

(19)



(10) **LT 2012 062 A**

(12) **PARAIŠKOS APRAŠYMAS**

(21) Paraiškos numeris: **2012 062**

(51) Int. Cl. (2013.01): **C12N 9/00**  
**C12Q 1/00**

(22) Paraiškos padavimo data: **2012 07 16**

(41) Paraiškos paskelbimo data: **2014 01 27**

(62) Paraiškos, iš kurios dokumentas išskirtas, numeris: —

(86) Tarptautinės paraiškos numeris: **PCT/US2011/021879**

(86) Tarptautinės paraiškos padavimo data: **2011 01 20**

(85) Nacionalinio PCT lygio procedūros pradžios data: —

(30) Prioritetas: **61/296630, 2010 01 20, US**  
**61/376932, 2010 08 25, US**

(71) Pareiškėjas:

**NEW ENGLAND BIOLABS, INC., 240 Country Road, Ipswich, Massachusetts**  
**01938, US**

(72) Išradėjas:

**Zhenyu ZHU, US**  
**Yu ZHENG, US**  
**Shengxi GUAN, US**  
**Hua WANG, US**  
**Aine QUIMBY, US**  
**Penghua ZHANG, US**  
**Lynne APONE, US**

(74) Patentinis patikėtinis/atstovas:

**Tatjana STERLINA, UAB „Intels“, Naugarduko g. 32/2, LT-03225 Vilnius, LT**

(54) Pavadinimas:

**Kompozicijos, būdai ir susiję panaudojimai, skirti modifikuotos DNR skaldymui**

(57) Referatas:

Išradime aprašyta kompozicijos, būdai ir rinkinys, susiję su nauja fermentų šeima, kurie pirmenybiškai rišasi prie hidroksimetilinto citozino arba gliukozilinto hidroksimetilinto citozino ir skaldo dvigrandę DNR nustatytame atstume nuo atpažinimo vietos 3' galo, gaunant suskaldytus fragmentus su 1-3 bazių lipniais galais.

## KOMPOZICIJOS, BŪDAI IR SUSIJĘ PANAUDOJIMAI, SKIRTI MODIFIKUOTOS DNR SKALDYMUI

### TECHNIKOS LYGIS

Pastaruoju metu buvo manoma, kad 5-hidroksimetilcitozinas (hmC) vaidina vaidmenį žinduolių genų ekspresijoje, ypačiai embrioninių kamieninių ląstelių ir neuronų ląstelių vystimesi kaip tarpinė pakopa demetilavimo procese. (Tahiliani *et al.*, *Science* 324 (5929): 930-5 (2009); Kriaucionis and Heinz, *Science* 324 (5929): 929-30 (2009)). Aprašyta TET fermentų šeima, kuri katalizuoja 5-metilinto citozino (mC) konversiją į hmC (WO 2010/037001). Hidroksimetilavimo aptikimas patvirtino sunkumus. Cheminiai būdai, dažniausiai natrio bisulfito sekvenavimas naudotas aptikti mC, neatskyrė mC ir hmC (Cokus *et al.*, *Nature* 452: 215-219 (2008)). PCT publikacija WO 2010/037001 aprašo antikūnų, kurie tiesiogiai rišasi prie hmC, panaudojimą. Produktas, kuris apima endonukleazų mišinį, leidžia aptikti hmC subtrakcijos būdu (EpiMark™, New England Biolabs, Inc. (NEB) ir PCT/US10/46632). Šis produktas naudoja T4 β-gliukoziltransferazę (BGT) gliukozilinti hmC, kuris tada yra atsparus endonukleazės skėlimui, leidžiančiam diferencijuotis nuo mC. Būtų pageidautina išplėtoti paprastus greitus ir tiesioginius metodus aptikti ir išanalizuoti hmC buvimą polinukleotido sekos kontekste, kuris papildomai įgalina atskirti hemi-hmC nuo simetrinio hmC taip, kad tiksliai identifikuotų citoziną (C), kuris yra hidroksimetilintas.

### IŠRADIMO ESMĖ

Išradimo įgyvendinime siūlomas preparatas, kuris apima: vieną arba daugiau išgrynintų rekombinantinių baltymų, kuriame baltymai yra ZZZY baltymų šeimos nariai; ir reakcijos buferį. Išradimo įgyvendinime, kiekvienas ZZZY baltymų šeimos narys turi katalitinį domeną ir surišimo domeną. Įgyvendinime, surišimo domenas DNR-je pirmenybiškai rišasi prie hmC ir gliukozilinto hidroksimetilinto citozino (ghmC). Įgyvendinime, katalitinis domenas skelia DNR nustatytą atstumą (3') nuo hmC arba ghmC. Nustatytas atstumas dar gali būti charakterizuotas kaip 11-13 nukleotidai grandinėje, turinčioje hmC arba ghmC ir 9-10 nukleotidai komplementarioje grandinėje (3'). Kiekvienas ZZZY baltymų šeimos narys dar gali būti charakterizuotas modifikuoto nukleotido atpažinimu, parinktu iš hmC arba ghmC DNR-je taip, kad ghmC arba hmC su mC skilimo santykis yra mažiausiai 8:1.

