

**VILNIAUS UNIVERSITETAS**

**TAUTVYDAS KARVELIS**

**II TIPO CRISPR-Cas SISTEMOS: NUO FUNDAMENTINIŲ  
TYRIMŲ LINK GENOMŲ REDAGAVIMO**

Daktaro disertacijos santrauka  
Fiziniai mokslai, biochemija (04 P)

**Vilnius, 2016**

Disertacija rengta 2011-2015 m. Vilniaus universiteto Biotechnologijos institute.

**Mokslinis vadovas** – prof. dr. **Virginijus Šikšnys** (Vilniaus universitetas, fiziniai mokslai, biochemija - 04 P).

**Disertacija ginama Vilniaus universiteto Biochemijos mokslo krypties taryboje:**

**Pirmininkas** – prof. habil. dr. **Saulius Klimašauskas** (Vilniaus universitetas, fiziniai mokslai, biochemija - 04 P).

**Nariai:**

prof. habil. dr. **Kęstutis Sasnauskas** (Vilniaus universitetas, fiziniai mokslai, biochemija - 04 P);

dr. **Vytautė Starkuvienė** (Heidelbergo universitetas, fiziniai mokslai, biochemija - 04 P);

prof. dr. **Saulius Serva** (Vilniaus universitetas, fiziniai mokslai, biochemija - 04 P);

dr. **Rolandas Meškys** (Vilniaus universitetas, fiziniai mokslai, biochemija - 04 P).

Disertacija bus ginama viešame Biochemijos mokslo krypties posėdyje 2016 m. gruodžio 15 d. 12 val. Vilniaus universiteto Jungtinio gyvybės mokslų centro R401 salėje (Saulėtekio al. 7, LT-10257, Vilnius, Lietuva).

Disertacijos santrauka išsiųsta 2016 lapkričio 14 d.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu:

[www.vu.lt/lt/naujienos/ivykiu-kalendorius](http://www.vu.lt/lt/naujienos/ivykiu-kalendorius)

**VILNIUS UNIVERSITY**

**TAUTVYDAS KARVELIS**

**TYPE II CRISPR-Cas SYSTEMS: FROM BASIC STUDIES  
TOWARDS GENOME EDITING**

Summary of doctoral dissertation  
Physical science, biochemistry (04 P)

**Vilnius, 2016**

The work presented in this doctoral dissertation has been carried out at the Institute of Biotechnology, Vilnius University during 2011-2015.

**Supervisor** – prof. dr. **Virginijus Šikšnys** (Vilnius University, physical sciences, biochemistry - 04 P).

### **Evaluation board of dissertation of Biochemistry trend**

**Chairman** - prof. habil. dr. **Saulius Klimašauskas** (Vilnius University, physical sciences, biochemistry - 04 P).

#### **Members:**

prof. habil. dr. **Kęstutis Sasnauskas** (Vilnius University, physical sciences, biochemistry - 04 P);

dr. **Vytautė Starkuvienė** (Heidelberg University, physical sciences, biochemistry - 04 P);

prof. dr. **Saulius Serva** (Vilnius University, physical sciences, biochemistry - 04 P);

dr. **Rolandas Meškys** (Vilnius University, physical sciences, biochemistry - 04 P).

The thesis defence will take place at the Vilnius university, Life Sciences Center, R401 (Sauletekio av. 7, LT-10257, Vilnius, Lithuania) on 15<sup>th</sup> of December, 2016, at 12:00 noon.

The summary of doctoral dissertation was sent on 14<sup>th</sup> of November, 2016.

The thesis is available at the Library of Institute of Vilnius University and at the VU internet link:

[www.vu.lt/lt/naujienos/ivykiu-kalendorius](http://www.vu.lt/lt/naujienos/ivykiu-kalendorius)

# TURINYS

SANTRUMPOS .....	7
ĮVADAS .....	8
1. TYRIMŲ METODIKA .....	11
2. REZULTATAI IR JŲ APTARIMAS .....	25
2.1. crRNR ir tracrRNR yra būtinos Cas9 sukeltai DNR interferencijai <i>Streptococcus thermophilus</i> bakterijose .....	25
2.1.1. tracrRNR koduojamos sekos identifikavimas <i>S. thermophilus</i> DGCC7710 CRISPR3-Cas sistemoje .....	25
2.1.1. tracrRNR yra būtina <i>S. thermophilus</i> CRISPR3-Cas sistemos sukeltai DNR interferencijai <i>in vivo</i> .....	26
2.1.2. tracrRNR gryninasi su Cas9 baltymu .....	27
2.1.3. Sth3 Cas9-crRNR-tracrRNR komplekso surinkimas <i>in vitro</i> .....	28
2.1.4. Cas9 įtaka pre-crRNR:tracrRNR duplekso formavimuisi .....	29
2.2. Cas9 komplekso perprogramavimas DNR kirpimui <i>in vitro</i> .....	31
2.2.1. Sth3 Cas9 komplekso kaip įrankio panaudojimas DNR manipuliacijose .....	31
2.2.2. Genominės DNR karpymas <i>in vitro</i> .....	33
2.3. Genų redagavimas transfekuoiant <i>in vitro</i> surinktą Sth3 Cas9 RNP kompleksą .....	34
2.4. Cas9 DNR taikinio atpažinimas .....	37
2.4.1. Tiesioginis R-kilpos susidarymo stebėjimas .....	37
2.4.2. PAM mutacijos trukdo R-kilpos susidarymui, bet neturi įtakos jos stabilumui .....	39
2.4.3. Proskirtuko galo trumpinimas destabilizuoja Cas9 suformuotas R-kilpas .....	40
2.4.4. Cas9 inicijuoto R-kilpos susidarymo ir disociacijos modelis .....	41
2.5. CRISPR-Cas9 PAM sekos identifikavimas .....	42
2.5.1. Cas9 PAM sekų nustatymas .....	43
2.5.2. <i>Streptococcus pyogenes</i> ir <i>Streptococcus thermophilus</i> (CRISPR3 ir CRISPR1 sistemų) Cas9 atpažįstamų PAM sekų nustatymas .....	44
2.5.3. PAM ir RNR komponentų identifikavimas <i>Brevibacillus laterosporus</i> Cas9 baltymui .....	45
2.5.4. Augalų genomo redagavimas naudojant Blat Cas9 kompleksą .....	47
IŠVADOS .....	49
MOKSLINIŲ DARBŲ SĄRAŠAS .....	50
FINANSINĖ PARAMA .....	52

CURRICULUM VITAE.....	53
PADĒKA .....	54
SUMMARY.....	55
LITERATŪROS SARAŠAS.....	57

## SANTRUMPOS

Ap	ampicilinas
Blat	<i>Brevibacillus laterosporus</i>
bp	bazių pora(os)
Cas	<i>angl.</i> CRISPR associated
CRISPR	<i>angl.</i> clustered regularly interspaced short palindromic repeats
crRNR	CRISPR RNR
DSB	<i>angl.</i> double strand brake
HDR	<i>angl.</i> homology directed repair
INDEL	<i>angl.</i> insertion and/or deletion
kPGR	kiekybinis PGR
NHEJ	<i>angl.</i> non-homologous end joining
nt	nukleotidas(ai)
PAGE	poliakrilamidinio gelio elektroforezė
PAM	<i>angl.</i> protospacer adjacent motif
pre-crRNR	pirminė crRNR
RNP	ribonukleoproteininis
sgRNR	dirbtinai sujungta crRNR su tracrRNR ( <i>angl.</i> single guide RNA)
Spy	<i>Streptococcus pyogenes</i>
Sth1	<i>Streptococcus thermophilus</i> CRISPR1
Sth3	<i>Streptococcus thermophilus</i> CRISPR3
TALEN	<i>angl.</i> transcription activator-like effector nuclease
tracrRNA	trans-aktyvuojanti crRNR
ZFN	<i>angl.</i> zinc finger nuclease

## ĮVADAS

Šiuo metu DNR sekvenavimo technologijos leidžia nustatyti pilną genomo seką per kelias dienas, tačiau mūsų gebėjimas modifikuoti DNR sekas *in vivo*, norint suprasti ar pakeisti biologinius požymius ar funkcijas, atsilieka dėl įrankių genomo manipuliacijoms trūkumo. Natūralios ar sukonstruotos introninės-inteininės endonukleazės, ZFN (*angl.* zinc finger nuclease) ar TALEN (*angl.* transcription activator-like effector nuclease) meganukleazės, įvedančios dvigrandinį trūki DNR molekulėje inicijuoja DNR pažaidų taisymo mechanizmus, kurių rezultatas gali būti geno „išjungimas“, ar naujų sekų įstatymas, pasirodė daug žadančiais įrankiais. Tačiau šių įrankių konstravimas, keičiant aminorūgščių kodą siekiant pakeisti specifiškumą DNR sekai, reikalauja didelių baltymų inžinerijos pastangų, o rezultatai dažnai būna nuviliantys (Carroll, 2014; Porteus, 2016).

Nauja technologija, paremta Cas9 baltymu, sukėlė revoliuciją genomo redagavimo srityje. Kaip ir su visais revoliuciniais atradimais, Cas9 genomo redagavimo technologija prasidėjo nuo pastangų atsakyti fundamentinius biologinius klausimus, nesusijusius su genomo manipuliacijomis. 2007 metais CRISPR (*angl.* clustered regularly interspaced short palindromic repeats)-Cas (*angl.* CRISPR associated) sistemos buvo atrastos kaip naujos bakterijų apsaugos nuo virusų ir plazmidžių sistemos (Barrangou et al., 2007). Jos, panaudodamos mažas RNR molekules, vadinamas CRISPR RNR (crRNR), nukreipia ribonukleoproteininius kompleksus svetimų nukleorūgščių atpažinimui ir sunaikinimui (Brouns et al., 2008). RNR paremtas svetimų nukleorūgščių atpažinimas, kur nukleorūgštys atpažįstamos sudarant Watson-Crick bazių poras, suteikė galimybę pritaikyti tokias sistemas genomo redagavimui. Realiausias kandidatas pasirodė II tipo CRISPR-Cas sistemos dėl sąlyginai paprastos sandaros (Makarova et al., 2015). Paaiškėjo, kad II tipo sistemose Cas9 baltymas surišęs crRNR ir tracrRNR yra vienintelis baltymas reikalingas svetimų nukleorūgščių sunaikinimui (Deltcheva et al., 2011; Sapranauškas et al., 2011). Tokiame komplekse crRNR nukreipia Cas9 baltymą prie taikinio DNR, kai tuo tarpu tracrRNR dalyvauja nuo RNazės III priklausomame pre-crRNR brendime.

Cas9 komplekso taikiny yra sudarytas iš dviejų dalių: DNR sekos, komplementarios skirtuko sekai crRNR molekulėje (proskirtuko) ir kelių nukleotidų ilgio sekos, esančios greta proskirtuko, vadinamos PAM, kurios yra būtinos efektyviam dvigrandinio trūkio įvedimui (Garneau et al., 2010; Sapranauškas et al., 2011). PAM seka yra atpažįstama Cas9 baltymo ir ji gali skirtis priklausomai nuo Cas9 baltymo kilmės.

Šie fundamentiniai II tipo CRISPR-Cas sistemų tyrimai bakterijose paruošė dirvą genomo redagavimo eksperimentams eukariotinėse ląstelėse. Tai sutampa ne tik su projektu,



aprašyto šioje disertacijoje pradžia, bet ir Cas9 sukulto CRISPR tyrimų bumo, kuris lėmė vieną iš didžiausių atradimų molekulinėje biologijoje.

### **Darbo tikslai ir uždaviniai:**

1. Identifikuoti *S. thermophilus* DGCC7710 CRISPR3-Cas sistemos Cas9 ribonukleoproteininio komplekso komponentus būtinus funkciniam aktyvumui;
2. Ištirti, ar Cas9 kompleksas gali būti panaudotas programuojamam DNR kirpimui *in vitro* ir *in vivo*;
3. Ištirti Cas9 sąveikos su DNR taikiniu ir sekos atpažinimo mechanizmą;
4. Sukurti eksperimentinį metodą Cas9 atpažįstamų PAM sekų identifikavimui;
5. Praktiškai išbandyti PAM nustatymo metodą, apibūdinant Cas9 baltymą iš *Brevibacillus laterosporus* SSP360D4 kamieno CRISPR-Cas sistemos.

### **Mokslinis naujumas.**

Šiame darbe, mes pirmieji identifikavome visus komponentus būtinus aktyvaus Cas9 komplekso iš *Streptococcus thermophilus* DGCC7710 CRISPR3-Cas sistemos surinkimui ir panaudodami surinktą kompleksą pademonstravome galimybę kirpti norimą DNR seką *in vitro* ir *in vivo*. Šiais eksperimentais buvo pirmą kartą parodyta, kad Cas9 kompleksas gali būti panaudotas kaip įrankis DNR klonavimo eksperimentuose, taip pat genomo redagavimui, *in vitro* surinktą kompleksą chemiškai transfekuojuojant į eukariotines ląsteles. Norėdami suprasti Cas9 DNR taikinio atpažinimo ir kirpimo mechanizmą, tyrėme Cas9 komplekso formuojamos R-kilpos susidarymo dinamiką. Mes pirmieji pateikėme tiesioginius įrodymus dėl kryptingo R-kilpos susidarymo mechanizmo, prasidedančio nuo PAM sekos atpažinimo ir kryptingai besitęsiančio link taikinio DNR 5'-galo. Suprasdami PAM svarbą, sukūrėme metodą leidžiantį eksperimentiškai charakterizuoti PAM seką naujai identifikuotiems Cas9 ir charakterizavome II-C tipo Cas9 baltymą iš *Brevibacillus laterosporus*, taip išplėsdami įrankių tinkamų genomo redagavimui pasirinkimą.

### **Praktinė reikšmė.**

Cas9 paremta CRISPR-Cas sistema tapo lanksčiu įrankiu genomo manipuliacijos eksperimentuose. Priešingai nei anksčiau naudotos introninės-inteininės endonukleazės, ZFN ar TALEN meganukleazės, kur DNR atpažinimas buvo paremtas baltymo-DNR sąveika, RNR nukreipiamos Cas9 nukleazės yra daug pranašesnės, nes DNR atpažinimui

yra panaudojamas Watson-Crick bazių poravimasis tarp crRNR ir taikinio DNR. Šiame darbe aprašomi tyrimai gali būti pritaikyti nuo CRISPR-Cas9 sistemos charakterizavimo iki genomo redagavimo eksperimentų. Taip pat mes sukūrėme universalų metodą, leidžiantį identifikuoti RNR ir PAM sekas, kas suteikia galimybę charakterizuoti ortogonalius Cas9 baltymus, taip išplėčiant genomo redagavimo įrankių pasirinkimą. Šie rezultatai prisideda prie šiuo metu vykstančio genomų manipuliavimo proveržio paremto Cas9 technologija.

### **Ginamieji disertacijos teiginiai.**

1. crRNR ir tracrRNR yra būtinos Cas9 komplekso iš *S. thermophilus* CRISPR3-Cas sistemos aktyvumui;
2. Iš atskirų komponentų *in vitro* „surinktas“ Cas9 ribonukleoproteininis kompleksas gali būti užprogramuotas DNR kirpimui norimoje vietoje *in vitro* ir *in vivo*;
3. PAM seka panaudojama kaip Cas9 komplekso prisirišimo ir R-kilpos formavimosi iniciacijos vieta kryptingai crRNR hibridizacijai su taikinio DNR seka;
4. Sukurtas PAM nustatymo metodas;
5. Cas9 baltymas iš *B. laterosporus* atpažįsta NNNNCNDD PAM seką ir gali būti panaudotas genomo redagavimui.

# 1. TYRIMŲ METODIKA

## Reagentai ir fermentai

Visi naudoti reagentai buvo aukščiausio grynumo laipsnio. Fermentai: T4 polinukleotidų kinazė, FastAP šarminė fosfatazė, Pfu ir DreamTaq polimerazės, T4 DNR ligazė, FastDigest restrikcijos endonukleazės, Proteinazė K, RNazė A ir RNazė III buvo įsigytos iš „Thermo Fisher Scientific“. Phusion DNR polimerazės – iš „Thermo Fisher Scientific“ ir „NEB“. Fermentai naudoti laikantis gamintojo rekomendacijų.

## Molekulinės biologijos rinkiniai

„Rapid DNA Ligation Kit“, „GeneJET Gel Extraction Kit“, „GeneJET PCR Purification Kit“, „GeneJET Plasmid Miniprep Kit“, „TranscriptAid T7 High Yield Transcription Kit“ ir „GeneJET RNA Cleanup and Concentration Micro Kit“ įsigyti iš „Thermo Fisher Scientific“. „miRNeasy Mini Kit“, „QIAquick PCR Purification Kit“ – iš „Qiagen“. „SURVEYOR Mutation Detection Kit for Standard Gel Electrophoresis“ – iš „Transgenomic“. Rinkiniai naudoti laikantis gamintojo rekomendacijų.

## Bakterijų kamienai

Klonavimo procedūroms buvo naudotas *Escherichia coli* DH5 $\alpha$  kamienas, plazmidžių interferencijos eksperimentams – *E. coli* RR1 ir *E. coli* BL21 (DE3), Cas9 baltymų grynimui – *E. coli* BL21 (DE3) ir *E. coli* DH10B kamienai.

## Ląstelių linijos

CHO-K1 – adhezinių ląstelių linija, išvesta iš kiniško žiurkėno kiaušidžių ląstelių. HEK293T – žmogaus embriono inkstų adhezinių ląstelių linija, ekspresuojanti SV40 T antigeną.

## Plazmidės

Visos naudotos plazmidės yra aprašytos 1 lentelėje.

## RNR

Visos naudotos RNR yra aprašytos 2 lentelėje.

## Oligonukleotidai

Tyrimuose naudoti oligonukleotidai ir pradmenys yra aprašyti 3 lentelėje.

