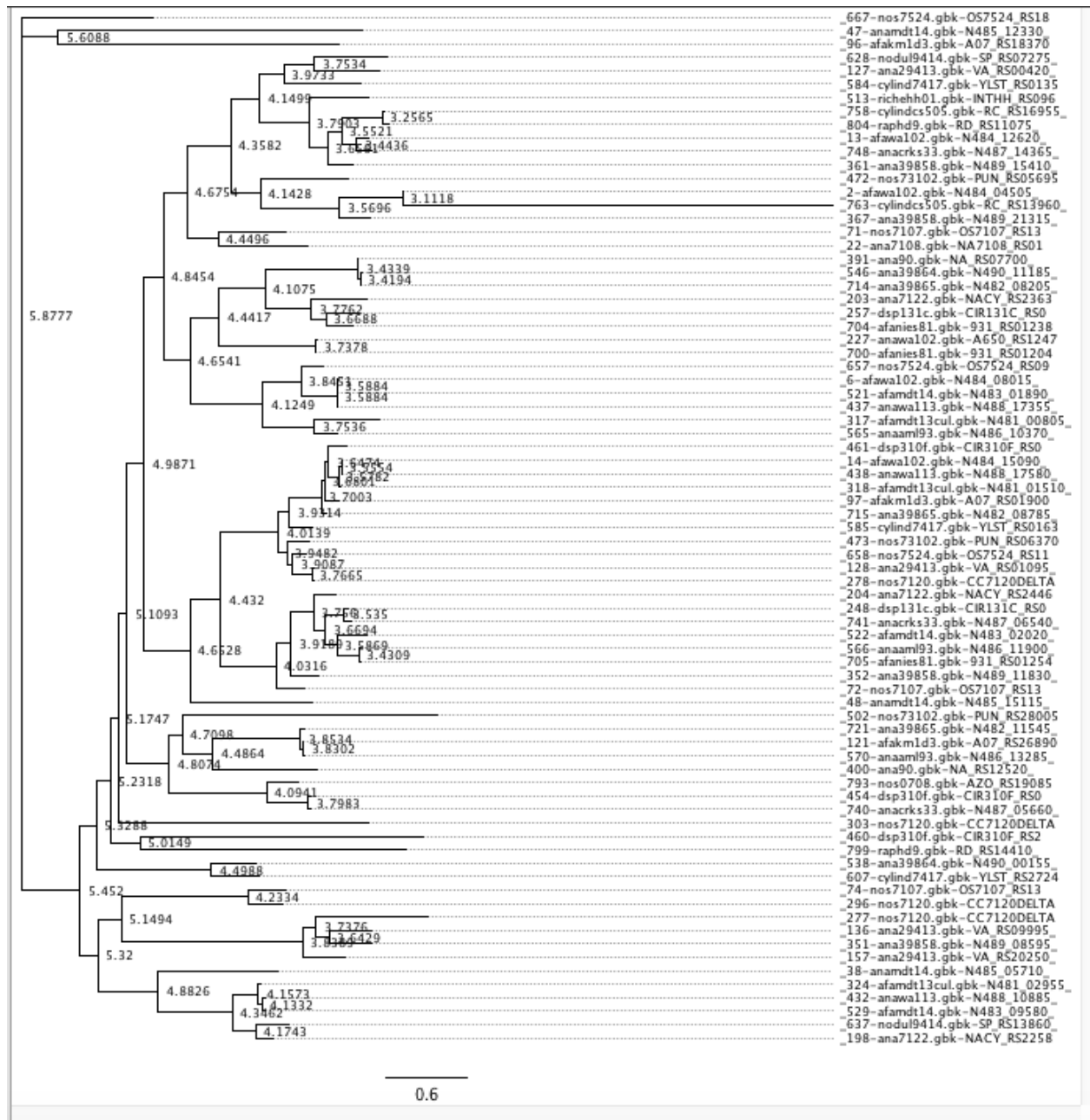
```
while ( my $line = <FILE2> ){  
    chomp $line;  
    my $name = substr $line, 0, 35;  
    my $start = substr $line, 37, 3;  
    my $end = substr $line, 44, 3;  
    my $seq = "";  
    my $find = 0;  
    $name =~ s/\s//g;  
    $start =~ s/\s//g;  
    $end =~ s/\s//g;  
  
    open( FILE1, "<", "duom.txt" ) or  
    die "Can't open input file 'bambam.txt' -- " . lcfirst($!);  
    while (my $line2 = <FILE1>){  
        if ($find == 1){  
            chomp $line2;  
            my $le = length $line2;  
            $seq = substr $line2, $start, $le;  
            print "$name>>>$start>>>$end|||||$le\n";  
            print FILE3 ">$name\n$seq\n";  
            $find = 0;  
            goto STEP;  
        }  
        if ($line2 =~ /$name/){  
            $find = 1;  
        }  
    }  
    }  
STEP:  
}
```

```
close (FILE1);  
close (FILE2);  
close (FILE3);
```

5. Priedas

Serino treonino baltymų kinazių grupių filogenetiniai medžiai. Sekos medžiuose įvardijamos pagal indekso numerį, ciano bakterijos genomo failo pavadinimą ir identifikatorių pagal sekos locus tag.