Išradimo įgyvendinime, kiekvienas ZZZY baltymų šeimos narys turi N-galinį domeną, apimančią aminorūgščių seką su mažiausiai 95 % aminorūgščių sekos homologija su  $RX_7KX_2EXYX_{18}QQX_{11-16}DLX_2PX_6EXDEX_2HX_{6-26}DX_2RX_3I$  (SEQ ID NO:14). Kitame įgyvendinime, kiekvienas ZZZY baltymų šeimos narys turi C-galinį domeną, apimančią aminorūgščių seką su mažiausiai 95 % aminorūgščių sekos homologija su  $WXNX_{30-40}AX_{12-13}FXGX_{16-18}R$  (SEQ ID NO:15) C-galiniame domene. Dar kitame įgyvendinime, mažiausiai vienas narys apima aminorūgščių seką  $RX_7KX_2EXYX_{18}QQX_{11-16}DLX_2PX_6EXDEX_2HX_{6-26}DX_2RX_3I$  (SEQ ID NO:14). Dar kitame įgyvendinime, mažiausiai vienas išgrynintas baltymas turi aminorūgščių seką su mažiausiai 95 % sekos homologija su SEQ ID NO: 7. Kitame įgyvendinime, mažiausiai vienas išgrynintas baltymas turi aminorūgščių seką, apimančią mažiausiai 95 % sekos homologija su fermentu, parinktu iš PvuRts1I, PpeHI, EsaSS310P, EsaRBORFBP, PatTI, YkrI, EsaNI, SpeAI, BbiDI, PfrCORF1I80P, PcoORF314P, BmeDI, AbaSDFI, AbaCI, AbaAI, AbaSI, AbaUMB3ORFAP ir Asp6ORFAP. Kitame įgyvendinime, mažiausiai vienas iš išgrynintų baltymų parinktas iš grupės, susidedančios iš: PvuRts1I, PpeHI, EsaSS310P, EsaRBORFBP, PatTI, YkrI, EsaNI, SpeAI, BbiDI, PfrCORF1I80P, PcoORF314P, BmeDI, AbaSDFI, AbaCI, AbaAI, AbaSI, AbaUMB3ORFAP ir Asp6ORFAP ir katalikiškai aktyvių mutantų ir jų darinių.

Išradimo įgyvendinime, buferis preparate gali apimti druską, turinčią anijoną, parinktą iš sulfato, fosfato, acetato arba citrato. Kitame įgyvendinime, buferis neturi chlorido, nitrato, karbonato arba imidazolo druskos. Druskos koncentracija gali būti 50-500 mM srityje.

Išradimo įgyvendinime, būdas siūlo hmC aptikimą genominiame DNR mėginyje. Būdas apima pridėjimą preparato, kaip anksčiau aprašyta, į genomines DNR mėginių; leidimą baltymui skelti genomine DNR skilimo saite; nustatymą DNR sekos bent jau vienoje skilimo saito pusėje; pasirinktinai, DNR sekos kartografavimą ant referencinės genomines DNR sekos; ir hmC aptikimą. Preparatas, naudojamas būde, gali apimti: vieną arba daugiau DNR polimerazių, pradmenų ir adaptorių.

Dar kitame įgyvendinime, būdas apima skilusios genomines DNR amplifikavimą prieš nustatant DNR seką. Aukščiau minėtuose būdo įgyvendinimuose, genomine DNR gali pirma sureaguoti su  $\beta$ -gliukoziltransferaze (BGT) arba pažymėta BGT arba BGT dariniu prieš sumaišant reakcijos inde genomine DNR su aukščiau apibūdintu preparatu. Įgyvendinime, būdas gali būti įvykdytas viename reakcijos inde arba mikrofluidiniame įtaise.

Papildomame išradimo įgyvendinime, siūlomas būdas išgryninti ZZZY šeimos baltymą, kuris apima: klonavimą ir ekspresavimą sulieto baltymo, apimančio ZZZY šeimos baltymą, inteiną ir afiniškai surišantį baltymą, pailiustruotą chitiną surišančiu domenu (CGD), maltozę

surišančiu domenu (MBP) arba bet kuriuo tinkamu baltymu, galinčiu pririšti prie matricos; priversti sulietam baltymui prisijungti prie matricos afiniškai surišančio baltymo priemonėmis; atskelti ZZZY šeimos baltymą nuo inteino; ir išgryninto baltymo regeneravimą eliuote. Viename įgyvendinime, afiniškai surišantis baltymas yra chitiną surišantis domenas.

Išradimo įgyvendinime, siūlomas fragmentų rinkinys, kuris apima mažiausiai vieną iš a) oligonukleotidų fragmentų, turinčių 20-23 nukleotidų dydį su patalpintu centre hmC arba ghmC; arba b) didelės DNR arba oligonukleotidų fragmentų, turinčių hmC arba ghmC, patalpintų DNR dvigubos spiralės vienguboje grandinėje 11-13 nukleotidų lokacijoje nuo grandinės 3' galo.