**1 lentelė. Tyrimams naudotos plazmidės.**

Plazmidė	Aprašymas	Nuoroda
pBAD24-Sth3-Cas9-CHis	<i>cas9</i> genas iš <i>S. thermophilus</i> CRISPR3-Cas sistemos įterptas į pBAD24-CHis raiškos vektorių	Karvelis et al., 2013
pBAD24-Sth1-Cas9-CHis	<i>cas9</i> genas iš <i>Streptococcus thermophilus</i> CRISPR1-Cas sistemos įterptas į pBAD24-CHis raiškos vektorių	Karvelis et al., 2015
pBAD24-Spy-Cas9-CHis	<i>cas9</i> genas iš <i>Streptococcus pyogenes</i> įterptas į pBAD24-CHis raiškos vektorių	Karvelis et al., 2015
pBAD24-Sth3-Cas9-NLS-CHis	Su SV40T branduolio lokalizacijos signaline seka (NLS) sulietas <i>cas9</i> genas iš <i>Streptococcus thermophilus</i> CRISPR3-Cas sistemos įterptas į pBAD24-CHis raiškos vektorių	Glemzaite et al., 2015
pBAD24-Blat-Cas9-CHis	<i>cas9</i> genas iš <i>Brevibacillus laterosporus</i> sistemos įterptas į pBAD24-CHis raiškos vektorių	Karvelis et al., 2015
pCRISPR3	Pilno ilgio <i>Streptococcus thermophilus</i> CRISPR3-Cas sistema įterpta į pACYC184 vektorių	Sapranauškas et al., 2011
pCRISPR3-Δt	pCRISPR3 plazmidė neturinti tracrRNR koduojančio geno	Karvelis et al., 2013
pCRISPR3-ΔtR	pCRISPR3 plazmidė turinti tracrRNR koduojančio geno fragmentą, apimančią sritį nuo transkripcijos pradžios vietos iki kartotiniui komplementarios srities pabaigos	Karvelis et al., 2013
p <tracrrna< td=""> <td>pCDF-DUET plazmidė su įterpta tracrRNR seka, kurios transkripcija yra priklausoma nuo T7 RNR polimerazės</td> <td>Karvelis et al., 2013</td> </tracrrna<>	pCDF-DUET plazmidė su įterpta tracrRNR seka, kurios transkripcija yra priklausoma nuo T7 RNR polimerazės	Karvelis et al., 2013
pS1	pUC18 plazmidė su įterptu 30 nt proskirtuku S1 (sutampančiu su S1 skirtuku <i>S. thermophilus</i> CRISPR3 srityje) ir NGGNG PAM seka (crRNR Cas9 komplekse hibridizuojasi tik su 20 nt iš 30 nt)	Sapranauškas et al., 2011
pS1'	pUC18 plazmidė su įterptu 20 nt proskirtuku S1' (atitinkančiu crRNR skirtuką) ir NGGNG PAM seka	Szczelkun et al., 2014
pS1'Δ1	pS1' plazmidė su 1 nt 5'-galo S1' proskirtuko patrupinimu (19 nt komplementarūs crRNR)	Szczelkun et al., 2014
pS1'Δ3	pS1' plazmidė su 3 nt 5'-galo S1' proskirtuko patrupinimu (17 nt komplementarūs crRNR)	Szczelkun et al., 2014

*1 lentelės tęsinys kitame puslapyje*

1 lentelės tęsinys

pS1'Δ5	pS1' plazmidė su 5 nt 5'-galo S1' proskirtuko patrupinimu (15 nt komplementarūs crRNR)	Szczelkun et al., 2014
pS1'Δ7	pS1' plazmidė su 7 nt 5'-galo S1' proskirtuko patrupinimu (13 nt komplementarūs crRNR)	Szczelkun et al., 2014
pS1'Δ9	pS1' plazmidė su 9 nt 5'-galo S1' proskirtuko patrupinimu (11 nt komplementarūs crRNR)	Szczelkun et al., 2014
pS1-G1C	pS1 plazmidė su G1C mutacija PAM sekoje (30 nt S1 proskirtukas)	Szczelkun et al., 2014
pS1-G2C	pS1 plazmidė su G2C mutacija PAM sekoje (30 nt S1 proskirtukas)	Szczelkun et al., 2014
pS1-G4C	pS1 plazmidė su G4C mutacija PAM sekoje (30 nt S1 proskirtukas)	Szczelkun et al., 2014
pMTC-RFP/EGFP reporterinė plazmidė	Reporterinė plazmidė, kurioje RFP genas yra perskirtas eGFP genu apsuptu 400 bp homoliginėmis sritimis (7 B pav.)	Glemzaite et al., 2015
pS1-5N-PAM	pTZ57R/T plazmidžių biblioteka su įterptu S1 proskirtuku ir 5 bp atsiktine PAM seka	Karvelis et al., 2015
pS1-7N-PAM	pTZ57R/T plazmidžių biblioteka su įterptu S1 proskirtuku ir 7 bp atsiktine PAM seka	Karvelis et al., 2015
pS1-GTCCCGAA	pUC18 plazmidė su įterptu S1 proskirtuku ir GTCCCGAA PAM seka	Karvelis et al., 2015
pS1-GTCaCGAA	pUC18 plazmidė su įterptu S1 proskirtuku ir GTCaCGAA PAM seka	Karvelis et al., 2015
pS1-GTCCtGAA	pUC18 plazmidė su įterptu S1 proskirtuku ir GTCCtGAA PAM seka	Karvelis et al., 2015
pS1-GTCCCGcA	pUC18 plazmidė su įterptu S1 proskirtuku ir GTCCCGcA PAM seka	Karvelis et al., 2015
pS1-GTCCCGAc	pUC18 plazmidė su įterptu S1 proskirtuku ir GTCCCGAc PAM seka	Karvelis et al., 2015
pS1-GTCCCGcc	pUC18 plazmidė su įterptu S1 proskirtuku ir GTCCCGcc PAM seka	Karvelis et al., 2015

**2 lentelė. Tyrimui naudotos RNR.**

<b>Pavadinimas</b>	<b>Seka (5'→3')</b>	<b>Gavimo būdas</b>
Sth3 tracrRNR3 (105 nt)	ggguaaauaaauugugguuuuaaaccauucgaaacaacacagcgaguuaaaauaaggcuuaguccguacucaacuugaaaagguggcaccgauucgguguuuuu	<i>In vitro</i> transkripcija
Sth1 tracrRNR1 (105 nt)	ggggguuaaagggggcgcuuacacaguuacuuaaaucuuugcagaagcuacaagauaaggcuucaugccgaaaucaacaccugucauuuuauggcaggguuuuu	<i>In vitro</i> transkripcija
Sth3 pre-crRNR3 (150 nt)	ggguagaaaagauauccuacgagguuuuagagcuguguuuuucgaaugguuccaaaacaaauucuaaacgcuaaaagggaagaggacaguuuagagcuguguuuuucgaaugguuccaaaacuacugcuguaauuagcuugguuguug	<i>In vitro</i> transkripcija
Sth3 pre-crRNR3 (94 nt)	gggaaauucuaaacgcuaaaagaggaagaggacaguuuagagcuguguuuuucgaaugguuccaaaacuacugcuguaauuagcuugguuguug	<i>In vitro</i> transkripcija
Sth1 pre-crRNR1 (94 nt)	ggguguuugacagcaaucaagauucgaauguguuuuuugucucucaagauuuuaguaacuguaacaauagacgaggagcuauuggcacaac	<i>In vitro</i> transkripcija
Sth3 tracrRNR (78 nt)	gggcgaaacaacacagcgaguuaaaauaaggcuuaguccguacucaacuugaaaagguggcaccgauucgguguuuuu	<i>In vitro</i> transkripcija
Sth3 tracrRNR (73 nt)	gggcgaaacaacacagcgaguuaaaauaaggcuuaguccguacucaacuugaaaagguggcaccgauucggug	<i>In vitro</i> transkripcija
Sth3 tracrRNR (68 nt)	gggcgaaacaacacagcgaguuaaaauaaggcuuaguccguacucaacuugaaaagguggcaccgauu	<i>In vitro</i> transkripcija
Sth3 tracrRNR (63 nt)	gggcgaaacaacacagcgaguuaaaauaaggcuuaguccguacucaacuugaaaagguggcac	<i>In vitro</i> transkripcija
Sth3 tracrRNR (58 nt)	gggcgaaacaacacagcgaguuaaaauaaggcuuaguccguacucaacuugaaaaggu	<i>In vitro</i> transkripcija
Sth3 tracrRNR (53 nt)	gggcgaaacaacacagcgaguuaaaauaaggcuuaguccguacucaacuugaa	<i>In vitro</i> transkripcija
Sth3 tracrRNR (48 nt)	gggcgaaacaacacagcgaguuaaaauaaggcuuaguccguacucaac	<i>In vitro</i> transkripcija
Sth3 tracrRNR (43 nt)	gggcgaaacaacacagcgaguuaaaauaaggcuuaguccguac	<i>In vitro</i> transkripcija
Sth3 tracrRNR (33 nt)	gggcgaaacaacacagcgaguuaaaauaaggcu	<i>In vitro</i> transkripcija
Sth3 crRNR (S1)	cgcuaaaagaggaagaggacaguuuagagcuguguuuuucg	Sintetinis oligonukleotidas

2 lentelės tęsinys kitame puslapyje

2 lentelės tęsinys

Sth3 pre-crRNR (N1)	ggguagaaaagauauccuacgagguuuuagagcuguguuguuucgaaugguuccaaaacugucaugauaaauaugguuucuuagacgucguuuuagagcuguguuguuucgaaugguuccaaaacggauccucuacgccggacgcaucgug	<i>In vitro</i> transkripcija
Sth3 pre-crRNR (N2)	ggguagaaaagauauccuacgagguuuuagagcuguguuguuucgaaugguuccaaaacacgagccggaagcauaaaguguaaaagccugguuuuagagcuguguuguuucgaaugguuccaaaacggauccucuacgccggacgcaucgug	<i>In vitro</i> transkripcija
Sth3 pre-crRNR ( $\lambda$ gDNR)	ggguagaaaagauauccuacgagguuuuagagcuguguuguuucgaaugguuccaaaaccgggaggggaagcugcaugaucggauguuauuuuagagcuguguuguuucgaaugguuccaaaacggauccucuacgccggacgcaucgug	<i>In vitro</i> transkripcija
Sth3 pre-crRNR ( <i>E. coli</i> gDNR)	ggguagaaaagauauccuacgagguuuuagagcuguguuguuucgaaugguuccaaaacucaagggagaauagaggcucucguugcauuuuuagagcuguguuguuucgaaugguuccaaaacggauccucuacgccggacgcaucgug	<i>In vitro</i> transkripcija
Sth3 crRNR (RASGEF1C)	gcucccggggcucgaugaagguuuuagagcuguguuguuucg	Sintetinis oligonukleotidas
Sth3 crRNR (ARL15)	ugaaucgugaaaucugcucaguuuuagagcuguguuguuucg	Sintetinis oligonukleotidas
Sth3 crRNR (NC)	cgcuaaagaggaagaggacaguuuuagagcuguguuguuucg	Sintetinis oligonukleotidas
Sth3 crRNR (EGFP-L1)	cuucagggucagcuugccguguuuuagagcuguguuguuucg	Sintetinis oligonukleotidas
Sth3 crRNR (EGFP-L2)	gcugaagggcaucgacuucaguuuuagagcuguguuguuucg	Sintetinis oligonukleotidas
Sth3 crRNR (DNMT3B)	gcugaauuacucacgccccaguuuuagagcuguguuguuucg	Sintetinis oligonukleotidas
Sth3 crRNR (PIIB)	guguauuuugaccuacgaauuuuagagcuguguuguuucg	Sintetinis oligonukleotidas
Spy crRNR (S1)	cgcuaaagaggaagaggacaguuuuagagcuauugcuguuuug	Sintetinis oligonukleotidas
Spy tracrRNR	gggaaacagcauagcaaguuaaaauaaggcuaguccguuaucacuugaaaaaguggcaccgagucggugcuuuuuuu u	<i>In vitro</i> transkripcija

2 lentelės tęsinys kitame puslapyje





**3 lentelė. Tyrimui naudoti oligonukleotidai ir pradmenys.**

<b>Oligonukleotidas</b>	<b>Seka (5'→3')</b>	<b>Aprašymas</b>
GG-322	gtacggactaagccttattttaactcgcgtgtgtgt	anti-tracrRNR zondas „Northern blot“ analizei
GG-321	tcgaaacaacacagctctaaaactgcctcttctcttagc	anti-crRNR zondas „Northern blot“ analizei
dir-pS1	gcgtaagtctcgagaactagttccgtaagatgctttctgtgact	Pradmenys naudoti 2,1 kbp fragmento padauginimui nuo pS1 plazmidės ir jos variantų
rev-pS1	gcgtaagtgcggccgcttccactgagcgtcaga	
dir-pUC19	gcgtaagtctcgagaactagtagaatagtgtatgcggcgacc	Pradmenys naudoti 1,0 kbp fragmento, pažymėto biotinu/digoksinu, padauginimui nuo pUC19 plazmidės
rev-pUC19	gcgtaagtgcggccgctgaccatgattacgccaagc	
TK-36	cagcaattataagagatgtatcagaagaagatgc	397 bp hibridizacijos zondo padauginimas nuo <i>E. coli</i> BL21(DE3) genomines DNR
TK-34	gcacctttattccaactgttctttgattaga	
pr-hs1f	tgctgctcgatgcacaggt	Pradmenys naudoti kPGR reakcijoje padauginti RASGEF1C (HS1) sritį
pr-hs1r	catcttcaccttctgctgag	
pr-hs2f	ccaattataagacagatgcctag	Pradmenys naudoti kPGR reakcijoje padauginti ARL15 (HS2) sritį
pr-hs2r	gccaacttctgtgaaactacact	
EGFP_L1/L2-f	agggcgaggagctgttcacc	eGFP DNR fragmento (592 bp) naudoto Surveyor nukleazės metode padauginimas
EGFP_L1/L2-r	tagtggtgtcgggcagcag	
DNMT3B-f	tgagaaggagccacttgctt	DNMT3B DNR fragmento (544 bp) naudoto Surveyor nukleazės metode padauginimas
DNMT3B-r	gaccaagaacgggaaagtca	
PPIB-f	gaacttaggctccgctcctt	PPIB DNR fragmento (505 bp) naudoto Surveyor nukleazės metode padauginimas
PPIB-r	ctctgcaggtcagtttgctg	
GG-821N	tgaccatgattacgaattcnnnnntgcctcttctcttagcgagc	Oligonukleotidai naudoti atsiktinių 5 bp PAM bibliotekos konstravimui (pS1-5N-PAM)
GG-820	aaggatccccgggtaccgagctgctgcgctaaagaggaagaggac	

3 lentelės tęsinys kitame puslapyje

3 lentelės tęsinys

GG-940-G	gtgcacgccggcgacgttgggtcaactnngnnnnntgcctcttctctttagcgtttag	Oligonukleotidai naudoti atsitiktinių 7 bp PAM bibliotekos konstravimui (pS1-7N-PAM)
GG-940-C	gtgcacgccggcgacgttgggtcaactnncnnnnntgcctcttctctttagcgtttag	
GG-940-A	gtgcacgccggcgacgttgggtcaactnnannnnntgcctcttctctttagcgtttag	
GG-940-T	gtgcacgccggcgacgttgggtcaactnntnnnnntgcctcttctctttagcgtttag	
GG-939	gactagacctgcaggggatcccgtcgacaattctaacgctaaagaggaagaggac	
TK-119	gagctcgctaaagaggaagagg	Pradmenys naudoti 5N PAM bibliotekos padauginimas PGR pagalba
pUC-dir	gccagggtttcccagtcacga	
TK-113	gaaattctaacgctaaagaggaagagg	Pradmuo (kartu su pUC-dir), naudotas 7N PAM bibliotekos padauginimas PGR pagalba
JKYS800.1	ctacacttttcctacacgacgctctccgatctaagtgagctcgctaaagaggaaga	Pradmenys skirti įvesti specifinę seką 5N padaugintos bibliotekos Illumina sekvenavimui
JKYS803	caagcagaagacggcatacagactctccgatctgaattcgagctcggtacct	
JKYS921.1	ctacacttttcctacacgacgctctccgatctggaataaacgctaaagaggaagagg	Pradmenys skirti įvesti specifinę seką 7N padaugintos bibliotekos Illumina sekvenavimui
JKYS812	caagcagaagacggcatacagactctccgatctcggcgacgttgggtc	
JKYS557	aatgatacggcgaccaccgagatctacacttttcctacacg	Universalūs pradmenys skirti antriniam PGR padauginimui Illumina sekvenavimui
JKYS558	caagcagaagacggcata	
TK-117	cggcattctgctgaaccgctctccgatct	Oligonukleotidai naudoti adapterio surinkimui
TK-111	gatcggaaagagcgggtcagcaggaatgccg	
JKYX1.1	ctacacttttcctacacgacgctctccgatctaaggggcgctggcctcctagtc	Pradmenys skirti <i>Ms45</i> egzono 1 taikinio padauginimui
JKYS178Rd	caagcagaagacggcatacagactctccgatctcgggctggcattgtctctg	
JKYS1083.1	ctacacttttcctacacgacgctctccgatctggaagggcaggttcgcaaacct	Pradmenys skirti <i>Ms45</i> egzono 4 taikinio padauginimui
JKYS1084	caagcagaagacggcatacagactctccgatctctccgagacaacaaactgcaggt	
JKYX2.1	ctacacttttcctacacgacgctctccgatctaagggggccggacgcggtgtt	Pradmenys skirti greta <i>liguleless-1</i> esančio taikinio padauginimui
JKYX3	caagcagaagacggcatacagactctccgatcttacctgcgaggtgcaaagtctac	

## Plazmidžių interferencija

Plazmidžių interferencijos tyrimas buvo atliktas kaip aprašyta anksčiau (Sapranauškas et al., 2011). *E. coli* RR1 kamienas buvo transformuojamas naudojant 1 ng, BL21 (DE3) kamienas – 25 ng ir HS115 kamienas – 100 ng pUC18 ar pS1 plazmidės (1 lentelė). Transformuotos bakterijos buvo auginamos ant LB terpės su atitinkamais antibiotikais ir 0,1 mM IPTG (BL21 (DE3) kamieno atveju). Visi transformacijos eksperimentai buvo pakartoti mažiausiai tris kartus.

## „Northern blot“ hibridizacija

„Northern blot“ hibridizacija buvo atliekama kaip aprašyta anksčiau (Gasiunas et al., 2012). Cas9 surištos RNR molekulės buvo išskirtos iš Strep-Tactin išgryninto Cas9-RNR komplekso (Gasiunas et al., 2012), panaudojant „miRNeasy Mini Kit“ (Qiagen). RNR buvo hibridizuojama su [ $\gamma$ -<sup>32</sup>P]ATP (Hartmann Analytic) pažymėtu 36 nt oligodeoksinukleotidu GG-322 (3 lentelė) komplementariu tracrRNR ar 42 nt ilgio žymėtu oligodeoksinukleotidu GG-321 (3 lentelė) komplementariu crRNR. RNR dydis buvo įvertintas lyginant su „Decade“ RNR ilgio standartais (Ambion) ir skirtingo ilgio RNR transkriptais.

## Cas9 baltymų raiška ir gryninimas

Sth1, Sth3, Sth3-NLS ir Spy Cas9 baltymai buvo ekspresuoti *E. coli* DH10B kamiene, Blat Cas9 – *E. coli* BL21 (DE3) kamiene. Bakterijos buvo auginamos LB terpėje su ampicilinu (100 mg/ml) 37°C. Optiniam tankiui (OD<sub>600</sub>) pasiekus 0,5, temperatūra buvo sumažinta iki 16°C. Cas9 baltymų ekspresija indukuota su arabinoze (0,2%) 20 val. Ląstelės buvo suardytos ultragarsu, o ląstelių nuolaužos pašalintos centrifuguojant. Cas9 gryninti skysčių chromatografijos metodu, naudojant Ni<sup>2+</sup> chelatinę ir heparino giminingumo kolonėles.

## RNR sintezė

Tyrimuose naudotos RNR molekulės buvo įsigytos kaip sintetiniai oligoribonukleotidai (Metabion) ar susintetinti *in vitro* transkripcijos pagalba, naudojant „TranscriptAid T7 High Yield Transcription Kit“ (Thermo Fisher Scientific) rinkinį. *In vitro* transkripcijai, T7 promotorius buvo įvedamas PGR pagalba, matrica naudojant plazmidės ar DNR fragmentus „surinktus“ iš oligonukleotidų. Po *in vitro* transkripcijos RNR buvo išgrynintos naudojant „GeneJET RNA Cleanup and Concentration Micro Kit“ (Thermo Fisher Scientific) rinkinį. Visų naudotų RNR sekos yra pateiktos 2 lentelėje.

RNR grandinių sulydymo aktyvumo tyrimui, tracrRNR, defosforilintos su FastAP fosfataze (Thermo Fisher Scientific) buvo pažymėtos naudojant [ $\gamma$ -<sup>33</sup>P]ATP (Hartmann Analytic) žymę T4 polinukleotidų kinazės (Thermo Fisher Scientific) pagalba.

### **Cas9 ribonukleoproteininių (RNP) kompleksų surinkimas**

Cas9 ribonukleoproteininiai (RNP) kompleksai buvo surenkami sumaišant Cas9 baltymą su sulydytu crRNR ir tracrRNR dupleksu 1:1 moliniu santykiu ir inkubuojant 37°C 1 val. Cas9 RNP komplekso susirinkimui naudojant pre-crRNR, gautas *in vitro* transkripcijos būdu nuo PGR produktų, Cas9 baltymas buvo maišomas su pre-crRNR ir tracrRNR transkriptais 1:1:2 moliniu santykiu ir inkubuojama 37°C 30 min. Po inkubacijos buvo pridėdama *E. coli* RNase III (Thermo Fisher Scientific) ir inkubuojama papildomas 30 min. 37°C.

### **Plazmidinės DNR karpymas**

Cas9 RNP komplekso DNR kirpimo reakcijos buvo pradamos superspiralizuotą plazmidinę DNR (3 nM) sumaišant su surinktu Cas9 RNP kompleksu (50 nM). Atitinkamais laiko intervalais iš reakcijos mišinio buvo išimama dalis tirpalo ir reakcija sustabdoma sumaišant su fenolio/chloroformo tirpalu. Vandeninė fazė nusiurbama ir sumaišoma su užnešimo dažu. Po elektroforezės agarozės gelis dažomas etidžio bromido tirpalu ir densitometriškai analizuojamas ImageJ (National Institutes of Health) programos pagalba, įvertinant superspiralizuotos (SC), nikuotos (OC) ir linijinės (FLL) DNR formų santykius. Reakcijų greičio konstantos apskaičiuojamos kaip aprašyta ankščiau (Zaremba et al., 2006).

