

VILNIAUS UNIVERSITETAS

Gyvybės mokslų centras



Molekulinės biologijos magistrantūros studijų programos II kurso studentas

Tomas VENCLOVAS

V tipo CRISPR-Cas sistemų integracijos tyrimas

Darbo vadovas:

Dr. Giedrius Sasnauskas

Vilnius, 2022 m.

Turinys

Įvadas	5
1. Literatūros apžvalga	6
1.1 Kova su virusais	6
1.1 CRISPR-Cas sistemos	7
1.2 Įvairovė ir kilmė	9
1.4 Adaptacija	11
1.4.1 Cas1-Cas2 mechanizmas	11
1.4.2 Cas1-Cas2 struktūra	12
1.4.3 Funkcijų įvairovė	14
2. Medžiagos ir Metodai	17
2.1 Medžiagos	17
2.2 Metodai	19
2.2.1 DNR amplifikavimas (PGR)	19
2.2.2 DNR elektroforezė	20
2.2.3 Plazmidžių transformacija	20
2.2.4 Plazmidžių išskyrimas	21
2.2.5 Baltymo raiška	21
2.2.6 Baltymo gryninimas	21
2.2.7 SDS-PAGE denatūruojanti baltymų elektroforezė	22
2.2.8 Lizatų ruošimas	22
2.2.9 Oligodupleksų (preskirtukų) paruošimas	23
2.2.10 <i>In vivo</i> integracijos reakcija	23
2.2.10 <i>In vitro</i> Integracijos reakcija	24
4. Rezultatai	25
4.1 genų sintezė ir amplifikavimas	25

4.2 Baltymų ekspresija ir gryninimas	26
4.3 <i>In vivo</i> integracija.....	29
4.4 <i>In vitro</i> integracija.....	30
4.5 Modelio generavimas ir analizė	33
5. Išvados	37
Santrauka.....	38
Spacer acquisition in a type V CRISPR-Cas system	39
Summary	39
PADĖKA	40
6. Literatūros sąrašas.....	40

Santrumpos

APS – amonio persulfatas;

C1 – Cas1;

C14 – Cas14;

C2 – Cas2;

C4 – Cas4;

Cas – angl. *CRISPR-associated (proteins)* – CRISPR asociuoti (baltymai);

CRISPR – angl. *Clustered regular interspaced palindromic repeats* – taisyklingai pertraukti trumpi susitelkę palindrominiai pasikartojimai;

crRNR – CRISPR RNR;

EDTA – etilendiamintetraacto rūgštis;

IHF – angl. *Integration host factor* – šeimininko integracijos faktorius;

IPTG – Isopropyl β- d-1-thiogalactopyranoside

LB – angl. *Lysogeny broth* – auginimo terpė;

NTP – nukleotidų trifosfatai;

PAM – angl. *Protospacer-adjacent motif* – preskirtuko gretutinis motyvas;

PGR – polimerazinė grandininė reakcija

REazė – Restrikcijos fermentas;

R-M – Restrikcijos–modifikacijos sistemos;

SDS, NDS – natrio dodecilsulfatas;

TAE – angl. *Tris-acetate-EDTA* – tirpalas, sudarytas iš Tris bazės, acto rūgšties ir EDTA;

TEMED – tetrametiletilendiaminas;

Tris – tris(hidroksimetil) aminopentanas

Ivadas

CRISPR-Cas sistemos yra adaptyvi imuninė sistema randama prokariotuose. Jos sudarytos iš unikalių baltymų ir CRISPR lokuso – pseudopalindrominių pasikartojimų, perskirtų unikaliais skirtukais. Adaptacija – pirmoji CRISPR-Cas imuninės sistemos veikimo stadija, kurios metu nauja genetinė medžiaga yra inkorporuojama į organizmo CRISPR lokusą, kuri vėliau naudojama gynybai prieš virusus. Nauji DNR fragmentai arba skirtukai (angl. *spacer*), yra įterpiami į genomą Cas1 ir Cas2 baltymų, į CRISPR lokuso pradžia, tarp lyderinės sekos ir pirmojo pasikartojimo.

Paskutiniu metu didelio susidomėjimo sulaukė V tipo CRISPR-Cas sistemos, kadangi jose yra randami mažieji Cas12f ir panašūs efektoriniai baltymai, Cas9 baltymo analogai, kurie gali būti panaudojami genų redagavimui. Šios sistemos buvo bioinformatiškai identifikuotos metagenomo sekų duomenų bankuose. Atitinkamai, šiuo metu mažai žinoma apie likusią šitų sistemų veiklą, įskaitant adaptaciją. Šiuo darbo metu pasirinkta tirti V-F potipio B17 sistema, kurioje randamas efektorinis baltymas Cas14B17; adaptacijos baltymai Cas1, Cas2 ir Cas4, bei CRISPR lokusas. Nors visais atvejais Cas1 ir Cas2 baltymai yra reikalingi naujų šablonų integracijai į genomą, dažnai taisyklingai adaptacijai būtini papildomi baltymai, kadangi kiekvienoje CRISPR sistemoje adaptacijos baltymai gali būti unikaliai išsivystę tas pačias funkcijas atlikti

Darbo tikslas: Ištirti V tipo CRISPR-Cas sistemos adaptacijos (arba naujų skirtukų įsistatymo) mechanizmą

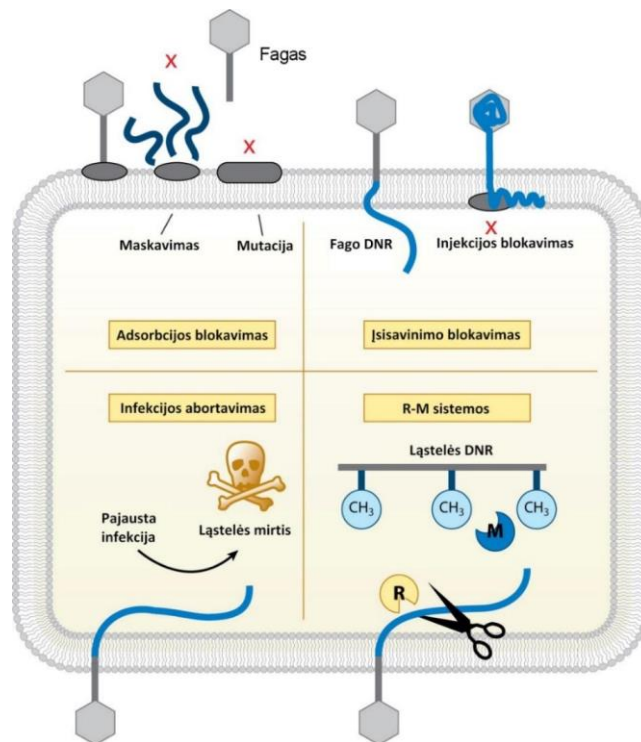
Darbo uždaviniai:

1. Klonuoti pasirinktos V-F potipio CRISPR-Cas sistemos baltymų genus į *E. coli* raiškos vektorius ir išgryninti adaptacijos baltymus;
2. Ištirti integracijos aktyvumą *in vivo* ir *in vitro*;
3. Sumodeliuoti ir išanalizuoti integracijos komplekso struktūrą.

1. Literatūros apžvalga

1.1 Kova su virusais

Gausiausia ir įvairiausia organizmų grupė Žemėje yra prokariotai – bakterijos ir archėjos. Tačiau šie vienaląsčiai organizmai yra taikinyi dar didesne įvairove ir gausybe pasižyminčių bakteriofagų arba trumpiau - fagų. Tikslus kiekis šių virusų ir jų tipų nėra žinomas, bet viename litre jūros vandens galime rasti apie 10 milijardų fagų ir spekuliuojama, kad kiekvienam prokariotui tenka virš 10 fagų (Bergh et al. 1989; Rohwer 2003; Suttle 2005; Weinbauer 2004). Bakteriofagai dauginasi užkrėsdami auką savo genetinė medžiaga, iš kurios, išnaudojant šeimininko molekulinę mašineriją, sukuriama daugybė viruso kopijų. Galutinis šio proceso rezultatas - ląstelės mirtis, o su ja į aplinką pasklinda fagų kopijos. Natūraliai iškyla klausimas – jei bakteriofagų yra daugiau, nei jų šeimininkų, ir su kiekvienu užkrėtimu virusų kopijų daugėja, tai kodėl bakterijos ir archėjos vis dar nėra išnaikintos? Prokariotai atsilaiko prieš fagų antplūdį, nes, kaip ir už juos sudėtingesni organizmai, turi imunitetą, kuris apsaugo ląstelę nuo infekcijos. Jų imuniteto pagrindas yra įvairūs mechanizmai ir sistemos, kurie evoliucionavo kartu su bakteriofagais, prieš kuriuos kovoja (Labrie, Samson, and Moineau 2010).



1.1 Pav. Prokariotinės ląstelės arba jų populiacijos apsaugos nuo bakteriofagų mechanizmai. Vaizdas adaptuotas iš (Westra et al. 2012).

Prokariotų antivirusinės sistemos yra labai įvairios, tačiau dalis jų turi bendrą, išskirtinę savybę – gebėjimą atskirti svetimą genetinę medžiagą nuo savos (Safari et al. 2020). Tam yra reikalingi baltymai, kurie gali specifiskai sąveikauti su DNR. Šitie baltymai yra sulaukę didelio susidomėjimo mokslo pasaulyje dėl savo pritaikymo genų inžinerijoje. Klasikinis pavyzdys: R-M sistemos. Tai yra antivirusinės sistemos, susidariusios iš dviejų komponentų: modifikacijos baltymų ir restrikcijos baltymų. Modifikacijos baltymai metilina specifinę, tipiška palindrominę, DNR seką. Jų veikimas yra suderintas ląstelėje, kad tik paties prokarioto DNR būtų modifikuojama. Restrikcijos baltymas, trumpiau REazė, atpažįsta tą pačią seką ir „perkerpa“ DNR molekulę, palikdamas dvigrandį trūkį, bet tik tuo atveju, jei seka nėra metilinta. Rezultatas - jei virusas suleidžia savo DNR į ląstelę ir joje yra R-M sistemos taikinio seka, tai restrikcijos baltymas ją perkerpa ir neutralizuoja, o savo paties genomą ignoruoja, nes jame esantys taikiniai yra metilinti modifikacijos baltymo (Roberts and Murray 1976). Dauguma REazių įvesdamos dvigubą DNR trūkį palieka vadinamus „lipnius galus“. Tai yra DNR trūkiai abiejose grandinėse, pasislinkę per keletą nukleotidų, atitinkamai, viena DNR grandinė yra ilgesnė už kitą nukirptame gale. Tokie galai specifiskai limpa vienas prie kito dėl nukleotidų komplementarumo. REazės ne tik leidžia perkirpti DNR molekulę specifinėje vietoje, bet ir tiksliai sujungti dvi atskiras molekules į vieną. Šios savybės nulėmė platų REazių kaip genų inžinerijos įrankių naudojimą (Hung and Wensink 1984).

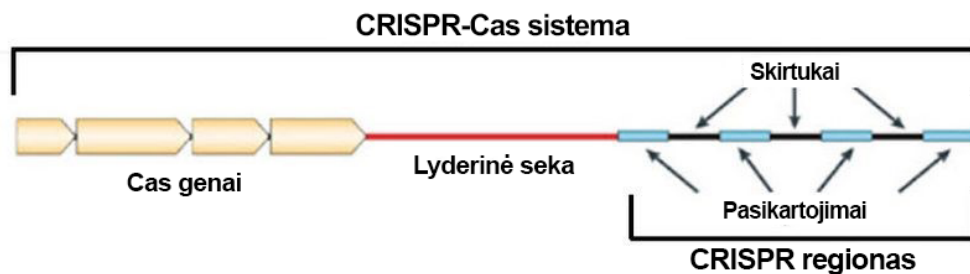
Duotas R-M sistemos pavyzdys turi vieną aiškų trūkumą, kuris taip pat būdingas daugeliui kitų antivirusinių sistemų. R-M sistemos taikiny yra nekintantis (Hyman and Abedon 2010). Sekos atpažinimas visiškai priklauso nuo baltymo struktūros – viena REazė atpažįsta vieną seką. Kiekvienos papildomos sekos atpažinimui reikalinga atskira R-M sistema. Jeigu viruso genome taikinio sekos nėra arba ji yra dingusi dėl mutacijos, R-M sistema yra visiškai neveikli ir neapsaugo ląstelės. Dėl šios priežasties galima teigti, jog tipinės antivirusinės sistemos yra eukariotuose esančio įgimto imuniteto atitikmuo.

1.1 CRISPR-Cas sistemos

Ilgą laiką buvo manoma, jog prokariotuose neegzistuoja įgyto imuniteto analogai, kol nebuvo atrastos CRISPR-Cas sistemos ir išaiškinta jų funkcija (Makarova et al. 2006). CRISPR-Cas yra antivirusinės sistemos tipas, leidžiantis prokariotui pasisavinti fragmentą iš svetimos genetinės medžiagos ir įterpti į savo paties genomą. Šie fragmentai yra naudojami kaip šablonai, skirti atpažinti jiems komplementarias sekas. Kitaip tariant, šablonai, skirti

atpažinti svetimą genetinę medžiagą, iš kurios paimtas fragmentas. Tai suteikia ląstelei atsparumą nuo pakartotinių infekcijų bakteriofagais (Sorek, Lawrence, and Wiedenheft 2013).

CRISPR-Cas imunitetas prieš naują taikinį susiformuoja trejais žingsniais. Pirmasis yra adaptacija arba naujo šablono įsistatymas į genomą. Šiame procese visada dalyvauja Cas1 ir Cas2 baltymai. Jie suformuoja kompleksą, kuris gali surišti naujam šablonui tinkamą DNR fragmentą ir integruoti jį į CRISPR regioną. DNR šablonai, integruoti į genomą, vadinami skirtukais (angl. *spacer*), o prieš integraciją – preskirtukais (angl. *protospacer* arba *prespacer*). CRISPR regionas sudarytas iš integruotų šablonų sekų, kurios yra atskirtos pasikartojančiomis, kai kuriais atvejais palindrominėmis, sekomis (angl. *repeats*) (Sorek, Kunin, and Hugenholtz 2008). CRISPR regiono pradžią žymi lyderinė seka, kuri gali dalyvauti adaptacijoje, nurodydama Cas1-Cas2 kompleksui tikslią integravimo poziciją (Rollie et al. 2015).

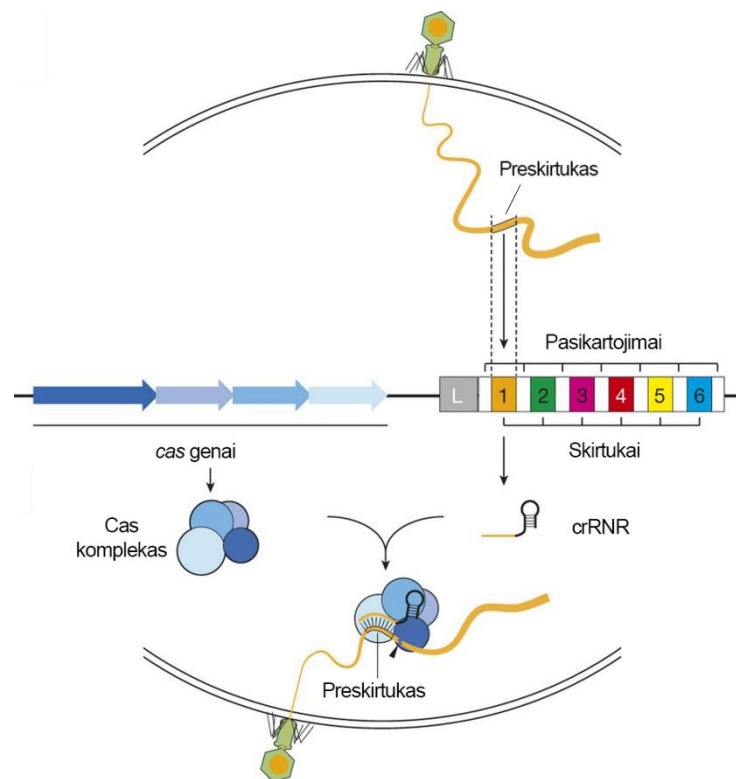


1.2 pav. CRISPR-Cas sistema, sudaryta iš Cas genų, lyderinės sekos bei pasikartojimų ir skirtukų eilės. Vaizdas adaptuotas iš (Sorek et al. 2008).