Išradimo įgyvendinime, siūlomas rinkinys, kuris apima: vieną arba daugiau išgrynintų ZZZY baltymų šeimos rekombinantinių baltymų, jų funkcinių darinių arba jų katalitinių fragmentų efektyviame buferyje, sudarančiame galimybę fermento aktyvumui; ir naudojimo instrukciją. Rinkinys dar gali turėti BGT ir UDP-gliukozę. Rinkinys dar gali turėti pradmenų. Rinkinys dar gali turėti adaptorių, skirtų panaudoti sekvenavimo platformose.

### **TRUMPAS BRĖŽINIŲ FIGŪRŲ APRAŠYMAS**

Fig. 1 parodo pavyzdinio fermento iš ZZZY šeimos išgrynintą produktą ant 10-20 % denatūruoto poliakrilamido gelio. 1 juosta yra dydžio žymuo. 2 juosta yra PvuRts1I, 3 juosta – PpeHI, 4 juosta – AbaSDFI, 5 juosta – PatTI, 6 juosta – EsaNI ir 7 juosta – YkrI.

Fig. 2 parodo DNR substrato (SEQ ID NO:1) seką, naudojamą 3-5 figūrose. Substratas pagamintas polimerazinės grandininės reakcijos (PGR) amplifikacija taip, kad kiekvienas citozinas buvo arba nemodifikuotas arba pakeistas per mC, hmC arba ghmC. mCs ir hmCs buvo įjungti į substratą PGR metu, to tarpu ghmC buvo gliukoziltransferazės reakcijos su hmCs produktas substrate. Pradmens seka tarp fragmento citozino neturi. Išorinis MfeI saitas buvo sukonstruotas į PGR seką tam, kad kontroliuotų gliukozilinimo statusą.

Fig. 3-5 parodo reakcijos produktą tarp pavyzdinių fermentų ZZZY šeimoje 2-kartinio serijinio praskiedimo ir substrato iš Fig. 2, kurioje visi Cs yra nemodifikuoti, visi Cs pakeisti į mCs, visi Cs pakeisti į ghmCs ( $\beta$ ghmCs) ir visi Cs pakeisti į hmCs. Pirmą ir paskutinę juostas turi PGR ženklą. Juostos tarp žymenų parodo atskiros fermento du-kartinį serijinį praskiedimą. Reakcijos sąlygos buvo 23°C, 20 min NEB4 buferyje (New England Biolabs, Inc., Ipswich, MA (NEB)) su galutine KOAc 250 mM koncentracija.

Fig. 3 parodo PvuRts1I rezultatus, kur santykis hmC:ghmC:mC:C = 2000:2000:8:1. Fermento kiekis pirmoje juostoje, gretimose kairės pusės žymeniui, yra 33  $\mu$ g.

Fig. 4 parodo PpeHI rezultatus, kur santykis  $hmC:ghmC:mC:C = 128:256:2:1$ . Fermento kiekis pirmoje juostoje, gretimose kairės pusės žymeniui, yra 123  $\mu g$ .

Fig. 5 parodo AbaSDFI rezultatus, kur santykis  $hmC:ghmC:mC:C = 500:8000:1:ND$  (neaptiktas). Fermento kiekis pirmoje juostoje, gretimose kairės pusės žymeniui, yra 52  $\mu g$ .

Fig. 6A ir 6B parodo, kad ZYZZ fermentų šeima skelia pilnai hidroksimetilintus oligosubstatus.

Fig. 6A parodo išgryninto rekombinantinio PpeHI aktyvumą 56 bazių porų sintetinio nukleotido substrate, turinčiame AhmCGT. 1 juosta: tik DNR, 2-8 juostos: fermento 2-gubą praskiedimą. Reakcijos buvo inkubuotos 37°C temperatūroje 1 valandai ir atskirtos 20 % TBE poliakrilamido gelyje.

Fig. 6B parodo schematinį oligonukleotido substrato skaldymo vaizdą, rodantį skėlimo produktus. PpeHI skelia bet kurioje atpažinimo saito (parodyta strėlėmis) pusėje, kad sukurtų 20 bazių porų fragmentą.

Fig. 7 parodo ZYZZ fermentų aktyvumą simetriškai hidroksimetilintuose oligonukleotiduose ir hemi-hidroksimetilintuose oligonukleotiduose. Gelis parodo rezultatus naudojant PpeHI. Skaldymo produktai buvo analizuoti 20 % poliakrilamido denatūruotame gelyje su 7M karbamilu. Sunumeruotos juostos gelyje atitinka skėlimo vaizdams, parodytiems piešiniu dešinėje gelio pusėje. Pavyzdžiui, 1 juosta atitinka skėlimo vaizdai, gautam iš hmCG/GhmC viršutinėje ir apatinėje grandinėse tik su pažymėta viršutine grandine.

Fig. 8A-B parodo ZYZZ šeimos fermentų skėlimo saitus kaip atskirai nustatytus dvigubos grandinės DNR kiekvienai grandinei, turinčiai vienoje grandinėje hmC arba ghmC.