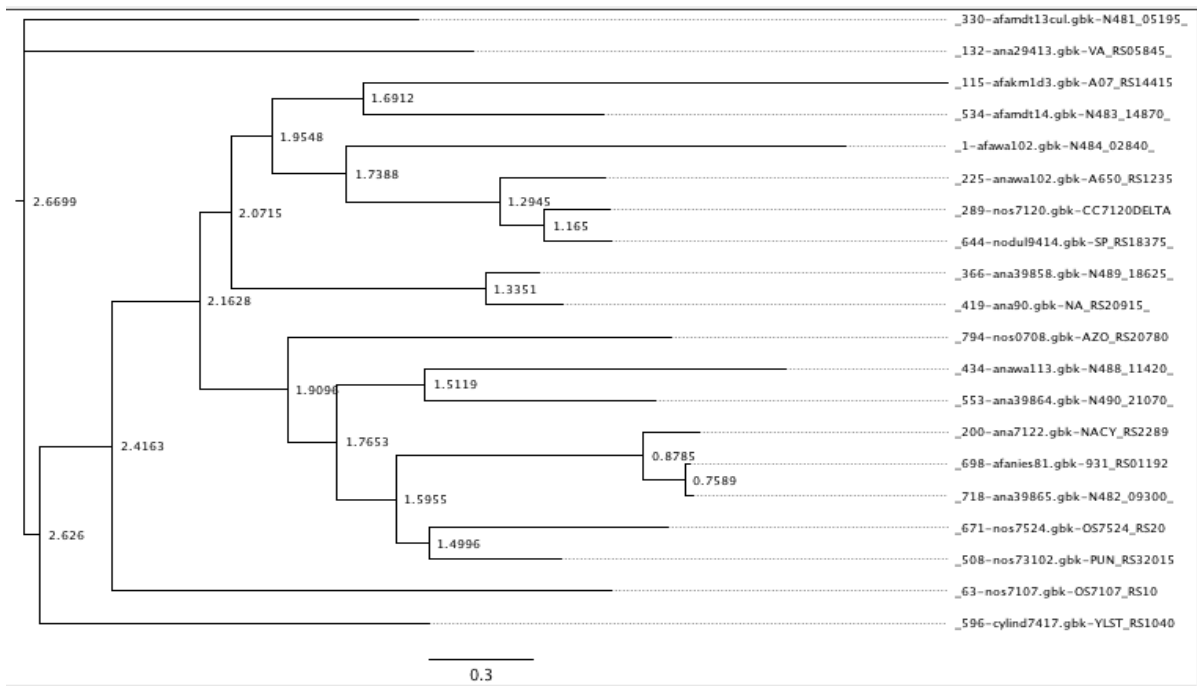
Grupė A



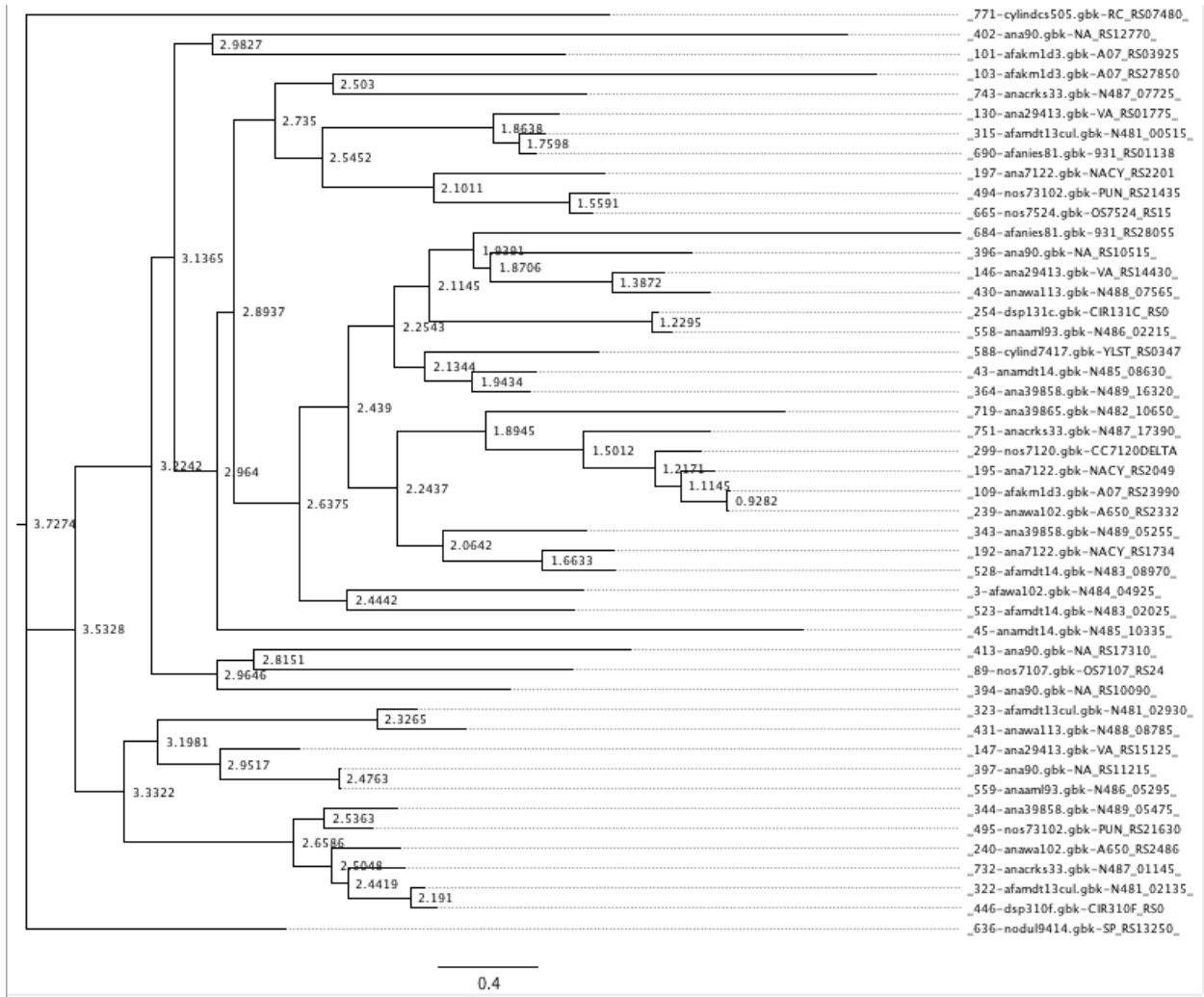
[illegible]

[illegible]