Antras žingsnis yra crRNR molekulių generavimas. Visas CRISPR regionas su skirtukais ir pasikartojimais yra transkribuojamas ir gauta RNR molekulė yra brandinama ir išskaidoma į mažesnes. Šie fragmentai turi po vieną skirtuko seką ir brandimo metu virsta į crRNR, kuri susijungia su efektoriniais baltymais (Charpentier, van der Oost, and White 2013).

Paskutinis imuniteto žingsnis yra interferencija, kurią vykdo kompleksas, sudarytas iš Cas baltymų (arba vieno Cas baltymo) ir crRNR. Skirtingai negu kiti baltymai, kurie sąveikauja su specifinėmis nukleotidų sekomis (pvz.: REazės), efektorinis CRISPR-Cas kompleksas atpažįsta taikinio seką pagal komplementarumą turimam crRNR. Tai reiškia, kad tie patys Cas baltymai gali specifiškai sąveikauti su skirtingomis sekomis. Jų taikinį nulemia

turimo skirtuko seka. Yra papildomas reikalavimas Cas kompleksui sėkmingai prisijungti prie nukleorūgščių grandinės. Tiesiai greta crRNR atpažįstamos sekos turi būti PAM seka, tipiška kelių nukleotidų ilgio, kuri tiesiogiai sąveikauja su Cas baltymu arba baltymais. Šis mechanizmas neleidžia efektoriniui kompleksui atpažinti skirtukų, esančių savo genome. Jei kompleksas sėkmingai prisijungia prie RNR arba DNR grandinės, tai (endo)nukleaziniai katalizės centrai grandinę perkerpa ir neutralizuoja puolantį fagą (Gasiunas et al. 2012).

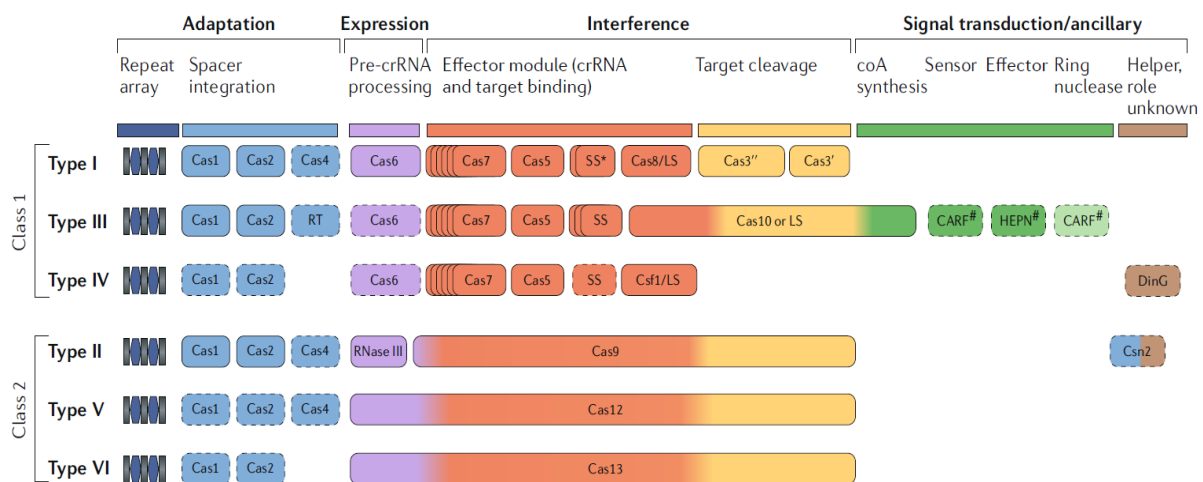


1.3. pav. CRISPR-Cas sistemos principo schema. Vaizdas adaptuotas iš (Marraffini 2015).

1.2 Įvairovė ir kilmė

Bendra žinoma CRISPR-Cas sistemų įvairovė yra stebinančiai didelė, kaip ir kitos prokariotų antivirusinės sistemos, kadangi jos yra veikiamos didelio evoliucinio spaudimo iš virusų pusės. Atitinkamai, klasifikacija yra apsunkinta. Nėra ne vieno geno, kuris būtų bendras visoms sistemoms (Makarova et al. 2020). Šiuo metu CRISPR-Cas sistemos skirstomos į dvi pagrindines klases: I ir II. I klasės sistemų efektorius yra baltymų kompleksas, tuo tarpu II klasės efektorius yra vienas multidomeninis baltymas. Šios klasės toliau skirstomos į tipus ir subtipus pagal konkrečius genus (1.4. Pav.). Žinios apie CRISPR-Cas sistemų įvairovę nuolat didėja vykdamas naujų sistemų paiešką didėjančiuose genomų ir

metagenomų duomenyse. Dalinai ši paieška yra skatinama tikslo naudoti CRISPR-Cas kaip biocheminį įrankį, (pvz.: naujų alternatyvų Cas9 baltymui ieškojimas) (Makarova et al. 2020). Dėl šios priežasties susidomėjimo sulaukė V tipas. Jam, kaip ir visoms II klasės sistemoms, yra būdingas vienas multidomeninis efektorinis baltymas. Tačiau palyginus su kitais variantais (pvz.: Cas9), kai kurių šių sistemų efektoriniai baltymai turi pranašumą: jie yra itin maži (400-700 aminorūgščių). Tai palengvina šių baltymų pritaikymą genų redagavime. Neskaitant jau minėtų efektorinių baltymų, V tipo sistemos yra mažai ištirtos (Harrington et al. 2018; Yan et al. 2019).



1.4. pav. CRISPR-Cas tipai ir jų pagrindiniai bruožai. Tipai toliau skirstomi į subtipus (pvz. I-A, I-B etc.) (Makarova et al. 2020).

Problema su dabartine CRISPR-Cas sistemų klasifikacija slypi tame, kad kol kas neišsprendžiama tinkamai atspindėti evoliucinių ryšių. Pavyzdžiui, II klasės efektoriniai baltymai pasižymi labai panašia struktūra ir mechanizmu, tačiau tarp jų genų nepastebima homologija. Tai yra konvergentinės evoliucijos rezultatas (Koonin, Makarova, and Zhang 2017). Viena pailgų evoliucinių ryšių priežastis yra savybė, būdinga antivirusinėms sistemoms. Tai yra tendencija horizontaliai genetinei pernašai. Nustatyta, kad genome gynybines sistemas labai dažnai eina greta transpozonų genų, kartais netgi jų apsuptos. Manoma, jog antivirusinės sistemos yra pernešamos kartu su mobiliomis genomu elementais, kas leidžia jiems lengviau pasklisti po prokariotų populiacijas. Šis reiškinys yra tiek paplitęs, kad antivirusinės sistemos formuoja vadinamąsias „gynybines salas“ prokarioto genome. Skirtingos antivirusinės

sistemos randamos šalia viena kitos genome labai aukštu dažniu, nes yra pernešamos tais pačiais transpozonais (Doron et al. 2018).

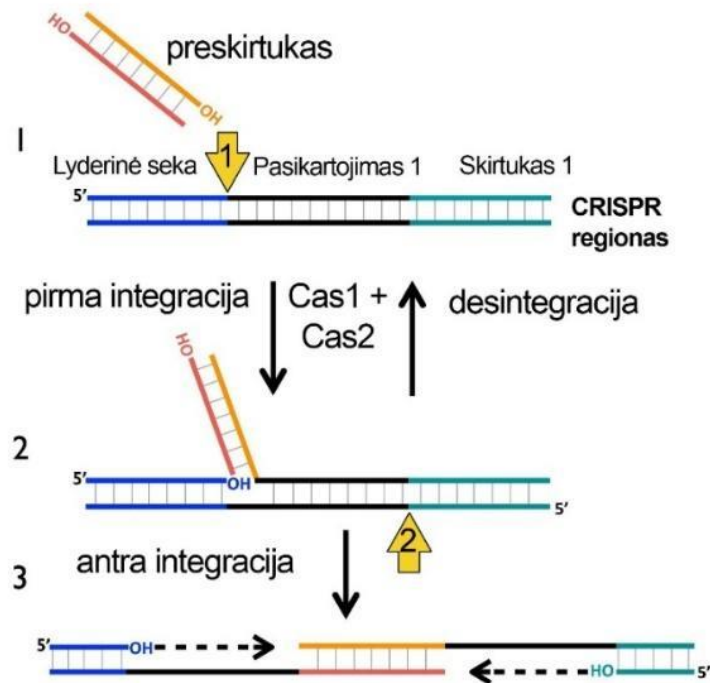
Transpozonai asocijuojami ne tik su CRISPR-Cas išplitimu, bet ir su jų atsiradimu. Esminiai baltymai, randami beveik kiekvienoje CRISPR-Cas sistemoje, yra Cas1 ir Cas2. Jie yra susiję su CRISPR-Cas kertine savybe – šablonų integravimu į genomą (Koonin et al. 2017). Pasirodo, kad egzistuoja Cas1 homologai, kurie nėra asocijuoti su CRISPR-Cas. Jie randami specifinėje šeimoje save sintetinačių transpozonų, dar vadinamų kaspozonais (angl. *casposons*). Šie mobilūs genetiniai elementai turi panašumų su eukariotiniais save transkribuojančiais transpozonais, nes turi invertuotus pasikartojimus savo galuose bei genus, homologiškus B šeimos polimerazėms. Kaspozonai yra unikalūs, nes juose DNR iškirpimą ir integraciją atlieka Cas1 homologas (Béguin et al. 2016; Krupovic et al. 2014). Manoma, jog CRISPR-Cas sistemos kilo adaptacijos moduliui, kuris išsivystė iš kaspozonų, susiliejus su protėviniais efektoriniais kompleksais (Koonin and Krupovic 2014).

1.4 Adaptacija

1.4.1 Cas1-Cas2 mechanizmas

Žinant, kad Cas1 ir Cas2 yra seniausi ir labiausiai paplitę baltymai tarp CRISPR-Cas sistemų, nėra stebėtina, kad jie yra atsakingi už pirmąją ir esminę stadiją CRISPR-Cas imuniteto susidaryme - naujų skirtukų integravimą į genomą. Abu šie baltymai yra katalitiškai aktyvūs. Cas1 yra nuo metalo jonų priklausoma DNR endonukleazė, o Cas2 yra, priklausomai nuo sistemos, DNR, RNR arba abiejų endonukleazė (Beloglazova et al. 2008; Dixit et al. 2016; Nam et al. 2012; Wiedenheft et al. 2009). Kaip ir baltymai, naujų skirtukų integravimo į ląstelės genomą mechanizmas yra konservatyvus tarp visų sistemų (Koonin et al. 2017; Yosef, Goren, and Qimron 2012). Cas1 ir Cas2 baltymai sudaro bendrą kompleksą, kuris gali surišti tinkamą substratą. Tipiškai tai yra 20–50 b.p. (kiekviena CRISPR sistema turi savo specifinį ilgį) dvigrandė DNR su 3' išsikišimais (Nuñez et al. 2015; Wang et al. 2015). Šie preskirtukai yra sugeneruojami kitų ląstelės baltymų kaip RecBCD komplekso ar kitų nukleazių ir atrenkami įvairiais mechanizmais, siekiant išvengti netinkamų substratų, pvz. iš pačios ląstelės genomo (Amitai and Sorek 2016; Fineran et al. 2014; Wang et al. 2015; Yosef et al. 2012). Tuomet, Cas1-Cas2 kompleksas katalizuoja dvi transesterifikacijos reakcijas, kurių metu preskirtuko 3' galai yra prijungiami prie CRISPR regiono pasikartojimo, greta lyderinės sekos, iš 5' galų. Pirma katalizuojama preskirtuko 3'-OH

nukleofilinė ataka į pirmo pasikartojimo 5' galą (tipiškai 5' galą esančią sandūroje su lyderine seka). Po pirmosios integracijos reakcijos yra vykdoma antroji reakcija į kitą pasikartojimo 5' galą, praskiriant DNR grandinę ties pasikartojimu. Susidarę tarpai užpildomi ir pataisomi ląstelės polimerazių ir ligazių, sudarant vieną naują skirtuką ir pasikartojimą (1.5. pav.) (Rollie et al. 2015).

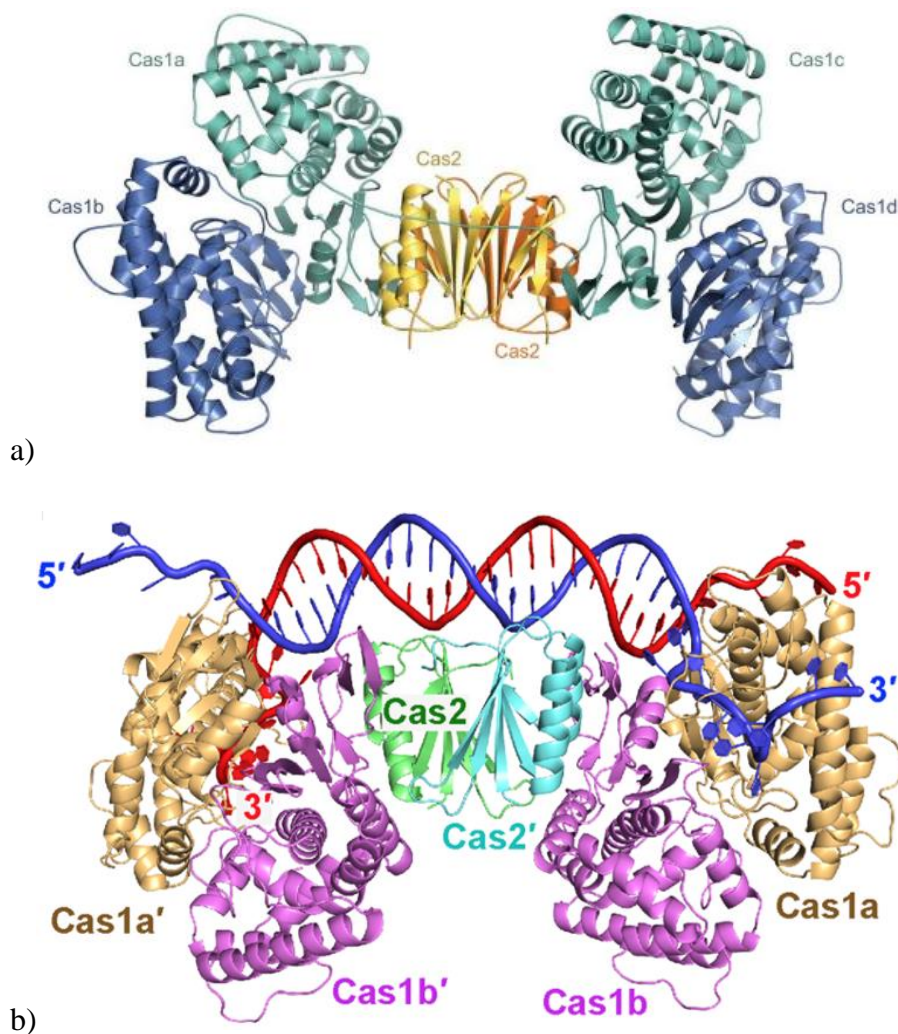


1.5 pav. Skirtuko integravimo modelis. Vaizdas adaptuotas iš (Rollie et al. 2015)

1.4.2 Cas1-Cas2 struktūra

Nors abu baltymai komplekse yra reikalingi adaptacijai ir pasižymi katalitiniais aktyvumais, jų funkcija skiriasi. Parodyta, kad adaptacijos kompleksai su Cas1 aktyviojo centro mutantais nebegali integruoti naujų skirtukų, tačiau kompleksai su Cas2 aktyviojo centro mutantais toliau sėkmingai vykdo integracijos reakcijas (Nuñez, Lee, Engelman, & Doudna, 2015). Šis ypatumas yra nulemtas unikalios Cas1-Cas2 komplekso struktūros. Komplexas yra sudarytas dviejų Cas1 homodimerų ir vieno Cas2 homodimero. Cas2 dimeras veikia kaip tiltelis, kuris jungia abu Cas1 dimerus. Surišto preskirtuko galai, kurie yra integruojami į genomą, lokalizuoti ties Cas1 baltymais. Cas2 tik padeda surišti substratą ir negali katalizuoti reakcijos, nes fiziškai nesiekia substrato galų. Nors Cas2 baltymas nėra

katalitiškai aktyvus integracijos komplekse, jis vis tiek yra svarbus taisyklingai adaptacijai. CRISPR lokuse ne tik pasikartojimai, bet ir skirtukai yra vienodo nukleotidų skaičiaus ilgio. Šis ilgis yra specifinis kiekvienai CRISPR-Cas sistemai. Kadangi preskirtukas yra surišamas Cas1 baltymais, surišto substrato ilgis tiesiogiai priklausys nuo atstumo tarp abiejų dimerų. Cas1-Cas2 kompleksas elgiasi kaip molekulinė liniuotė, kuri užtikrina vienodą skirtukų ilgį (Xiao et al. 2017). Įdomu, kad CRISPR-Cas sistemos, kurios pasižymi trumpiausiais žinomais skirtukų ilgiais (~18b.p.), neturi Cas2 – adaptacijos kompleksas sudarytas iš dviejų Cas1 dimerų, tad atstumas tarp jų katalitinių centrų yra mažesnis nei tipiniu atveju (Wright et al. 2019).