Fig. 8A parodo skėlimo produktą dviguboje grandinėje, komplementarios grandinei, turinčiai hmC, atitinkančiai 3 ir 4 juostoms gelyje. Naudojamų oligonukleotidų sekos yra parodytos gretimai geliui. Gelis yra toks pat kaip parodytas Fig. 7. Sintetiniai oligonukleotidų žymenys (M1, M5 ir M2, M3) siūlo gidą skėlimo saitų aptikimui. Oligonukleotidų substratai ir sintetiniai žymenys buvo pažymėti FAM fluorescuojančia grupe ir atskirti denatūruoto 20 % poliakrilamido gelyje su 7M karbamilu. Skėlimo fragmentai atitinka M2 ir M3 (3 juosta) ir M1 ir M5 (4 juosta) ir pažymėti oligonukleotido sekoje (SEQ ID NO:2), sekančioje geliui.

Fig. 8B parodo skėlimo produktą dviguboje grandinėje, turinčioje ghmC. Substratas (R2), turintis parodytas sekas (SEQ ID NO:3 ir 4), kiekviename gale (\*) buvo pažymėtas su alfa- $^{33}P$  dATP. Žymenys (M1 ir M2) pažymėti su gama- $^{33}P$  ATP naudojant polinukleotidkinazę.

Fig. 8C parodo, kad ZYZZ baltymas skelia DNR, turinčią ghmC arba hmC tame pačiame 3' atstume nuo modifikuoto nukleotido. Gliukozilimas neveikia skėlimo atstumo. Reakcijoje buvo naudojami skirtingi substratų rinkiniai, arba su hmC (1 ir 3 juostos) arba su ghmC (2 ir 4

juostos). Fluorescentiniai žymenys (FAM) Figūroje parodyti kaip užpildyti apskritimai. Laukiamos skėlimo vietos pažymėtos linijomis. Čia buvo naudotas fermentas AbaSDFI. Lyginant 3 ir 4 juostas, gali būti padaryta išvada, kad skėlimo atstumai modifikacijų dviejų tipų priešingose grandinėse yra tokie pat.

Fig. 9 parodo hmC vietų identifikavimo būdą žiurkių smegenyse genominėje DNR. Genominė DNR buvo paveikta su T4 BGT, kad gliukozilintų bet kuriuos hmCs ir po to skaldoma ZZZY fermentu (pvz. AbaSDFI). Skaldyti produktai buvo liguoti su sintetiniu dvigubos grandinės oligonukleotido adaptoriumi, kaip pailiustruota. Ligatabilus galas turi 2 bp degeneruotų nukleotidų. Liguoti fragmentai buvo atrinkti pagal dydį nuo maždaug 1 kb-3 kb. Šie buvo PGR amplifikuoti naudojant įprastą pradmenį, specifinį adapteriui. Po to amplifikuoti PCR produktai gali būti klonuoti į įprastą klonavimo vektorių ir sekvenuoti, kad nustatytų, kuris iš gretimų citozinų yra hidroksimetilintas.

Fig. 10 parodo klonuotų fragmentų iš žiurkių smegenų genominės DNR, gautos pagal Fig. 9, pavyzdinę logo seką. Visi klonuoti fragmentai atitiko pavyzdiniam žiurkės genomui. Skaldymo saito 40 bp langai kairėje ir dešinėje buvo išekstrahuoti ir įvertinti, kaip pailiustruota Figūroje.

Fig. 11 aprašo didelio našumo sekvenavimo protokolą, kad nustatytų, kuris iš citozinų, gretimuose CpG motyvuose yra hidroksimetilinti ir tarnauja kaip substratas ZZZY šeimos narių skėlmui. (A): DNR fragmento laisvieji galai apklausai yra užblokuoti. Skėlimas įvyksta saite, esančiame tarp dviejų identifikuotų citozinų su ZZZY fermentu, besibaigiančiu 1-3 baze lipniame gale. (B): Lipnus DNR galas pašalinamas DNR fragmento buku galu. (C): Žinomos sekos adapteris yra liguotas prie abiejų skėlimo fragmentų bukų galų, kur adapteris turi nemodifikuotą CC/GG viename gale ir žymenį X - kitame. Neliguotas adapteris pašalinamas sukinio kolonėle. (D): Adapteris, liguotas DNR, yra skeliamas antrą kartą su ZZZY fermentu, kad išlaisvintų adapterį, liekantį už CC kaip lipnusias galas nuo fragmento, turinčio hmC. Fragmentas, kuris turi nemodifikuotą citoziną, yra neskeliamas. Yra naudojama sukinio kolonėlė, kad pašalintų skeltą adapteri. (E): Antrasis adapteris pridedamas skelti fragmentą su CC lipniu galu. Antras adapteris turi Y žymenį. Citozinas, kuris yra modifikuotas, gali būti nustatytas iš sekvenavimo fragmento su Y žymeniu. Nemodifikuotas citozinas gali būti patvirtintas sekvenuojant fragmentą su X žymeniu. (F): Nesuskaldytas fragmentas sekvenuojamas transkripcijos kryptimi nuo hmC (5-hmC) saitų. (G): hemi-hmC (5-hmC) yra kartografuotas ant referencinio geno.