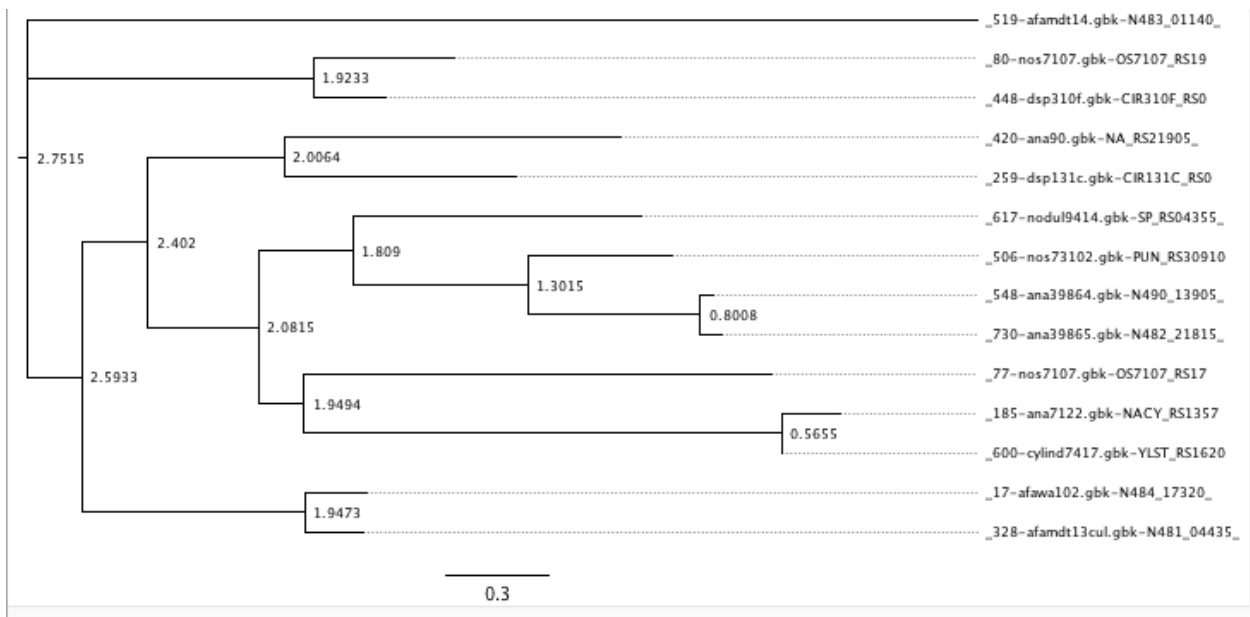
Grupè D



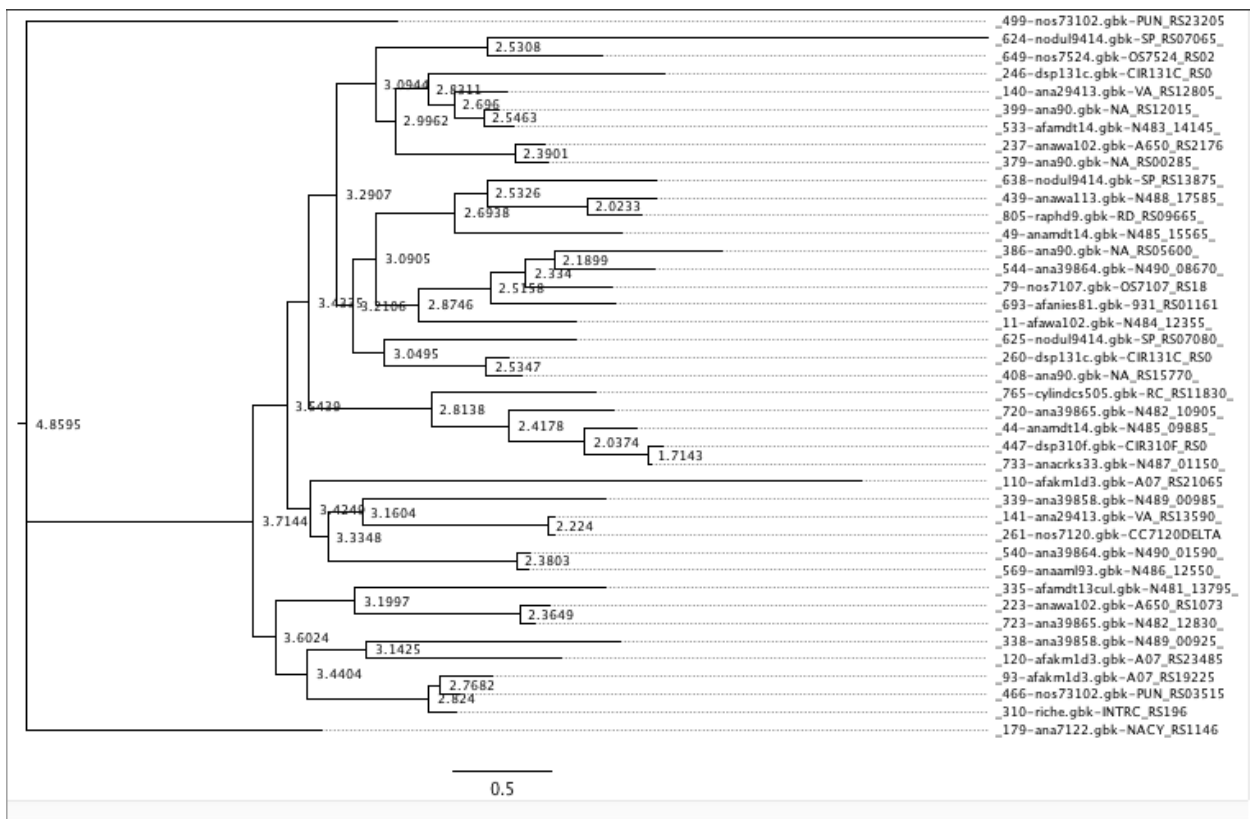
Grupè E



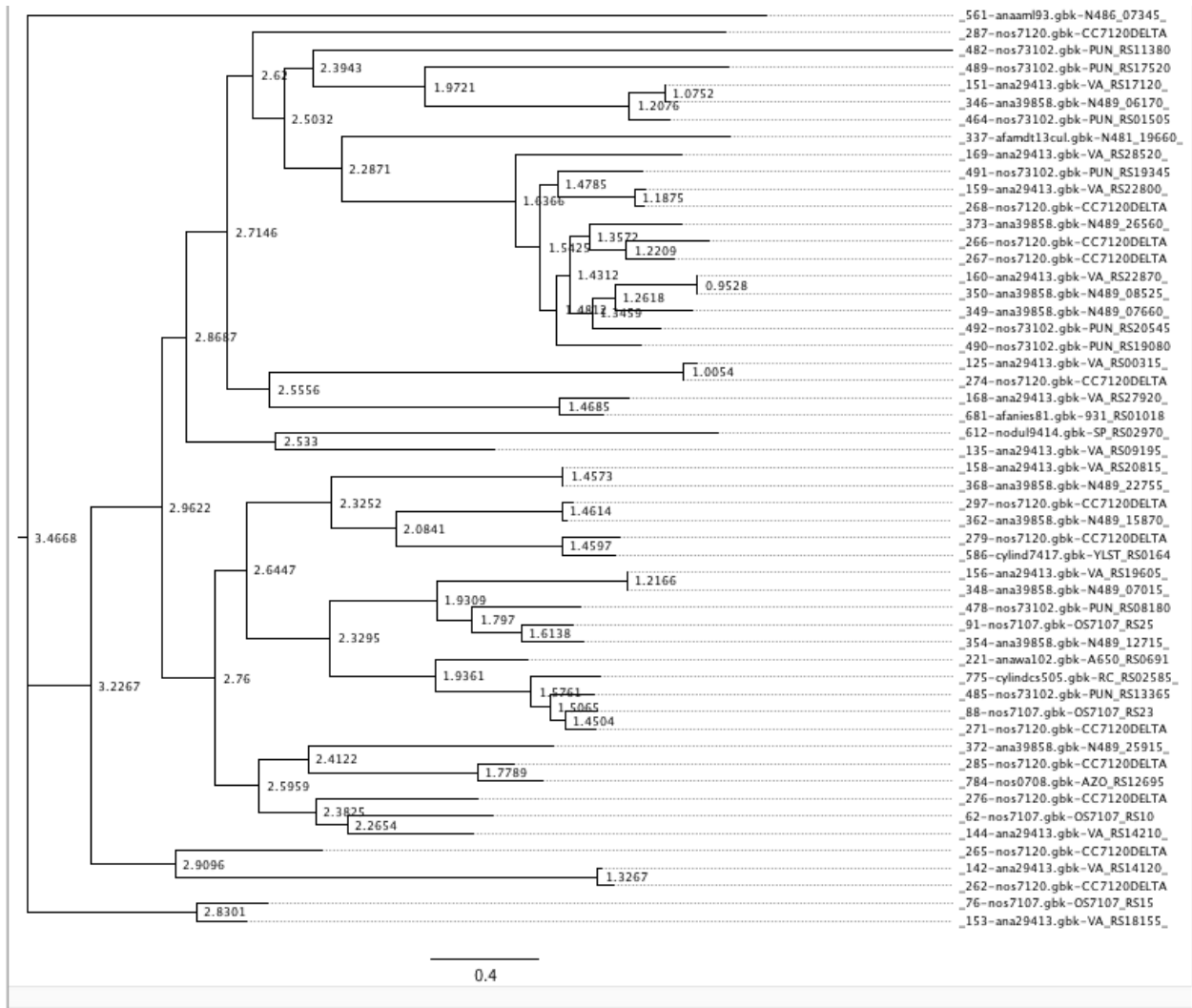
Grupè F



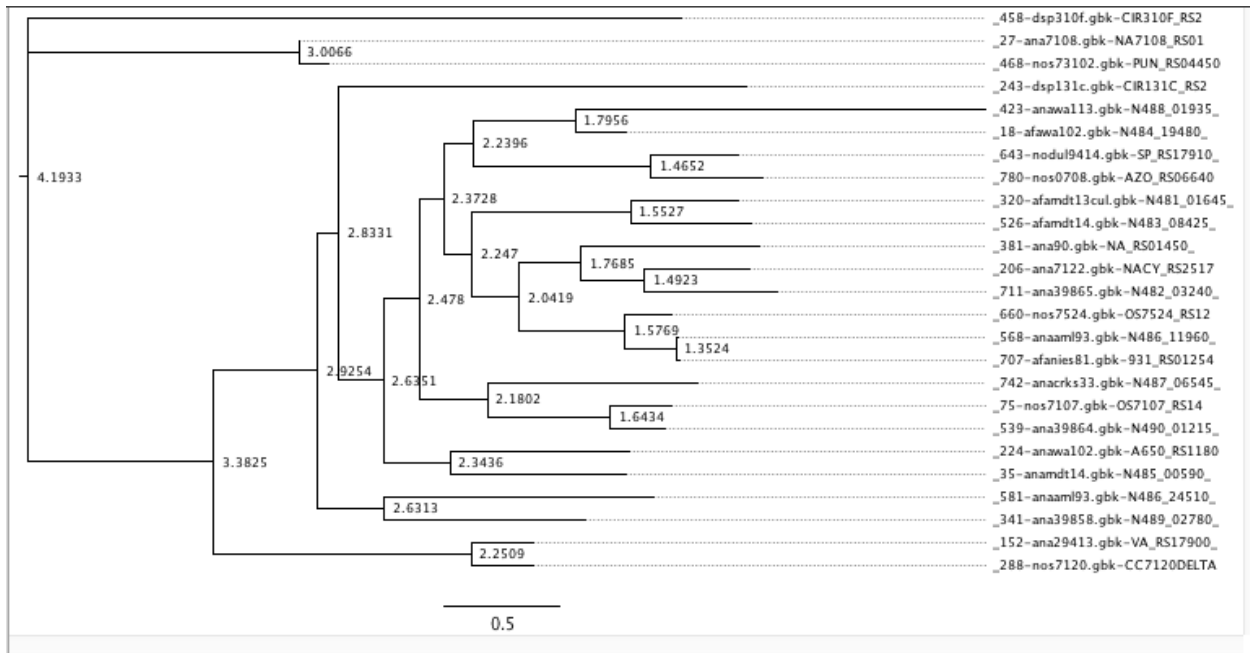
Grupè G