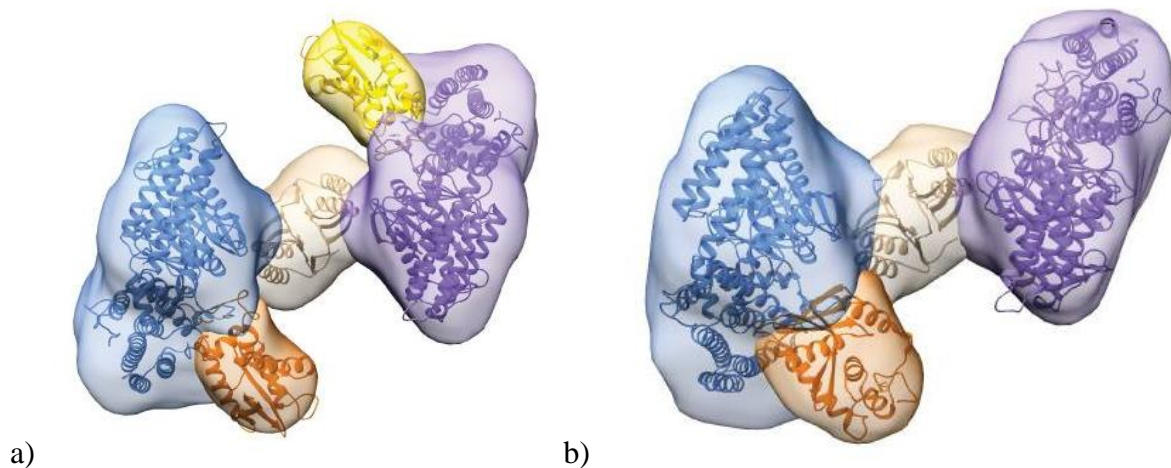
1.6 pav. a) Cas1-Cas2 kompleksas (Nuñez et al. 2014). b) Cas1-Cas2 kompleksas surišęs preskirtuką (Wang et al. 2015).

1.4.3 Funkcijų įvairovė

Dar viena esminė funkcija adaptacijos metu yra PAM atpažinimas. Tik sekos, kurios randamos tiesiai greta PAM sekos, suteiks atsparumą jas integravus kaip skirtuką. Šis mechanizmas pasižymi labai didele įvairove tarp skirtingų sistemų. I-E CRISPR-Cas sistemos Cas1-Cas2 kompleksas, randamas *Escherichia Coli*, šią funkciją atlieka savarankiškai. Cas1 baltymas turi papildomą kišenę, skirtą PAM sekos atpažinimui, kuris reikalingas sėkmingai integracijai (Wang et al. 2015). Pagal struktūrinius panašumus manoma, kad kitose I-E potipio sistemose egzistuoja panašus mechanizmas (Yoganand et al. 2019). Tačiau dažnai Cas1-Cas2 kompleksas nesugeba atlikti šios funkcijos ir reikalingi papildomi baltymai PAM atpažinimui. Cas4 yra endonukleazė, kuri priskiriama prie adaptacijos modulio, ir visose kitose I tipo sistemose ji perima šį vaidmenį. (Kieper et al. 2018; Shiimori et al. 2018; Yoganand et al. 2019). PAM atpažinimą nebūtinai turi atlikti adaptacijos baltymas. Kai kuriose II klasės sistemose PAM nurodymą integracijos metu atlieka efektorinis baltymas Cas9 (Heler et al. 2015). Tose pačiose sistemose gali būti randamas ir Cas4 baltymas, tačiau šiais atvejais dar nėra aiški jo funkcija (Makarova et al. 2020).

CRISPR-Cas, kaip ir dauguma antivirusinių sistemų tipiškai yra savarankiška. Visas specifines funkcijas ji atlieka be kitų ląstelės baltymų pagalbos. Tai parodo, kad visi elementai atsakingi už, pavyzdžiui, adaptaciją, yra randami pačioje sistemoje (Koonin et al. 2017; Stern and Sorek 2011). Tačiau yra išimčių. Taisyklingai adaptacijai vykti, būtina, kad Cas1-Cas2 kompleksas integruotų į specifinę vietą – pirmo CRISPR pasikartojimo ir lyderinės sekos sandūrą. Priklausomai nuo sistemos tipo, Cas1-Cas2 kompleksas specifiškai sąveikauja su šią sandūrą arba atpažindamas ją pagal nukleotidų seką, arba pagal tretinę DNR struktūrą (Babu et al. 2011; Moch et al. 2016; Rollie et al. 2015). I-E CRISPR-Cas sistemoje randama *E. coli* Cas1-Cas2 negali savarankiškai atpažinti šios sekos. Pasirodo, reikalingas specifinis *E. coli* baltymas IHF, kurio atpažįstamas motyvas yra sistemos lyderinėje sekoje. IHF (kuris jau buvo žinomas kaip būtinas kofaktorius λ bakteriofago rekombinacijai *E. coli* (Xin and Feiss 1993)) susiriša su DNR ir išlenkia ją 160° kampu, smarkiai deformuodamas ją. Šias tretines struktūras atpažįsta Cas1-Cas2, ir pagal tai nustato lyderinės sekos-CRISPR sandūrą (Nuñez et al. 2016). Tai rodo, kad nepaisant CRISPR-Cas savarankiškumo, kiekvienu atveju yra įmanomas unikalus evoliucinis išsivystymas, kuris gali priklausyti nuo specifinio organizmo.

Nors įprastas „kanoninis“ integracijos kompleksas yra sudarytas iš dviejų Cas1 dimerų, sujungtų Cas2 dimeru, galimi sudėtingesni variantai, į kurių sudėtį įeina minėti papildomieji baltymai. I-C *Bacillus halodurans* sistemoje parodyta, kad adaptacijos kompleksas sudarytas iš trijų baltymų: Cas4-Cas1-Cas2, kur Cas4 baltymai lokalizuoti ties Cas1 aktyviaisiais centrais. Įdomu tai, kad šie kompleksai turėjo simetrinius ir asimetrinius variantus. Vienu atveju, kai Cas4 lokalizuotas ties abejais Cas1 dimerais, o kitu, kai Cas4 lokalizuotas tik ties vienu Cas1 dimeru. Asimetriniai variantai buvo surišę preskirtuką ir substrato galas su PAM seka buvo lokalizuotas komplekso pusėje su Cas4 (1.7 pav) (Lee, Dhingra, and Sashital 2019). Sudėtingiausias žinomas variantas randamas *Streptococcus thermophilus* II-A tipo sistemoje, kuri susideda iš keturių Cas1 dimerų, dviejų Cas2 dimerų ir dviejų Csn2 (Cas4 atitikmuo) tetramerų. Šis kompleksas sudaro monomerą, kuris gali toliau sudaryti sudėtingesnes struktūras, į kurių sudėtį įeina netgi efektorinis Cas9 baltymas. (Wilkinson et al. 2019).



1.7 pav. Mėlyna/violetinė – Cas1; Pilka – Cas2; Geltona/oranžinė – Cas4. a) Simetrinė Cas1-Cas2-Cas4 struktūra b) Asimetrinė Cas1-Cas2-Cas4 struktūra (Lee et al. 2019).

Panašiai kaip komplekso struktūros, adaptacijos mechanizmai pasižymi didele įvairove nepaisant Cas1 ir Cas2 konservatyvumo. Paruošto skirtuko integracija į CRISPR lokusą katalizuojant Cas1 vyksta tuo pačiu principu visais žinomais atvejais. Tačiau preskirtuko paruošimo stadijos gali smarkiai skirtis. I-D CRISPR-Cas sistema, randama *Synechocystis* sp. turi įprastą adaptacijos baltymų komplektą – Cas1, Cas2 ir Cas4. Šioje sistemoje pasirodo, jog Cas2 ir Cas4 konkuruoja tarpusavyje susijungti su Cas1 baltymu. Pirma Cas1 ir Cas4 sudaro stabilų kompleksą, kuris gali procesuoti potencialų substratą į integracijai tinkamą preskirtuką. Kai yra surištas preskirtukas, Cas2 sąveikauja stipriau su

Cas1, ir išstumia Cas4, sudarant įprastą Cas1-Cas2 integracijos kompleksą. Šiuo atveju nėra momento, kai Cas4 ir Cas2 įeina į kompleksą kartu, jie susijungia su Cas1 skirtinguose adaptacijos žingsniuose (Wu et al. 2021). Visiškai skirtingai šis procesas vyksta I-G subtype, randamame *Geobacter sulfurreducens*. Šioje sistemoje yra unikalus Cas1, kuris yra susiliejęs kartu su Cas4 į vieną baltymą. Šios sistemos integracijos kompleksas Cas4/Cas1-Cas2 suriša tik DNR dupleksą su PAM turinčiu 3' išsikišimu. Cas4 suriša su PAM, bet ne hidrolizuoja. Priešingas duplekso galas, neturintis PAM, nėra surištas, todėl gali būti nugriaužtas nespecifinių nukleazių. Šis galas gali tada būti integruojamas į CRISPR lokusą, ir dalinė integracija aktyvuoja Cas4, kuris nukerpa PAM seką ir disocijuoja nuo duplekso. Galiausiai, įvykdoma antra integracijos reakcija, sudarant naują skirtuką (Hu et al. 2021). Šiame pavyzdyje preskirtuko apdorojimas – PAM pašalinimas – vyksta tik įvykus pirmai integracijos reakcijai, kas yra priešinga praeitam atvejui, kur PAM pašalinimas įvyksta pats pirmas.

Iš visų šių pavyzdžių matome, kad nepaisant konservatyvaus mechanizmo, skirtukų integravimas pasižymi stebėtina įvairove ir prisitaikymais palaikyti efektyvią adaptaciją. Tai leidžia ląstelėms toliau sėkmingai gintis nuo bakteriofagų. Šiame darbe tirtame V-F tipo sistemoje randami visi įprasti adaptacijos baltymai Cas1, Cas2 ir Cas4, bei efektorinis baltymas Cas14. Cas14 PAM seka žinoma – 5'-TTR (Karvelis et al. 2020). Šiuo metu nėra žinoma, kas atpažįsta PAM integracijos metu, kokias etapus vyksta adaptacija, ir kokius kompleksus šie baltymai gali sudaryti.

2. Medžiagos ir Metodai

2.1 Medžiagos

Prietaisai: -20°C ir +4°C šaldytuvai, centrifuga, „MiliQ“ distiliavimo aparatas, termocikleris „Eppendorf“, termostatas, svarstyklės, purtyklė, mėgintuvėlių maišyklė, DNR elektroforezės aparatas, baltymų elektroforezės aparatas, „ÄKTA avant“ chromatografijos sistema, svarstyklės, spektrofotometras „Nanodrop 2000“, CyclonetM skeneris (Packard Instrument), elektropravimo aparatas „Bio-Rad genepulser“ (Bio-Rad *E.coli* pulser).

Indai ir kitos priemonės: Matavimo kolbos, matavimo cilindrai, petri lėkštelės, graduotos stiklinės, stikliniai mėgintuvėliai, plastikiniai mėgintuvėliai, dializei skirta celiuliozinė žarnelė, elektroforezės stikliukai, nitrilinės, pirštinės, vienkanalės ir daugiakanalės pipetės, vienkartiniai antgaliai pipetėms, mėgintuvėlių stovai, organinio stiklo ekranas, Geigerio skaitiklis, dozimetras, fosforescenciniai ekranai „FujiFilm“, HisTrap kolonėlė, StrepTrap kolonėlė, 10 kDa koncentruojamosios kolonėlės, ledo vonelė.

Fermentai	Gamintojas
“Phusion” DNR polimerazė	ThermoFisher Scientific
“DreamTaq” DNR polimerazė	ThermoFisher Scientific
“FastDigest” REazės fermentai (REazės)	ThermoFisher Scientific
T4 DNR ligazė	ThermoFisher Scientific
“FastAP” šarminė fosfatazė	ThermoFisher Scientific

Rinkiniai	Gamintojas
“GeneJET PCR purification” rinkinys	ThermoFisher Scientific
“GeneJET Plasmid purification” rinkinys	ThermoFisher Scientific

Medžiagos	Gamintojas
“TopVision agarose” tabletės	ThermoFisher Scientific
“PageRuler Unstained” baltymų molekulinės masės standartas	ThermoFisher Scientific
Fenilmetilsulfonilfluoridas (PMSF)	ThermoFisher Scientific
“PageBlue” baltymų dažymo tirpalas	ThermoFisher Scientific
40% akrilamidas/bisakrilamidas (37.5:1)	Carl Roth
Amonio persulfatas (APS)	Carl Roth
Tetrametiletildiaminas (TEMED)	Carl Roth

Buferinis tirpalas	Sudėtis
“FastDigest” buferinis tirpalas	Tiekia ThermoFisher Scientific
“DreamTaq” buferinis tirpalas	Tiekia ThermoFisher Scientific
“Phusion” buferinis tirpalas	Tiekia ThermoFisher Scientific
TAE 50x	Tiekia ThermoFisher Scientific
Buferinis tirpalas A	20 mM Tris-HCl 500 mM NaCl 5 mM β-merkaptioetanolis
Buferinis tirpalas B	20 mM Tris-HCl 500 mM NaCl 5 mM β-merkaptioetanolis 500 mM imidazolas
Saugojimo buferinis tirpalas	20 mM Tris-HCl 500 mM NaCl 50% glicerolio
NBE-EtBr buferinis tirpalas	6.2 g/L boro rūgštis 1.2 g/L NaOH 0.744 g/l EDTA 850 ul/L acto rūgštis 4.25 mg/L EtBR
SDS buferinis tirpalas	0.025 M Tris-HCl 0.2 M glicinas 10. M NDS
Integracijos buferinis tirpalas	40 mM HEPES-NaOH 300 mM KCl 20% dimetilsulfoksidas (DMSO) 20 mM MgCl ₂

Auginimo terpė	Sudėtis
LB auginimo terpė	10g/L triptonas 10 g/L NaCl 5 g/L mielių ekstraktas

2.2 Metodai

2.2.1 DNR amplifikavimas (PGR)

1. Pasirenkama DNR polimerazė „Phusion“ arba „DreamTaq“ ir paruošiamas norimas kiekis (20 µL arba 50 µL) PGR mišinių (sudėtis pateikta lentelėse). Reakcijos ruošiamos ledo vonelėje.