Fig. 12 parodo sekvenavimo schemas pavyzdį hidroksimetilomos dekodavimui. Genominė DNR yra pirmiausiai paveikiama su BGT, o po to skaldoma su ZZZY šeimos fermentu (pvz. AbaSDFI). Atskiri adapteriai, arba su 2-degeneruotos bazės 3' lipniu galu arba su 3-degeneruotos bazės 3' lipniu galu, tada yra liguojami prie suskaldytų produktų. DNR adapteris turi žymę (tokią

kaip biotinas), kad palengvintų gryninimą transkripcijos kryptimi. Liguota DNR yra sukarpoma į mažesnius fragmentus. Šie fragmentai su adapteriais, liguotais prie galų, yra ištempiami naudojant avidino rutuliukus. Po to fragmentai liguojami su antru sekvenuojančiu adapteriu. Liguoti produktai yra PGR-amplifikuoti naudojant sekvenavimui specifinius pradmenis.

Fig. 13 parodo Promals'o (Pei & Grishin, *Bioinformatics*, 23(7):802-808 (2007)) naujos fermentų šeimos narių pirminę struktūros sulyginamąją analizę, čia identifikuotą kaip ZZZYZ šeimą (SEQ ID NO: 5-13). Parodyti devyni šeimos nariai, kur kiekvienas fermentas turi hmC arba ghmC pirmenybinį skėlimo aktyvumą, palyginus su mC arba nemodifikuotu citozinu dvigrandėje DNR (dsDNR).

Fig. 14 parodo, kad PvuRts1I gali išskirti DNRs su hmC skirtingais lygiais skirtingose ląstelėse. Čia hmC lygiai embrioninėse kamieninėse ląstelėse buvo sulyginti su fibroblastais. E-14 genomine DNR buvo suskaldyta į lipnią klampią medžiagą, tuo tarpu kai NIH3T3 DNR buvo suskaldyta tik truputį, patvirtinant žymų hemi-metilinto hmC buvimą E-14 ląstelėse ir žymiai mažesnius bet kurio hmC lygius fibroblastuose.

### **DETALUS ĮGYVENDINIMU APRAŠYMAS**

Čia aprašyti fermentų šeimos unikalūs požymiai yra jų rišamosios savybės ir specifinis skėlimas, kuris pirmiausiai siūlo tiesioginį fermentinį būdą aptikti hmC genome, apimant hemi-hidroksimetilintuosės saituose, kiekybiškai ir sekų kontekste.

Čia aprašyti naujai apibrėžtų ZZZYZ šeimos fermentai yra ypatingo dėmesio dėl priežasčių, kurios apima jų gebėjimą atskirti hmC (taip pat ir ghmC) liekanas nuo mCs arba nemetilinto citozino, kad gautų DNR fragmentų rinkinį, kur skėlimas įvyksta pagrindinai fiksuoto atstumo transkripcijos kryptimi nuo fermento atpažinimo saito DNR-je. Fragmentų narių dydis fragmentų rinkinyje, atsiradusių iš skėlimo su ZZZYZ fermentais, gali būti pilnai heterogeninis. Kai DNR turi hemi-hmC arba hemi-ghmC, dviguba grandinė trūksta tik ant modifikuoto nukleotido vienos pusės(3'), kad generuotų kintamo ilgio dsDNR fragmentus. dsDNR fragmentai bus generuojami 20-23 nukleotidų dydžio, kai hmC arba ghmC yra simetriškai išsidėstę CpG saituose.

Tam tikri terminai aprašyti žemiau.

„Didelė DNR“ skirtas nurodyti bet kuriai natūralios kilmės arba sintetinei DNR, turinčiai dydį, didesnę negu 100 nukleotidų iki genomo dydžio.

„Panašus dydis“ su nuoroda į fragmentų „rinkinys“ skirtas nurodyti fragmentams, kurie kinta ne daugiau negu apie  $\pm 5$  nukleotidai pagal ilgį.

„Patalpintas centre“ skirtas nurodyti modifikuoto nukleotido atitikimui lokacijai vienoje grandinėje, kuri yra apytiksliai centruota dsDNR fragmento toje pačioje grandinėje. Lokacija pagrindinai yra tarp 5 centre nustatytų nukleotidų skaičiuojant nukleotidus nuo fragmento bet kurio galo.

„N-galinis domenas“ nurodo į sritį, ištęstą iki baltymo aminorūgščių sekos apie 50 %.

„C-galinis domenas“ nurodo į sritį, ištęstą iki baltymo aminorūgščių sekos apie 50 %.

„Matrica“ skirtas įjungti bet kurią struktūrą, turinčią paviršių, tinkamą imobilizuoti molekulę, kuri turi giminingumą matricai ir apima, pavyzdžiui, granules (rutuliukus), kolonėles, plokščius paviršius, tokius kaip popierius arba stiklas, viso paviršiaus vidus arba tuščia forma, ir t.t.