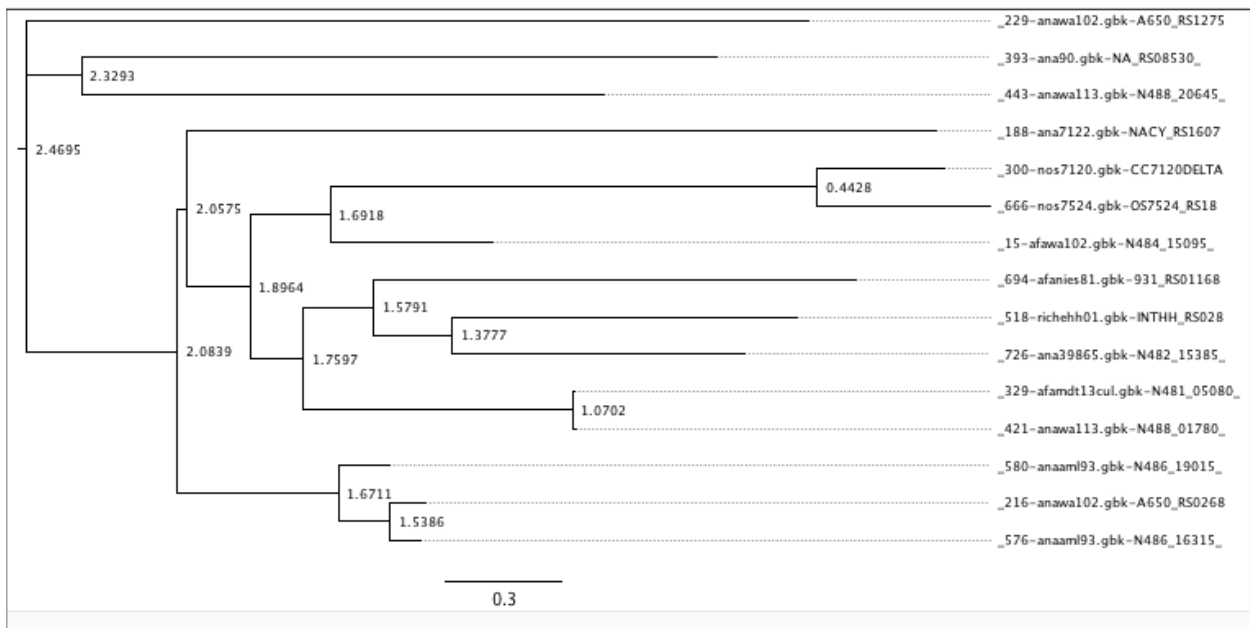
Grupè H



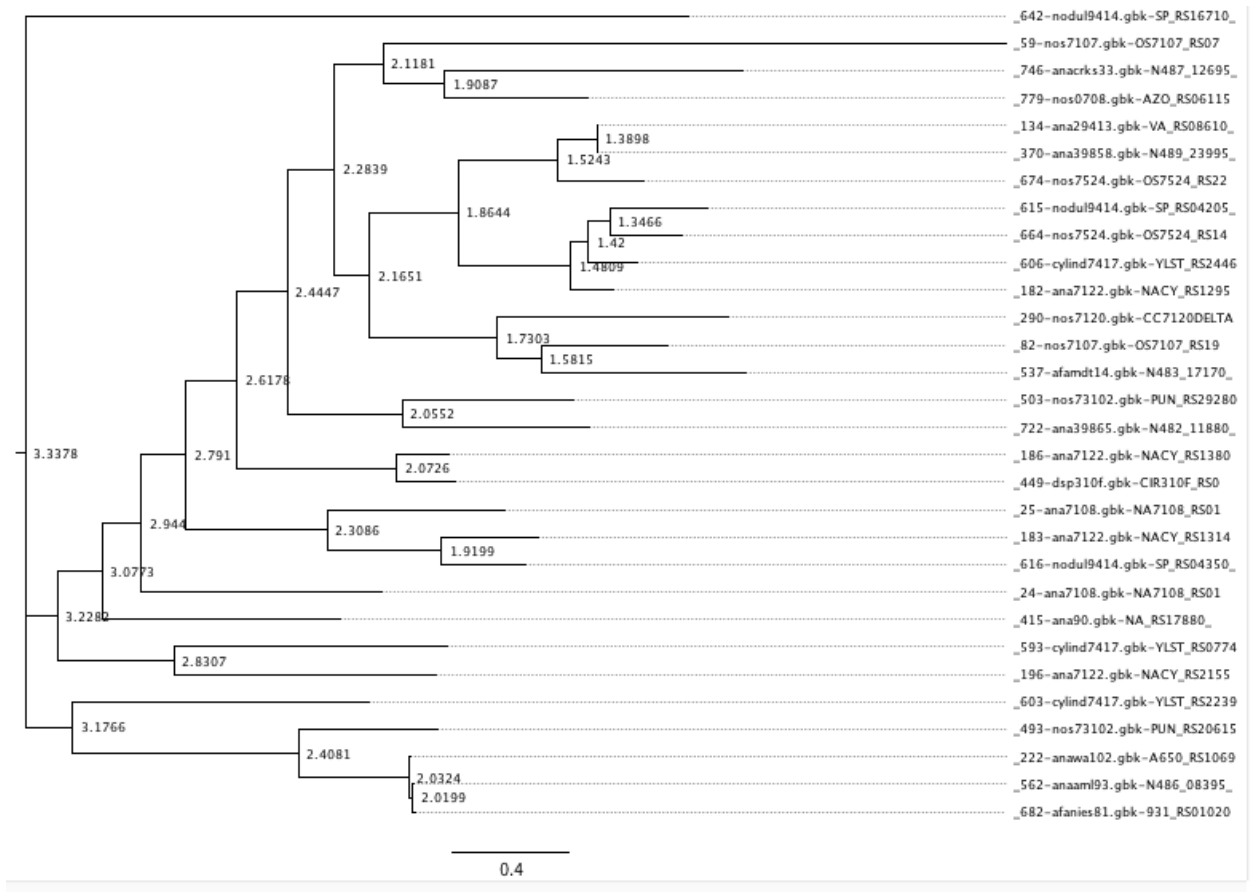
Grupè I



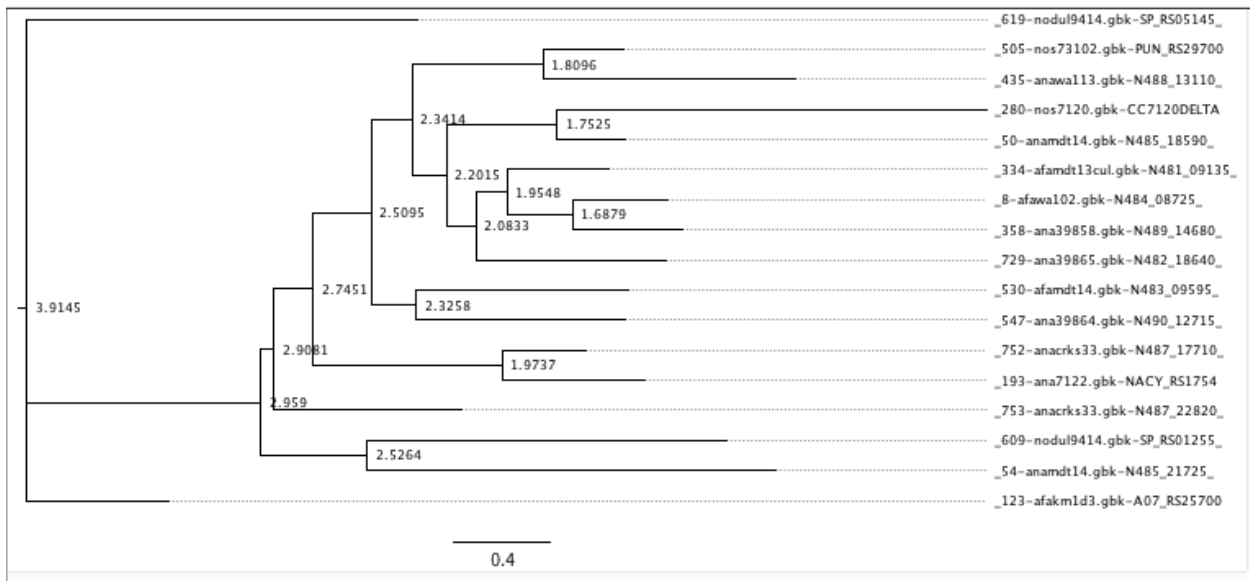
Grupè J



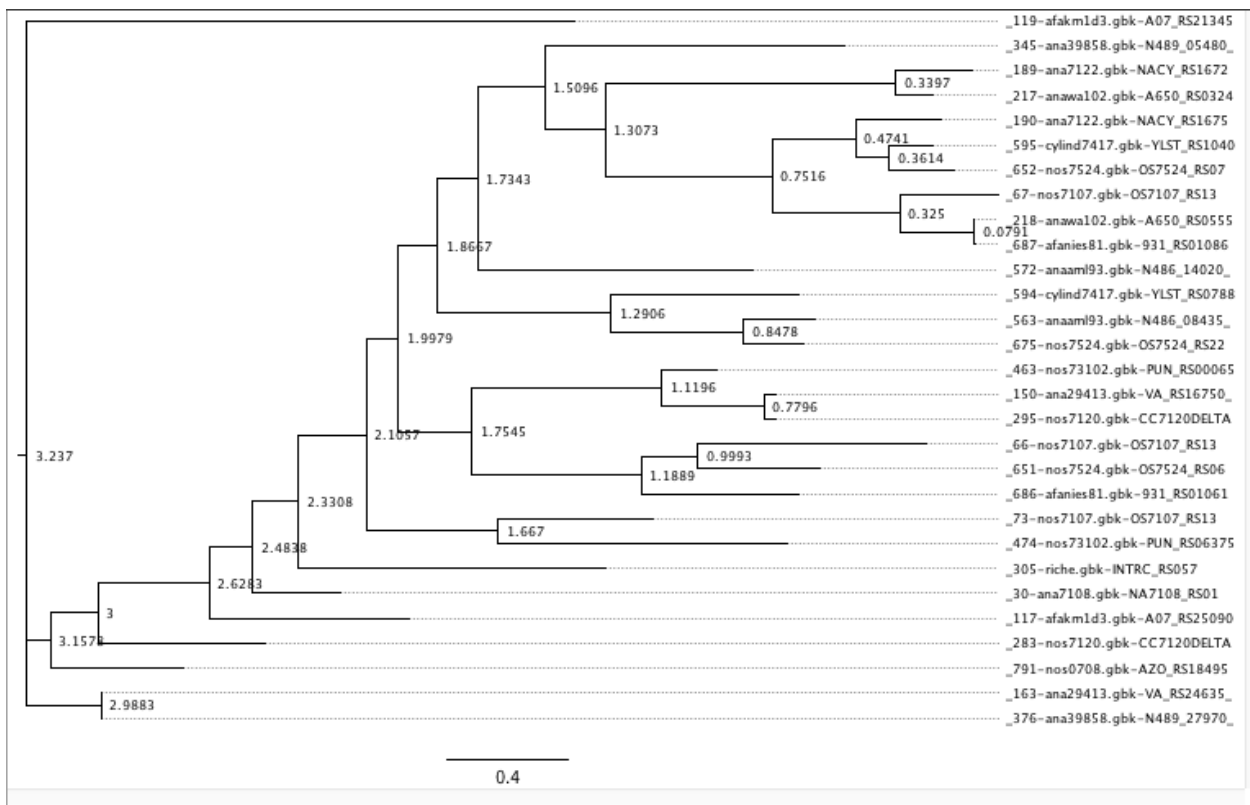
Grupè K