### **RNR grandinių sulydymo aktyvumo tyrimas**

Cas9 grandinių sulydymo aktyvumas buvo tiriamas 10  $\mu$ l reakcijos tūryje, sumaišant 5 nM pre-crRNR (94 nt) su įvairios koncentracijos Cas9 baltymu ir 5 nM tracrRNR (105 nt) pažymėtos radioaktyve žyme ir inkubuojant 10 min. 37°C. Reakcijos mišinyje taip pat buvo 1  $\mu$ g nespecifinės 2 kb ilgio RNR. Po reakcijų sustabdymo panaudojant Proteinazę K ir nežymėtą tracrRNR, reakcijos produktai buvo analizuojami nedenatūruojančios PAGE elektroforezės pagalba.

### **$\lambda$ DNR karpymas**

Reakcijos buvo atliekamos sumaišant  $\lambda$  DNR (Thermo Fisher Scientific) su surinktu Cas9 kompleksu (1:1 v/v santykiu) ir inkubuotos 37°C. Galutiniame reakcijos mišinyje (100  $\mu$ l) buvo 2  $\mu$ g  $\lambda$  DNR ir 50 nM Cas9 komplekso. Atitinkamais laiko intervalais iš reakcijos mišinio buvo išimama dalis tirpalo ir reakcija sustabdoma sumaišant su

fenolio/chloroformo tirpalu. Vandeninė fazė nusiurbiamą ir sumaišoma su užnešimo dažų. Po elektroforezės agarozės gelis dažomas etidžio bromido tirpalu.

### ***E. coli* genominės DNR karpymas**

Reakcijos buvo atliekamos *E. coli* (BL21) genominę DNR sumaišant su surinktu Cas9 kompleksu (1:1 v/v santykiu) ir inkubuojant 3 val. 37°C. Galutiniame reakcijos mišinyje (300 µl) buvo 30 µg genominės DNR ir 1 µM Cas9 komplekso. Po inkubacijos buvo pridėdama PstI (Thermo Fisher Scientific) restrikcijos endonukleazės ir mėginiai laikomi 16 val. 37°C. Reakcijos buvo stabdomos pridėdant Proteinazės K (0.5 mg/ml, Thermo Fisher Scientific) ir SDS (0,5%, w/v) inkubuojant 30 min. 55°C. RNR pašalinimui, reakcijos mėginiai buvo papildomai inkubuojami kambario temperatūroje 30 min. su RNaze A (0,25 mg/ml, Thermo Fisher Scientific). Paveikus reakcijos mišinius fenoliu/chloroformu ir DNR išsodinusi izopropanoliu, reakcijos produktai buvo analizuojami „Southern blot“ metodo pagalba.

### **Žmogaus genominės DNR karpymas**

Žmogaus genominė DNR išskirta iš smegenų audinio buvo gauta iš Dr. Artūro Petronio (Psichikos ir priklausomybės ligų centras, Torontas, Kanada). Cas9-HS1 ir Cas9-HS2 kompleksai buvo atitinkamai nukreipti į RASGEF1C ar ARL15 sritis. 1 µg DNR buvo sumaišoma su 100 nM kiekvieno Cas9 komplekso 100 µl reakcijos tūryje. DNR kirpimo reakcijos produktai buvo analizuojami kiekybinės PGR (kPGR) pagalba.

### **„Southern blot“ hibridizacija**

„Southern blot“ hibridizacija buvo atliekama kaip aprašyta (Sambrook et al., 1989), su tam tikrais pakeitimais. Elektroforezės pagalba išfrakcionuota DNR (10 µg) buvo perkelta iš 1% agarozės gelio ant „SensiBlot Plus“ nailoninės membranos (Thermo Fisher Scientific) pusiau sauso pernešimo būdu. Po DNR denatūracijos ir pririšimo prie membranos, buvo atliekama prehibridizacija su 6× SSC buferiniu tirpalu, turinčiu 0,5% SDS ir 100 µg/ml nespecifinės DNR 65°C 1 val. Hibridizacija buvo atliekama pridėdant radioaktyviai žymėtą DNR zondą, gautą PGR pagalba padauginus atitinkamą fragmentą (397 bp) nuo *E. coli* BL21 (DE3) genominės DNR (pradmenys TK-36/TK-34, 3 lentelė). Po 16 val. hibridizacijos 65°C, membrana buvo plaunama 2× SSC, 0,5% SDS ir 2× SSC, 0,1% SDS tirpalais kambario temperatūroje ir išdžiovinta.

### **DNR kiekio įvertinimas kPGR metodu**

Žmogaus DNR kirpimo reakcijos produktai buvo analizuojami kiekybinės PGR (kPGR) pagalba naudojant pradmenis surašytus 3 lentelėje. 30 ng DNR buvo naudojama

kaip matrica kPGR reakcijoje (25  $\mu$ l) su SYBR Green reagentu (Maxima SYBR Green Master Mix 2 $\times$ , Thermo Fisher Scientific). Santykinis DNR kiekio įvertinimas buvo atliekamas panaudojant  $2^{-\Delta\Delta C_t}$  metodą (Livak and Schmittgen, 2001).

### **Cas9 kompleksų transfekcija į ląsteles**

Dieną prieš transfekciją CHO-K1 ar HEK293T ląstelės buvo užsėjamos į 24 šulinėlių plokštelę  $7 \times 10^4$  ląstelių 1 ml RPMI-1640 terpės (su 10% FBS). Prieš transfekcijos eksperimentus (transfekuojuojant baltymus) ląstelių mitybinė terpė pakeičiama į terpę be serumo. DNR transfekcijos atveju, plazmidinė DNR buvo sumaišoma su TurboFect *in vitro* DNR transfekcijos agentu (Thermo Fisher Scientific) laikantis gamintojo siūlomų rekomendacijų. Transfekcijos efektyvumas buvo vertinamas „Guava EasyCyte 8HT“ (Millipore) tėkmės citometro pagalba, duomenis analizuojant „Guava CytoSoft 5.3“ programine įranga (Millipore).

DNR ir Cas9 komplekso kotransfekcijos atveju, 0,5  $\mu$ g DNR ir 1  $\mu$ g Cas9 komplekso (papildomai pridedant 300 ng crRNR:tracrRNR duplekso efektyvesniam komplekso formavimuisi su polimeriniu katijoniniu TurboFect reagentu) sumaišoma skirtinguose mėgintuvėliuose su beserumine mitybine terpe (50  $\mu$ l). Pridėjus po 0,5  $\mu$ l TurboFect transfekcijos agento, mėginiai 15-20 min. inkubuojami kambario temperatūroje ir tada, tuo pačiu metu, sumaišomi su ląstelėmis. Po 3 val. beseruminė mitybinė terpė pakeičiama pilna mitybine terpe ir ląstelės auginamos ir analizuojamos po 48 val.

Cas9 kompleksų kotransfekcijos atveju, 1  $\mu$ g Cas9 komplekso (papildomai pridedant 300 ng crRNR:tracrRNR duplekso) sumaišoma su beserumine mitybine terpe (100  $\mu$ l). Pridėjus 1  $\mu$ l TurboFect transfekcijos agento mėginiai 15-20 min. inkubuojami kambario temperatūroje ir tada sumaišomi su ląstelėmis. Po 3 val. beseruminė mitybinė terpė pakeičiama pilna mitybine terpe ir ląstelės auginamos ir analizuojamos po 48 val.

Dvigubos transfekcijos atveju, DNMT3B ir/ar PPIB genų modifikavimui, po 1  $\mu$ g atitinkamo Cas9 komplekso (papildomai pridedant po 300 ng atitinkamo crRNR:tracrRNR duplekso) 100  $\mu$ l PBS tirpale buvo sumaišoma su 2,5  $\mu$ l TurboFect transfekcijos agento. Mėginiai 15-20 min. inkubuojami kambario temperatūroje ir tada sumaišomi su ląstelėmis. Po 16 val. beseruminė mitybinė terpė pakeičiama pilna mitybine terpe ir ląstelės auginamos ir analizuojamos po 48 val.

Stabili ląstelių linija, ekspresuojanti RFP/EGFP reporterinę sistemą, buvo gauta naudojant „cGPS® CHO-K1 Full Kit DD“ (Cellecctis Bioresearch) rinkinį. Plazmidžių transfekcijai į CHO-K1 ląsteles buvo panaudotas TurboFect *in vitro* DNR transfekcijos agentas.

## **DNR mutacijų nustatymas Surveyor nukleazės pagalba**

Iš CHO-K1 ir HEK293T ląstelių DNR buvo išskiriama naudojant „GeneJET genomic DNA purification kit“ (Thermo Fisher Scientific) rinkinį. Sritys, supančios Cas9 kompleksų taikinius, buvo padauginamos PGR pagalba naudojant pradmenis aprašytus 3 lentelėje. Mutacijų, įvestų dvigrandininio trūkio NHEJ reparacijos metu, dažnis buvo kiekybiškai įvertinamas „Surveyor“ metodo pagalba, naudojant „SURVEYOR Mutation Detection Kit for Standard Gel Electrophoresis“ (Transgenomic) rinkinį ir laikantis gamintojo rekomendacijų.

## **Pavienių molekulių eksperimentai**

Pavienių molekulių magnetinių pincetų eksperimentai su Cas9 buvo atliekami kaip aprašyta anksčiau (Revyakin et al., 2005; Seidel et al., 2005). Naudoti DNR fragmentai buvo gauti T4 DNR ligazės pagalba sujungiant PGR būdu gautą 2,1 kb ilgio DNR fragmentą turintį Cas9 taikinį (proskirtuką ir PAM) su 1,0 kbp ilgio biotinu/digoksigenu pažymėtais fragmentais. DNR substrato digoksigeno modifikuotas galas buvo „prištas“ prie tėkmės kameros paviršiaus padengto anti-digoksigeninu (Roche), o biotinu modifikuotas galas – prie streptavidinu padengtos magnetinės dalelės (MyOne, Invitrogen). Eksperimentai buvo vykdomi kambario temperatūroje, pridėjus 1 nM Cas9 komplekso. Magnetiniame lauke įtempta DNR molekulė buvo neigiamai ar teigiamai superspiralizuojama. Eksperimentų metu, buvo fiksuojami DNR ilgio pokyčiai, atsiradę dėl Cas9 komplekso ir substrato sąveikos, susiformuojant R-kilpai.

## **Atsitiktinių 5 bp PAM bibliotekos konstravimas**

5 bp PAM biblioteka buvo konstruojama į pTZ57R/T vektorių (Thermo Fisher Scientific) įterpiant DNR fragmentą gautą PGR pagalba, naudojant GG-821N (5N PAM pozicijoje) ir GG-820 oligonukleotidus (3 lentelė). Po transformacijos, *E. coli* DH5 $\alpha$  ląstelės auginamos ant LB terpės. Transformacijos efektyvumas buvo įvertinamas skaičiuojant kolonijas skirtingų skiedimų lėkštelėse. Ląstelės buvo auginamos iš kolonijų resuspenduotų skystoje LB terpėje. Plazmidinė DNR buvo išskiriama naudojant „GeneJET Plasmid Miniprep Kit“ (Thermo Fisher Scientific) rinkinį.

## **Atsitiktinių 7 bp PAM bibliotekos konstravimas**

7 bp PAM biblioteka buvo konstruojama kaip ir 5N biblioteka, tačiau su tam tikrais pakeitimais. Atsitiktinės sekos (6N PAM pozicijoje) buvo įvedamos GG-940-G, GG-940-C, GG-940-A ir GG-940-T oligonukleotidų (3 lentelė) pagalba, atliekant atskirus PGR su GG-939 (3 lentelė). Gauti DNR fragmentai buvo įterpiami į pTZ57R/T vektorių (Thermo

Fisher Scientific). Po plazmidžių transformacijos, plazmidės buvo atskirai išskiriamos iš kiekvienos transformacijos reakcijos ir apjungiamos, gaunant 7N atsitiktinę PAM seką turinčią biblioteką.

### **PAM bibliotekų sekvenavimas**

Norint patvirtinti, kad gautose bibliotekose PAM pozicijoje nukleotidai yra išsidėstę atsitiktinai, buvo atliekamas PGR (15 ciklų) nuo 5N ir 7N bibliotekų, naudojant atitinkamai TK-119/pUC-dir ir TK-113/pUC-dir pradmenų poras (3 lentelė). Gauti DNR fragmentai, buvo pažymimi Illumina sekvenavimui papildomai padauginant (10 ciklų) su JKYS800.1/JKYS803 ir JKYS921.1/JKYS812 pradmenų poromis (3 lentelė). Pažymėti fragmentai buvo dar kartą papildomai padauginami (10 ciklų) su universaliais JKYS557/JKYS558 pradmenimis (3 lentelė) ir sekvenuojami bei analizuojami.

### **PAM sekų identifikavimas**

1  $\mu$ g (5,6 nM) plazmidžių bibliotekos su atsitiktiniais nukleotidais PAM pozicijoje (5N ar 7N) buvo inkubuojama su 0,5 nM ir 50 nM koncentracijos Cas9 kompleksu 37°C 1 val. 100  $\mu$ l. Po karpymo, į reakcijos mišinį buvo pridedama 2,5 aktyvumo vienetai DreamTaq DNR polimerazės (Thermo Fisher Scientific) ir 0,5  $\mu$ l 10 mM dATP (ar dNTP) ir inkubuojama papildomas 30 min. 72°C. Tokiu būdu prie atsiradusių po kirpimo laisvų DNR galų buvo pridedama 3' dA. Reakcijos produktai buvo valomi „GeneJET PCR Purification Kit“ (Thermo Fisher Scientific) rinkinio pagalba ir liguojami su DNR adapteriu, gautu sulydžius TK-117 ir fosforilintą TK-111 oligonukleotidus (3 lentelė). Norint specifiškai praturtinti perkirptus DNR fragmentus su PAM sekomis, PGR pagalba buvo padauginama DNR, naudojant pUC-dir ir TK-117 pradmenis. Kitame žingsnyje gauti DNR fragmentai buvo pažymimi Illumina sekvenavimui, dar kartą papildomai padauginami ir sekvenuojami (naudoti pradmenys aprašyti 3 lentelėje).



## 2. REZULTATAI IR JŲ APTARIMAS

### 2.1. crRNR ir tracrRNR yra būtinos Cas9 sukeltai DNR interferencijai *Streptococcus thermophilus* bakterijoje

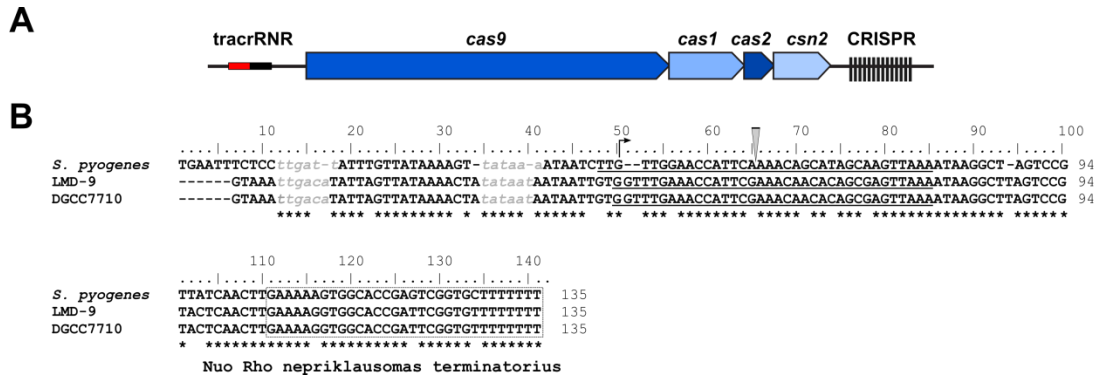
*S. thermophilus* DGCC7710 kamienas turi keturias skirtingas CRISPR-Cas sistemas: CRISPR1, CRISPR2, CRISPR3 ir CRISPR4 (Horvath and Barrangou, 2010). CRISPR2 ir CRISPR4 atitinkamai priklauso III ir I tipams, kai CRISPR1 ir CRISPR3 – II tipui (Horvath and Barrangou, 2010; Makarova et al., 2015). II tipo sistemos yra santykinai paprastos. II-A tipo sistemos yra sudarytos iš *tracrRNR* ir keturių *cas* genų, esančių greta CRISPR srities (1 A pav.). II tipo CRISPR-Cas sistemoms būdingas *cas9* genas koduoja multidomeninį baltymą, kuris yra atsakingas už DNR taikinio kirpimą (Gasiunas et al., 2012; Jinek et al., 2012; Sapranauškas et al., 2011).

Anksčiau buvo pademonstruota, kad efektorinis *S. thermophilus* CRISPR3-Cas sistemos Cas9 kompleksas turi surišęs 42 nt ilgio crRNR (Gasiunas et al., 2012). Nors *tracrRNR* svarba Cas9 komplekso aktyvumui iš *S. pyogenes* buvo nustatyta (Jinek et al., 2012), tačiau *S. thermophilus* CRISPR3-Cas sistemos atveju nebuvo tirta.

#### 2.1.1. *tracrRNR* koduojamos sekos identifikavimas *S. thermophilus* DGCC7710 CRISPR3-Cas sistemoje

DNR sekos, esančios prieš *cas9* geną *S. thermophilus* DGCC7710 CRISPR3-Cas sistemoje analizė parodė, kad kaip ir *S. pyogenes* CRISPR-Cas sistemoje, numanoma *tracrRNR* koduojanti seka yra prieš *cas* operoną (1 A pav.). *S. pyogenes* atveju, RNR sekvenavimas atskleidė, kad *tracrRNR* yra 171 nt ir 89 nt ilgio – transkribuojamos nuo dviejų skirtingų promotorių, tačiau turinčios bendrą terminacinę seką (Deltcheva et al., 2011). *S. thermophilus* LMD-9 kamieno CRISPR3-Cas sistemos, beveik identiškos *S. thermophilus* DGCC7710 CRISPR3-Cas sistemai, „Northern blot“ tyrimo rezultatai leido identifikuoti ~100 nt ilgio *tracrRNR* transkriptą (Deltcheva et al., 2011).

Norint nustatyti spėjamą *tracrRNR* transkripcijos pradžios poziciją *S. thermophilus* DGCC7710 CRISPR3-Cas sistemoje, mes palyginome aprašytų CRISPR-Cas sistemų DNR sekas (Deltcheva et al., 2011), turinčias komplementarias sritis kartotinių sekoms (1 B pav.). Paaiškėjo, kad *tracrRNR* koduojančios sekos yra labai panašios (87% identiškų nukleotidų) ir turi nuo Rho nepriklausomas terminacijos sekas. Taigi mes padarėme

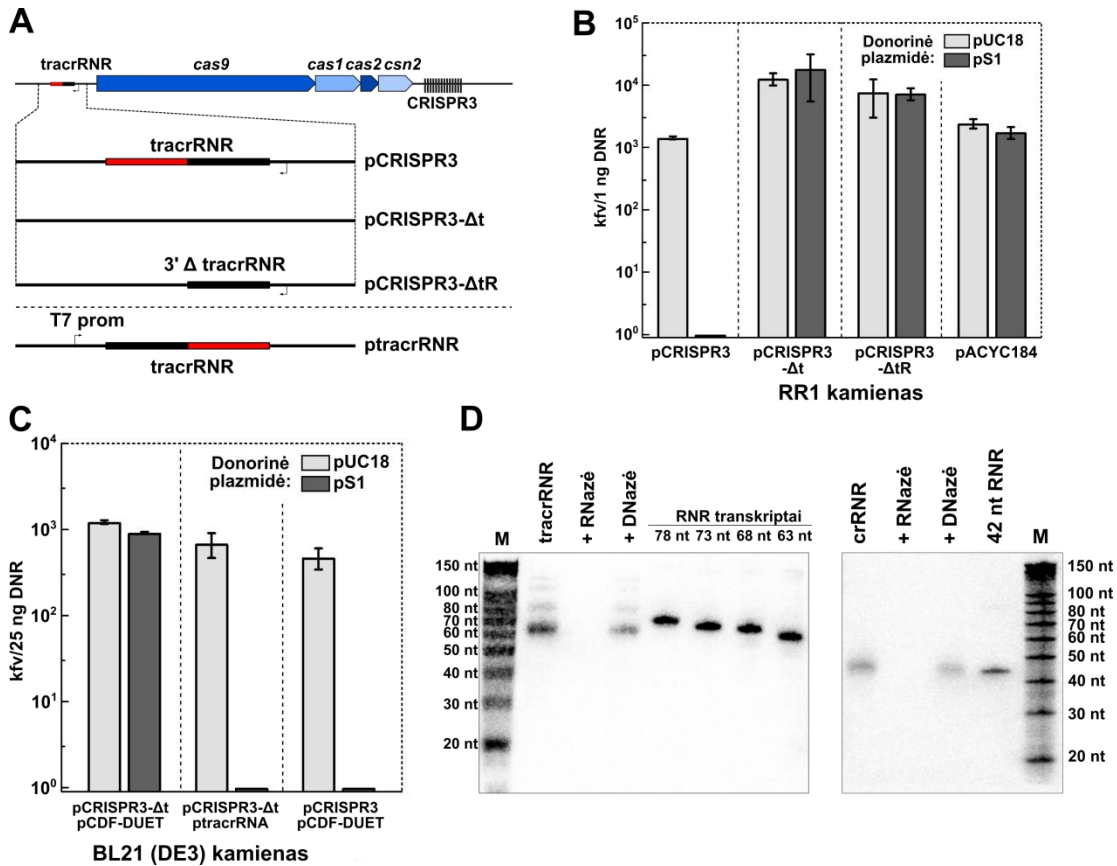


**1 pav. II-A tipo CRISPR-Cas sistemų sandara *S. pyogenes*, *S. thermophilus* LMD-9 ir DGCC7710 bakterijų kamienuose.** (A) Keturi *cas* genai yra išsidėstę greta skirtukų-kartotinių (CRISPR) srities. *tracrRNR* yra koduojama prieš *cas* genų operoną. (B) *S. pyogenes*, *S. thermophilus* LMD-9 ir DGCC7710 CRISPR-Cas sistemų *tracrRNR* koduojančių DNR sričių palyginimas. Sekos komplementrios kartotinių sekoms yra pabrauktos. Numatomos promotoriaus sekos yra vaizduojamos pilka spalva. Nuo Rho nepriklausomi terminatoriai yra apibraukti. RNazės III kirpimo ir transkripcijos pradžios vietos yra atitinkamai pažymėtos trikampiu ir rodykle.

prelaidą, kad *tracrRNR* transkripcija *S. thermophilus* DGCC7710 kamiene prasideda toje pačioje vietoje kaip ir *S. pyogenes* ir *tracrRNR* ilgis galėtų būti ~100 nt, kas sutaptų su nustatytu *S. thermophilus* LMD-9 *tracrRNR* ilgiu (Deltcheva et al., 2011).