„Phusion“ PGR mišinys	Bendras tūris 20 µL	Bendras tūris 50 µL
„Phusion“ buferinis tirpalas	4 µL	10 µL
2mM dNTP	2 µL	5 µL
Fwd pradmuo	1 µL	2,5 µL
Rev pradmuo	1 µL	2,5 µL
DNR matrica	X µL	X µL
„Phusion“ DNR polimerazė	0,2 µL	0,5 µL
„MiliQ“ filtruotas vanduo	Iki 20 µL	Iki 50 µL

„DreamTaq“ PGR mišinys	Bendras tūris 20 µL	Bendras tūris 50 µL
DreamTaq buferinis tirpalas	4 µL	10 µL
2mM dNTP	2 µL	5 µL
Fwd pradmuo	1 µL	2,5 µL
Rev pradmuo	1 µL	2,5 µL
DNR matrica	X µL	X µL
DreamTaq DNR polimerazė	0,2 µL	0,5 µL
„MiliQ“	Iki 20 µL	Iki 50 µL

2. Lydymosi temperatūros apskaičiuojamos naudojantis *Thermofisher Tm calculator* įrankiu. PGR sąlygos optimizuojamos PGR reakcijas atliekant prilydimo temperatūros gradientu. Reakcijos atliekamos pagal šią lentelę:

Ciklo žingsnis	„Phusion“	„DreamTaq“
Pradinė denatūracija	+98 °C, 30s	+95 °C, 1-3 min
Denatūracija	+98 °C, 5-10s	+95 °C, 30s
Sulydimas	X °C, 30s	X °C, 30s
Prailginimas	+72°C, 15-30s/kb	+72 °C, 1 min/kb
Galutinis prailginimas	+72 °C, 5-10 min	+72 °C, 5-15 min
Laikymas	+4 °C, neribotai	+4 °C, neribotai

3. Amplifikuota DNR laikoma +4 °C temperatūroje.
4. Reakcijos išeigai įvertinti atliekama PGR mėginių DNR elektroforezė (2.2.2).

2.2.2 DNR elektroforezė

1. Kolboje ištirpinama viena „TopVision“ agarozės tabletė ir 50 mL NBE-EtBr buferinis tirpalas. Agarozės tirpalas išlydomas mikrobangų krosnelėje.
2. Ištirpintas agarozės tirpalas pilamas į elektroforezės formą ir įdedamos specialios šukos. Laukiama, kol tirpalas sustings (~20 min).
3. Paruošiami DNR mėginiai: DNR tirpalas ~5 µL, dažas ~5 µL, „MiliQ“ vanduo - iki 15 µL.
4. Agarozės tirpalui sustingus į gelį, jis įdedamas į elektroforezės aparatą, užpildytą NBE-EtBr buferiniu tirpalu. Į pasirinktus takelius suleidžiamas „Massruler mix“ standartas (ThermoFisher Scientific). Į likusius takelius suleidžiami paruošti DNR mėginiai. Aparatas įjungiamas, 100V, 300mA, ~45 min.
5. Pasibaigus elektroforezei, DNR juostelės yra vaizdinamos UV transiliumatoriuje ir fotografuojamos.

2.2.3 Plazmidžių transformacija

1. Į mėgintuvėlį supilama 4 mL LB terpė, tada užsėjamas ir per naktį (~16 h) paliekamas augti pasirinktas *E. coli* kamienas, +37 °C temperatūroje, purtant 220 aps./min.
2. Kitą dieną pirminis inokuliatas 30 kartų praskiedžiamas šviežia LB terpe ir toliau auginamas 37 °C temperatūroje, purtant 220 aps./min, kol optinis tankis (OD₆₀₀) pasiekia ~0,6-0,8.
3. Ląstelės centrifuguojamos 5 min, 3000g, +4 °C temperatūroje. Supernatantas nupilamas.
4. Ląstelės suspenduojamos 4 mL plovimo tirpale ir vėl centrifuguojamos 5 min, 3000g, +4 °C temperatūroje. Supernatantas nupilamas.
5. Ląstelės suspenduojamos 5 mL transformacijos tirpale ir 1 h inkubuojamos ledo vonelėje. Ląstelės dar kartą centrifuguojamos, taip pat nupilamas supernatantas.
6. Kompetentinės ląstelės tuomet suspenduojamos 100-250 µL transformacijos tirpalo ir sumaišomos su plazmidine DNR. Mišinys inkubuojamas ledo vonelėje 30 min.
7. Mišinys 0,5-1 min perkeliama į +42 °C temperatūrą (terminis šokas), o po to 5 min atvėsinaamas ledo vonelėje. Į mišinį pridedamas 1 mL šviežios LB terpės ir ląstelės inkubuojamos +37 °C temperatūroje purtant 1 h.
8. Į atvėsintą iki +50 °C agarizuotą LB terpę iki rekomenduojamosios koncentracijos pridedamas atitinkamas antibiotiko (pvz. ampicilino) kiekis ir terpė išpilstoma į Petri lėkšteles.
9. Ląstelės centrifuguojamos 5 min, 3000 aps./min, kambario temperatūroje. Dalis supernatanto nupilama, ląstelės resuspenduojamos ir išsėjamos į Petri lėkšteles. *E. coli* bakterijos lėkštelėse auginamos +37 °C, per naktį (~16h).
10. Pasirinktinaai patikrinama, ar susitransformavo norima plazmidė, atliekant kolonijų PGR. Atliekama „DreamTaq“ PGR reakcija, naudojant bakterijas (jose esančias plazmides) kaip DNR matricą.

2.2.4 Plazmidžių išskyrimas

1. Į mėgintuvėlį supilama 4 mL LB terpė, tada užsėjamas ir per naktį (16-18h) auginamas transformuotas *E. coli* DH5 α kamienas su norima plazmide ir atitinkamu rekomenduojamos koncentracijos antibiotiku, +37 °C temperatūroje, purtant 220 aps./min.
2. Kitą dieną ląstelės centrifuguojamos 5 min, 3000g, kambario temperatūroje, supernatantas nupilamas.
3. Ląstelių plazmidinė DNR izoliuojama naudojant „GeneJET Plasmid Miniprep“ (ThermoFisher Scientific) rinkinį, dirbant pagal gamintojo nurodymus.
4. Išgrynintos plazmidinės DNR koncentracija ir grynumas įvertinamas naudojant spektrofotometrą.

2.2.5 Baltymo raiška

1. Kolboje su 200 mL LB terpe užsėjamas ir per naktį (~16h) auginamas pasirinktas *E. coli* raiškos kamienas su norima plazmide, 1% koncentracijos gliukoze ir atitinkamo antibiotiko rekomenduojama koncentracija, +37 °C temperatūroje, purtant 220 aps./min.
2. Kitą dieną 15mL pirminio inokulianto, yra perkeliama į 2L kolbą su 400 mL autoklavuotos LB terpės. Taip pat vėl įdedamas antibiotikas su tokia pačia koncentracija, kaip pirmame žingsnyje. Ląstelės auginamos 37 °C temperatūroje, purtant 200 aps./min. kol optinis tankis (OD₆₀₀) pasiekia ~0,4-0,6
3. Kolbos perkeliama į +4 °C šaldytuvą, kol jos yra atvėsintos iki mažiausiai +16 °C.
4. Į kolbas suleidžiama 80 μ L 1M IPTG tirpalo (galutinė koncentracija 0,2mM). Ląstelės toliau auginamos 16-18h, +16 °C temperatūroje purtant 220 aps./min.
5. Praėjus laikui ląstelės nucentrifuguojamos, supernatantas nupilamas. Ląstelės saugomos -20 °C.

2.2.6 Baltymo gryninimas

1. Resuspenduojama X g bakterinės biomasės (kiek užauginta 2.2.6 skyriuje) 4X mL lizės buferyje.
2. Pridedama 5X μ L PMSF proteazių inhibitoriaus į lizės tirpalą.
3. Lizavimo mišinys ardomas ultragarsu 5X min (pulsas: 10 s įjungtas, 15 s pauzė) 55% amplitudėje. Maksimali temp. +22 °C
4. Nucentrifuguojamas lizatas ~18000g 1h esant 4°C.
5. Lizatas užnešamas tiesiai ant norimos chromatografinės kolonėlės. Chromatografija vykdoma laikantis rekomenduojamo protokolo skirtai kolonėlei. Tipiniu atveju (pvz. Su His-Trap arba Strep-Trap kolonėle) naudoti buferiniai tirpalai A ir B.

6. Atliekama norimų frakcijų SDS-PAGE denatūruojanti baltymų elektroforezė.
7. Frakcijos laikomos +4°C temperatūroje. Negalima užšaldyti.
8. Pasirinktas frakcijas dializuojama prieš saugojimo buferinį tirpalą, į kurio sudėti įeina 50% glicerolio. Iškeitus baltymų buferį, mėginiai laikomi -20°C

2.2.7 SDS-PAGE denatūruojanti baltymų elektroforezė

1. Sumontuojama gelio kasetė. Paruošiamas 12% skiriamasis gelis – 1.2 mL 40% AA/BAA (37:5:1), 1 mL 4x Tris-HCl pH 8.8, 1.8 mL „MiliQ“. Suleidžiama 20 µL APS ir 15 µL TEMED. Gelis greitai įpilamas į stikliuką ir ant viršaus užpilamas sluoksniu vandens. Laukiama, kol gelis sustings (~20min).
2. Nuo skiriamojo gelio nupilamas vanduo ir paruošiamos šukos. Paruošiama 1.5 mL koncentruojamojo gelio tirpalo: 30 % bisakrilamido tirpalas, 0,5 M TRIS-HCl pH 6.8, 10 % NDS, „MiliQ“. Suleidžiama 20 µL APS ir 15 µL TEMED. Gelis greitai įpilamas į stikliuką ir įdedamos šukos. Laukiama, kol sustings (~20min).
3. Paruošiami mėginiai: 5 mL SDS, ~5 mL lizato/frakcijos/baltymo (tūris keičiamas pagal spėjamą koncentraciją), „MiliQ“ - iki 30 mL. Baltymai denatūruojami 5 min ties +95°C.
4. Sumontuojamas elektroforezės aparatas, įstatomas stikliukas su geliu, užpilama SDS buferiniu tirpalu, kad gelis būtų panardintas.
5. Suleidžiami mėginiai ir įjungiamas šaltinis – 200V 15A ~45min.
6. Pasibaigus elektroforezei išmontuojamas aparatas ir vonelėje išplaunamas gelis su 100 mL vandens švelniai purtant 10 min. Išpilamas vanduo. Plovimas atliekamas 3 kartus.
7. Gelis panardinamas į „PageBlue“ dažą (~20 mL) ir dažomas švelniai purtant 60 min.
8. Gelis vėl plaunamas kaip aprašyta 6 punkte, kol išplaunamas su baltymais nesusirišęs dažas.

2.2.8 Lizatų ruošimas

1. Resuspenduojama ~3g bakterinės biomasės 10 mL lizės buferiniame tirpale.
2. Pridėdama 60 µL 0.2M PMSF proteazių inhibitoriaus.
3. Sonikuojamas lizavimo mišinys ant ledo vonelės 20 min (pulsas: 10 s įjungtas, 15 s pauzė) 55% amplitudėje. Maksimali temperatūra: +22 °C
4. Nucentrifūguojamas lizatas ~18000g 1 h ties +4°C.
5. Supernatantas filtruojamas naudojant švirkštą ir užmaunamas 0.2-0.22 µm filtru. Sterilumas nebūtinai.
6. Pasirinktinai, lizatai, skirti integracijos eksperimentams, sukcentruojami. Naudojamas 10 kDa centrifuginės koncentravimo kolonėlės. 500 µL sukcentruojami iki 25 µL.
7. Lizatas/frakcijos laikomos +4°C temperatūroje. Negalima užšaldyti.

2.2.9 Oligodupleksų (preskirtukų) paruošimas

1. Paruošiamas mėginys su dvejais (dalinais) komplementariais vienodos koncentracijos DNR oligonukleotidais, kurie galėtų sudaryti dupleksą.
2. Pasirinktinais, Oligonukleotidų 5' galai fosforilinami radioaktyviu ^{32}P naudojant ATP, [$\gamma\text{-}^{32}\text{P}$]. Laikomasi visų apsaugos priemonių darbui su radioaktyviomis medžiagomis.

Fosforiliacijos reakcija	Bendras tūris 10 μL
PNK A buffer, 10x	1 μL
$\gamma\text{-ATP}$, 10mM	1 μL
Oligonukleotidas1, 100 μM	1 μL
Oligonukleotidas2, 100 μM	1 μL
PNK	1 μL
H ₂ O	5 μL

2.2.10 *In vivo* integracijos reakcija

1. X mL LB terpėje mėgintuvėlyje užsėjamas ir per naktį (~16h) auginamas pasirinktas *E. coli* kamienas, +37 °C temperatūroje, purtant 220 aps./min. Kamienas turi būti transformuotas integracijos Cas baltymų, ekspresijos vektoriais ir plazmide, turinčia atitinkamą CRISPR regioną + lyderinę seką (gali būti įstatyta į genomą).
2. Sekančią dieną 30 μL naktinės kultūros perkeliama į 3 mL šviežios LB terpės su atitinkamais antibiotikais ir induktoriumi (IPTG).
3. Ląstelės auginamos 2 h +37 °C temperatūroje, purtant 220 aps./min.
4. Nuo šio žingsnio viskas atliekama +4 °C (arba ant ledo vonelės). Kiekvienai atskirai reakcijai paimama 1 mL kultūros, ji nusukama (5 min ties 3000g), ir ląstelės suspenduojamos 1 mL „MiliQ“ vandeniui. Nusukimo ir resuspendavimo žingsnis kartojamas papildomus 2 kartus. Po trečio suspendavimo ląstelės vėl nusukamos.
5. Ląstelės suspenduojamos 50 μL 5 μM oligoduplekso (preskirtuko) mišinyje. Kontrolinis variantas suspenduojamas „MiliQ“ vandenyje.
6. Mišinys perkeliamas į atšaldytą 1mm kiuvetę ir elektroporuota ties 1.8 kV ir 25 μF su pulso kontrole ties 200 Ω , laiko konstanta > 4.0 ms.
7. Iškart po elektroporacijos, ląstelės perkeltos į 3mL LB terpės su atitinkamais antibiotikais ir induktoriais (IPTG) ir auginama 2h.
8. Praėjus laikui 1 μL kultūros pridedama į PGR mišinį analizei (2.2.1).

2.2.10 *In vitro* Integracijos reakcija

1. Pasirinkti oligonukleotidai sulydomi į DNR dupleksą (2.2.9).
2. Atliekama *in vitro* integracijos reakcija, sudėtis nurodyta lentelėje. Inkubuojama 30 min +37 °C. Pasibaigiu reakcijai mėginiai saugomi -20 °C.