„Fragmentų rinkinys“ gaunamas skeliant didelę dsDNR su ZZZY šeimų fermento nariu. Fragmentų rinkinys apima dvigubos grandinės oligonukleotidus, kurie yra didelės DNR skėlimo produktas simetrinėje hmC arba ghmC (hmCG/GhmC) abiejose pusėse ir kintamo dydžio fragmentus, kuriuose hemi-hmC arba hemi-ghmC yra išsidėstę fiksuotoje padėtyje nuo skaldymo saito (pvz., 9-13 nukleotidai). Jeigu genomas, kuris susideda iš daugybinių didelių DNR (pvz., chromosomų), yra suskaldytas, tai kiekviena didelė DNR turės duoti daug fragmentų rinkinių. Fragmentų mišinys, gautas iš viso žmogaus genomo, gali būti laikomas kaip daugybė fragmentų rinkinių, kai kiekvienas rinkinys išvestas iš chromosomos arba kaip fragmentų vienas rinkinys, priklausomai nuo konteksto. Įgyvendinime, fragmentų rinkinys apima mažiausiai 6 nukleotidų fragmentus su skirtingomis DNR sekomis. Pavyzdžiui, nukleotidų rinkinys gali apimti mažiausiai 10 fragmentų su skirtingomis sekomis arba mažiausiai 20 fragmentų su skirtingomis sekomis. Viename įgyvendinime, nukleotidų rinkinys apima vieną arba daugiau panašaus dydžio fragmentų su centre patalpintu hmC arba ghmC.

„Fermento preparatas“ skirtas nurodyti reagentui ir niekam kitam, kilusiam jo natūralioje būsenoje *in vivo*.

„X“, kai naudojamas aminorūgščių sekoje, reiškia bet kurią aminorūgštį.

Čia identifikuota DNR skelianti nauja šeima kaip ZZZY šeima, kur šeimos nariai pirminybiškai atpažįsta hmC ir ghmC ir nemetilintą (mC) arba nemodifikuotus citozinius dsDR-je ir po to skelia DNR neatstikitiniame atstume transkripcijos kryptimi (3' kryptimi) nuo hmC arba ghmC nukleotido kiekvienoje DNR grandinėje, kuri turi hmC arba ghmC. Šios šeimos nariai apima, bet yra neapriboti bakterinių restrikcijos endonukleazių.

Šeimos nariai gali būti apibūdinti struktūriškai skirtingu skėlimo domenu ir surišimo domenu ir ypatingai N-galiniu konservatyviu domenu su didesniu negu 85 % aminorūgščių sekos homologiškumu, pavyzdžiui, didesniu negu 90 % aminorūgščių sekos homologiškumu,

pavyzdžiui, didesniu negu 98 % aminorūgščių sekos homologiškumu su  $RX_7KX_2EXYX_{18}QQX_{11-16}DLX_2PX_6EXDEX_2HX_{6-26}DX_2RX_3I$  (SEQ ID NO:14) N-galiniame domene, ir/arba daugiau negu 85 % arba 90 % arba 98 % aminorūgščių sekos homologiškumu su  $WXNX_{30-40}AX_{12-13}FXGX_{16-18}R$  (SEQ ID NO:15) C-galiniame domene. Vienaime įgyvendinime, baltymas (pailiustruotas per AbaSDFI) su  $RX_7KX_2EXYX_{18}QQX_{11-16}DLX_2PX_6EXDEX_2HX_{26}DX_2RX_3I$  (SEQ ID NO:16) su N-galine seka, turi didelį selektyvumo laipsnį tarp ghmC/hmC/ ir mC.

ZZYZ šeimos nariai (pavyzdžiui, žiūr. Fig. 13) apima rekombinantinius baltymus, variantus ir darinius. Variantų pavyzdžiai apima mutantus, kuriuose, pavyzdžiui, katalitinis domenas buvo modifikuotas arba pašalintas iš baltymo ir liko tik DNR surišantis domenas, skirtas hmC ir ghmC. Variantų papildomi pavyzdžiai yra sulieti baltymai. Šie sulieti baltymai gali tarnauti kaip reagentai arba gali būti tarpiniai junginiai fermentų gryninime arba hidroksimetilintų nukleotidų kartografijoje.

Sulietų baltymų variantai ZZZY baltymų šeimoje gali būti sulieti su antru baltymu arba daugybe baltymų, kurie tarnauja kaip žymė, žymelė arba žymuo vizualizacijai *in situ* baltymui, prijungtam prie hmC arba ghmC arba afininiam gryninimui. Sulietų partnerių pavyzdžiai apima inteinus (JAV patentas Nr. 5643758), maltozę surišančius baltymus (MBPs) (JAV patentas Nrr. 5643758 ir 7825218; ir PCT publikacija WO 2010/114532), SNAP-TAG® (JAV publikacija Nr. US-2004-0115130) arba gali turėti pakaitą, kuris veikia kaip žymė, pavyzdžiui, pakeičiant cisteiną į selenocisteiną (žiūr. JAV patentas Nr. 7141366). Alternatyviai, sulietas baltymas gali apimti nukleorūgšties aptamerą žymėjimui arba gryninimui (pavyzdžiui, žiūr. JAV patentus 5670637; 5696249; 5874557; 5693502).