### 2.1.1. *tracrRNR* yra būtina *S. thermophilus* CRISPR3-Cas sistemos sukeltai DNR interferencijai *in vivo*

pCRISPR3 plazmidė, turinti pilną *S. thermophilus* DGCC7710 CRISPR3-Cas sistemą (2 A pav.), užtikrina ląstelių atsparumą transformacijai pS1 plazmidė, turinčiai proskirtuką identišką S1 skirtukui CRISPR srityje su greta esančia 5'-NGGNG-3' PAM seka (Gasiunas et al., 2012; Sapranauškas et al., 2011). Norėdami išsiaiškinti ar *tracrRNR* yra būtina *S. thermophilus* CRISPR3-Cas sukeltai plazmidžių interferencijai, mes sukonstravome pCRISPR3 plazmidės variantus su pakeitimais *tracrRNR* koduojančioje sekoje (2 A pav.). pCRISPR3-Δ*t* plazmidėje yra pašalinta *tracrRNR* koduojanti seka, tuo tarpu pCRISPR3-Δ*tR* plazmidėje yra palikta sritis nuo transkripcijos pradžios vietos iki kartotiniui komplementarios sekos pabaigos. Atlikus *E. coli* ląstelių transformacijos pS1 plazmidė eksperimentus paaiškėjo, kad šios aukščiau paminėtos plazmidės neužtikrino DNR interferencijos (2 B pav.). Norint patvirtinti, kad DNR interferencija buvo prarasta dėl *tracrRNR* pašalinimo/sutrumpinimo, pabandėme ląsteles papildomai transformuoti plazmidėmis turinčiomis pilno ilgio *tracrRNR* koduojančią seką – pateikti *tracrRNR in trans*. Tam į pCDF-DUET plazmidė su už T7 promotorinės sekos įterpėme pilno ilgio



**2 pav. tracrRNR yra būtina DNR interferencijai.** (A) Plazmidės naudotos ląstelių transformacijos ribojimo eksperimentuose. pCRISPR3-Δt plazmidė koduoja CRISPR3-Cas sistemą be tracrRNR. pCRISPR3-ΔtR plazmidėje, tracrRNR turi tik sritį komplementarią kartotinio sekai. p plazmidė buvo gauta įterpus pilno ilgio tracrRNR koduojančią seką už T7 RNR polimerazės promotoriaus į pCDF-DUET plazmidę. (B) tracrRNR trumpinimas ar delecija inaktyvuoja CRISPR3-Cas sistemą. (C) tracrRNR gali būti pateikta *in trans* atskiroje plazmidėje. (D) Nukleorūgščių, išskirtų iš aktyvaus Sth3 Cas9 komplekso “northern blot” analizė. Sth3 Cas9 gryninasi su ~65 nt tracrRNR ir ~42 nt crRNR.

tracrRNR koduojančią seką (2 B pav., p plazmidė). *E. coli* ląstelės, turinčios pCRISPR3-Δt plazmidę buvo transformuojamos p plazmide. Atlikus ląstelių transformacijos pS1 plazmide eksperimentus paaiškėjo, kad pateikus tracrRNR koduojančią seką *in trans*, DNR interferencijos aktyvumas atsistatė (2 C pav.). Taigi šie rezultatai patvirtina, kad tracrRNR yra būtina *S. thermophilus* CRISPR3-Cas sistemos sukeltai DNR interferencijai *in vivo*.

### 2.1.2. tracrRNR gryninasi su Cas9 baltymu

*S. thermophilus* CRISPR3 (Sth3) Cas9 yra vienintelis baltymas užtiktinantis *S. thermophilus* CRISPR3-Cas sistemos sukeltą DNR interferenciją (Gasiunas et al., 2012;

Sapranauškas et al., 2011). Sth3 Cas9 baltymas gryninasi su ~42 nt crRNR (2 D pav.) (Gasiunas et al., 2012). Norėdami išsiaiškinti ar ir tracrRNR gryninasi kartu su Cas9 ir crRNR, mes atlikome „Northern blot“ hibridizaciją naudodami 36 nt ilgio anti-tracrRNR oligodeoksinukleotidą. Nukleorūgščių, išskirtų iš išgryninto Cas9 komplekso preparato (Gasiunas et al., 2012) analizė parodė, kad Cas9 baltymas, kartu su crRNR buvo surišęs ir ~65 nt ilgio tracrRNR (2 D pav.). Taigi, kaip ir *S. pyogenes* (Spy) efektorinio komplekso atveju (Jinek et al., 2012), *S. thermophilus* DGCC7710 CRISPR3-Cas sistemos efektorinis kompleksas sukeliantis DNR interferenciją yra sudarytas iš Cas9, crRNR ir tracrRNR.

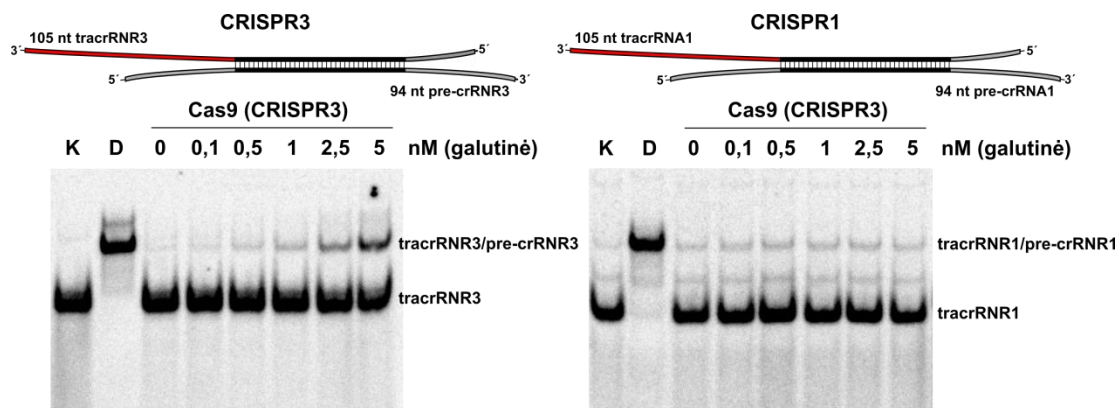
### 2.1.3. Sth3 Cas9-crRNR-tracrRNR komplekso surinkimas *in vitro*

Nustačius, kad Sth3 efektorinis kompleksas yra sudarytas iš Cas9 baltymo, ~42 nt crRNR ir ~65 nt tracrRNR, buvo siekiama surinkti aktyvų kompleksą *in vitro* iš šių trijų komponentų. Tuo tikslu Cas9 baltymas buvo inkubuojamas su ekvimoliariniais kiekiais sintetinės 42 nt crRNR ir *in vitro* transkripcijos būdu gautos 78 nt tracrRNR, atitinkančios subrendusią tracrRNR formą. Papildomi nukleotidai (78 nt vs. 65 nt) buvo pridėti dėl galimybės susintetinti tracrRNR *in vitro* transkripcijos būdu. Surinkto komplekso DNR kirpimo aktyvumas buvo stebimas *in vitro* karpant pS1 plazmidę. Esant Sth3 Cas9 kompleksui, pS1 plazmidė buvo verčiama į linijinę formą (3 A pav.). Nesant vieno iš komplekso komponentų (Cas9, crRNR ar tracrRNR), DNR nukleazinis aktyvumas nebuvo stebimas. pUC18 plazmidė, neturinti proskirtuko, komplementaraus crRNR skirtuko sekai, taip pat nebuvo karpoma. Taigi, aktyvus Sth3 Cas9 kompleksas gali būti surinktas *in vitro* sumaišant individualius komplekso komponentus.

Ekspirimentai su Spy Cas9 kompleksu atskleidė, kad tracrRNR gali būti sutrumpinta nuo 3'-galo išlaikant komplekso aktyvumą *in vitro* (Jinek et al., 2012). Norėdami nustatyti minimalų tracrRNR 3'-galo ilgį reikalingą DNR kirpimui Sth3 Cas9 kompleksui, mes trumpinome tracrRNR nuo 3'-galo kas 5 nt (3 B pav.). Sth3 Cas9 kompleksas surinktas naudojant pačias trumpiausias tracrRNR molekules (33 nt ir 38 nt) nebuvo aktyvus (3 C pav.). Tuo tarpu 43, 48 ir 53 nt ilgio tracrRNR variantai leido išlaikyti aktyvumą, tačiau jis buvo mažesnis nei naudojant 58 nt ir ilgesnes tracrRNR molekules. Antrinės struktūros tracrRNR analizė parodė, kad tracrRNR galėtų formuoti tris plaukų segtuko struktūras tracrRNR 3'-gale. Remiantis tokia prielaida ir atsižvelgiant



et al., 2012; Jinek et al., 2012). Netiesioginiai įrodymai gauti atliekant eksperimentus *in vivo* *S. pyogenes* bakterijose parodė, kad Cas9 taip pat yra būtinas crRNR brendimui (Deltcheva et al., 2011) ir galėtų dalyvauti pre-crRNR:tracrRNR duplekso susidaryme (Deltcheva et al., 2011), tačiau tiesioginių eksperimentinių duomenų nepakanka. Norėdami eksperimentiškai patikrinti ar Sth3 Cas9 baltymas dalyvauja sulydant pre-crRNR ir tracrRNR, mes susintetinome 105 nt tracrRNR3 ir 94 nt pre-crRNR (turinčią 36 nt pilno ilgio kartotinio seką) iš CRISPR3 sistemos ir stebėjome tracrRNR3:pre-crRNR3 duplekso formavimąsi, esant Sth3 Cas9 baltymui (4 pav.). Taip pat žiūrėjome, ar homologinės *S. thermophilus* DGCC7710 II tipo CRISPR1 sistemos tracrRNA1 (105 nt) ir pre-crRNA1 (94 nt) yra sulydomos CRISPR3 sistemos Cas9 baltymo. Radioaktyviai pažymėta tracrRNR buvo inkubuojama su pre-crRNR, esant nespecificinei 2 kb RNR, ir skirtingiems Cas9 kiekiams 10 min. Reakcijos buvo stabdomos pridendant proteinazės K ir nežymėtos tracrRNR. Mėginiai buvo analizuojami nedematūruojančios PAGE pagalba (4 pav.). Šiomis sąlygomis, nesant Cas9, tracrRNR:pre-crRNR dupleksas nesiformavo. Didėjantis Sth3 Cas9 kiekis reakcijos mišiniuose skatino tracrRNR3:pre-crRNR3 duplekso formavimąsi, bet neturėjo įtakos tracrRNR1:pre-crRNR1 duplekso susidarymui (4 pav.). Apibendrinant, šie rezultatai rodo, kad Sth3 specifiskai skatina tracrRNR3 ir pre-crRNR3 susilydimą ir duplekso formavimąsi.



**4 pav. tracrRNR ir pre-crRNR susilydimas esant Cas9.** 105 nt tracrRNR ir 94 nt pre-crRNR naudotos sulydimo eksperimentuose yra vaizduojamos virš gelių. Komplementarios sekos pažymėtos juoda spalva. Nedenatūruojančios PAGE geliai rodantys duplekso formavimąsi tarp tracrRNR3 ir pre-crRNR3 iš CRISPR3 sistemos (kairė dalis) ir tracrRNR1 ir pre-crRNR1 iš CRISPR1 (dešinė dalis) esant įvairios koncentracijos Cas9 iš CRISPR3. K – takelis, kuriame buvo leidžiama tik žymėta tracrRNR, D – tracrRNA:pre-crRNA dupleksas, gautas pakaitinus tracrRNR:pre-crRNR mišinį (1:100 molinis santykis) iki 95°C ir lėtai atvėsinus iki kambario temperatūros.

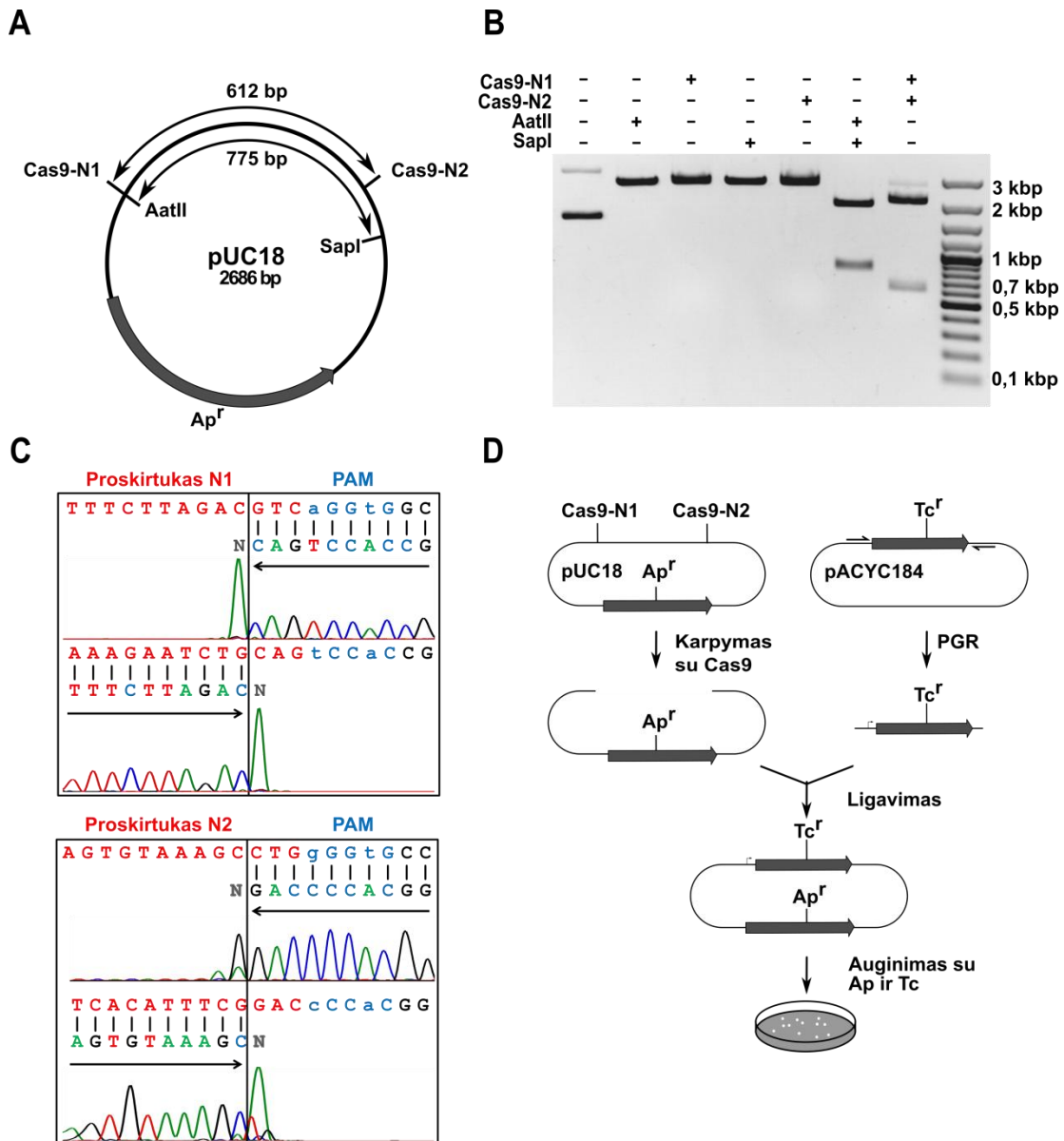
## 2.2. Cas9 komplekso perprogramavimas DNR kirpimui *in vitro*

Paprasta modulinė Cas9 komplekso sandara, kur DNR specifiskumas yra užkoduotas crRNR sekoje, o DNR kirpimas atliekamas paties Cas9 baltymo, suteikia galimybę naudoti Cas9 kompleksą kaip universalų įrankį DNR kirpimui. Keisdami crRNR seką Cas9 komplekse, mes siekėme panaudoti Cas9 kompleksą kaip įrankį manipuluoti DNR *in vitro*.

### 2.2.1. Sth3 Cas9 komplekso kaip įrankio panaudojimas DNR manipuliacijose

Restrikcijos endonukleazės atpažįsta 4-8 bp ilgio DNR sekas ir jas kerpa. Dėl tokių savybių, šie fermentai tapo nepakeičiami įrankiai manipuliacijose su DNR (Roberts, 2005). Tačiau pasirinkimo galimybės, norint tiksliai įterpti pasirinktą fragmentą į plazmidinį vektorių, yra ribojamos restrikcijos endonukleazių atpažįstamų sekų. Norėdami pademonstruoti, kad Sth3 Cas9 gali būti naudojama kaip universali nuo RNR priklausoma DNR endonukleazė, kurios specifiskumas yra apspręstas crRNR sekos, nusprendėme pritaikyti Cas9 klonavimo eksperimente. Pasirinkome pUC18 plazmidėje du taikinius (N1 ir N2), esančius greta AatII ir SapI restrikcijos endonukleazių atpažįstamų sekų (5 A pav.). *In vitro* transkripcijos pagalba nuo PGR fragmentų turinčių N1 ir N2 koduojančius skirtukus apsuptus kartotinių sekomis susintetiname pre-crRNR molekules.

Teorinis DNR fragmento ilgis tarp Cas9-N1 ir Cas9-N2 taikinių yra 612 bp (5 A pav.). pUC18 kirpimo produktų su Cas9-N1 ir Cas9-N2 kompleksais analizė parodė, kad buvo gauti du fragmentai (~2000 and ~600 bp) (5 B pav.), kas rodytų, kad plazmidė buvo perkirpta pasirinktuose N1 ir N2 taikiniuose. Kirpimo vietų sekvenavimas patvirtino, kad Cas9 kompleksai kirpo abiejuose DNR taikiuose 3 nt prieš PAM sekas, palikdami bukus DNR galus (5 C pav.). PGR pagalba pasidauginę pACYC184 plazmidės sritį, koduojančią promotorių ir antibiotikui tetraciklinui atsparumą koduojantį geną (5 D pav.) pabandėme įterpti gautą fragmentą į su Cas9 paruoštą pUC18 vektorių. Po ligavimo, rekacijos mišiniu transformavome *E. coli* ląsteles. Išsiskyrus iš augančių ant lėkštelės su antibiotikais ampicilinu ir tetraciklinu ląstelių plazmidės ir jas nusekvenavus paaiškėjo, kad PGR fragmentas buvo įterptas norimoje vietoje be nepageidaujamų mutacijų. Apibendrinant, šis eksperimentas parodė, kad Cas9 endonukleazė yra tinkama klonavimo eksperimentams. Taip pat, DNR gali būti karpoma naudojant kelis Cas9 kompleksus vienu metu.