Integracijos mišinys	Bendras tūris 20 µL
Preskirtukas, 10µM	1 µL
CRISPR regioną + lyderinę seką turinti plazmidė (100ng/µl)	1.5 µL
2x Integracijos buferis	10 µL
Lizato/Frakcijų/„MiliQ“	Iki 20 µL

3. Mėginiai skiedžiami 20 kartų „MiliQ“ vandeniu, ir naudojami PGR analizei.

Arba

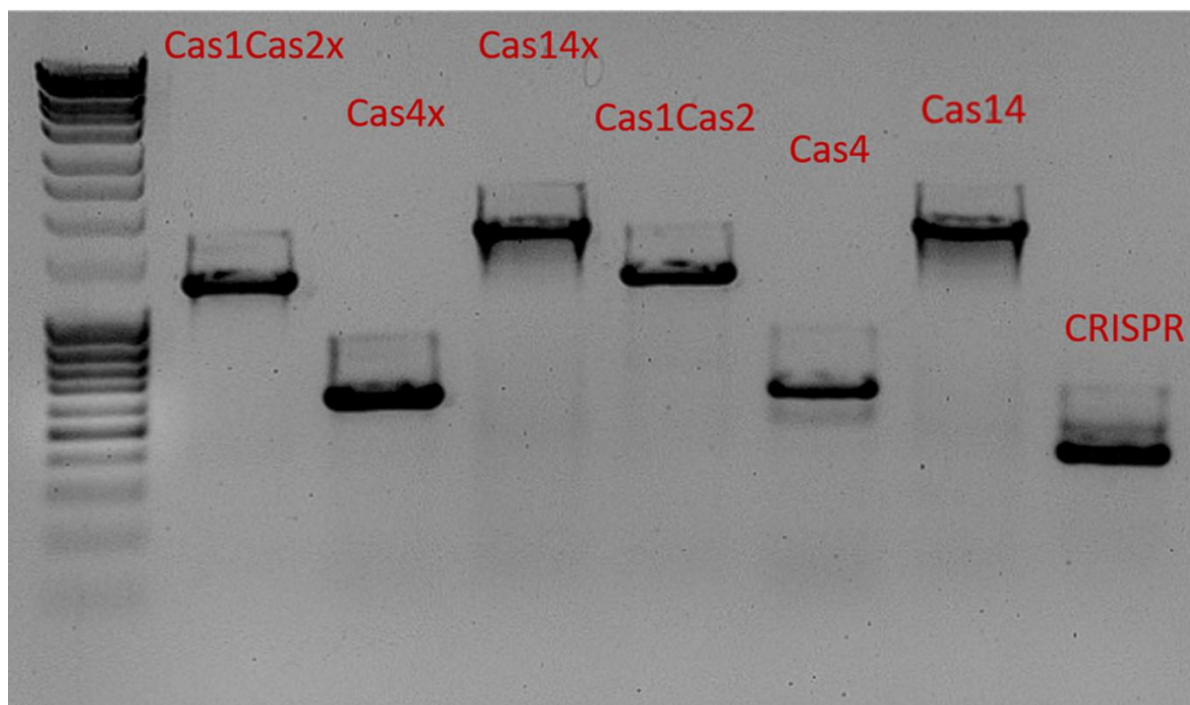
4. Jei naudoti radioaktyviai žymėti dupleksai, reakcija sustabdoma pridendant 20 µL fenolio-chloroformo mišinio. Gerai sumaišoma.
5. Mėginiai centrifuguojami 5 min ties 14000g.
6. Atsargiai nusiurbiamo viršutinė, vandenyje tirpi fazė su radioaktyvia DNR.
7. Kolboje ištirpinama 1 „TopVision“ agarozės tabletė 50 mL TAE buferio (be EtBR). Agarozės tirpalas išlydomas mikrobangų krosnelėje.
8. Ištirpintas agarozės tirpalas pilamas į elektroforezės kasetę ir įdedamos šukos. Laukiama kol sustings (~20 min).
9. Paruošiami DNR mėginiai: radioaktyvus DNR tirpalas ~5 µL, dažas ~5 µL, „MiliQ“ iki 15 µL.
10. Sustingus geliui jis įdedamas į elektroforezės aparatą užpildytą TAE buferiu. Į takelius suleidžiami paruošti DNR mėginiai ir įjungiamas aparatas, 100V, 300mA, ~45 min.
11. Pasibaigus, agarozės gelis pilnai išdžiovinamas. Išdžiūvus, ant gelio uždedamas fosforescencinis ekranas „FujiFilm“ ir po ~16h ekranas nuskenuojamas.

4. Rezultatai

4.1 genų sintezė ir amplifikavimas

Tyrimams pasirinkta Cas14B17 efektorinį baltymą koduojanti V-F tipo CRISPR-Cas sistema (toliau – B17 sistema) buvo aptikta metagenominių mėginių sekoskaitos duomenyse. Tyrimams jos baltymus koduojantys genai (optimizuoti *E. coli* raiškai) ir CRISPR regiono fragmentas buvo užsakyti sintezei ir pristatyti plazmidžių pavidalu.

Visi B17 sistemos baltymai ir CRISPR lokuso fragmentas, buvo sėkmingai amplifikuoti PGR metodu. Genai buvo amplifikuojami įvedant REazių taikinius, ir kiekvieno geno sukurti du variantai – vienas, tinkamas, suliejimui su vektoriaus koduojamais gryninimo inkarais, ir kitas, pritaikytas klonavimui be gryninimo inkarų (4.1 pav.)



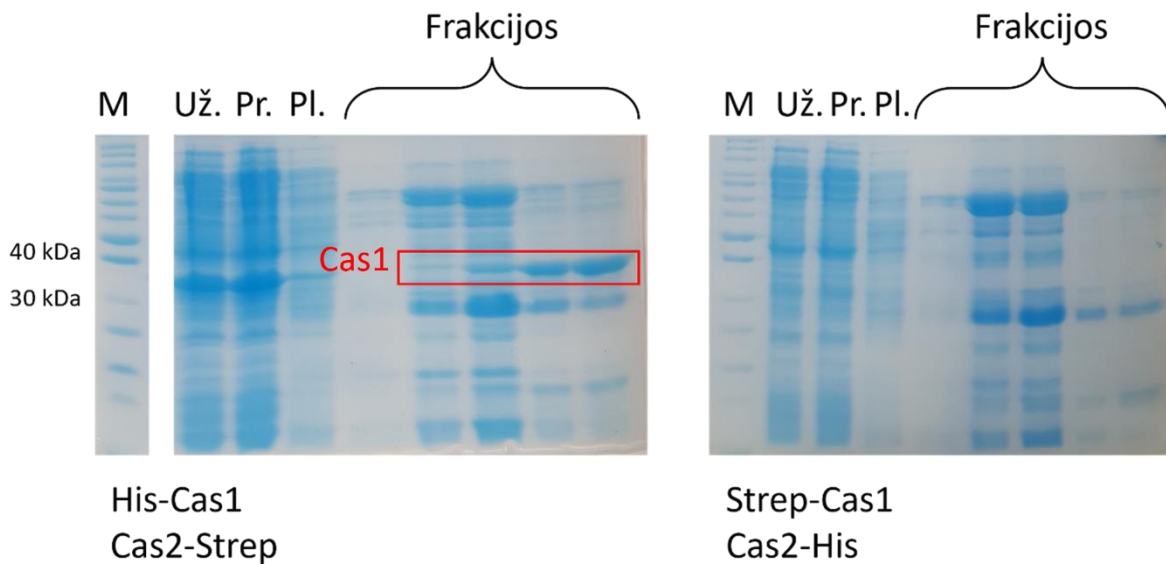
4.1 Pav. Amplifikuoti B17 sistemos genai. Cas1 ir Cas2 buvo amplifikuoti kartu. X nurodo žymenų nebuvimą (suligavus su vektorium).

4.2 Baltymų ekspresija ir gryninimas

Amplifikuoti genai buvo suliguoti į ekspresijos vektorius ir sutransformuoti į *E. coli*. Klonai patikrinti PGR metodu naudojant „DreamTaq“ polimerazę. Plazmidės buvo išgrynintos, jų sekos patikrintos sekoskaita. Sukonstruoti ekspresijos vektoriai:

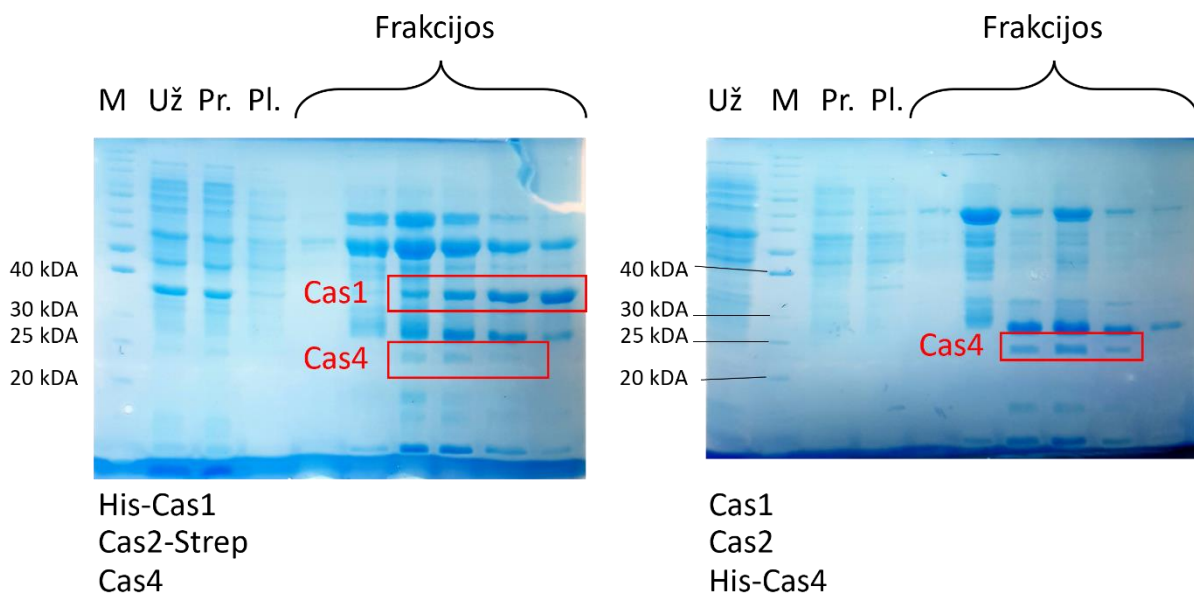
Pavadinimas	Vektoriaus pagrindas	Inkaras sulietas su baltymo N-galu	Inkaras sulietas su baltymo C-galu	Antibiotiko atsparumas
pET-His-Cas1-Cas2-Strep	pETDuet-1	6xHis (Cas1)	Streptag (Cas2)	Ampicilinas/ Karbenicilinas
pET-Strep-Cas1-Cas2-His	pETDuet-1	Streptag (Cas1)	6xHis (Cas2)	Ampicilinas/ Karbenicilinas
pET-Cas1-Cas1	pETDuet-1	x	x	Ampicilinas/ Karbenicilinas
pCDF-His-Cas4-Strep	pCDFDuet-1	6xHis	Streptag	Streptomycin
pCDF-His-Cas4	pCDFDuet-1	6xHis	x	Streptomycin
pCDF-Cas4	pCDFDuet-1	x	x	Streptomycin
pCOLA-His-Cas14	pCOLADuet-1	6xHis	x	Kanamycin
pCOLA-Cas14	pCOLADuet-1	x	x	Kanamycin

Prieš baltymų gryninimą patikrinta jų raiška. Užtikrinus deramą baltymų raišką, buvo auginamos *E. coli* biomasės baltymų gryninimui (2.2.5). Cas1 ir Cas2 baltymų raiška buvo vykdoma kartu ir gryninti naudojant HisTrap kolonėlę (Ni⁺² afiniškumo chromatografija) (4.2 pav.).



4.2 Pav. Gryninti Cas1 ir Cas2 baltymai naudojant HisTrap kolonėlę. Kairėje gryninta per Cas1 histidino inkarą, dešinėje per Cas2 inkarą. Pažymėta Cas1 baltymo juosta (38kDa)

Gryninant His-Cas1 (38 kDa) Cas2-Strep (11.9 kDa) baltymus kartu, frakcijos buvo praturtintos Cas1, tačiau be Cas2. Gryninant Strep-Cas1 Cas2-His nebuvo nei vieno norimo baltymo. Gryninant baltymus naudojant StrepTrap kolonėlę (Streptacino afiniškumo chromatografija) visais atvejais nematome Cas2 baltymo. Cas2 nesigrynina arba krenta į nuosėdas, potencialiai dėl sulieto inkaro, kuris neleidžia tinkamai mažam baltymui susivynioti, arba dėl Cas4 trūkumo, kuris potencialiai reikalingas sudaryti stabilų Cas1-Cas2-Cas4 kompleksą. Atitinkamai, Cas1 ir Cas2 raiška buvo vykdoma kartu su Cas4, ir visi baltymai gryninti vienu metu (4.3 pav.).



4.3. Pav. Gryninti Cas1, Cas2 ir Cas4 baltymai naudojant HisTrap kolonėlę. Kairėje gryninta per Cas1 histidino inkarą, dešinėje per Cas4 inkarą.

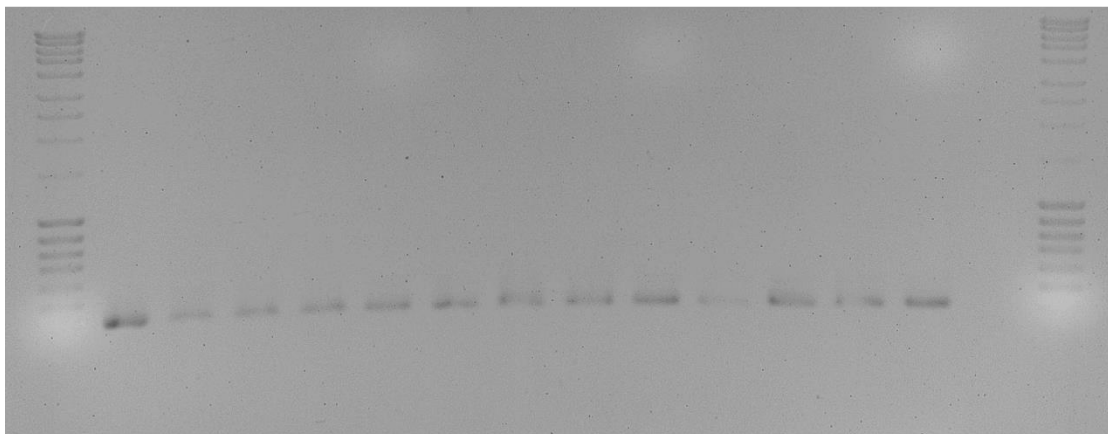
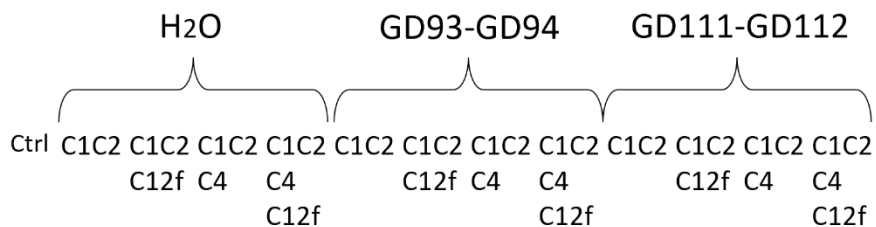
Gryninant His-Cas1 (38 kDa) Cas2-Strep (11.9 kDa) ir Cas4 (23kDa) baltymus kartu, frakcijos buvo praturtintos Cas1, labai mažu kiekiu Cas4. Gryninant Cas1 ir Cas2, neturinčius inkarų kartu su His-Cas4, stebime tik labai mažą Cas4 kiekį be Cas1 ir Cas2 baltymų. Matome, kad yra problemų su baltymų ekspresija arba Cas1-Cas2-Cas4 kompleksu susirinkimu. Potenciali priežastis – inkaro buvimas. Kadangi šie baltymai potencialiai sudaro sudėtingą kompleksą, inkaras, įterptas į netinkamą vietą, gali suardyti baltymų sąveikos paviršių.

4.3 *In vivo* integracija

Siekiant išsiaiškinti šių baltymų veiklą, buvo pereita prie metodų, kuriems nereikalingas baltymų gryninimas. *In vivo* integracijos tyrimui į *E. coli* BL21(DE3) kamienas buvo transformuotas plazmidė, turinčią CRISPR regioną kartu su lyderine seka. CRISPR regionas su lyderine seka buvo įklonuotas į mažakopijinį vektorių (pACYC pagrindu), kuriame atvirkštiniu PGR būdu buvo pašalinti T7 promotoriaus ir terminatoriaus sritys, kad neįvyktų CRISPR regiono ekspresija. Kamienas, turintis CRISPR plazmidę, buvo transformuotas vektoriais su adaptacijos baltymais ir Cas14 efektoriniu baltymu, be sulietų inkarų. Transformuoti keturi baltymų ekspresijos variantai: Cas1-Cas2, Cas1-Cas2 + Cas14, Cas1-Cas2 + Cas4, Cas1-Cas2 + Cas4 + Cas14. Pagal 2.2.8 protokolą atlikti *in vivo* integracijos eksperimentai šiuose kamieniuose naudojant 2 skirtingus DNR dupleksus ir kontrolę („MiliQ“ vanduo) (4.4 pav.).