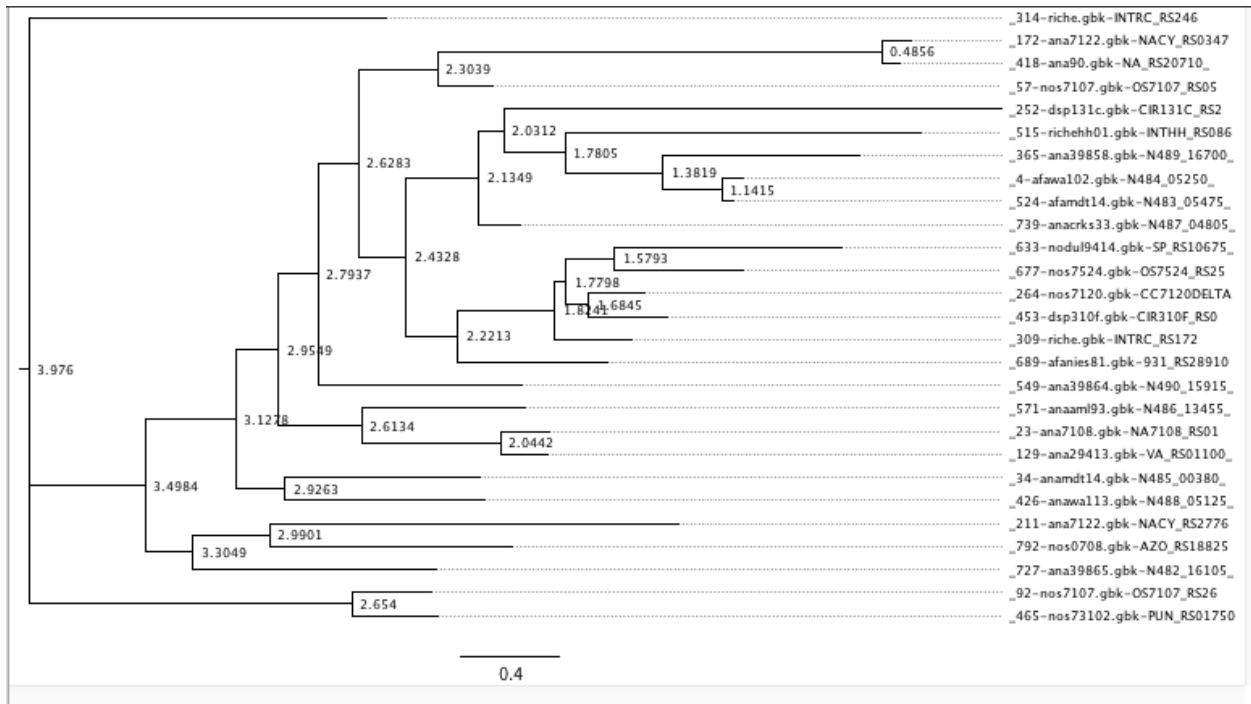
Grupè L



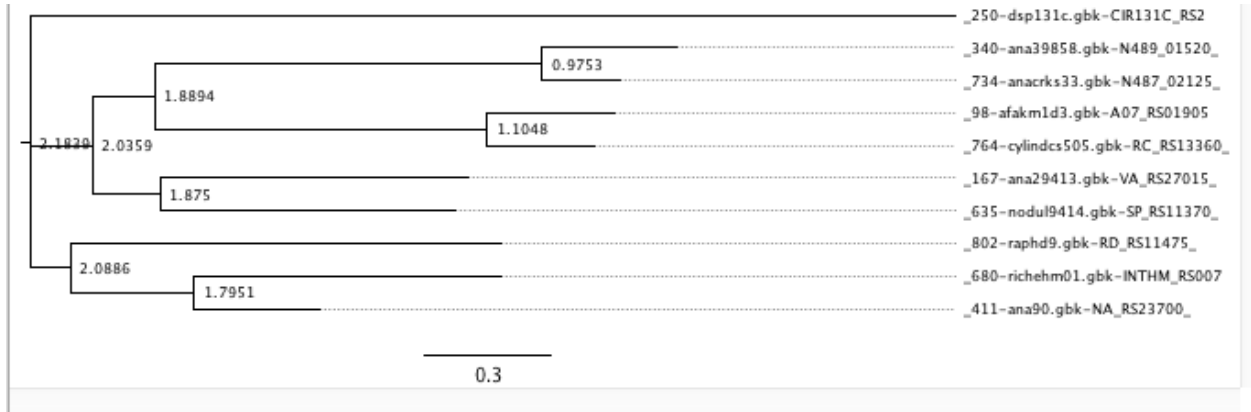
Grupè M



Grupè N



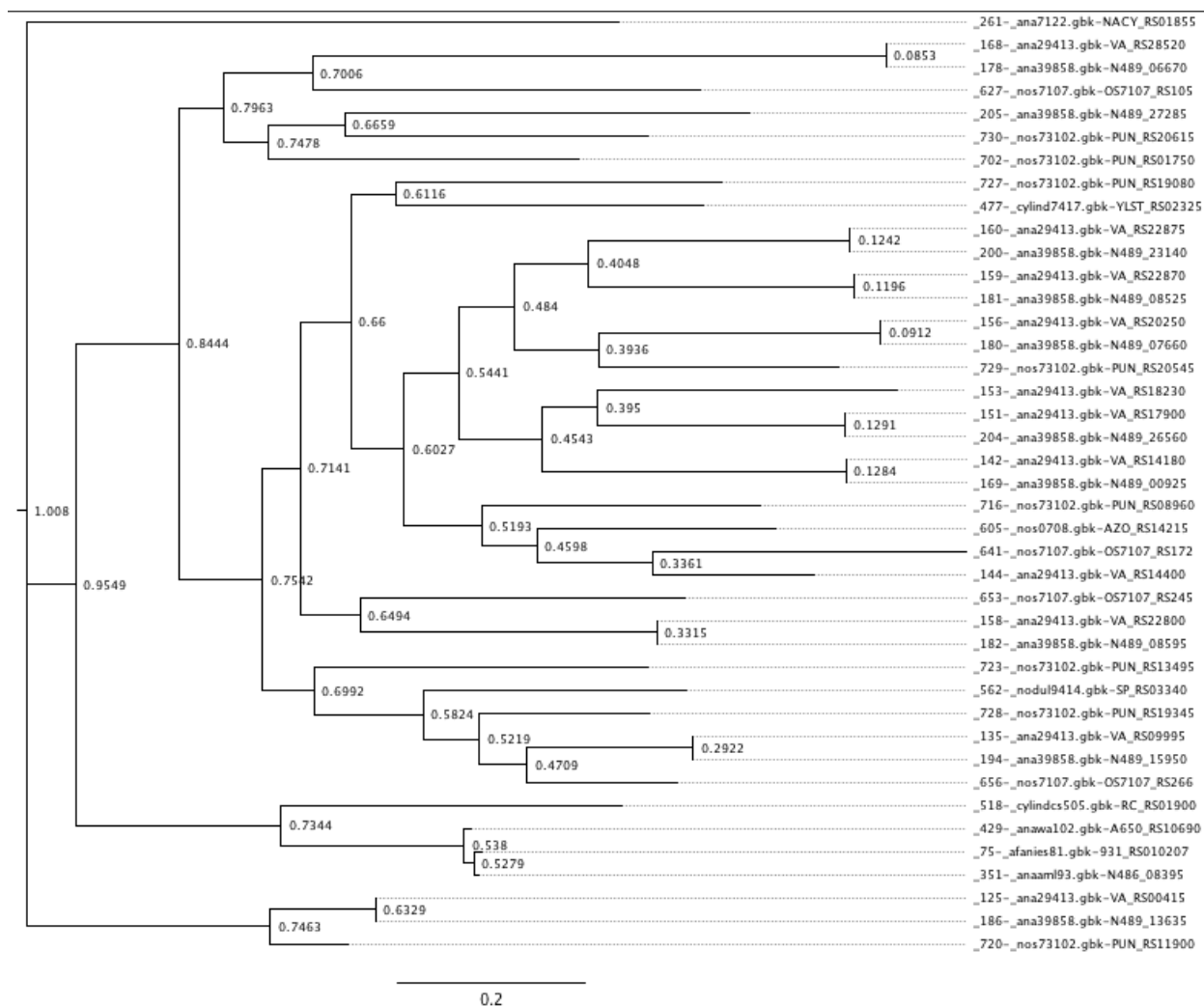
Grupè O



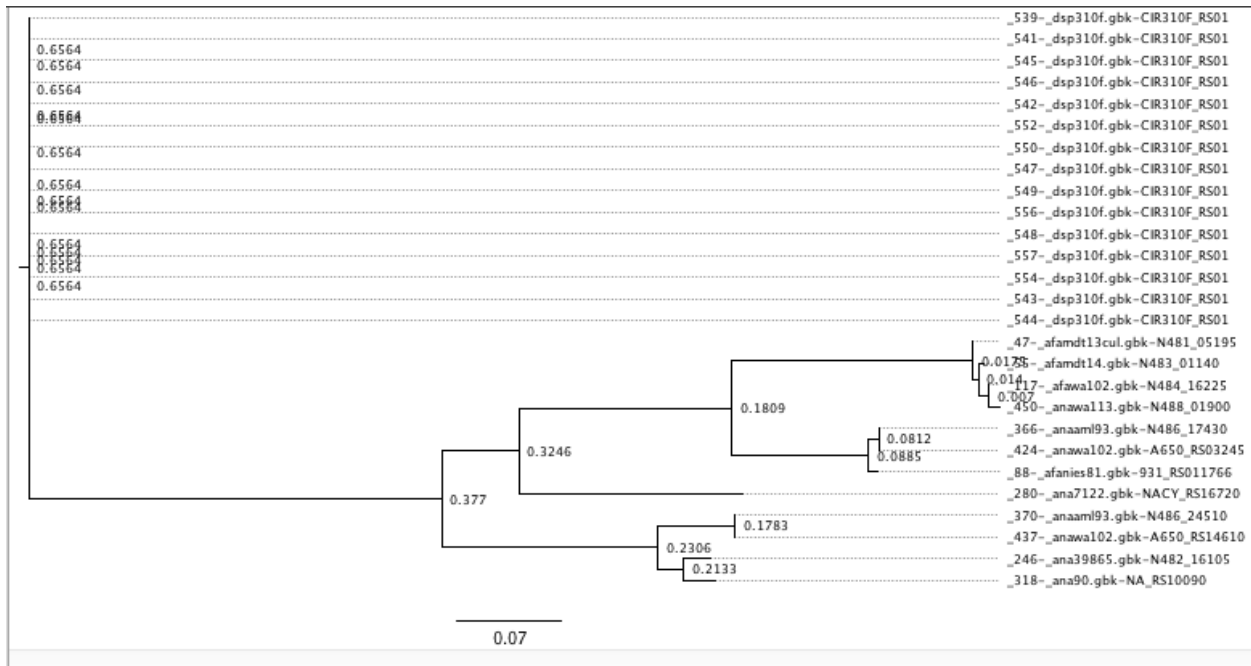
6. Priedas

Serino treonino baltymų kinazių identifikuotų domenų grupių filogenetiniai medžiai. Sekos medžiuose įvardijamos pagal indekso numerį, ciano bakterijos genomo failo pavadinimą ir identifikatorių pagal sekos locus tag.