Šeimos nariai dalijasi vienas su kitu mažiausiai 30 % aminorūgščių sekos identiškumu, pavyzdžiui mažiausiai 40 %, pavyzdžiui mažiausiai 50 %, pavyzdžiui mažiausiai 60 %, pavyzdžiui mažiausiai 70 %, pavyzdžiui mažiausiai 80 %, pavyzdžiui mažiausiai 90 %, arba mažiausiai 95 % aminorūgščių sekos tapatumo kaip pareikšta Promals'o pirminės struktūros suliginamojoje analizėje (Pei & Grishin, *Bioinformatics*, 23(7):802-808 (2007)) ir apima PvuRts1I, PpeHI, AbaSDF1, EsaSS310P, EsaRBORFBP, PatTI, YkrI, EsaNI, SpeAI, BbiDI, PfrCORF1I80P, PcoORF314P ir BmeD1, kai kurie iš jų aprašyti 1 lentelėje ir Fig. 13. Pavyzdžiui, vienas šeimos narys yra AbaSDF1, kur variantai AbaCI (GenBankas:ACQB1000053), AbaI (GenBankas: ABXK01000045), AbaSI (GenBankas: NC\_009085), AbaUMB3ORFAP (GenBankas: AEP01000006) ir Asp6ORFAP (GenBankas: ACYS1000207) dalijasi mažiausiai 95 % sekos tapatumu su AbaSDF1. Šie fermentai neįtraukti į 1 lentelę. Jų savybės yra iš esmės tokios pat kaip ir AbaSDFI. Nuoroda į šeimos atskirus narius yra skirtas apimti jo darinius ir katalikiškai aktyvius fragmentus.

Antikūnai gali būti pakelti iki ZZZY fermentų šeimos narių naudojant standartines technikas, skirtas sukurti monokloninius arba polikloninius antikūnus arba antikūnų fragmentus. Šie antikūnai arba jų fragmentai gali būti panaudoti ZZZY fermento šeimos nario *in situ* žymėjimui, prijungtam prie hidroksimetilintos arba gliukozilintos hidroksimetilintos didelės DNR.

Vienas ZZZY fermentų šeimos narys yra PvuRts1I. Ši restrikcijos endonukleazė pirmiausia buvo aprašyta Ishaq & Kaji (*Biological Chemistry* 255(9):4040-4047 (1980)) ir parodyta, kad ji yra hmC-specifinė restrikcijos endonukleazė, koduota plazmide Rts1. PvuRts1I genas buvo klonuotas ir ekspresuotas (Janosi and Kaji, *FASEB J.* 6:A216 (1992); Janosi *et al.*, *Journal of Molecular Biology* 242: 45-61 (1994)) ir Rts1 plazmidė buvo pilnai sekvenuota (Murata *et al.*, *Journal of Molecular Biology* 184 (12):3194-202 (2002)).

Tačiau nėra atlikta šio fermento gilesnės studijos arba publikacijos. Be to, po pirmųjų publikacijų, šiuo fermentu buvo menkai domimasi. Tai ryškiausiai buvo atsispindėta PCT publikacijoje WO 2010/037001, kuri turėjo skyrių apie hmC ir ghmC aptikimo būdus, bet nebuvo jokių užuominų apie šį fermentą. WO 2010/037001 aprašo būdus, kurie yra cheminės prigimties arba apima baltymų, tokių kaip antikūnai, surišimą. Kur fermentai yra paminėti, tai tokie iliustruoti išimtinai gliukoziltransferazėmis.

Tam, kad būtų ištirtos šios šeimos savybės kaip specifinės, reikėjo fermentus išgryninti. Po rūpestingų analizių ir išplėstinių eksperimentų atrasta, kad daug įprastai gryninimui naudojamų joninių reagentų, tokių kaip NaCl, fermentus išaktyvuoja. Dar atrasta, kad skirtingų tipų druskų ir jų koncentracijų panaudojimas paveikia aktyvumą. Rasta, kad ZZZY fermentų šeima buvo išaktyvuota dalyvaujant druskoms, tokioms kaip chloridai, nitratai, karbonatai arba imidazolo druskos. Todėl buvo padaryta išvada, kad šių druskų reikia vengti. Pasirodė, kad kitos druskos, tokios kaip sulfatai, fosfatai, acetatai ir citratai aiškiai slopina fermentų aktyvumą esant didelėms koncentracijoms ir todėl geriau fermentus laikyti arba naudoti 50-500 mM koncentracijų ribose.

Kitas svarbus svarstymas buvo fermentus skiriančios savybės hmC nuo mC ir C. Nuo tada, kai nustatyta, kad genome hmC pasitaiko labai retai, maži skėlimo lygiai prie mC arba C galėtų išplaukti į didelę genetinę apsuptį, kuri paveiktų galimybę aptikti hmC.