GD93								5'	A	T	T	T	A	C	T	A	C	T	C	G	T	T	C	T	G	G	T	G	T	T	T	C	T	C	G	A	T	T	3'													
GD94	3'	T	T	A	G	G	G	T	T	A	A	A	T	G	A	T	G	A	G	C	A	A	G	A	C	C	A	C	A	A	A	5'																				
GD111																					5'	G	A	A	A	C	A	C	C	A	G	A	A	C	G	A	G	T	A	G	T	A	A	A	T	G	T	G	A	T	T	3'
GD112	3'	A	A	C	G	A	G	T	A	G	T	A	A	A	T	G	T	T	A	T	C	T	T	T	G	T	G	G	T	C	T	T	G	C	T	C	A	T	C	A	T	T	T	A	5'							

4.4 Pav. Naudotų DNR dupleksų schema. Raudoni nukleotidai rodo atpažįstama PAM seką (ji yra komplementari Cas14 atpažįstamai PAM). Mėlyni nukleotidai rodo oligonukleotidų komplementarią sritį, kuri turėtų atitikti būsimą skirtuką sėkmingos integracijos atveju.



4.5 Pav. *In vivo* integracijos CRISPR regionai amplifikuoti PGR būdu. Teigiama kontrolė (Ctrl) – PGR nuo plazmidės su CRISPR regionu.

In vivo Integracijos reakcijos detektuojamos PGR būdu, pilnas CRISPR regionas yra amplifikuojamas (4.5. pav.). Jei yra sėkmingai įsistatęs naujas skirtukas, CRISPR regionas išsiplečia. Matome, kad visomis sąlygomis skirtukas neįsistatė, visais atvejais CRISPR regiono dydis nepasikeitė (lyginant su kontrole.) Adaptacijos baltymai integruoja per mažu efektyvumu, kad adaptacija būtų užfiksuojama, arba šie baltymai negali tinkamai funkcionuoti *E. coli* ląstelėse.

4.4 *In vitro* integracija

Toliau siekiant nustatyti baltymų funkcionalumą pereita prie integracijos reakcijos *in vitro*. Lygiagrečiai bandytos trys sąlygos. Pirma, vienu metu grynintų Cas1, Cas2 ir Cas4 baltymų frakcijos (apjungta kartu) (4.3 pav.). Antra, BL21(DE3) kamieno, kuriame ekspresuoti baltymai, lizatas (tirpi frakcija ekspresuotų pagal 2.2.6 metodą, bet negrynintų baltymų). Trečia, tas pats lizatas, koncentruotas 10kDa koncentratoriumi, siekiu padidinti adaptacijos baltymų koncentracija (4.6 pav.). Mėginiai buvo analizuoti PGR būdu (4.7 pav.).

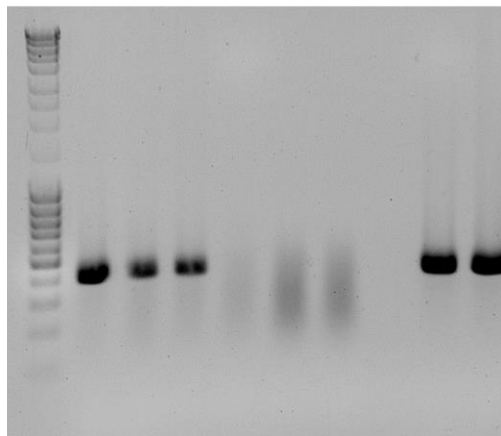
Integracijos mišinys	Bendras tūris 20 μ L
Preskirtukas, 10 μ M	1 μ L
CRISPR regioną + lyderinę seką turinti plazmidė (100ng/ μ l)	1.5 μ L
Lizatas	iki 20 mL

Integracijos mišinys (k.)	Bendras tūris 20 μ l
Preskirtukas, 10 μ M	1 μ L
CRISPR regioną + lyderinę seką turinti plazmidė (100ng/ μ l)	1.5 μ L
2x Integracijos buferis	10 μ L
Koncentruotas lizatas	iki 20 μ L

Integracijos mišinys	Bendras tūris 20 μ l
Preskirtukas, 10 μ M	1 μ L
CRISPR regioną + lyderinę seką turinti plazmidė (100ng/ μ l)	1.5 μ L
2x Integracijos buferis	10 μ L
Grynintas Cas1 Cas2 Cas4	iki 20 μ L

4.6 Pav. A *In vitro* integracijos reakcijų sąlygos. Naudotas GD111-112 dupleksas, turintis PAM seką.

BL. Bl. + Bl. + BL. Bl. + Bl. + F1 F2 Ctrl
C1C2 C1C2 k. C1C2 C1C2
+ C4 k. + C4
k.

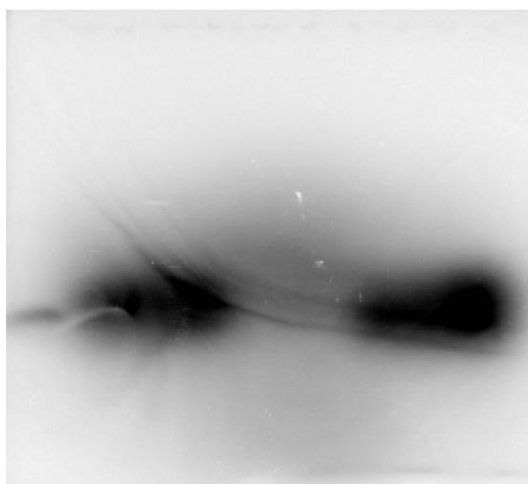


4.7 Pav. A *In vitro* integracijos PGR analizė. BL – BL21(DE3) kamieno lizatas. K. – koncentruotas lizatas. F1 – His-Cas1+ Cas2-Strep + Cas4. F2 – Cas1+ Cas2 + His-Cas4. Ctrl – integracijos reakcija tik su buferiu ir vandeniu.

Visais atvejais matome, kad CRISPR regionas neišsiplečia, ir sėkmingos integracijos reakcijos neužfiksuojame. PGR analizės reakcija nuo mėginių su ir be ATP nepavyksta su koncentruotais lizatais ir F1 takeliu (HisCas1 + Cas2-Strep + Cas4 frakcija). Koncentruotų lizatų atveju tikėtina, kad *E. coli* baltymai pakankamai sukonzentruoti, kad degraduoja DNR. F1 nepavyko, nors nėra priežasties manyti, kad DNR degraduojama. Įmanoma, kad įvyksta pirma arba dalinė integracija ir trūkstant ląstelės baltymų CRISPR regionas nėra užtaisomas. Tai neleisėtų susidaryti PGR produktui, aprėpiančiam visą CRISPR regioną. Patikrinti, ar

nevyksta dalinė integracija, *in vitro* reakcija buvo pakartota su tais pačiais mėginiais, naudojant radioaktyviai žymėtą dupleksą (4.8 pav.).

BL.	Bl. +	Bl. +	BL.	Bl. +	Bl. +	F1	F2	Ctrl
	C1C2	C1C2	k.	C1C2	C1C2			
		+ C4		k.	+ C4			
					k.			

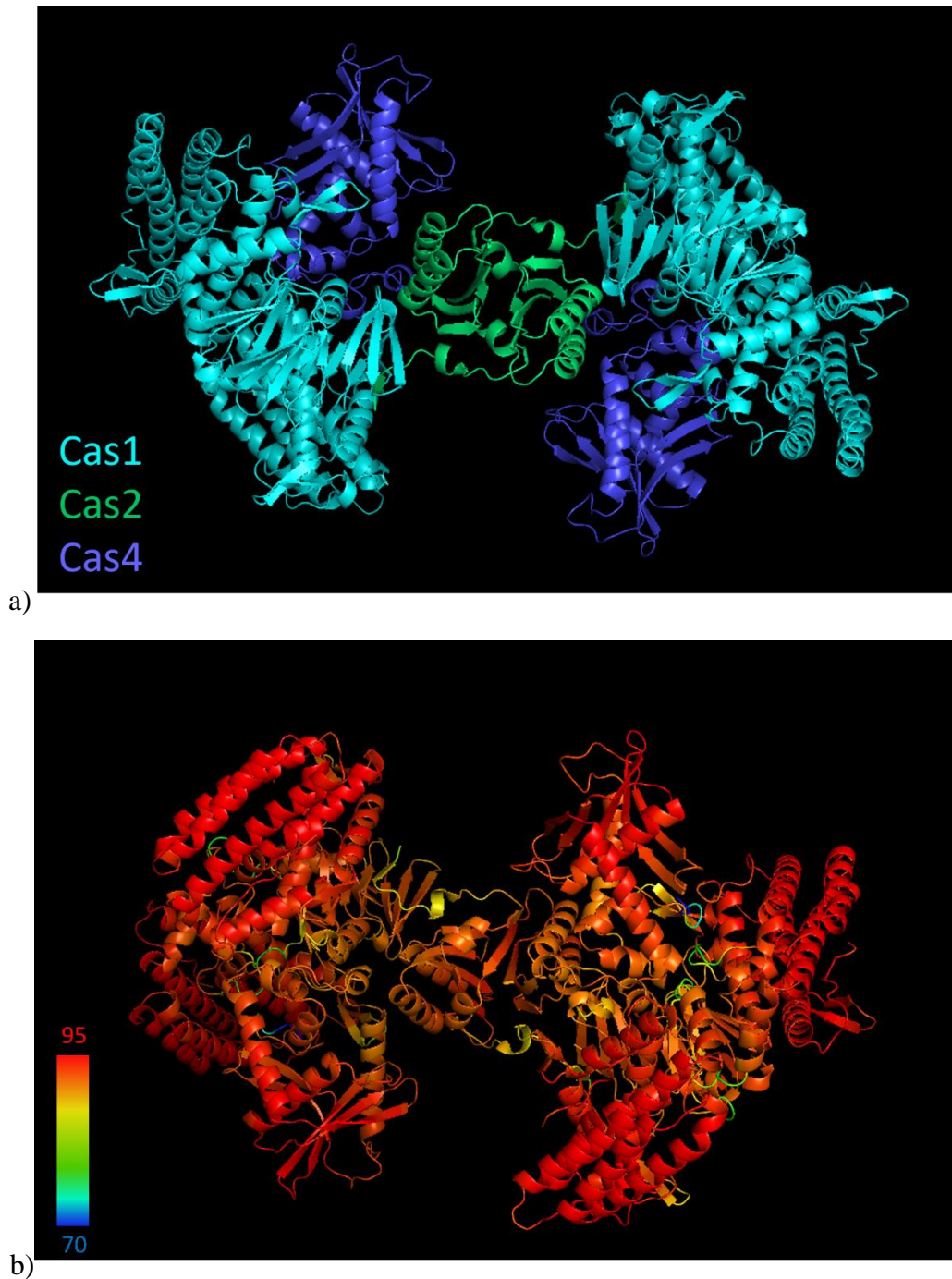


4.8 Pav. *In vitro* integracijos analizė. BL. – BL21(DE3) kamieno lizatas. K. – koncentruotas lizatas. F1 – His-Cas1+ Cas2-Strep + Cas4. F2 –Cas1+ Cas2 + His-Cas4. Ctrl – integracijos reakcija tik su buferiu ir vandeniu.

Sėkmingos integracijos metu, radioaktyviai žymėtas proskirtukas būtų kovalentiškai prijungiamas prie plazmidės CRISPR regione. Atitinkamai, matytųsi žymėti DNR fragmentai, kurie judėtų gelyje gerokai lėčiau nei žymėtas proskirtukas. Šiuo atveju matome, kad visomis sąlygomis integracijos nematome (lyginant su kontrole). V-F sistemos Cas1 ir Cas2 negeba integruoti substratą, jeigu juos ekspresuojame *E. coli* kamiene. Arba Cas1-Cas2 ekspresiją reikia toliau optimizuoti, arba kompleksui yra reikalingas papildomas faktorius, randamas šeimininko organizme, kuris nėra aptinkamas *E. coli*.

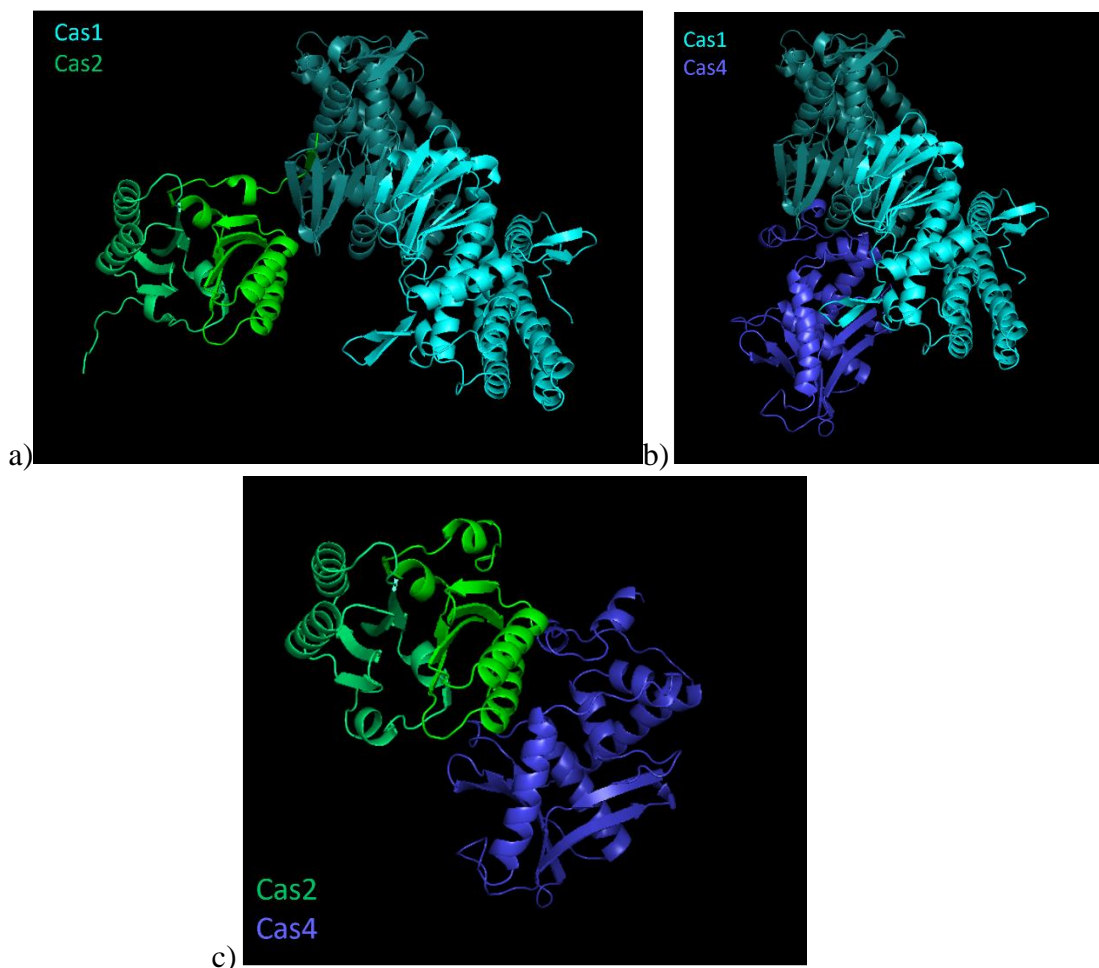
4.5 Modelio generavimas ir analizė

Šiuo darbo metu išėjo vieša prieiga prie AlphaFold2, kuri galėjome panaudoti sugeneruoti modelį tiriamam Cas1-Cas2-Cas4 kompleksui (4.9 pav. a)) (Jumper et al. 2021).