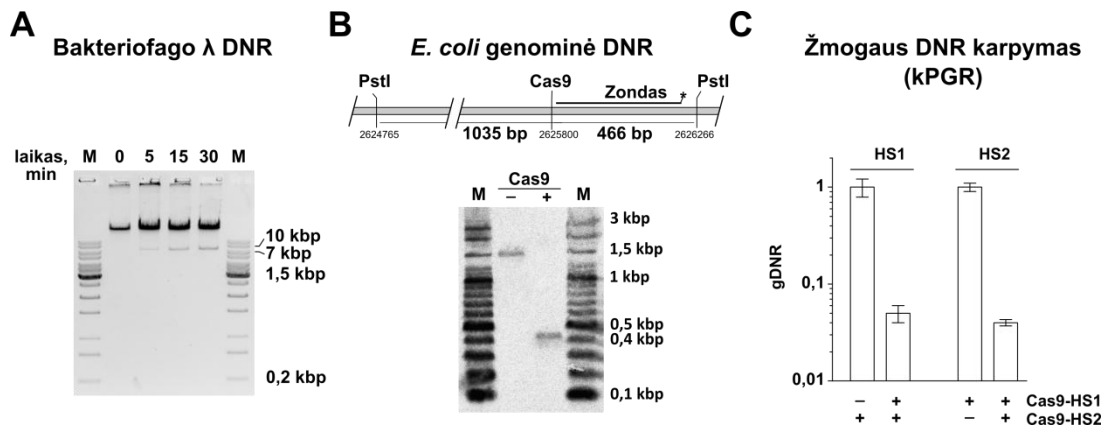
**5 pav. Sth3 Cas9 kompleksas kaip įrankis DNR manipuliacijose. (A)** Taikiniai pUC18 plazmidėje. **(B)** pUC18 plazmidės karpymo rezultatai panaudojus Cas9-N1 ir Cas9-N2 kompleksus. **(C)** pUC18 sukarpytos su Cas9-N1 (viršutinė dalis) ir Cas9-N2 (apatinė dalis) kompleksais sekvenavimo rezultatai. Plazmidė taikiniuose buvo kerpama 3 nt iki PAM sekos. **(D)** Klonavimo eksperimento schema. Vektorius paruoštas su Cas9-N1 ir Cas9-N2 kompleksais buvo liguotas su PGR fragmentu turinčiu promotorių ir atsparumą tetraciklinui koduojantį geną. Transformuotos ląstelės buvo auginamos ant terpės su tetraciklinu (Tc) ir ampicilinu (Ap).



### 2.2.2. Genominės DNR karpymas *in vitro*

Norėdami išsiaiškinti ar Cas9 geba surasti ir kirpti taikinius genominės DNR kontekste, mes nukreipėme Cas9 kompleksus į unikalius taikinius bakteriofago  $\lambda$  (49 kbp), *E. coli* (4,6 Mbp) ir žmogaus (3,2 Gbp) genominėje DNR. Bakteriofago  $\lambda$  genominės DNR analizė po kirpimo su Cas9 kompleksu parodė, kad DNR buvo kirpta norimoje vietoje (susidarė 42 kbp ir 7 kbp produktai) (6 A pav.). Norėdami aptikti Cas9 kirpimo produktus *E. coli* genominėje DNR, panaudojome “Southern blot” hibridizaciją. Po kirpimo su Cas9 kompleksu, *E. coli* genominė DNR buvo inkubuojama su PstI restrikcijos endonukleaze. Išfrakcionavus DNR agaroziniame gelyje, atlikome “Southern blot” hibridizaciją naudodami zondą, kuris hibridizuojasi tarp PstI ir Cas9 kirpimo vietų (6 B pav.). Atstumas tarp PstI taikinių yra ~1500 bp, kai tuo tarpu tarp Cas9 ir PstI taikinių ~500 bp. Po genominės DNR poveikio su Cas9 kompleksu, mums pavyko aptikti tik ~500 bp DNR fragmentą, kas reikštų, kad *E. coli* genominė DNR buvo perkirpta norimoje vietoje.

Kitame etape mes analizavome žmogaus genominės DNR karpymą su Cas9-HS1 ir Cas9-HS2 kompleksais, atitinkamai nutaikytais prieš RASGEF1C (HS1) ir ARL15 (HS2)



**6 pav. Genominės DNR karpymas *in vitro* panaudojant Cas9.** (A)  $\lambda$  DNR karpymas. Fago  $\lambda$  DNR buvo inkubuojama su Cas9 kompleksu ir reakcijos produktai analizuojami elektroforezės pagalba. Cas9 yra nukreiptas į seką, esančią 8 kb nuo *cos* sekos. (B) *E. coli* genominės DNR karpymas. Prieš fragmentaciją su PstI, *E. coli* BL21 genominė DNR buvo inkubuojama 30 min. su Cas9 kompleksu. Cas9 buvo nukreiptas tarp PstI kirpimo vietų. Jei genominė DNR yra kerpama Cas9, po zondo hibridizacijos turėtume aptikti 466 bp fragmentą; nesant kirpimo, turėtų būti aptinkamas 1499 bp ilgio fragmentas. (C) Žmogaus genominės DNR karpymas. Po karpymo su Cas9-HS1 ir Cas9-HS2 kompleksais, kPGR pagalba buvo santykinai vertinamas neperkirtų DNR taikinių kiekis (HS1 ir HS2).

sritis genominėje DNR. Kirpimo produktai buvo analizuojami kiekybinės PGR pagalba (kPGR) (6 C pav.). Po poveikio Cas9-HS1 ir Cas9-HS2 kompleksais DNR taikinių kiekis abejose srityse sumažėjo daugiau nei 25 kartus, kas rodo kad DNR buvo kerpama norimose vietose. Apibendrinant šie rezultatai rodo, kad Sth3 Cas9 efektyviai randa ir kerpa taikinio DNR įvairaus kompleksiško genominės DNR kontekste.

### **2.3. Genų redagavimas transfekuoiant *in vitro* surinktą Sth3 Cas9 RNP kompleksą**

Pastaraisiais metais Cas9 baltymas buvo pritaikytas kaip įrankis genomo redagavimo eksperimentuose. Nukreiptas Cas9 kompleksas randa taikinio DNR seką ir ją kerpa padarydamas dvigrandininį trūkį (DSB) (Gasiunas et al., 2012; Jinek et al., 2012). Eukariotinėse ląstelėse DSB ištaisomi panaudojant du mechanizmus: mutacijas įvedančiu nehomologinių galų sujungimu (NHEJ) (Lieber, 2010) bei homologinės rekombinacijos (HDR) (Moynahan and Jasin, 2010) pagalba. Pastarasis kelias gali būti išnaudotas tiksliam genomo modifikavimui ar naujos genetinės informacijos įvedimui (Carroll, 2011; Perez-Pinera et al., 2012; Urnov et al., 2010).

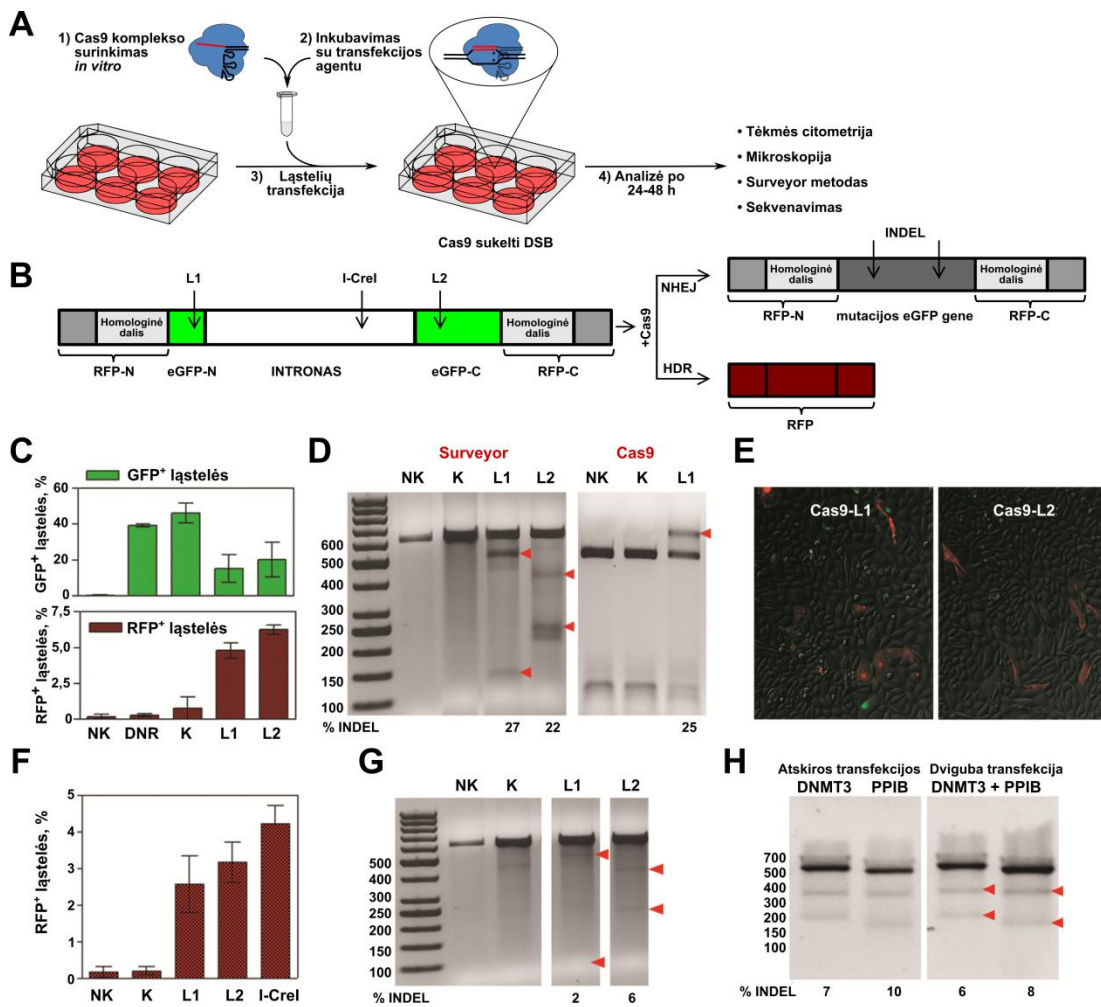
Šiuo metu genomo redagavimo eksperimentuose dažniausiai yra naudojamas Spy Cas9. Įprastu atveju, eukariotinėms ląstelėms optimizuotas Cas9 ir RNR koduojanti seka yra įterpiami į DNR raiškos vektorius, kuris transfekuojamas į ląsteles (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013). Dėl tokio Cas9 patekimo į ląsteles, DNR kirpimo efektyvumas eukariotinėse ląstelėse priklauso nuo daugybės veiksnių, tokių kaip raiškos vektoriaus tipas, transfekcijos efektyvumas, ląstelės tipas, komplekso susirinkimo efektyvumas ląstelėje ir t.t. (Hsu et al., 2013). Dėl šių problemų, būtina atlikti daugybės eksperimentinių sąlygų optimizavimą. Taip pat, toks Cas9 pristatymas į ląstelės būdas gali sukelti nepageidaujamą raiškos vektoriaus integraciją į genominę DNR ir dažnai yra neefektyvus ir sukelia ląstelėms stresą (Gabriel et al., 2011). Norint išvengti šių problemų, mes pabandėme pritaikyti alternatyvią strategiją, paremtą *in vitro* surinkto Sth3 Cas9 RNP komplekso chemine transfekcija į ląsteles (7 A pav.).

Sth3 Cas9 kompleksas su prilietu branduolio lokalizacijos signalu (NLS) buvo surinktas *in vitro* naudojant crRNR:tracrRNR dupleksą ar dirbtinai sujungtas crRNR ir tracrRNR molekules, vadinamas sgRNR. Komplekso transfekcijai buvo naudojamas TurboFect™ transfekcijos agentas. Norint stebėti DNR kirpimo efektyvumą ląstelėse buvo sukonstruota reporterinė sistema, turinti raudoną ir žalią fluorescuojančius baltymus (atitinkamai RFP ir eGFP) koduojančius genus (7 B pav.). eGFP genas, turi dviejų Sth3 Cas9 kompleksų taikinius, L1 ir L2. I-CreI nukleazės taikinytis taip pat buvo įterptas į reporterinę sistemą (Jurica et al., 1998). Nesant Cas9, po introno procesavimo *in vivo*

turėtų būti stebima eGFP fluorescencija. Dvigrandiniai trūkiai (DSB) L1 ar L2 srityse turėtų sukelti DNR pažaidų taisymą NHEJ ir/ar HDR būdais. Dėl pažaidų taisymo NHEJ būdu, atsiradusios mutacijos eGFP gene sukeltų fluorescencijos sumažėjimą ar išnykimą. HDR būdu taisomi trūkiai, turėtų inicijuoti RFP koduojančio geno sekos atsistatymą per įterptas homologines sritis ir tai leistų kiekybiškai įvertinti HDR efektyvumą ląstelių populiacijoje stebint RFP signalo atsiradimą. Taigi, tokios reporterinės sistemos įterpimas į vektorių leido gauti reporterinę plazmidę, o į CHO-K1 ląstelių genomine DNR – stabilią reporterinę ląstelių liniją.

Norint pademonstruoti Sth3 Cas9 kompleksų funkcinį aktyvumą *in vivo*, reporterinė plazmidė buvo transfekuojama kartu su *in vitro* surinktais Sth3 Cas9 kompleksais. DNR plazmidės ir Cas9 kompleksai buvo sumaišomi su transfekcijos agentu skirtinguose mėgintuvėliuose, tačiau transfekcija atliekama tuo pačiu metu. Po 48 val. tėkmės citometrijos pagalba buvo kiekybiškai vertinama, kokioje dalyje ląstelių atsirado eGFP signalas. Kai ląstelės buvo transfekuojamos tik reporterine plazmide, ~40% ląstelių buvo stebimas eGFP signalas (7 C pav.). Kartu transfekavus Sth3 kompleksą, eGFP ekspresuojančių ląstelių sumažėjo iki 5-15%. Kotransfekcija su Cas9 kompleksu, neturint taikinio reporterinėje plazmidėje, neturėjo jokios įtakos ląstelėms. Surveyor nukleazės metodo (Guschin et al., 2010) pagalba gauti rezultatai parodė, kad L1 ir L2 taikiniuose insercijos ir delecijos (INDELS) buvo atitinkamai aptinkamos ~27% ir ~22% ląstelių. Tai rodo, kad dvigrandiniai trūkiai buvo taisomi NHEJ būdu. Šie skaičiai buvo patvirtinti ir *in vitro* karpymo metodo pagalba, kurio esmė buvo PGR pagalba padaugintą L1 sritį eGFP gene karpyti *in vitro* su Cas9 kompleksu. Tokiu būdu, padauginti PGR fragmentai su insercijomis ir delecijomis nebūtų karpomi, kai tuo tarpu nemodifikuoti produktai būtų perkerpami. Paaiškėjo, kad ~25% DNR fragmentų nebuvo karpoma, kas sutampa su dydžiu, gautu Surveyor metodo pagalba (7 D pav.). Apibendrinant, šie rezultatai rodo, kad *in vitro* surinkti Cas9 kompleksai transfekuoti į ląsteles sukelia dvigrandinius trūkius taikinio vietose, kurie gali būti ištaisomas NHEJ būdu. Tokiu būdu sukeltų DSB efektyvumas buvo palyginamas su efektyvumu gautu transfekuojant Cas9 ir sgRNR koduojančias plazmides (duomenys nepateikti).

Norint išsiaiškinti ar dvigrandiniai trūkiai buvo taisomi HDR būdu, tėkmės citometrijos pagalba buvo siekiama nustatyti po transfekcijos atsiradusį RFP signalą. 5-8% ląstelių, transfekuotų prieš L1 ir L2 taikinius nukreiptais Cas9 kompleksais, buvo stebimas atsiradęs RFP signalas (7 C pav.). Taip pat tokias ląsteles buvo galima identifikuoti fluorescencinės mikroskopijos pagalba (7 E pav.).



**7 pav. Genų redagavimas panaudojant *in vitro* surinktą *S. thermophilus* Cas9 kompleksą. (A)** Ląstelių transfekcijos schema. **(B)** Reporterinė sistema stebėti NHEJ ir HDR taisymo atvejus. L1 ir L2 rodo Cas9 taikinių vietas. **(C)** GFP<sup>+</sup> ir RFP<sup>+</sup> ląstelių dalis po transfekcijos Cas9-L1 ir Cas9-L2 kompleksais nustatyta tėkmės citometrijos (FACS) pagalba. **(D)** Reporterinės plazmidės kirpimo analizė. Cas9 taikinių sritys reporterinėje plazmidėje PGR pagalba buvo padaugintos (592 bp produktai) ir po perlydimio karpytos Surveyor nukleazės pagalba ar prieš L1 taikinį nukreiptu Cas9 kompleksu. Raudonos rodyklės rodo 100+492 bp ir 363+229 bp hidrolizės produktus atitinkamai L1 ir L2 taikiniams, ir nekirtą 592 bp fragmentą. Skaičiai gelių apačioje rodo insercijų/delecijų (INDEL) dažnį. **(F)** RFP<sup>+</sup> ląstelių dalis reporterinėje CHO-K1 ląstelių linijoje po transfekcijos Cas9 ar I-CreI nukleazėmis. **(E)** RFP<sup>+</sup> ląstelių detekcija fluorescentinio mikroskopo pagalba. **(G)** L1 ir L2 taikinių, esančių CHO-K1 ląstelių chromosomoje analizė po transfekcijos. PGR produktai (592 bp) buvo karpomi Surveyor nukleaze. Raudonos rodyklės rodo 100+492 bp (L1) ir 363+229 bp (L2) karpymo produktus. **(H)** HEK293T ląstelės buvo tranfekuotos Cas9 kompleksais nukreiptais prieš taikinius, esančius DNMT3B ir/ar PPIB genuose. Po padaugintų taikinių sričių karpymo Surveyor nukleaze buvo gauti atitinkamai 335+209 bp ir 330+174 bp karpymo produktai.

Taip pat, buvo siekiama pademonstruoti, kad Sth3 Cas9 kompleksas geba kirpti DNR taikinius esančius chromosominėje DNR. Šiam tikslui į CHO-K1 ląstelių linijos genomine DNR buvo įterpta aukščiau aprašyta reporterinė sistema. Po Cas9 kompleksų transfekcijos, Surveyor metodo pagalba, buvo nustatyta, kad 2-6% ląstelių buvo įvestos insercijos ar delecijos taikinių vietose (7 G pav.). Tėkmės citometrijos metodo pagalba atlikta analizė taip pat patvirtino, kad 3-4% ląstelių buvo aptinkamas RFP signalas (7 F pav.). Po I-CreI nukleazės transfekcijos, RFP signalas buvo aptinkamas ~4,5% ląstelių (7 F pav.). Tai rodo, kad panašus kirpimo efektyvumas gali būti pasiektas naudojant skirtingas nukleazes, jas chemiškai transfekuojant į ląsteles.

Parodžius, kad *in vitro* surinktas Cas9 kompleksas gali būti panaudotas genomo redagavimo eksperimentuose, nusprendėme patikrinti, ar Cas9 gali būti panaudotas genomo redagavimui keliose vietose, vieno eksperimento metu. Taikiniaus pasirinkome endogeninius DNMT3B ir PPIB genus HEK293T ląstelėse. Po transfekcijos kiekvienu kompleksu atskirai ir kartu, Surveyor metodo pagalba tikrinome mutacijų dažnį kiekviename taikinyje. Mutacijų dažnis DNMT3B ir PPIB genuose atitinkamai siekė 7% ir 10% po transfekcijos kiekvienu kompleksu atskirai, bei 6% ir 8% atlikus transfekciją su abiem kompleksais kartu (7 H pav.). Apibendrinant, šie duomenys rodo, kad *in vitro* surinkti Cas9 kompleksai gali būti sėkmingai pritaikyti genomo redagavimo eksperimentams, taip pat parodo galimybę tai atlikti keliose genomo srityse vienu metu.