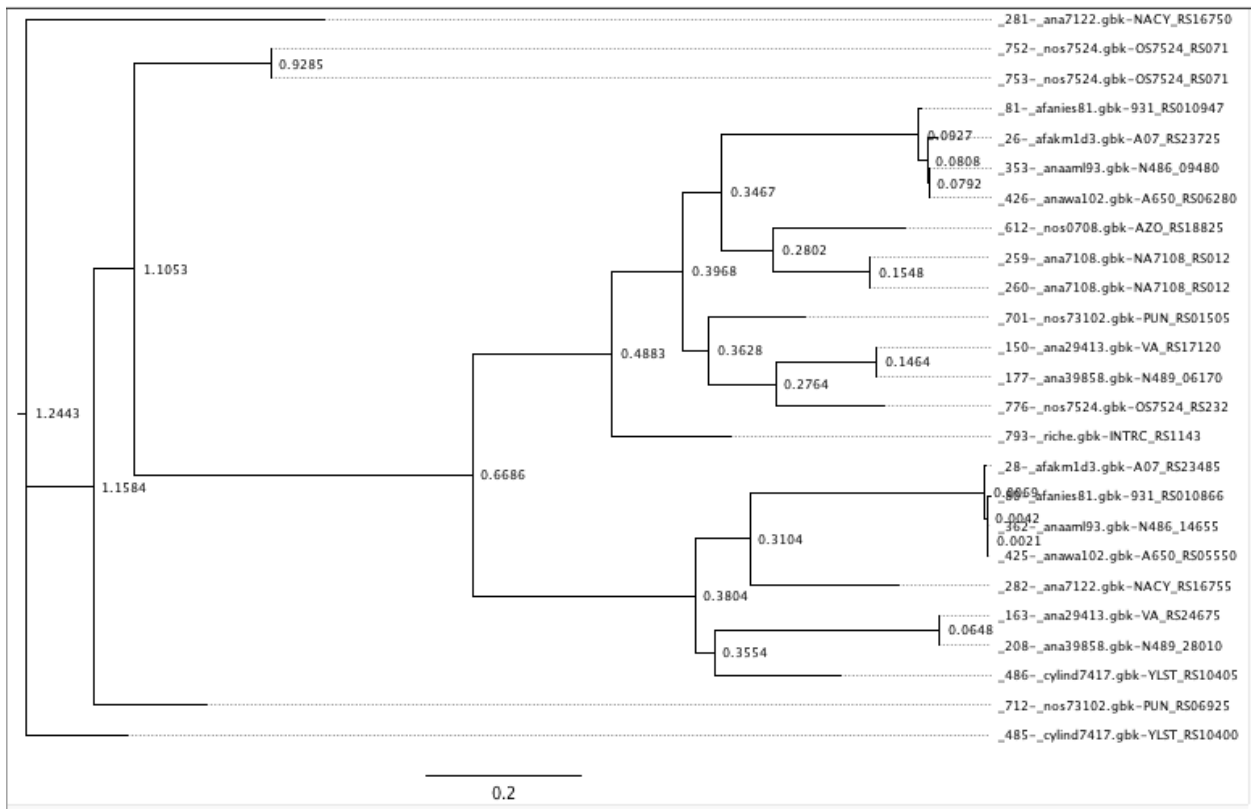
Grupė A



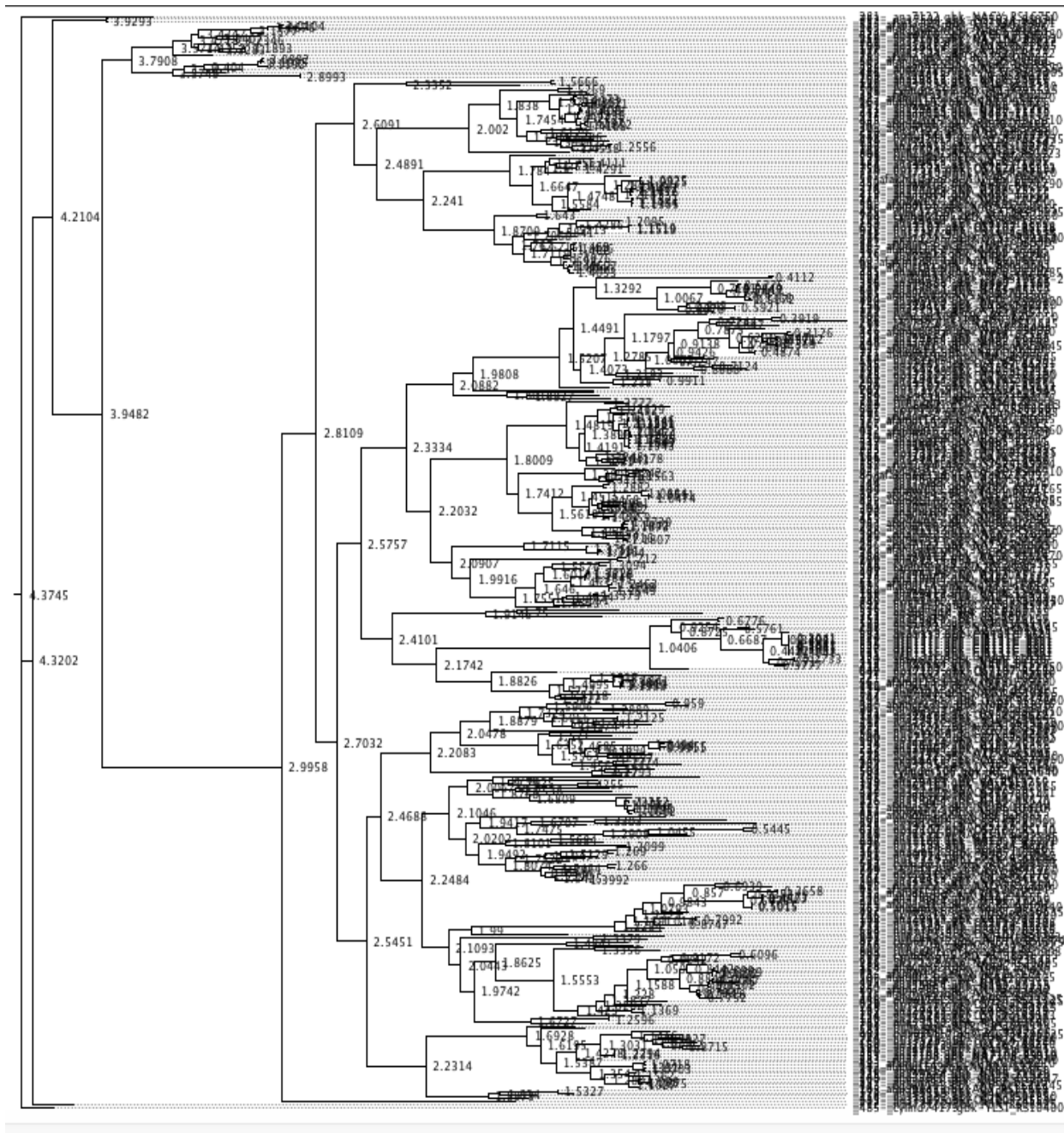
Grupè B



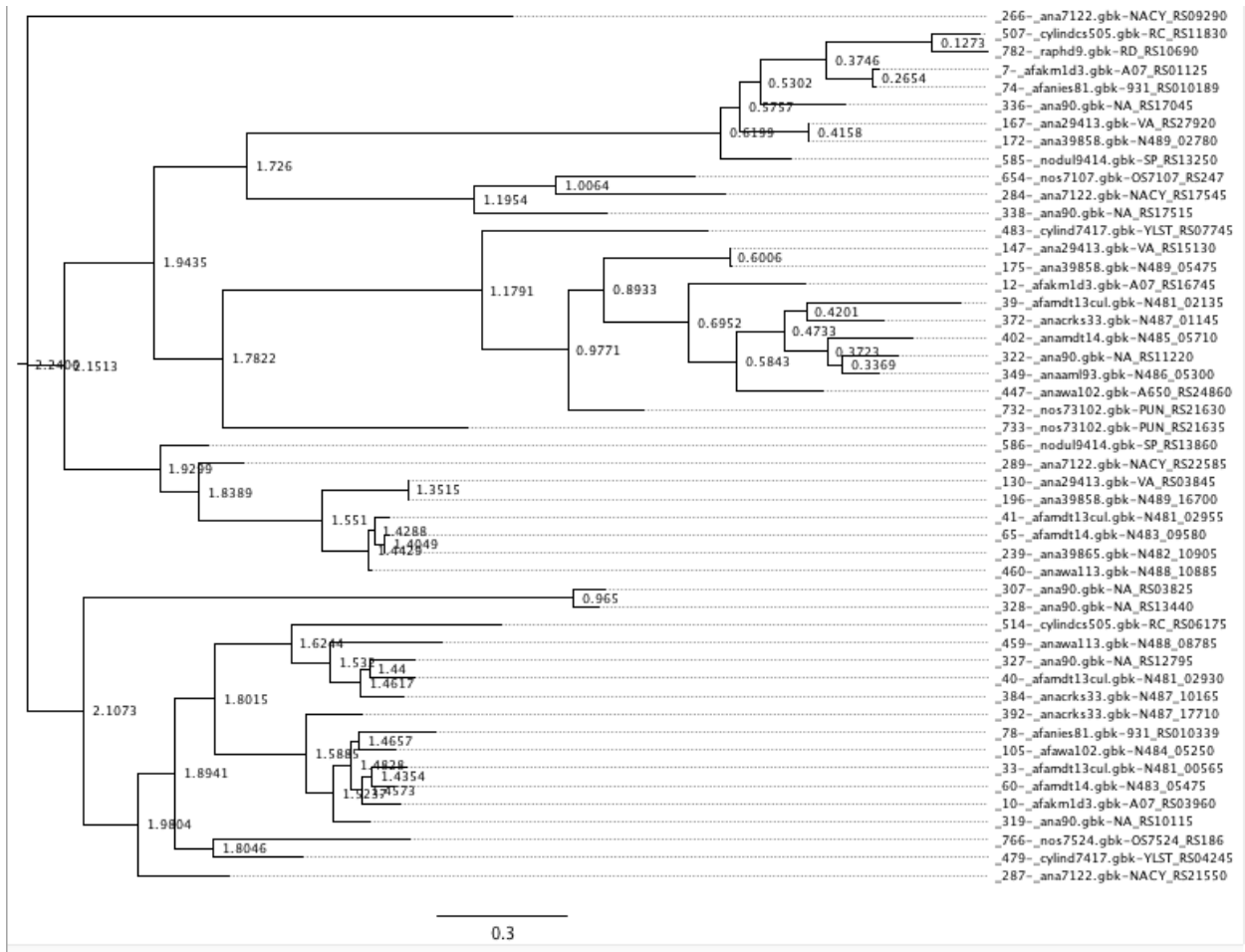
Grupè C



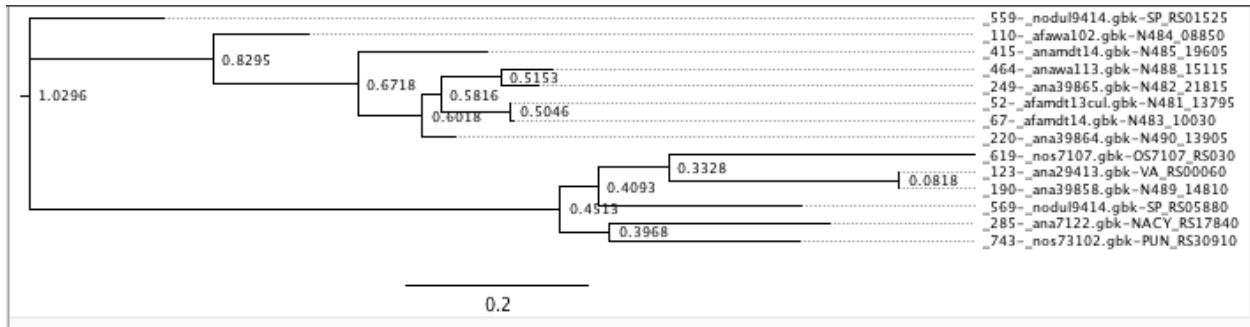
Grupè D



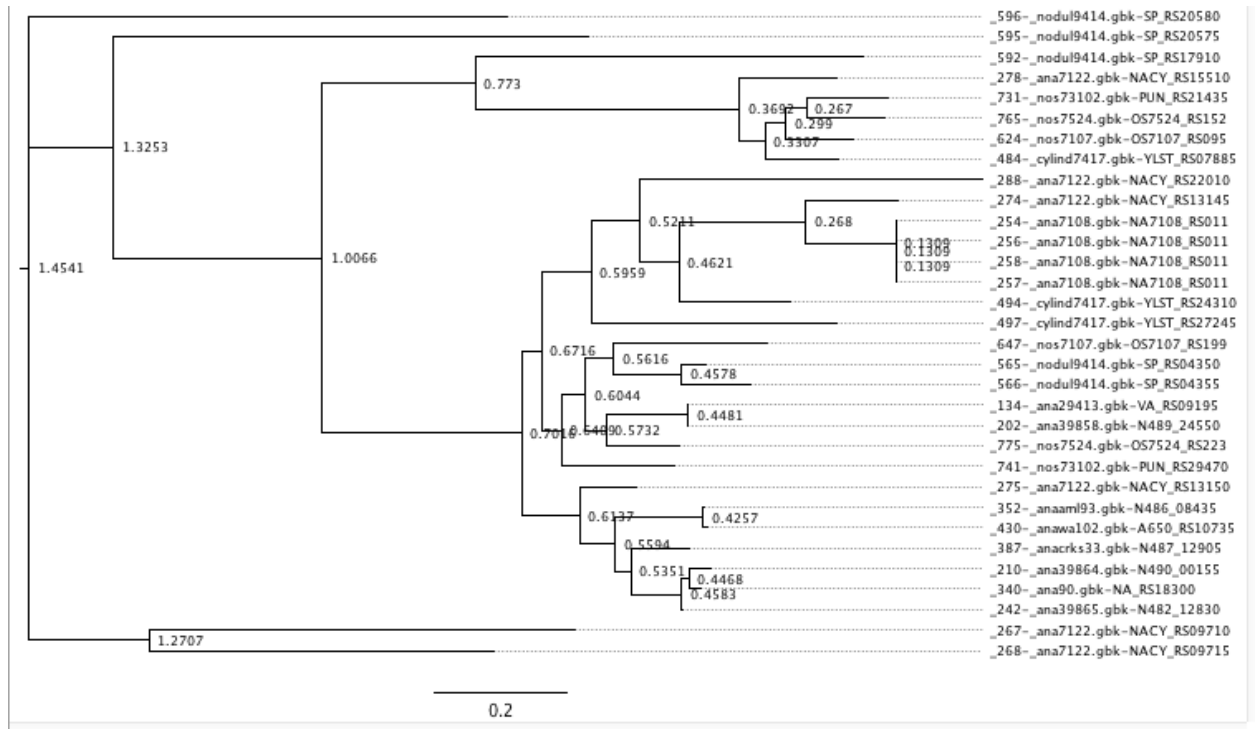
Grupè E



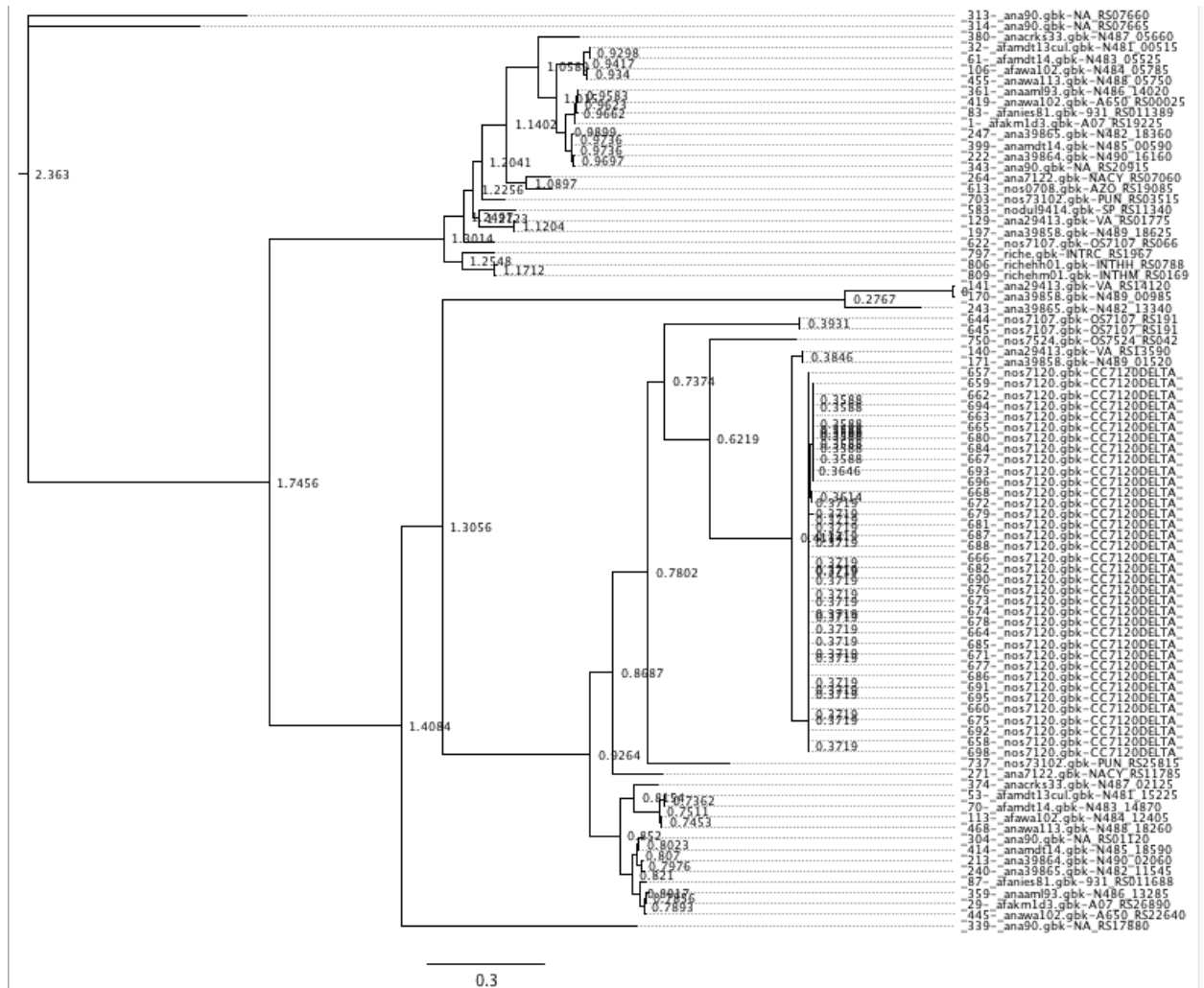
Grupè F



Grupè G



Grupè H



7. Priedas

Serino treonino baltymų kinazių identifikuotų visų domenų bendrąs filogenetiniai medžiai. Sekos medžiuose įvardijamos pagal indekso numerį, ciano bakterijos genomo failo pavadinimą ir identifikatorių pagal sekos locus tag.