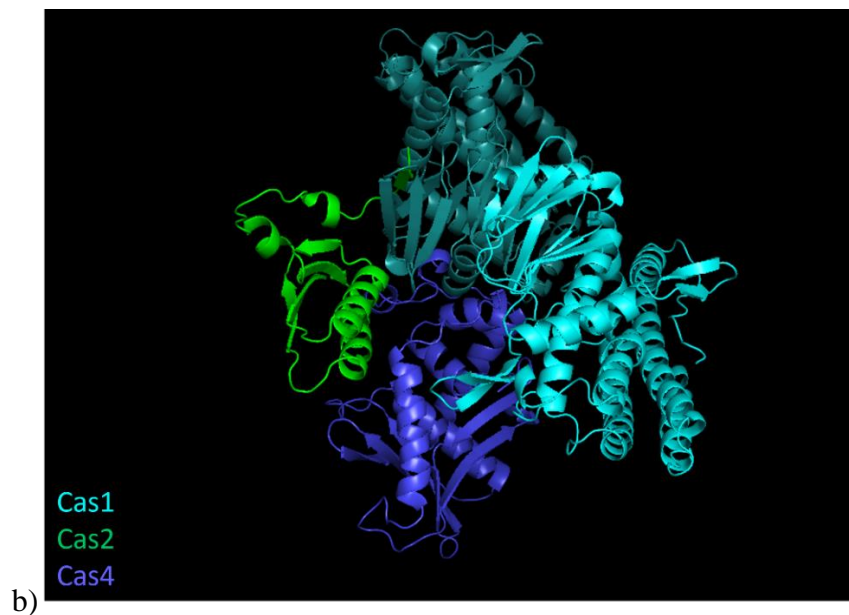
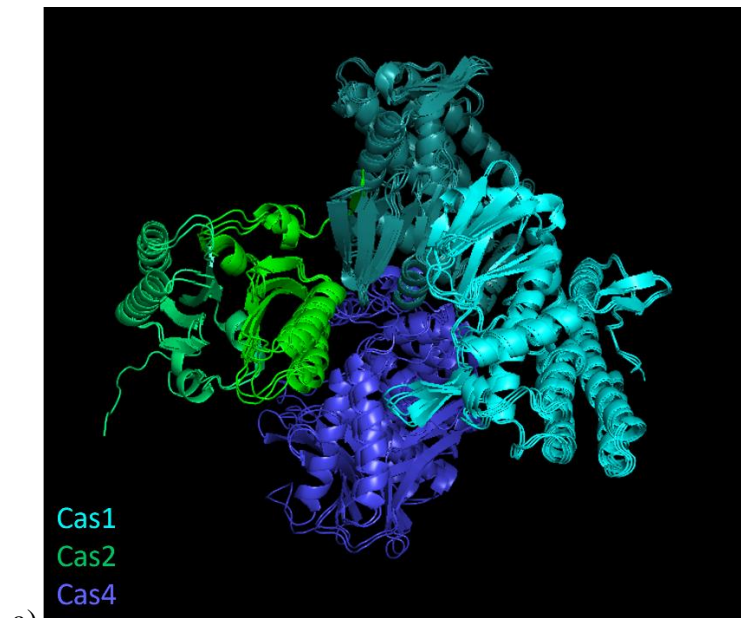
4.9 Pav. a) AlphaFold2 sugeneruotas Cas1-Cas2-Cas4 kompleksas. b) Cas1-Cas2-Cas4 modelis nuspalvintas pagal patikimumo lygį balais.

Sugeneruotas modelis rodo, kad Cas1-Cas2-Cas4 sudaro simetrinę struktūrą iš visų trijų baltymų. Tai yra struktūra, atitinkanti žinomus šių trijų baltymų išsidėstymų variantus kitose sistemose. Šis modelis padėjo nustatyti tinkamą inkarų poziciją siekiant gryninti kompleksą ateityje. Cas2 ir Cas4 yra maži baltymai ir jų N bei C galai yra labai arti sąveikos paviršių, todėl ant jų yra rizikinga dėti inkarus. Tai potenciali priežastis, kodėl Cas2 nebuvo aptinkamas išgrynintuose frakcijose. Mažiausią tikimybę suardyti struktūrą turi inkaras ant Cas1 C-galo. Papildomai matome, kad modelis pasižymi aukštu patikimumo lygiu (4.9 pav. b)). Patikimumas matuojamas balais nuo 0 iki 100. Nuo 70 balų modelis laikomas patikimu, nuo 90 – itin patikimu. Mažiau nei 70 balų surinkusios modelio dalys laikomos nepatikimomis, o žemiau nei 50 balų koreliuoja su visiškai neorganizuota struktūra (arba organizuota tik specifiniais atvejais kaip komplekse). Įdomu, kad mažiausiai balų surinko struktūros, jungiančios atskirus baltymus kaip Cas1 dimerą ir Cas4. Kadangi šis rezultatas gali koreliuoti su padidėjusiu lankstumu, tai atspindi turimas žinias, kad Cas4 elgiasi dinamiškai adaptacijos komplekse pagal PAM buvimą substrate arba ne.



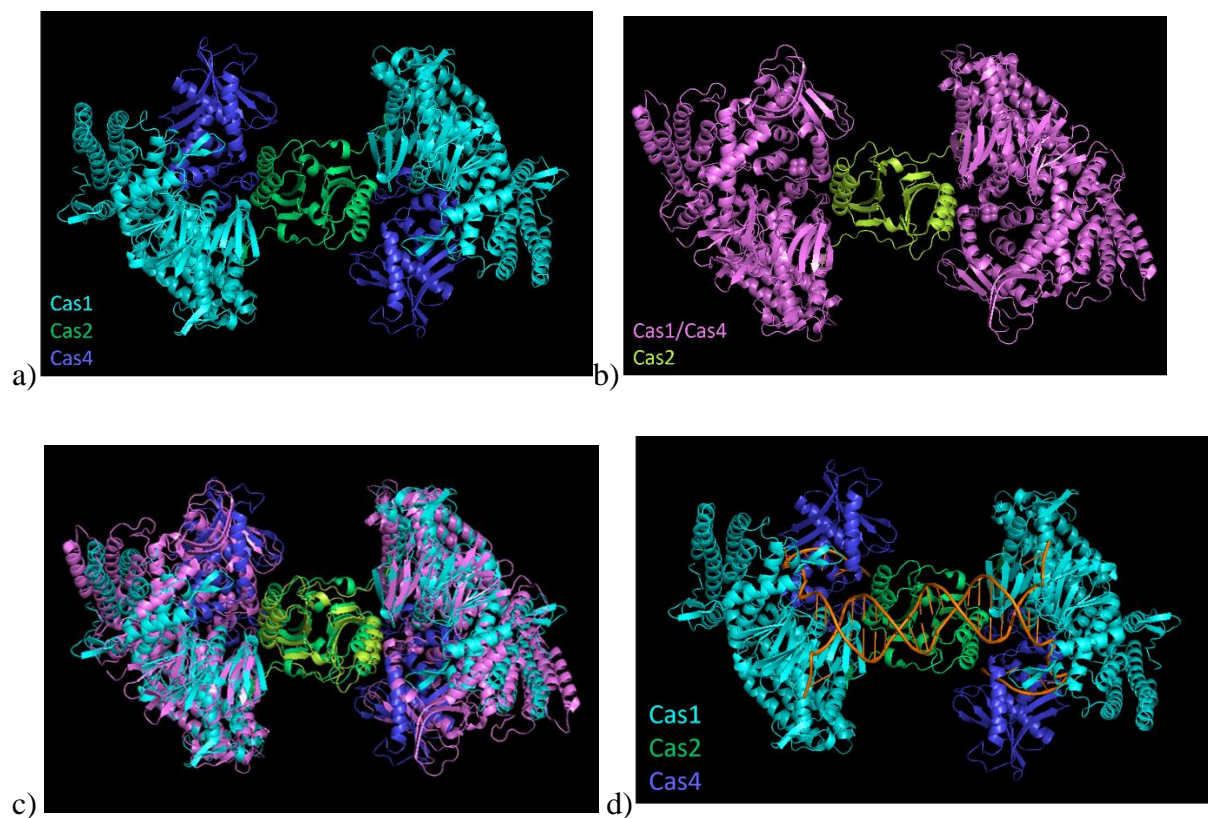
4.10 Pav. AlphaFold2 sugeneruoti subkompleksai. a) Cas1 dimeras + Cas2 dimeras. b) Cas1 dimeras + Cas4. c) Cas2 dimeras + Cas4.

Toliau, jei modeliuojame atskirus įmanomus subkompleksus (4.10 pav.), matome, kad visi baltymai išlaiko tas pačias pozicijas ir konformaciją vienas kito atžvilgiu. Tai leidžia manyti, kad Cas1-Cas2-Cas4 sudaro stabilų kompleksą ir nesutampa su modeliais, kur Cas4 ir Cas2 konkuruoja tarpusavyje. Tai yra aiškiai matyti, kai visi modeliai yra perdengiami kartu (4.11 pav.). Persidengiantis vaizdas beveik nesiskiria nuo (pusės) pilno modelio su visais baltymais.



4.11 Pav. AlphaFold2 sugeneruoti subkompleksai. a) Visų subkompleksų persidengiantis vaizdas b) Cas1 dimeras + Cas2(pusė dimero) + Cas4

Iš modeliuotų struktūrų galima manyti, kad PAM atpažinimą vykdo Cas4 baltymas. Jis ne tik sudaro darnią struktūrą su kitais baltymais, bet ir yra tinkamoje pozicijoje sąveikauti su preskirtuko galais. Tai galima aiškiai pamatyti tiesiogiai lyginant B17 sistemos modelį su jau žinoma *G. sulfurreducens* sistemos struktūra (4.12 pav.). Todėl tikėtina, kad tirtos B17 sistemos skirtukų integracijos mechanizmas yra panašus į kitų Cas4 turinčių CRISPR-Cas sistemų kaip *G. sulfurreducens* sistemos, kuriose Cas4 atpažįsta PAM. Cas4 baltymui surišant preskirtuką, tas duplekso galas yra blokuojamas integracijai, kas savo ruožtu nurodo, kuria kryptimi integruoti skirtuką. Atlikus pirmą integraciją, Cas4 nukleazinis aktyvumas galėtų pašalinti PAM, kas leistų integruoti procesuoto preskirtuko antrąjį galą.



4.12 Pav. a) B17 sistemos Cas1-Cas2-Cas4 modelis b) *G. sulfurreducens* Cas1/Cas4-Cas2 kompleksas (PDB ID: 7mi4). c) B17 ir *G. sulfurreducens* persidengiantys kompleksai d) B17 kompleksas su įstatytu preskirtuko iš *G. sulfurreducens* komplekso struktūros (PDB ID: 7mi4).

5. Išvados

1. Darbo metu sėkmingai sukonstruoti V-F potipio CRISPR-Cas sistemos baltymų raiškai *E. coli* ląstelėse skirti vektoriai. Tačiau dėl nepakankamos tirpių baltymų raiškos baltymų išgryninti nepavyko. Cas1 ir Cas4 buvo gauti mažo grynumo preparatai, o Cas2 baltymo ar Cas1-Cas2-Cas4 sudaromų kompleksų gryninimo metu neaptikta.
2. Naudojant adaptacijos baltymų raiškos vektoriais transformuoto *E. coli* kamieno ekstraktus naujų skirtukų įstatymo aktyvumo *in vitro* taip pat neaptikta. Naujų skirtukų įsistatymo neaptikta ir adaptacijos baltymų raiškos vektoriais bei CRISPR regioną turinčia plazmide transformuotose *E. coli* ląstelėse.
3. Naudojantis AlphaFold2 algoritmu buvo sudaryti aukštos kokybės V tipo Cas1-Cas2-Cas4 komplekso ir įvairių galimų subkompleksų struktūriniai modeliai. Jie atskleidė, kad visi 3 tiriamos sistemos adaptacijos baltymai turėtų sudaryti stabilų $(Cas1)_4(Cas2)_2(Cas4)_2$ integraciją vykdantį kompleksą, kurio veikimo mechanizmas turėtų būti artimas struktūriškai charakterizuotam *Geobacter sulfurreducens* CRISPR-Cas sistemos adaptacijos kompleksui.

Santrauka

CRISPR-Cas yra antivirsinė sistema suteikianti adaptyvaus imuniteto atitikmenį prokariotams. Ši sistema leidžia ląstelei pasisavinti svetimą genetinę medžiagą šabloną, kuris bus naudojamas atpažinti tą patį taikinį, ir įsistatyti jį į savo paties genomą. Adaptacija yra kertinė šios sistemos stadija, kada yra surišamas, atsirenkamas ir integruojamas šablonas į genomą esantį CRISPR lokusą. Integracijai būtini, bet ne visada pakankamai Cas1 ir Cas2 baltymai. Dažnai reikalingi papildomieji baltymai, kaip Cas4, kurių funkcija gali smarkiai skirtis tarp konkrečių sistemos tipų. Iš aibės CRISPR-Cas sistemų paskutiniu metu didelio susidomėjimo sulaukė V tipo sistemos, kadangi jose randami maži efektoriniai baltymai – Cas9 analogai. Šiuo metu apie šias sistemas, ypač apie adaptaciją, yra mažai žinoma. Šiame darbe buvo nagrinėjama V-F subtipo sistema, atrasta metagenominiuose duomenyse, turinti adaptacijos baltymus Cas1, Cas2 ir Cas4, bei efektorinį baltymą Cas14. Nebuvo žinoma, ar Cas4 atlieką panašią funkciją kaip ir kitose sistemose, ar turi unikalią funkciją. Išanalizavus adaptacijos komplekso struktūrą, atrodo tikėtina, jog V-F sistemos Cas4 turi labai daug panašumų su analogu iš kitų CRISPR-Cas sistemų.

Spacer acquisition in a type V CRISPR-Cas system

Summary

CRISPR-Cas is an antiviral system that provides adaptive immunity to prokaryotes. This system allows a cell to sample a fragment from invasive genetic material and incorporate it into its own genome. Adaptation is the core stage in providing immunity, when the sample is bound, processed and integrated into the CRISPR locus of the genome. Cas1 and Cas2 are proteins which are essential, but not necessarily sufficient for integration. Often, additional proteins are required, such as Cas4, which performs a vast array of functions throughout the different types of systems. Recently, type V CRISPR-Cas systems have received increased interest due to the presence of small effector proteins – Cas9 analogs. Currently, little is known about these systems, especially adaptation. In this work, the V-F subtype CRISPR-Cas system was analyzed, which was discovered within metagenomic data, and possesses the adaptation proteins Cas1, Cas2 and Cas4 alongside the small effector protein Cas12f. It was unknown if Cas4 acted like its analogues found in other systems, or if it exhibited unique functions. Analysis of the adaptation complex structural models revealed that the V-F subtype Cas4 protein similarly to those found in some other Cas4-encoding CRISPR-Cas systems is responsible for PAM recognition during spacer acquisition, and forms a stable integration complex with Cas1 and Cas2 proteins.

PADĖKA

Norėčiau labai nuoširdžiai padėkoti skyriaus vedėjui prof. Dr. Virginijui Šikšniui už suteiktą galimybę atlikti darbus BNST skyriuje, darbo vadovui Dr. Giedriui Sasnauskui už itin kantriai suteiktą pagalbą ir konsultacijas ruošiant darbą bei visiems skyriaus darbuotojams už nepriekaištingą darbo aplinką.