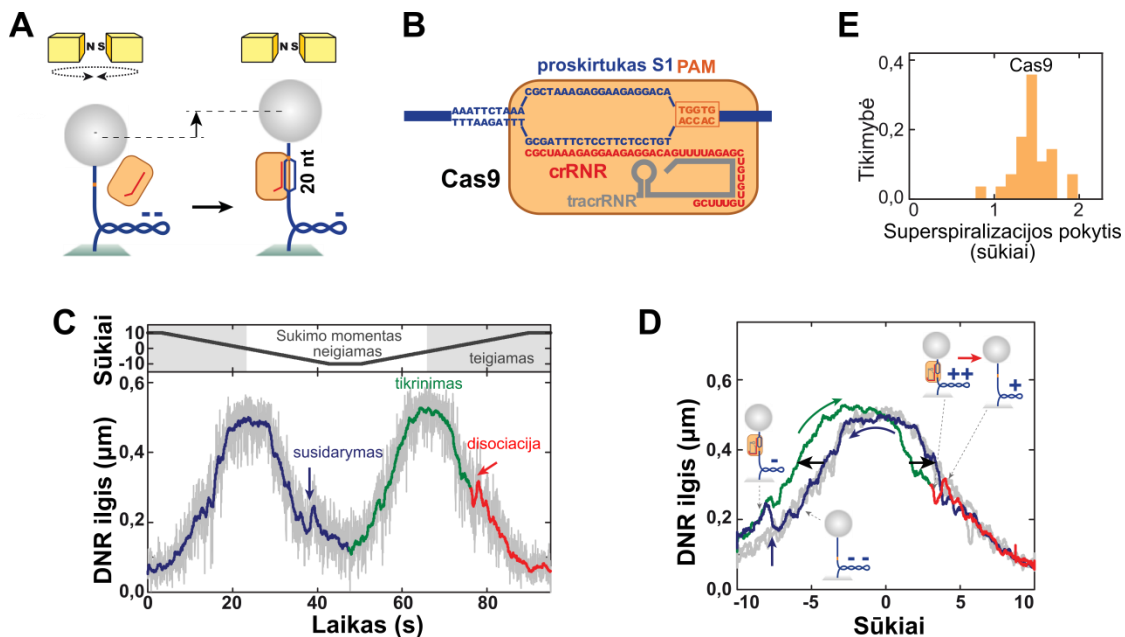
## **2.4. Cas9 DNR taikinio atpažinimas**

Prieš DNR nukreiptose CRISPR-Cas sistemose, DNR taikinyus yra atpažįstamas RNR komponento susidarant R-kilpai su komplementaria grandine ir „išstumiant“ nekomplementarią grandinę. Susidarius tokiai struktūrai toliau vyksta DNR degradacija. Nors mums pavyko šias reakcijas atstatyti *in vitro* (2.1-2.3 skyriai), tačiau tikslus R-kilpos formavimosi mechanizmas nėra žinomas.

### **2.4.1. Tiesioginis R-kilpos susidarymo stebėjimas**

Siekiant tiesiogiai stebėti ir kiekybiškai vertinti R-kilpos formavimąsi ir disociaciją Sth3 Cas9 kompleksui pavienių molekulių superspiralizacijos eksperimentuose, buvo panaudoti magnetiniai pincetai (Brutzer et al., 2010; Mosconi et al., 2009) (8 A pav.). 2,1 kbp DNR fragmentas turintis proskirtuką S1 ir PAM seką buvo viename gale „prikabintas“ prie magnetinės dalelės, o kitu galu – prie tėkmės kameros paviršiaus. Panaudojant magnetus DNR buvo ištempama ir juos sukant superspiralizuojama, ir šio

proceso metu buvo matuojamas DNR ilgis (Klaue and Seidel, 2009). Vykdamt superspiralizacijos eksperimentus, esant pastoviai jėgai, DNR ilgis nekinta. Pasiekus kritinį sukimo momentą, DNR ilgis pradeda trumpėti dėl susidarančių plektoneminių superspiralių. Gauti charakteringi sukimo kreivės ir sukimo momento profiliai leidžia kiekybiškai vertinti R-kilpos susidarymo dinamiką (Forth et al., 2008; Kauert et al., 2011; Mosconi et al., 2009; Oberstrass et al., 2012). Jei baltymui prisijungus prie DNR vyksta lokalus DNR išvyniojimas (pvz., formuojantis R-kilpai), keičiasi DNR molekūlės superspiralizacija ir tai sukelia visos sukimo kreivės poslinkį ar DNR ilgio pokytį (Howan et al., 2012) (8 A pav.). Norėdami stebėti Cas9 sukeltą R-kilpos susidarymą, mes atlikome „R-kilpos ciklą“ eksperimentus (8 B pav.). Pirmiausia DNR molekūlė buvo nestipriai išvyniojama (sukeliama neigiama superspiralizacija) esant nedidelei jėgai, norint palengvinti R-kilpos susidarymą (mėlynos kreivės 8 C ir D pav.). Tada, vykdamt DNR susukimą (sukeliama teigiama superspiralizacija) buvo tikrinamas R-kilpos iširimas (žalios kreivės 8 C ir D pav.). Mes stebėjome efektyvų Cas9 inicijuotą R-kilpos

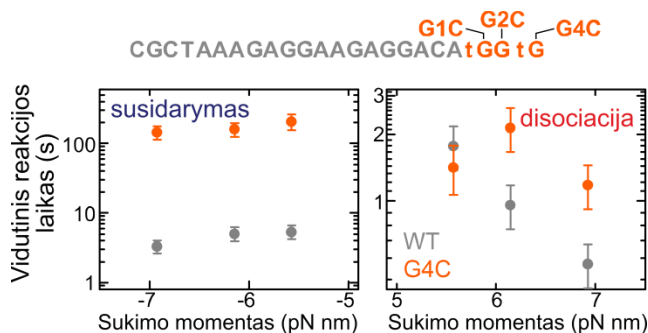


**8 pav. Cas9 komplekso R-kilpos susidarymas ir iširimas pavienių molekulių sukimo eksperimentuose.** (A) DNR sukimo eksperimentų magenitinių pincetų pagalba schema. R-kilpos susidarymas, esant superspiralizuotai DNR molekūlei, sukelia lokalių DNR išvyniojimą ir superspiralizacijos laipsnio pakitimą. Dėl šio proceso kinta DNR ilgis. (B) Cas9 komplekso sudaromos R-kilpos (20 bp) schema. (C ir D) R-kilpos susidarymo ciklai naudojant 1 nM Cas9 kompleksą. DNR turinti proskirtuką ir PAM seką yra negatyviai superspiralizuojama veikiant 0,31 pN jėgai, skatinant R-kilpos susidarymą (mėlyna kreivės dalis). Teigiamai superspiralizuojant yra tikrinimas R-kilpos buvimas (žalia kreivės dalis) ir disociacija (raudona kreivės dalis). R-kilpa disocijuoja esant mažam teigiamam sukimo momentui. Mėlyna ir raudona rodyklės žymi atitinkamai R-kilpos susidarymo ir disociacijos įvykius. (E) Cas9 sukeltas superspiralizacijos kreivių poslinkis.

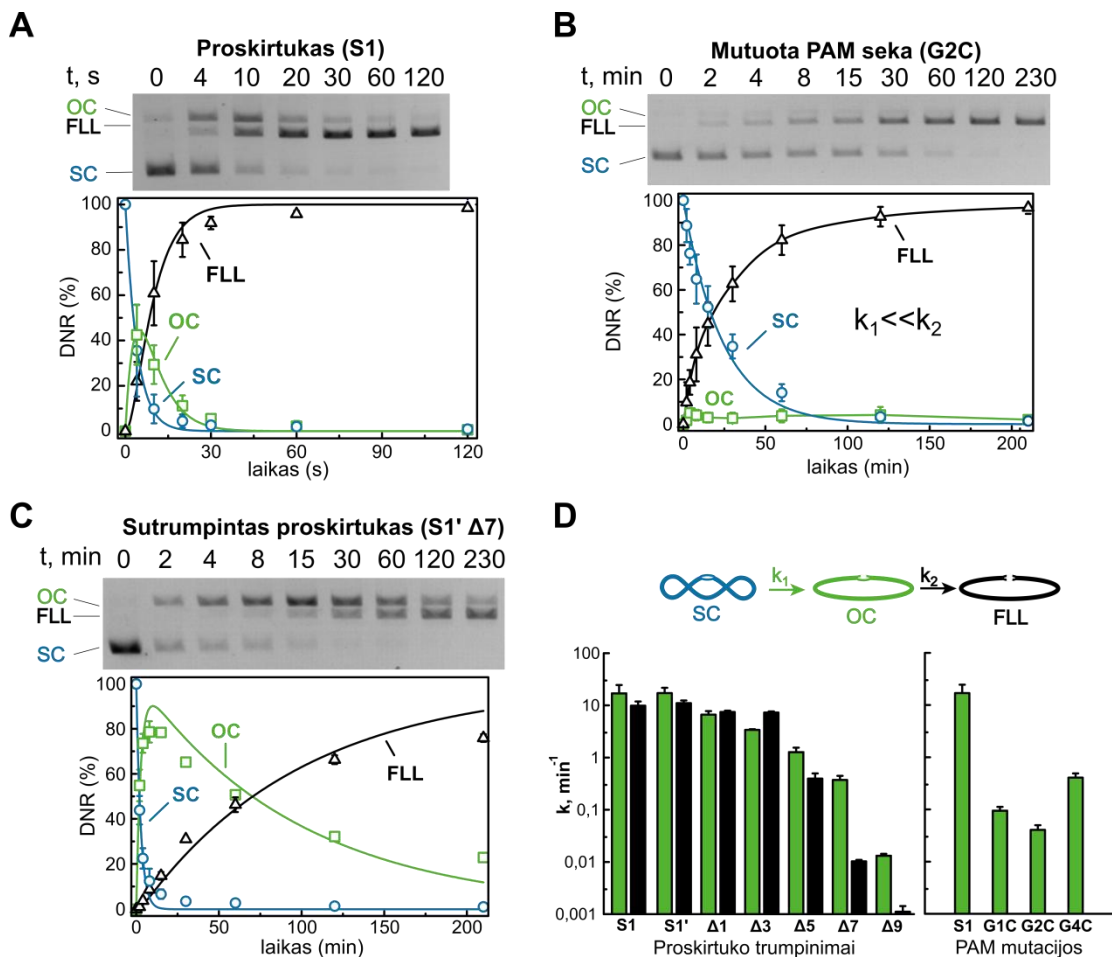
susidarymą (100% visų atvejų, N = 50), kas atitinka sukimo kreivės kairios pusės poslinkį link neigiamų apsisukimų, lyginant su kreive gauta atliekant eksperimentus be baltymo (8 D pav.). Cas9 suformuotos R-kilpos iširdavo esant nedidelei teigiamai superspiralizacijai (8 C ir D pav.). Cas9 atveju, superspiralizacijos kreivės poslinkis atitiko  $1,45 \pm 0,05$  sūkio (8 E pav.). Šis dydis yra šiek tiek mažesnis nei prognozuojamas (1,9 sūkio įvertinant tai, kad spėjamas R-kilpos ilgis yra 20 nt ir DNR spiralės žingsnis 10,5 bp). Toks skirtumas galėjo atsirasti dėl DNR superspirališkumą kompensuojančio Cas9 sukkelto DNR lenkimo.

## 2.4.2. PAM mutacijos trukdo R-kilpos susidarymui, bet neturi įtakos jos stabilumui

Norėdami nustatyti PAM sekos įtaką R-kilpos susidarymo greičiui ir jos stabilumui, magnetinių pincetų metodo pagalba tyrėme R-kilpos susidarymo ir iširimo dinamiką naudodami DNR substratus su mutacijomis PAM sekoje. Nors R-kilpų susidarymas buvo stebimas naudojant substratus su G4C PAM mutacija, tačiau tokios struktūros susidarė gerokai lėčiau nei naudojant laukinio tipo substratą (NGGNG) (9 pav.). Įdomu tai, kad susiformavus R-kilpai, jos stabilumas išliko nepakitęs lyginant su laukinio tipo substratu (9 pav.). Naudojant G1C ir G2C substratus R-kilpos susidarymas buvo dar lėtesnis, o PAM delecijos atveju R-kilpos nesiformavo net sukeltos DNR denatūracijos metu (duomenys nepateikti). Patvirtindami duomenis gautus magnetinių pincetų metodu, DNR kirpimo eksperimentai leido daryti prielaidą, kad R-kilpos formavimasis buvo lėtesnis, naudojant substratus su PAM mutacijomis (mutacijų įtaka pasireiškė tokia tvarka G2C>G1C>G4C>>WT), tačiau jau susidariusios R-kilpos kirpimui, mutacijų įtaka buvo mažesnė (10 A, B ir D pav.). Taigi, šie rezultatai rodo, kad PAM seka reguliuoja R-kilpos susidarymo greitį, tačiau neturi įtakos jos stabilumui.



**9 pav. PAM mutacijų įtaką R-kilpos susidarymui ir suirimui.** Vidutinio R-kilpų susidarymo ir disociacijos laiko priklausomybė nuo sukimo momento lyginant DNR substratus turinčius kanoninę PAM seką su G4C PAM mutaciją turinčiais substratais.

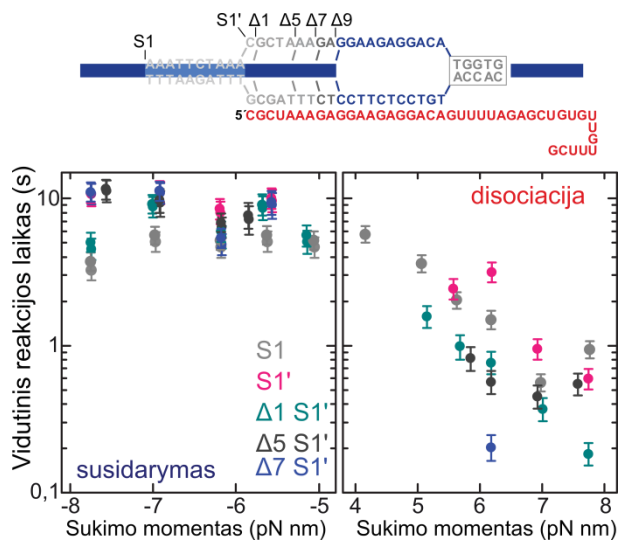


**10 pav. Plazmidinės DNR karpymas su Cas9 kompleksu naudojant substratus su įvairaus ilgio proskirtuku bei mutacijomis PAM sekoje.** (A) Plazmidės, turinčios pilno ilgio proskirtuką ir NGGNG PAM seką karpymas su Cas9 kompleksu. (B) Plazmidės, turinčios pilno ilgio proskirtuką ir mutuotą G2C PAM seką karpymas su Cas9 kompleksu. (C) Plazmidės, turinčios sutrumpintą proskirtuką (S1' Δ7 bp) ir NGGNG PAM seką karpymas su Cas9 kompleksu. (D) Pirmos ir antros DNR grandinių kirpimo greičiai, naudojant substratus su sutrumpintais proskirtukais (kairė dalis) ar mutacijomis PAM sekoje (dešinė dalis).  $k_1$  yra nikuoto produkto susidarymo greičio konstanta, o  $k_2$  yra linijinio DNR fragmento susidarymo greičio konstanta. Karpant substratus su PAM mutacijomis, nikuoto produkto kaupimasis nebuvo stebimas (C), kas leidžia daryti prielaidą, kad  $k_1 \ll k_2$ .

### 2.4.3. Proskirtuko galo trumpinimas destabilizuoja Cas9 suformuotas R-kilpas

Atliekant reakcijas su Cas9, R-kilpų susidarymas buvo stebimas trumpinant proskirtuką iki 7 bp jų formavimosi greičiui beveik nekintant (11 pav.). Esant 9 bp sutrumpinimui R-kilpos nesiformavo. 1 ar 5 bp sutrumpinimai nežymiai sumažino R-kilpos stabilumą ir DNR kirpimo greitį (11 ir 10 pav.). Esant 7 bp sutrumpinimui



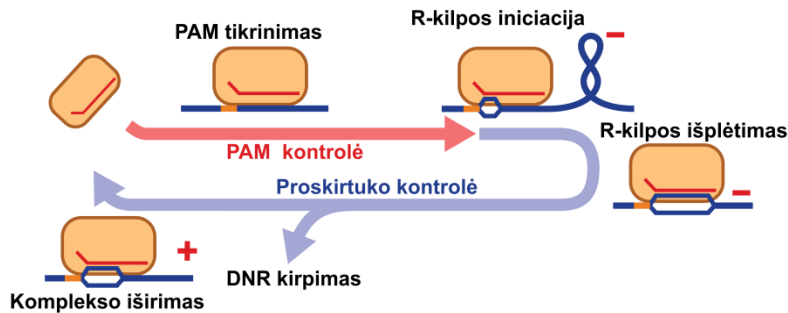


**11 pav. Proskirtuko galo trumpinimas mažina R-kilpų stabilumą.** Vidutinio R-kilpų susidarymo ir disociacijos laiko priklausomybė nuo sukimo momento, esant S1' proskirtuko galo trumpinimams.

R-kilpos buvo aptinkamos (11 pav.), tačiau jos pasižymėjo lėtesniu, nuo sukimo momento nepriklausomu disociacijos greičiu. Tai galėtų būti paaiškinama, kad mutacijos tarp crRNR ir proskirtuko gali sukelti Cas9 RNP komplekso persitvarkymus į užslopintą formą. Trumpinant proskirtuką DNR kirpimo greitis lėtejo. Taip pat, trumpėjant proskirtukui buvo stebimas nikuotos DNR formos kaupimasis (10 C ir D pav.).

#### 2.4.4. Cas9 inicijuoto R-kilpos susidarymo ir disociacijos modelis

Duomenys, gauti atlikus eksperimentus aprašytus aukščiau, leidžia teigti, kad PAM seka yra svarbi R-kilpos susidarymui, tačiau praktiškai neturi įtakos jos stabilumui. Mutacijos proskirtuko 5'-gale turėjo įtakos R-kilpos stabilumui, tačiau jos susiformavimo greitis išliko nepakitęs. Tai rodo, kad R-kilpa formuojasi kryptingai nuo PAM sekos (12 pav.). Pirmiausia, Cas9 kompleksas prisijungia prie PAM sekos ir destabilizuoja DNR. Kadangi R-kilpos formavimosi greitis priklauso nuo Cas9 komplekso koncentracijos (duomenys nepateikti), tai leidžia teigti, kad 3D difuzija yra svarbi Cas9 kompleksui ieškant PAM sekos. Atpažinus PAM, dėka vandenilinių ryšių susidarymo tarp crRNR ir taikinio DNR, pradeda kryptingai formotis R-kilpa. Esant nepalankioms energetinėms sąlygoms (DNR veikiant teigiamam sukimo momentui ar mutacijoms tarp crRNR ir DNR), R-kilpa suyra nepaisant atpažintos PAM sekos (12 pav.).



**12 pav. R-kilpos susidarymo ir disociacijos modelis.** (+) ir (-) simboliai rodo atitinkamos superspiralizacijos efektą vaizduojamai stadijai.

Cas9 gali efektyviai kirpti proskirtuką, esant 7 bp 5'-galo sutrumpinimui, nors ir lėčiau nei esant pilno ilgio proskirtukui (10 pav.). Esant proskirtukui 11 bp (9 bp 5'-galo sutrumpinimas), R-kilpos nesiformavo, kas rodytų, kad Cas9 baltymas "tikrina" koks yra R-kilpos ilgis, taip užtikrindamas, kad tik "tinkami" DNR taikiniai būtų kerpami. Struktūriškai tai būtų galima bandyti paaiškinti tuo, kad Cas9 baltymas sudaro daugybę nespecifinių kontaktų su crRNR ir taikinio DNR heterodupleksu šioje srityje (Nishimasu et al., 2014), kas leidžia jam patikrinti ar R-kilpa yra susidariusi korektiškai.

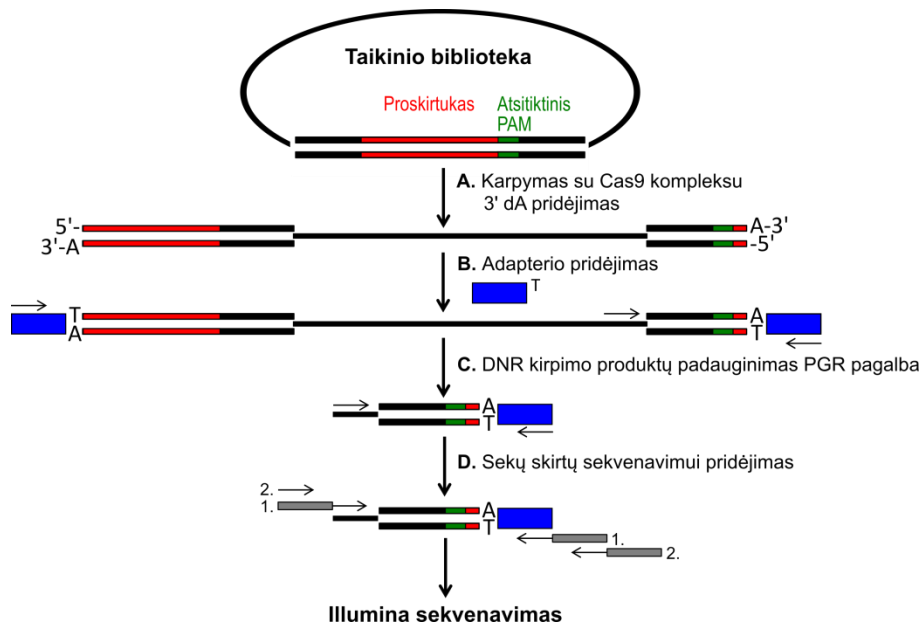
## 2.5. CRISPR-Cas9 PAM sekos identifikavimas

Cas9 nukleazės taikiny yra sudarytas iš proskirtuko ir greta esančios PAM sekos. Kaip aprašyta aukščiau ir Sternberg et al., 2014, taikinio atpažinimo mechanizmas susideda iš Cas9 sąveikos su PAM seka ir crRNR sąveikos su komplementaria proskirtuko seka. Taigi, PAM seka yra svarbiausias veiksnys ribojantis DNR taikinio parinkimo galimybes. Norint išspręsti šią problemą buvo panaudotos kelios strategijos: PAM sekos atpažinimo specifiškumo keitimas modifikuojant Cas9 baltymus (Anders et al., 2016; Hirano et al., 2016a; Kleinstiver et al., 2015a, 2015b), naujų Cas9 baltymų paieška, ar abiejų strategijų pritaikymas. Taip pat, ortologiniai Cas9 baltymai leistų ne tik padidinti sekų, į kurias būtų galima nukreipti Cas9 baltymą skaičių, bet kartu galėtų turėti kitų savybių, naudingų genomo modifikavimo eksperimentuose.

Taigi, norint charakterizuoti naujus Cas9 baltymus, būtina sukurti metodą, leidžiantį tiksliai, greitai ir tiesiogiai nustatyti jo atpažįstamą seką.

### 2.5.1. Cas9 PAM sekų nustatymas

Atsitiktines PAM sekas turinčios bibliotekos buvo sukonstruotos ir panaudotos empiriniam Cas9 atpažįstamų PAM sekų nustatymui (11 pav.). Tokiose bibliotekose, proskirtuko sekai esant fiksuotai, PAM srityje esantys atsitiktiniai nukleotidai yra panaudojami kaip substratas tiesioginiam Cas9 PAM sekos specifiškumui nustatyti. Šiam tikslui buvo sukonstruotos dvi bibliotekos besiskiriančios savo kompleksiskumu: atsitiktinių 5 bp PAM biblioteka (1024 PAM sekų kombinacijos) ir 7 bp PAM biblioteka (16384 PAM sekų kombinacijos). Šios plazmidžių bibliotekos buvo naudojamos kaip substratas karpymo su Cas9 eksperimentuose *in vitro*. Jei PAM seka atpažįstama, tokia plazmidė perkerpama ir tolimesniame etape prie susidariusių laisvų plazmidės galų yra pridedamas adapteris (11 pav. A ir B etapai). Norint padidinti šio etapo efektyvumą, prie DNR 3'-galų yra pridedami dA nukleotidai (adapteris atitinkamai prie 3'-galo turi papildomą dT nukleotidą). Tokie DNR fragmentai (gauti tik esant plazmidės kirpimui) PGR pagalba padauginami (11 pav. C etapas) ir paruošiami Illumina sekvenavimui (11 pav. D etapas). Po sekvenavimo ir analizės gauti rezultatai yra vizualizuojami kaip WebLogo (Crooks, 2004).

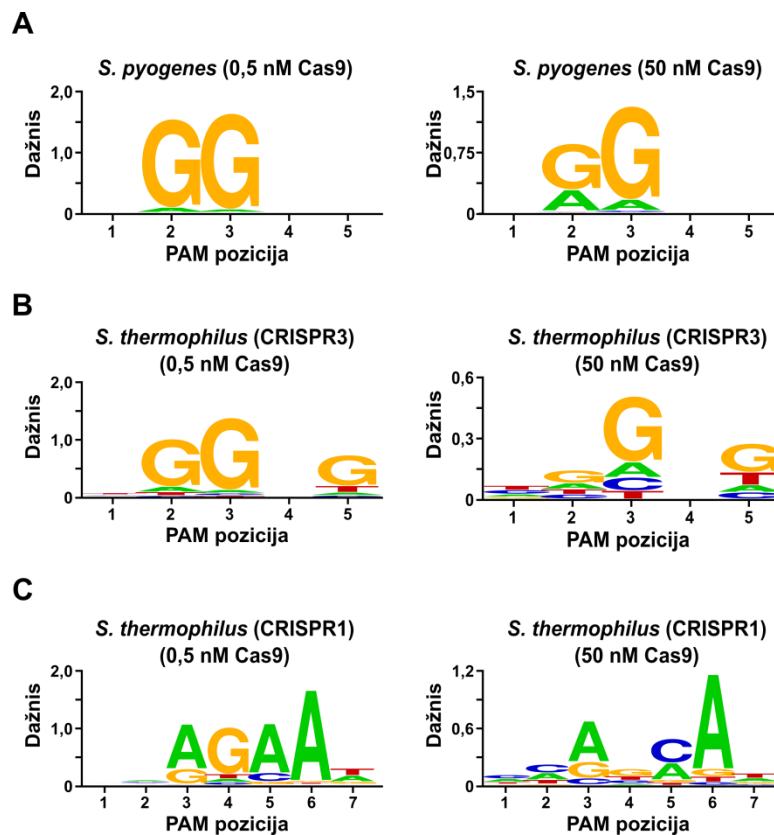


**11 pav. Cas9 atpažįstamos PAM sekos nustatymo schema.** A. Plazmidžių biblioteka su atsitiktinėmis PAM sekomis yra karpoma su Cas9 kompleksu ir prie atsiradusių laisvų DNR galų yra pridedama dA. B. Pridedami adapteriai. C. PGR pagalba padauginama perkirptų plazmidžių PAM sritis. D. DNR fragmentai paruošiami Illumina sekvenavimui.

## 2.5.2. *Streptococcus pyogenes* ir *Streptococcus thermophilus* (CRISPR3 ir CRISPR1 sistemų) Cas9 atpažįstamų PAM sekų nustatymas

Norėdami įsitikinti, kad mūsų metodas yra tinkamas PAM sekų charakterizavimui, pasirinkome nustatyti PAM sekas Spy ir Sth3 baltymams, kurių PAM sekos yra nustatytos kitais metodais (Gasiunas et al., 2012; Horvath et al., 2008; Jinek et al., 2012; Sapranauškas et al., 2011). Karpymo eksperimentuose buvo naudojamas 1 µg (5,6 nM) 5 bp PAM bibliotekos ir 0,5 ar 50 nM Cas9 kompleksai. Reakcija buvo atliekama 1 val. 100 µl reakcijos tūryje. Plazmidžių turinčių Spy ir Sth3 Cas9 atpažįstamų PAM sekų (atitinkamai NGG ir NGGNG) 5 bp PAM bibliotekoje koncentracija buvo atitinkamai 0,40 nM ir 0,11 nM.

Po sekvenavimo ir analizės gauti rezultatai parodė (12 A ir B pav.), kad abiejų baltymų atvejais buvo identifikuoti PAM, anksčiau nustatyti kitais metodais (Spy Cas9 – NGG (Jinek et al., 2012) ir Sth3 Cas9 – NGGNG (Gasiunas et al., 2012; Horvath et al., 2008; Sapranauškas et al., 2011)). Nors kanoninės PAM sekos buvo sėkmingai



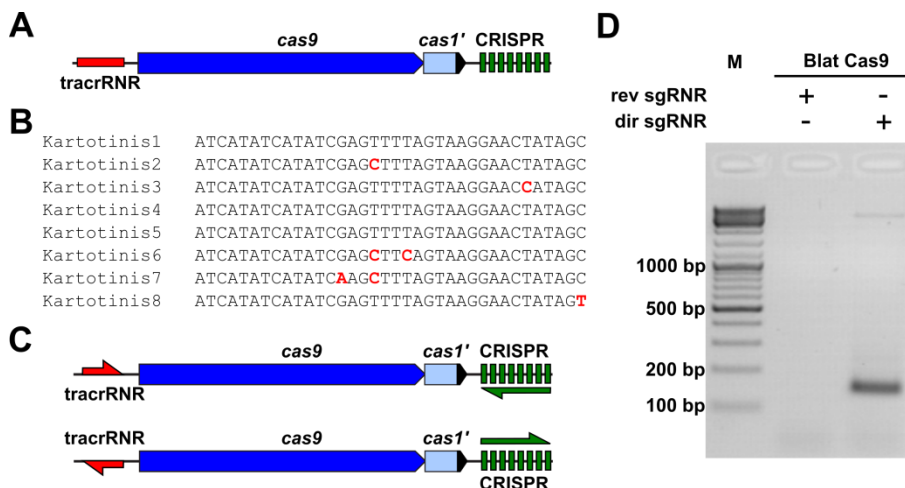
12 pav. *S. pyogenes* (A), *S. thermophilus* CRISPR3 (B) ir *S. thermophilus* CRISPR1 (C) Cas9 nustatytos atpažįstamos PAM sekos. Charakterizuotos PAM sekos vaizduojamos kaip WebLogo (Crooks, 2004).

identifikuotos naudojant skirtingas baltymų koncentracijas (0,5 nM ir 50 nM), tačiau esant didesnei baltymo koncentracijai stebimas PAM sekos specifiškumo sumažėjimas.

Sudėtingesniam metodo patikrinimo atvejui buvo pasirinktas Cas9 baltymas iš *Streptococcus thermophilus* CRISPR1 sistemos (Sth1), kurio atpažįstama PAM seka yra „išsitiesusi“ iki 7 bp nuo proskirtuko 3'-galo (Fonfara et al., 2014; Horvath et al., 2008). Karpymo reakcijos buvo atliekamos kaip ir Spy ir Sth3 Cas9 atvejais, tačiau buvo naudota 7 bp PAM biblioteka. Šiuo atveju plazmidžių turinčių Sth1 Cas9 PAM seką (NNAGAAW) (Horvath et al., 2008) koncentracija buvo 0,01 nM. Tokia PAM seka ir buvo nustatyta, naudojant 0,5 nM koncentracijos Sth1 Cas9 kompleksą reakcijos mišinyje (12 C pav.). Esant didesnei baltymo koncentracijai, kaip ir Spy ar Sth3 Cas9 atvejais, PAM sekos atpažinimo specifiškumas buvo mažesnis. Šie rezultatai sutampa su duomenimis gautais iš *in vivo* eksperimentų, kurie parodė, kad Cas9 koncentracijos ar karpymo laiko ribojimas ženkliai sumažina nepageidaujamo kirpimo dažnį ne taikinio sekose (Lin et al., 2014; Pattanayak et al., 2013). Apibendrinant, mūsų metodas leidžia charakterizuoti nuo Cas9 komplekso koncentracijos priklausomas PAM sekas *in vitro*.

### **2.5.3. PAM ir RNR komponentų identifikavimas *Brevibacillus laterosporus* Cas9 baltymui**

Norėdami nustatyti PAM seką iki šiol dar necharakterizuotam baltymui, pasinaudojome DuPont Pioneer vidine duomenų baze ir pasirinkome Cas9 iš *Brevibacillus laterosporus* SSP360D4 (Blat) II-C tipo CRISPR-Cas sistemos (13 A ir B pav.). Negalėdami tokioje sistemoje identifikuoti tracrRNR ir crRNR transkripcijos krypties, nusprendėme patikrinti abu galimus variantus (13 C pav.). Dirbtinai sulietų tracrRNR ir crRNR molekulių (sgRNR) variantas, kuriame tracrRNR yra transkribuojama ta pačia kryptimi kaip ir Cas9 genas buvo pavadinta tiesiogine sgRNR (dir sgRNR), kai tuo tarpu atvirkštinis variantas – atvirkštine sgRNR (rev sgRNR) (13 C pav.). 50 nM koncentracijos Blat Cas9 kompleksas, surinktas naudojant dir ar rev sgRNR molekules buvo inkubuojamas su 1 µg (5,6 nM) 7 bp PAM biblioteka. Po karpymo reakcijos, adapterių pridėjimo ir PGR padauginimo, DNR produktų analizė agaroziniame gelyje parodė, kad tik esant dir sgRNR variantui, Blat Cas9 kompleksas buvo aktyvus (13 D pav.).

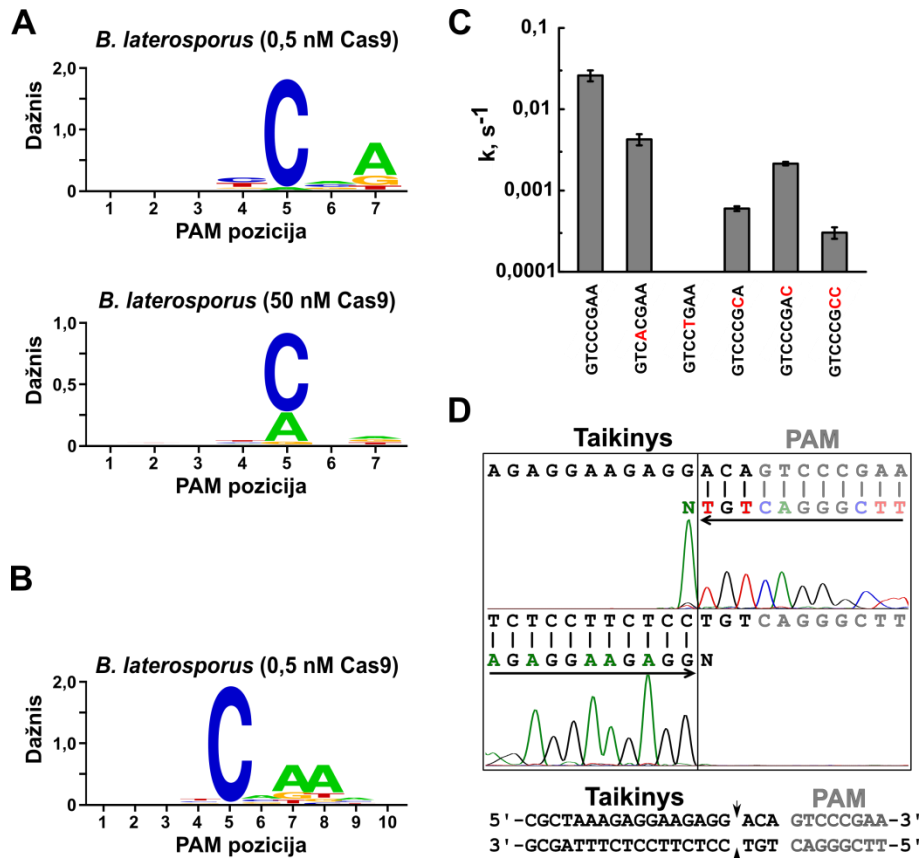


**13 pav. *Brevibacillus laterosporus* SSP360D4 CRISPR-Cas sistemas charakterizavimas.** (A) *Brevibacillus laterosporus* SSP360D4 II-C tipo CRISPR-Cas sistema. (B) CRISPR srities kartotinių sekų palyginimas. (C) Tiesioginis ir atvirkštinis tracrRNR ir CRISPR srities transkripcijos scenarijai. (D) Gelis su padaugintais plazmidžių bibliotekos karpymo produktais, rodantis, kad tik tiesioginio transkripcijos scenarijaus atveju (dir sgRNR) yra gaunamas aktyvus *Brevibacillus laterosporus* SSP360D4 Cas9 kompleksas (Blat Cas9).

Taigi, nustačius su koku sgRNR variantu Blat Cas9 baltymas yra aktyvus, buvo siekiama charakterizuoti jo atpažįstamą PAM seką kaip aprašyta Spy, Sth3 ir Sth1 Cas9 atvejais ir naudojant 7 bp PAM biblioteką. Kaip parodyta 14 A pav. atpažįstama PAM seka, esant 0,5 nM Blat Cas9 komplekso koncentracijai buvo NNNNCND (D=A, G ar T), esant stipriai išreikštam C nukleotidui 5 pozicijoje. Efektyviausiai buvo atpažįstama PAM seka, kurioje 7 pozicijoje buvo A, 4 pozicijoje C ar T, o 6 pozicijoje G, C ar A nukleotidai. Kaip ir Spy, Sth3 ir Sth1 Cas9 atvejais koncentracijos padidinimas (iki 50 nM) sumažino PAM specifiškumą – 5 pozicijoje buvo galima aptikti A nukleotidą.

Kadangi Blat Cas9 atpažįstamoje PAM sekoje pirmi trys nukleotidai nėra svarbūs (14 A pav.), mes nusprendėmė „paslinkti“ skirtuką 3 nt 5' kryptimi, taip išplėsdami PAM nustatymo galimybes nuo 7 bp iki 10 bp. Po bibliotekos karpymo su Blat Cas9 kompleksu ir analizės paaiškėjo, kad prie aukščiau aprašytų svarbių PAM nukleotidų, taip pat yra svarbus 8 PAM pozicijoje esantis A nukleotidas (14 B pav.).

Plazmidžių karpymo eksperimentai, kuriuose konservatyviuose Blat Cas9 PAM pozicijose buvo įvestos mutacijos, patvirtino, kad nustatyti nukleotidai yra svarbūs PAM atpažinimui ir efektyviam plazmidžių kirpimui (14 C pav.). Taip pat, nusekvenavus



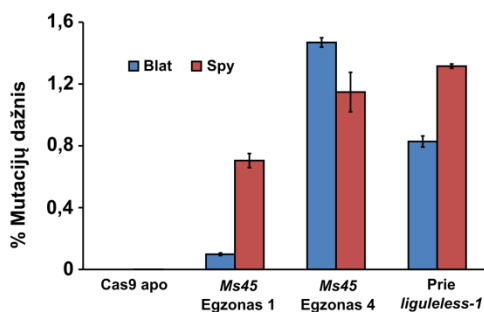
**14 pav. Blat Cas9 PAM sekos ir kirpimo vietos identifikavimas.** (A) Blat Cas9 PAM seka buvo nustatoma, kai 1  $\mu$ g (5,6 nM) PAM bibliotekos buvo karpoma su 0,5 nM ir 50 nM koncentracijos Cas9 kompleksu. (B) PAM seka, nustatyta paslinkus skirtuką 3 nt 5' kryptimi. (C) Plazmidžių, turinčių mutacijas PAM sekoje (GTCCCGAA) karpymo greičių palyginimas. (D) Blat Cas9 komplekso plazmidės kirpimo vietos nustatymas.

Plazmidės, turinčios Blat Cas9 komplekso atpažįstamą PAM seką, kirpimo produktus paaiškėjo, kad Blat Cas9 kompleksas kirpo plazmidę proskirtuko srityje 3 nt nuo PAM sekos (14 D pav.), kaip ir Spy, Sth3 ar Sth1 Cas9 baltymai (Garneau et al., 2010; Gasiunas et al., 2012; Jinek et al., 2012).