6. Literatūros sąrašas

1. Amitai, Gil and Rotem Sorek. 2016. "CRISPR–Cas Adaptation: Insights into the Mechanism of Action." *Nature Reviews Microbiology* 14:67.
2. Babu, Mohan, Natalia Beloglazova, Robert Flick, Chris Graham, Tatiana Skarina, Boguslaw Nocek, Alla Gagarinova, Oxana Pogoutse, Greg Brown, Andrew Binkowski, Sadhna Phanse, Andrzej Joachimiak, Eugene V Koonin, Alexei Savchenko, Andrew Emili, Jack Greenblatt, Aled M. Edwards, and Alexander F. Yakunin. 2011. "A Dual Function of the CRISPR–Cas System in Bacterial Antivirus Immunity and DNA Repair." *Molecular Microbiology* 79(2):484–502.
3. Béguin, Pierre, Nicole Charpin, Eugene V Koonin, Patrick Forterre, and Mart Krupovic. 2016. "Casposon Integration Shows Strong Target Site Preference and Recapitulates Protospacer Integration by CRISPR-Cas Systems." *Nucleic Acids Research* 44(21):10367–76.
4. Beloglazova, Natalia, Greg Brown, Matthew D. Zimmerman, Michael Proudfoot, Kira S. Makarova, Marina Kudritska, Samvel Kochinyan, Shuren Wang, Maksymilian Chruszcz, Wladek Minor, Eugene V Koonin, Aled M. Edwards, Alexei Savchenko, and Alexander F. Yakunin. 2008. "A Novel Family of Sequence-Specific Endoribonucleases Associated with the Clustered Regularly Interspaced Short Palindromic Repeats." *Journal of Biological Chemistry* 283(29):20361–71.
5. Bergh, Øivind, Knut Yngve Børshheim, Gunnar Bratbak, and Mikal Heldal. 1989. "High Abundance of Viruses Found in Aquatic Environments." *Nature* 340(6233):467–68.
6. Charpentier, Emmanuelle, John van der Oost, and Malcolm F. White. 2013. "CrRNA Biogenesis BT - CRISPR-Cas Systems: RNA-Mediated Adaptive Immunity in Bacteria and Archaea." Pp. 115–44 in, edited by R. Barrangou and J. van der Oost. Berlin, Heidelberg: Springer Berlin Heidelberg.
7. Dixit, Bhuvan, Karukriti Kaushik Ghosh, Gary Fernandes, Pankaj Kumar, Prerana Gogoi, and Manish Kumar. 2016. "Dual Nuclease Activity of a Cas2 Protein in CRISPR–Cas Subtype I-B of *Leptospira interrogans*." *FEBS Letters* 590(7):1002–16.
8. Doron, Shany, Sarah Melamed, Gal Ofir, Azita Leavitt, Anna Lopatina, Mai Keren, Gil Amitai, and Rotem Sorek. 2018. "Systematic Discovery of Antiphage Defense Systems in the Microbial Pangenome." *Science (New York, N.Y.)* 359(6379).
9. Fineran, Peter C., Matthias J. H. Gerritzen, María Suárez-Diez, Tim Künne, Jos Boekhorst, Sacha A. F. T. van Hijum, Raymond H. J. Staals, and Stan J. J. Brouns. 2014. "Degenerate Target Sites Mediate Rapid Primed CRISPR Adaptation." *Proceedings of the National Academy of Sciences* 111(16):E1629 LP-E1638.
10. Gasiunas, Giedrius, Rodolphe Barrangou, Philippe Horvath, and Virginijus Siksnys.

2012. “Cas9–CrRNA Ribonucleoprotein Complex Mediates Specific DNA Cleavage for Adaptive Immunity in Bacteria.” *Proceedings of the National Academy of Sciences* 109(39):E2579 LP-E2586.
11. Harrington, Lucas B., David Burstein, Janice S. Chen, David Paez-Espino, Enbo Ma, Isaac P. Witte, Joshua C. Cofsky, Nikos C. Kyrpides, Jillian F. Banfield, and Jennifer A. Doudna. 2018. “Programmed DNA Destruction by Miniature CRISPR–Cas14 Enzymes.” *Science* 362(6416):839 LP – 842.
 12. Heler, Robert, Poulami Samai, Joshua W. Modell, Catherine Weiner, Gregory W. Goldberg, David Bikard, and Luciano A. Marraffini. 2015. “Cas9 Specifies Functional Viral Targets during CRISPR–Cas Adaptation.” *Nature* 519:199.
 13. Hu, Chunyi, Cristóbal Almendros, Ki Hyun Nam, Ana Rita Costa, Jochem N. A. Vink, Anna C. Haagsma, Saket R. Bagde, Stan J. J. Brouns, and Ailong Ke. 2021. “Mechanism for Cas4-Assisted Directional Spacer Acquisition in CRISPR–Cas.” *Nature* 598(7881):515–20.
 14. Hung, Mien-Chie and Pieter C. Wensink. 1984. “Different Restriction Enzyme-Generated Sticky DNA Ends Can Be Joined in Vitro.” *Nucleic Acids Research* 12(4):1863–74.
 15. Hyman, Paul and Stephen T. B. T. *Advances in Applied Microbiology* Abedon. 2010. “Chapter 7 - Bacteriophage Host Range and Bacterial Resistance.” Pp. 217–48 in *Advances in Applied Microbiology*. Vol. 70. Academic Press.
 16. Jumper, John, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Židek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstern, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. 2021. “Highly Accurate Protein Structure Prediction with AlphaFold.” *Nature* 596(7873):583–89.
 17. Karvelis, Tautvydas, Greta Bigelyte, Joshua K. Young, Zhenglin Hou, Rimante Zedaveinyte, Karolina Budre, Sushmitha Paulraj, Vesna Djukanovic, Stephen Gasior, Arunas Silanskas, Česlovas Venclovas, and Virginijus Siksnys. 2020. “PAM Recognition by Miniature CRISPR–Cas12f Nucleases Triggers Programmable Double-Stranded DNA Target Cleavage.” *Nucleic Acids Research* 48(9):5016–23.
 18. Kieper, Sebastian N., Cristóbal Almendros, Juliane Behler, Rebecca E. McKenzie, Franklin L. Nobrega, Anna C. Haagsma, Jochem N. A. Vink, Wolfgang R. Hess, and Stan J. J. Brouns. 2018. “Cas4 Facilitates PAM-Compatible Spacer Selection during CRISPR Adaptation.” *Cell Reports* 22(13):3377–84.
 19. Koonin, Eugene V and Mart Krupovic. 2014. “Evolution of Adaptive Immunity from Transposable Elements Combined with Innate Immune Systems.” *Nature Reviews Genetics* 16:184.
 20. Koonin, Eugene V, Kira S. Makarova, and Feng Zhang. 2017. “Diversity, Classification and Evolution of CRISPR–Cas Systems.” *Current Opinion in Microbiology* 37:67–78.
 21. Krupovic, Mart, Kira S. Makarova, Patrick Forterre, David Prangishvili, and Eugene V Koonin. 2014. “Casposons: A New Superfamily of Self-Synthesizing DNA Transposons at the Origin of Prokaryotic CRISPR–Cas Immunity.” *BMC Biology* 12(1):36.
 22. Labrie, Simon J., Julie E. Samson, and Sylvain Moineau. 2010. “Bacteriophage Resistance Mechanisms.” *Nature Reviews Microbiology* 8(5):317–27.

23. Lee, Hayun, Yukti Dhingra, and Dipali G. Sashital. 2019. “The Cas4-Cas1-Cas2 Complex Mediates Precise Prespacer Processing during CRISPR Adaptation” edited by A. Ke, C. Wolberger, and M. F. White. *ELife* 8:e44248.
24. Makarova, Kira S., Nick V Grishin, Svetlana A. Shabalina, Yuri I. Wolf, and Eugene V Koonin. 2006. “A Putative RNA-Interference-Based Immune System in Prokaryotes: Computational Analysis of the Predicted Enzymatic Machinery, Functional Analogies with Eukaryotic RNAi, and Hypothetical Mechanisms of Action.” *Biology Direct* 1:7.
25. Makarova, Kira S., Yuri I. Wolf, Jaime Iranzo, Sergey A. Shmakov, Omer S. Alkhnbashi, Stan J. J. Brouns, Emmanuelle Charpentier, David Cheng, Daniel H. Haft, Philippe Horvath, Sylvain Moineau, Francisco J. M. Mojica, David Scott, Shiraz A. Shah, Virginijus Siksnys, Michael P. Terns, Česlovas Venclovas, Malcolm F. White, Alexander F. Yakunin, Winston Yan, Feng Zhang, Roger A. Garrett, Rolf Backofen, John van der Oost, Rodolphe Barrangou, and Eugene V Koonin. 2020. “Evolutionary Classification of CRISPR–Cas Systems: A Burst of Class 2 and Derived Variants.” *Nature Reviews Microbiology* 18(2):67–83.
26. Marraffini, Luciano A. 2015. “CRISPR-Cas Immunity in Prokaryotes.” *Nature* 526:55.
27. Moch, Clara, Michel Fromant, Sylvain Blanquet, and Pierre Plateau. 2016. “DNA Binding Specificities of Escherichia Coli Cas1–Cas2 Integrase Drive Its Recruitment at the CRISPR Locus.” *Nucleic Acids Research* 45(5):2714–23.
28. Nam, Ki Hyun, Fran Ding, Charles Haitjema, Qingqiu Huang, Matthew P. DeLisa, and Ailong Ke. 2012. “Double-Stranded Endonuclease Activity in Bacillus Halodurans Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Associated Cas2 Protein.” *Journal of Biological Chemistry* 287(43):35943–52.
29. Nuñez, James K., Lawrence Bai, Lucas B. Harrington, Tracey L. Hinder, and Jennifer A. Doudna. 2016. “CRISPR Immunological Memory Requires a Host Factor for Specificity.” *Molecular Cell* 62(6):824–33.
30. Nuñez, James K., Philip J. Kranzusch, Jonas Noeske, Addison V Wright, Christopher W. Davies, and Jennifer A. Doudna. 2014. “Cas1–Cas2 Complex Formation Mediates Spacer Acquisition during CRISPR–Cas Adaptive Immunity.” *Nature Structural & Molecular Biology* 21:528.
31. Nuñez, James K., Amy S. Y. Lee, Alan Engelman, and Jennifer A. Doudna. 2015. “Integrase-Mediated Spacer Acquisition during CRISPR–Cas Adaptive Immunity.” *Nature* 519:193.
32. Roberts, Richard J. and Kenneth Murray. 1976. “Restriction Endonuclease.” *CRC Critical Reviews in Biochemistry* 4(2):123–64.
33. Rohwer, Forest. 2003. “Global Phage Diversity.” *Cell* 113(2):141.
34. Rollie, Clare, Stefanie Schneider, Anna Sophie Brinkmann, Edward L. Bolt, and Malcolm F. White. 2015. “Intrinsic Sequence Specificity of the Cas1 Integrase Directs New Spacer Acquisition” edited by T. W. Nilsen. *ELife* 4:e08716.
35. Safari, Fatemeh, Mehrdad Sharifi, Safar Farajnia, Bahman Akbari, Mohammad Karimi Baba Ahmadi, Manica Negahdaripour, and Younes Ghasemi. 2020. “The Interaction of Phages and Bacteria: The Co-Evolutionary Arms Race.” *Critical Reviews in Biotechnology* 40(2):119–37.
36. Shiimori, Masami, Sandra C. Garrett, Brenton R. Graveley, and Michael P. Terns. 2018. “Cas4 Nucleases Define the PAM, Length, and Orientation of DNA Fragments Integrated at CRISPR Loci.” *Molecular Cell* 70(5):814-824.e6.
37. Sorek, Rotem, Victor Kunin, and Philip Hugenholtz. 2008. “CRISPR — a Widespread System That Provides Acquired Resistance against Phages in Bacteria

- and Archaea.” *Nature Reviews Microbiology* 6:181.
38. Sorek, Rotem, C. Martin Lawrence, and Blake Wiedenheft. 2013. “CRISPR-Mediated Adaptive Immune Systems in Bacteria and Archaea.” *Annual Review of Biochemistry* 82(1):237–66.
 39. Stern, Adi and Rotem Sorek. 2011. “The Phage-Host Arms Race: Shaping the Evolution of Microbes.” *BioEssays* 33(1):43–51.
 40. Suttle, Curtis A. 2005. “Viruses in the Sea.” *Nature* 437(7057):356–61.
 41. Wang, Jiuyu, Jiazhi Li, Hongtu Zhao, Gang Sheng, Min Wang, Maolu Yin, and Yanli Wang. 2015. “Structural and Mechanistic Basis of PAM-Dependent Spacer Acquisition in CRISPR-Cas Systems.” *Cell* 163(4):840–53.
 42. Weinbauer, Markus G. 2004. “Ecology of Prokaryotic Viruses.” *FEMS Microbiology Reviews* 28(2):127–81.
 43. Westra, Edze R., Daan C. Swarts, Raymond H. J. Staals, Matthijs M. Jore, Stan J. J. Brouns, and John van der Oost. 2012. “The CRISPRs, They Are A-Changin’’: How Prokaryotes Generate Adaptive Immunity.” *Annual Review of Genetics* 46(1):311–39.
 44. Wiedenheft, Blake, Kaihong Zhou, Martin Jinek, Scott M. Coyle, Wendy Ma, and Jennifer A. Doudna. 2009. “Structural Basis for DNase Activity of a Conserved Protein Implicated in CRISPR-Mediated Genome Defense.” *Structure* 17(6):904–12.
 45. Wilkinson, Martin, Gediminas Drabavicius, Arunas Silanskas, Giedrius Gasiunas, Virginijus Siksnys, and Dale B. Wigley. 2019. “Structure of the DNA-Bound Spacer Capture Complex of a Type II CRISPR-Cas System.” *Molecular Cell* 75(1):90-101.e5.
 46. Wright, Addison V, Joy Y. Wang, David Burstein, Lucas B. Harrington, David Paez-Espino, Nikos C. Kyrpides, Anthony T. Iavarone, Jillian F. Banfield, and Jennifer A. Doudna. 2019. “A Functional Mini-Integrase in a Two-Protein Type V-C CRISPR System.” *Molecular Cell* 73(4):727-737.e3.
 47. Wu, Chengyong, Dongmei Tang, Jie Cheng, Daojun Hu, Zejing Yang, Xue Ma, Haihuai He, Shaohua Yao, Tian-Min Fu, Yamei Yu, and Qiang Chen. 2021. “Mechanisms of Spacer Acquisition by Sequential Assembly of the Adaptation Module in *Synechocystis*.” *Nucleic Acids Research* 49(5):2973–84.
 48. Xiao, Yibei, Sherwin Ng, Ki Hyun Nam, and Ailong Ke. 2017. “How Type II CRISPR–Cas Establish Immunity through Cas1–Cas2-Mediated Spacer Integration.” *Nature* 550:137.
 49. Xin, Winnie and Michael Feiss. 1993. “Function of IHF in λ DNA Packaging: I. Identification of the Strong Binding Site for Integration Host Factor and the Locus for Intrinsic Bending in CosB.” *Journal of Molecular Biology* 230(2):492–504.
 50. Yan, Winston X., Pratyusha Hunnewell, Lauren E. Alfonse, Jason M. Carte, Elise Keston-Smith, Shanmugapriya Sothiselvam, Anthony J. Garrity, Shaorong Chong, Kira S. Makarova, Eugene V Koonin, David R. Cheng, and David A. Scott. 2019. “Functionally Diverse Type V CRISPR-Cas Systems.” *Science* 363(6422):88 LP – 91.
 51. Yoganand, Kakimani Nagarajan, Manasasri Muralidharan, Siddharth Nimkar, and Baskaran Anand. 2019. “Fidelity of Prespacer Capture and Processing Is Governed by the PAM-Mediated Interactions of Cas1-2 Adaptation Complex in CRISPR-Cas Type I-E System.” *Journal of Biological Chemistry* 294(52):20039–53.
 52. Yosef, Ido, Moran G. Goren, and Udi Qimron. 2012. “Proteins and DNA Elements Essential for the CRISPR Adaptation Process in *Escherichia Coli*.” *Nucleic Acids Research* 40(12):5569–76.