#### 2.5.4. Augalų genomo redagavimas naudojant Blat Cas9 kompleksą

Nustačius Blat Cas9 atpažįstamą seką, siekėme pademonstruoti ar šis baltymas gali būti pritaikytas augalų genomų redagavimui. Cas9 ir sgRNR genų sekų atliktas optimizavimas efektyviai raiškai kukurūzo ląstelėse, bei tokių konstrukto biolistinė transformacija į nesubrendusias kukurūzų embrionines ląsteles yra detalai aprašyta

	Taikinys ↓	PAM	Dažnis	Dažnis (Cas9 apo)
	CCAGGATAATGAGGTACTGGCT	GGAAGGCCCAAGAGCGGGCGAGGT		
Mutation 1	CAGGATAATGAGGTACTGGCT	tGGAAGGCCCAAGAGCGGGCGAGGT	4,553	0
Mutation 2	CAGGATAATGAGGTACTGGCT	aGGAAGGCCCAAGAGCGGGCGAGGT	1,110	0
Mutation 3	CAGGATAATGAGGTACTGGCT	gGGAAGGCCCAAGAGCGGGCGAGGT	281	0
Mutation 4	CAGGATAATGAGGTACTGGCT	cGGAAGGCCCAAGAGCGGGCGAGGT	120	0
Mutation 5	CCAGGATAATGAGGTACTGGC	-GGAAGGCCCAAGAGCGGGCGAGGT	119	0
Mutation 6	AGGATAATGAGGTACTGGCT	tGGAAGGCCCAAGAGCGGGCGAGGT	112	0
Mutation 7	CAGGATAATGAGGTACTGGC	gTGAAGGCCCAAGAGCGGGCGAGGT	111	0
Mutation 8	CCAGGATAATGAGGTACTGGCT	--AAGGCCCAAGAGCGGGCGAGGT	87	0
Mutation 9	CCAGGATAAT	-----GAGGT	79	0
Mutation 10	CCAGGATAATGAGGTACT	-----GGCCCAAGAGCGGGCGAGGT	63	0
	Sekvenavimo gilumas:		450,873	755,448



**15 pav. Blat Cas9 sukelia NHEJ mutacijas *in vivo*.** (A) 10 dažniausiai pasitaikančių mutacijų *Ms45* geno 4 egzone po poveikio Blat Cas9 kompleksu. Mažosios raidės rodo insercijas, „-“ – delecijas. (B) Spy ir Blat Cas9 kompleksų mutacijų sukėlimo dažnio palyginimas (sekvenavimo rezultatai).

Karvelis et al., 2015. Šiam tikslui buvo pasirinkti trys taikiniai kurkurūzo genome: 1 ir 4 *Ms45* geno egzonuose ir srityje esančioje greta *liguleless-1* geno. Po ląstelių su DNR raiškos vektoriais, turinčiais sgRNR ir Cas9 genus, transformacijos, ląstelės buvo auginamos 3 dienas. Toliau buvo išskiriama genomine DNR ir PGR pagalba buvo padauginamos taikinius turinčios DNR sritys ir sekvenuojama. Insercijos ir delecijos buvo aptinkamos Blat Cas9 kirpimo vietose (15 A pav.) ir jų dažnis buvo palyginamas su Spy Cas9 sukeltų mutacijų dažniu (15 B pav.).

Mūsų sukurtas metodas prisideda prie pastangų charakterizuoti naujus Cas9 baltymus ir performuluoja PAM sekos supratimą, kaip seką, kurios atpažinimo specifiškumas priklauso nuo Cas9 komplekso koncentracijos. Pritaikę šį metodą, charakterizavome netirtą Blat Cas9 baltymą ir pademostravome jo funkcinį aktyvumą *in vivo*, taip parodydami metodo galimybes naujų Cas9 charakterizavimui, bei išplėsdami Cas9 paremtų įrankių pasirinkimą.



## IŠVADOS

1. Funkciškai aktyvus Cas9 ribonukleoproteininis kompleksas iš *S. thermophilus* CRISPR3-Cas sistemos gali būti surinktas *in vitro*, sumaišant Cas9 baltymą su tracrRNR ir crRNR.
2. Cas9 kompleksas gali būti panaudotas kaip molekulinis įrankis perkirpti taikinio DNR *in vitro*.
3. *In vitro* surinktas ribonukleoproteininis Cas9 kompleksas cheminės transfekcijos pagalba gali būti perneštas į eukariotines ląsteles ir panaudotas genomo redagavimui *in vivo*.
4. Cas9 komplekso atpažįstama PAM seka naudojama kaip R-kilpos iniciacijos vieta kryptingai crRNR hibridizacijai su DNR taikiniu.
5. Sukurtas PAM nustatymo metodas leidžia eksperimentiškai nustatyti Cas9 atpažįstamas PAM sekas.
6. Cas9 iš *B. laterosporus* SSP360D4 atpažįsta NNNNCNDD PAM seką ir yra funkciškai aktyvus augalų ląstelėse.

## MOKSLINIŲ DARBŲ SĄRAŠAS

### Disertacijoje pateikta medžiaga paskelbta šiuose moksliniuose staipsniuose:

1. **Karvelis, T.**, Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P., and Siksnys, V. (2013). crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. RNA Biol. 10, 20–19.
2. **Karvelis, T.**, Gasiunas, G., and Siksnys, V. (2013). Programmable DNA cleavage *in vitro* by Cas9. Biochem. Soc. Trans. 41, 1401–1406.
3. Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., **Karvelis, T.**, Pschera, P., Siksnys, V., and Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proc. Natl. Acad. Sci. U. S. A. 111, 9798–9803.
4. Glemzaite, M., Balciunaite, E., **Karvelis, T.**, Gasiunas, G., Grusyte, M.M., Alzbutas, G., Jurcyte, A., Anderson, E.M., Maksimova, E., Smith, A.J., et al. (2015). Targeted gene editing by transfection of *in vitro* reconstituted *Streptococcus thermophilus* Cas9 nuclease complex. RNA Biol. 12, 1–4.
5. **Karvelis, T.**, Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., Cigan, M., and Siksnys, V. (2015). Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. Genome Biol. 16, 253.

### Patentinės paraiškos:

1. Cigan A. M., Gasiunas G., **Karvelis T.**, Siksnys V., Young J. K. BB2475-US-PSP2, 62/196535. 2015.07.24
2. Cigan A. M., Gasiunas G., **Karvelis T.**, Siksnys V., Young J. K. BB2539-US-PSP, 62/162377. 2015.05.15
3. Cigan A. M., Gasiunas G., **Karvelis T.**, Siksnys V., Young J. K. BB2475-US-PSP, 62/162353. 2015.05.15
4. Siksnys V., Gasiunas G., **Karvelis T.**, Lubys A., Zaliauskiene L., Glemzaite M., Smith A. WO/2013/142578, PCT/US2013/033106. 2013.03.20.
5. Šikšnys V., Gasiūnas G., **Karvelis T.** WO/2013/141680, PCT/LS2013/000006. 2013.03.15.

### Disertacijoje pateikta medžiaga pristatyta šiose mokslinėse konferencijose:

#### Stendiniai pranešimai:

1. **Karvelis T.**, Gasiunas G., Miksys A., Barrangou R., Horvath P. and Siksnys V. Cas9-mediated DNA interference in *Streptococcus thermophilus*: composition and assembly of the silencing complex. CRISPR: Evolution, Mechanisms and Infection; University of St Andrews, UK; 2013.06.17-19.
2. Gasiunas G., **Karvelis T.** and Siksnys V. Cas9 – a programmable DNA endonuclease for *in vitro* applications. CRISPR: Evolution, Mechanisms and Infection; University of St Andrews, UK; 2013.06.17-19.

3. Gasiunas G., **Karvelis T.**, Barrangou R., Horvath P. and Siksnys V. Cas9 – a programmable RNA-guided DNA endonuclease from the bacterial adaptive immune system. 13th Young Scientists Forum (YSF); Saint Petersburg, Russia; 2013.07.03-06.
4. **Karvelis T.**, Gasiunas G., Miksys A., Grusyte M. M., Siksnys V. Molecular determinants of DNA cleavage by Cas9 endonuclease. CRISPR 2014 - The Prokaryotic Immune System CRISPR/Cas; Berlin, Germany; 2014.05.13-16.
5. Szczelkun M. D., Gasiunas G., **Karvelis T.**, Seidel R., Siksnys V. Direct observation of R-loop formation by RNA-guided Cas9 effector complex. CRISPR 2014 - The Prokaryotic Immune System CRISPR/Cas; Berlin, Germany; 2014.05.13-16.
6. Gasiunas G, **Karvelis T**, Glemzaite M, Zaliauskiene L and Siksnys V. Cas9 is able to replace the restriction enzymes. 7<sup>th</sup> NEB Meeting on DNA Restriction and Modification. Gdansk, Poland; 2015.08.24-29.
7. **Karvelis, T.**, Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., Cigan, M., and Siksnys, V. Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. CRISPR 2016; Rehovot, Israel; 2016.05.23-25.

#### **Žodinis pranešimas:**

1. **Karvelis, T.**, Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., Cigan, M., and Siksnys, V. Rapid characterization of CRISPR-Cas9 protospacer adjacent motif (PAM) sequence elements. XIV<sup>th</sup> Conference of the Lithuanian Biochemical Society; Druskininkai, Lithuania; 2016.06.28-30.

## **FINANSINĖ PARAMA**

Disertacijoje aprašyti darbai buvo finansuoti Visuotinės dotacijos (R100), Mokslo, inovacijų ir technologijų agentūros (MITA) (#31v-13 ir #31v-36) projektų lėšomis. Stipendijas už mokslinius pasiekimus skyrė Lietuvos mokslo taryba.

## CURRICULUM VITAE

**Vardas, Pavardė** Tautvydas Karvelis

**Gimimo data** 1987-04-16

**Darbo adresas** Baltymų-nukleorūgščių sąveikos tyrimų skyrius  
Vilniaus universiteto Biotechnologijos institutas  
Saulėtekio al. 7 LT-10222 Vilniu

**Telefonas** +370-673-40011 (asmeninis)

**E-paštas** tautvydas.karvelis@bti.vu.lt

### Išsilavinimas

2005-2009 Biochemijos bakalauras  
Vilniaus Universitetas

2009-2011 Biochemijos magistras  
Vilniaus Universitetas

### Darbo patirtis

2011-dabar Jaunesnysis mokslo darbuotojas Vilniaus universiteto  
Biotechnologijos instituto Baltymų-nukleorūgščių sąveikos tyrimų  
skyriuje

## **PADEKA**

Esu dėkingas savo vadovui prof. dr. Virginijui Šikšniui už galimybę dirbti šioje srityje ir taip pat už diskusijas, patarimus ir pagalbą dirbant su projektu aprašytu šioje disertacijoje.

Dėkoju dr. Giedriui Gasiūnui už bendrą darbą, diskusijas, ir pagalbą, kuri tęsiasi iki dabar.

Norėčiau padėkoti dr. M. D. Szczelkun (Bristolio Universitetas) už atliktus pavienių molekulių eksperimentus, M. Glemžaiti ir dr. L. Zaliauskienei (Thermo Fisher Scientific Baltics) už atliktus ląstelių transfekcijos eksperimentus, J. Young ir dr. M. Cigan (Dupont Pioneer) už pagalbą kuriant PAM nustatymo metodą, mėginių sekvenavimą ir genų inžinerijos eksperimentus kukurūzo ląstelėse. Taip pat visiems publikacijų bendraautoriams už bendradarbiavimą be kurio nebūtų šios disertacijos.

Esu dėkingas visam Baltymų-nuklorūgščių sąveikos tyrimo skyriaus kolektyvui už diskusijas, patarimus, pagalbą atliekant eksperimentus ir gerą darbo atmosferą.

Taip pat dėkoju dr. Rolandui Meškiui ir dr. S. Servai už kritišką šios disertacijos skaitymą ir komentarus.

## SUMMARY

Nowadays, DNA sequencing technologies enable rapid sequencing of complete genomes within a few days. However, our abilities to manipulate DNA sequences *in vivo* in order to understand or change biological traits or functions are lagging behind due to lack of available molecular tools for genome manipulation. Naturally occurring or engineered homing endonucleases, ZFN or TALEN meganucleases that introduce double stranded breaks (DSBs) in DNA and trigger repair mechanisms that may result in gene knock-outs/knock-ins appeared as first promising genome editing tools. However, development of these proteins for precise genome manipulation requires extensive protein engineering for reprogramming of the amino acid code for a specific DNA sequence and often results in tools with low reproducibility, efficiency and specificity (Carroll, 2014; Porteus, 2016).

The recent development of the Cas9 technology revolutionized the genome editing field. As with the most revolutionary discoveries, Cas9 nuclease genome editing technology started by trying to answer fundamental biological questions that were not directly related to genome manipulation. In 2007, CRISPR (clustered regularly interspaced short palindromic repeats)-Cas (CRISPR associated) systems emerged as a novel bacterial defense system against viruses and plasmids (Barrangou et al., 2007). Soon it was revealed that CRISPR-Cas systems use small RNA molecules called CRISPR RNA (crRNA) to guide ribonucleoprotein complexes to foreign nucleic acids and trigger their degradation (Brouns et al., 2008). RNA-based recognition of foreign nucleic acids through the Watson-Crick type basepairing, offered unprecedented opportunity for adaptation of such systems for genome editing. Type II CRISPR-Cas systems emerged as most promising candidates, because of the relatively simple CRISPR-Cas locus composition (Makarova et al., 2015). It turned out that Type II systems require a single Cas9 protein bound to a dual crRNA:tracrRNA (trans-activating crRNA) molecules for destruction of foreign DNA (Deltcheva et al., 2011; Sapranaukas et al., 2011). The crRNA is used as guide for DNA targeting while tracrRNA is involved in bacterial RNase III dependent pre-crRNA maturation.

DNA target recognized by the Cas9 complex is composite and is comprised of DNA sequence (protospacer) complementary to the spacer component of crRNA and short adjacent sequence termed PAM (protospacer adjacent motif) that is absolutely required for efficient DSB formation (Garneau et al., 2010; Sapranaukas et al., 2011). PAM sequence is recognized by the Cas9 protein and is uniquely associated with each Cas9 protein. These fundamental studies of Type II CRISPR-Cas systems in bacteria set the stage for the genome editing experiments in eukaryotes. This coincides with the start and focus of the study described in this PhD thesis.

In this work, for the first time we identified all molecular components required for an assembly of the functional Cas9 complex of *S. thermophilus* CRISPR3-Cas system

and demonstrated targeted DNA cleavage by the reconstituted Cas9 complex *in vitro* and *in vivo*. More specifically, we showed for the first time that Cas9 complex can be used as a tool for DNA cloning and targeted genome editing by chemical transfection of the *in vitro* assembled Cas9 ribonucleoprotein complex into the cells. Aiming to understand the molecular mechanism governing DNA recognition and cleavage by the Cas9 complex, we analyzed mechanism of R-loop formation and provided first direct evidence for directional R-loop formation, starting from PAM recognition and expanding toward the distal protospacer end. Realizing the importance of PAM sequence for Cas9 function, we developed an assay for rapid PAM identification for newly identified Cas9 proteins and characterized Type II-C Cas9 protein of *Brevibacillus laterosporus*, expanding Cas9 toolbox for genome editing applications.



## LITERATŪROS SĄRAŠAS

1. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712.
2. Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E. V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960–964.
3. Brutzer, H., Luzziatti, N., Klaue, D., and Seidel, R. (2010). Energetics at the DNA Supercoiling Transition. *Biophys. J.* 98, 1267–1276.
4. Carroll, D. (2011). Genome engineering with zinc-finger nucleases. *Genetics* 188, 773–782.
5. Carroll, D. (2014). Genome engineering with targetable nucleases. *Annu. Rev. Biochem.* 83, 409–439.
6. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* 339, 819–823.
7. Crooks, G.E. (2004). WebLogo: A Sequence Logo Generator. *Genome Res.* 14, 1188–1190.
8. Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z. a, Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602–607.
9. Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K.S., Lécrivain, A.L., Bzdrenga, J., Koonin, E. V., and Charpentier, E. (2014). Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* 42, 2577–2590.
10. Forth, S., Deufel, C., Sheinin, M.Y., Daniels, B., Sethna, J.P., and Wang, M.D. (2008). Abrupt Buckling Transition Observed during the Plectoneme Formation of Individual DNA Molecules. *Phys. Rev. Lett.* 100, 148301.
11. Gabriel, R., Lombardo, A., Arens, A., Miller, J.C., Genovese, P., Kaepfel, C., Nowrouzi, A., Bartholomae, C.C., Wang, J., Friedman, G., et al. (2011). An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* 29, 816–823.
12. Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D. a, Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71.
13. Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci.* 109, 2579–2586.
14. Glemzaitė, M., Balciunaite, E., Karvelis, T., Gasiunas, G., Grusyte, M.M., Alzbutas, G., Jurcyte, A., Anderson, E.M., Maksimova, E., Smith, A.J., et al. (2015). Targeted gene editing by transfection of *in vitro* reconstituted *Streptococcus thermophilus*

- Cas9 nuclease complex. *RNA Biol.* *12*, 1–4.
15. Hofacker, I.L., Fontana, W., Stadler, P.F., Bonhoeffer, L.S., Tacker, M., and Schuster, P. (1994). Fast folding and comparison of RNA secondary structures. *Monatshefte Fur Chemie Chem. Mon.* *125*, 167–188.
  16. Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* *327*, 167–170.
  17. Horvath, P., Romero, D.A., Coûté-Monvoisin, A.-C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C., and Barrangou, R. (2008). Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* *190*, 1401–1412.
  18. Howan, K., Smith, A.J., Westblade, L.F., Joly, N., Grange, W., Zorman, S., Darst, S.A., Savery, N.J., and Strick, T.R. (2012). Initiation of transcription-coupled repair characterized at single-molecule resolution. *Nature* *490*, 431–434.
  19. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* *31*, 827–832.
  20. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* *337*, 816–821.
  21. Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. *Elife* *2*, 1–9.
  22. Jurica, M.S., Monnat, R.J., and Stoddard, B.L. (1998). DNA recognition and cleavage by the LAGLIDADG homing endonuclease I-CreI. *Mol. Cell* *2*, 469–476.
  23. Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P., and Siksnys, V. (2013). crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biol.* *10*, 20–19.
  24. Karvelis, T., Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., Cigan, M., and Siksnys, V. (2015). Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. *Genome Biol.* *16*, 253.
  25. Kauert, D.J., Kurth, T., Liedl, T., and Seidel, R. (2011). Direct Mechanical Measurements Reveal the Material Properties of Three-Dimensional DNA Origami. *Nano Lett.* *11*, 5558–5563.
  26. Klaue, D., and Seidel, R. (2009). Torsional Stiffness of Single Superparamagnetic Microspheres in an External Magnetic Field. *Phys. Rev. Lett.* *102*, 28302.
  27. Lieber, M.R. (2010). The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annu. Rev. Biochem.* *79*, 181–211.
  28. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the  $2^{-\Delta\Delta CT}$  Method. *Methods* *25*, 402–408.
  29. Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., et al. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nat. Rev. Microbiol.* *13*, 722–

- 736.
30. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-Guided Human Genome Engineering via Cas9. *Science* *339*, 823–826.
  31. Moynahan, M.E., and Jasin, M. (2010). Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat. Rev. Mol. Cell Biol.* *11*, 196–207.
  32. Mosconi, F., Allemand, J.F., Bensimon, D., and Croquette, V. (2009). Measurement of the Torque on a Single Stretched and Twisted DNA Using Magnetic Tweezers. *Phys. Rev. Lett.* *102*, 78301.
  33. Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* *156*, 935–949.
  34. Oberstrass, F.C., Fernandes, L.E., and Bryant, Z. (2012). Torque measurements reveal sequence-specific cooperative transitions in supercoiled DNA. *Proc. Natl. Acad. Sci.* *109*, 6106–6111.
  35. Perez-Pinera, P., Ousterout, D.G., and Gersbach, C.A. (2012). Advances in targeted genome editing. *Curr. Opin. Chem. Biol.* *16*, 268–277.
  36. Porteus, M. (2016). Genome Editing: A New Approach to Human Therapeutics. *Annu. Rev. Pharmacol. Toxicol.* *56*, 163–190.
  37. Revyakin, A., Ebright, R.H., and Strick, T.R. (2005). Single-molecule DNA nanomanipulation: Improved resolution through use of shorter DNA fragments. *Nat. Methods* *2*, 127–138.
  38. Roberts, R.J. (2005). How restriction enzymes became the workhorses of molecular biology. *Proc. Natl. Acad. Sci.* *102*, 5905–5908.
  39. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
  40. Sapranaukas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* *39*, 9275–9282.
  41. Seidel, R., Bloom, J.G., van Noort, J., Dutta, C.F., Dekker, N.H., Firman, K., Szczelkun, M.D., and Dekker, C. (2005). Dynamics of initiation, termination and reinitiation of DNA translocation by the motor protein Eco R124I. *EMBO J.* *24*, 4188–4197.
  42. Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* *507*, 62–67.
  43. Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V., and Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl. Acad. Sci.* *111*, 9798–9803.

44. Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., and Gregory, P.D. (2010). Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* *11*, 636–646.
45. Zarembo, M., Sasnauskas, G., Urbanke, C., and Siksnys, V. (2006). Allosteric Communication Network in the Tetrameric Restriction Endonuclease Bse634I. *J. Mol. Biol.* *363*, 800–812.