ZZYZ šeimos narių gryninimas sudarytų galimybę fermentų šeimos skėlimo savybių studijai. Pirminės pastangos išgryninti gamtinį arba rekombinantinį PvuRts1I buvo problematinės, nes visos standartinės kolonėlės, naudojamos gryninimui, vedė į žymų fermento aktyvumo praradimą. Rekombinantinis PvuRts1I buvo gautas kaip sulietas baltymas su 6xHis žymele PvuRts1I N-gale, turėjo fermento labai didėjančio skėlimo aktyvumo ant nemodifikuoto substrato nepageidautiną efektą, kas sumažino fermento pritaikymą aptikti hmC tiesiogiai. C-galinės 6xHis

žymelės pridėjimas nepakeitė santykinio aktyvumo skirtingų metilinimo statusų ant kiekvieno substrato, bet fermento specifinis aktyvumas buvo ženkliai sumažintas.

Rekombinantinis PvuRts1I buvo taip pat sulietas su intein-CBD (IMPACT™ rinkiniu, NEB), kuris po to buvo skeliamas, kad atpalaiduotų gryną fermentą. Šiuo keliu galėjo būti gauta aktyvaus fermento didelė išeiga (žiūr. 1 pavyzdį).

Rekombinantinis baltymas arba vienas, modifikuotas arba sulietas su žyme gali būti pažymėtas vaizdo gavimo tikslais panaudojant spektroskopinį žymėjimą, tokį kaip fluorescentinį arba chemiluminescentinį žymėjimą, radioaktyvų žymėjimą arba kitą reakcingą cheminį agentą, arba žymėtą cukrų, žymėtą antikūną arba kitą žymėtą baltymo žymę (apimant pavyzdžiui SNAP-TAG®, NEB) (taip pat žiūr. Chung-Xiao Song *et al.*, *Nature Biotechnology* 29:68-72 (2011)).

Tyrimo taikymas, parodytas Fig. 3-5, fermentų koncentracijų nustatymui, prie kurių atsiranda maksimalus ghmC- ir hmC-specifinis skėlimas ir aptiktas minimalus mC ir C skėlimas įgalino identifikuoti ZYZ šeimos narių skėlimo savybes. Rezultatai, gauti iš specifinių analizių (pavyzdžiui, žiūr. 1 lentelę) parodė, kad šios šeimos nariai pateikė žymiai pagerintą ir paprastesnį būdą specifiniam hmC nustatymui ir kartografavimui DNR-je.

Efektyvesnei hmC analizei genome, reagentų požymiai apima reagentų atskyrimų savybes hmC ir/arba ghmC nuo mC ir C. Taigi, buvo nustatyti santykiai fermento skėlimo specifiškumui. Santykių pavyzdžiai skirtingiems fermentams ZYZ šeimoje pateikti 1 lentelėje.

Rasta, kad išgryninti fermentai ZYZ šeimoje turi hmC arba ghmC:mC minimalų skėlimo santykį mažiausiai 8:1, pavyzdžiui, 50:1, pavyzdžiui, mažiausiai 100:1, pavyzdžiui mažiausiai 200:1 arba pavyzdžiui mažiausiai 250:1; ir hmC arba ghmC:C minimalų skaldymo santykį mažiausiai 50:1, pavyzdžiui, mažiausiai 100:1, pavyzdžiui mažiausiai 200:1 arba pavyzdžiui mažiausiai 250:1. Šie santykiai gali būti nustatyti pagal tiriamąjį pavyzdį, aprašytą 1 pavyzdyje ir Fig. 3-5.

Tiriamąjo pavyzdžio konsistencijos pagerinimui, PGR DNRs buvo naudoti kaip substratas vietoje fago T4 ir λDNR, hmC-, mC- ir C-, turintys DNR substratus buvo gauti simetrinių oligonukleotidų, gautų komerciškai, PCR būdu (Integruotos DNR technologijos (IDT), Coralville, IA), kuriame hidroksimetilintas deoksicitidino trifosfatas (hmdCTP) arba metilintas deoksicitidino trifosfatas (mdCTP) buvo pakeistas į dCTP kaip aprašyta. hmC PGR fragmentas toliau buvo gliukozilintas naudojant BGT, kad taptų ghmC PGR DNR fragmentu. Šio tyrimo rezultatai parodė, pavyzdžiui, kad PvuRts1I turi hmC:ghmC:5-mC:C=2000:2000:8:1 aktyvumo santykį; PpeHI jis yra 12:250:2:1; AbaSDFI jis yra 500:8000:1:ND. ND čia reiškia, kad skėlimas nestebimas net prie didžiausių fermento koncentracijų.

ZZYX šeimos narių požymis yra tas, kad tikslus hmC arba ghmC skėlimo saitas gali kisti 3' transkripcijos kryptimi nuo hmC arba ghmC ne daugiau kaip 5 nukleotidais, pavyzdžiui, ne daugiau kaip 3 nukleotidais. Pavyzdžiui, naudojant sintetinius substratus (žiūr., pavyzdžiui, 3 pavyzdį ir Fig. 6-8) parodyta, kad nors ZZYX šeimos nariai daugiausia skėlė dsDNR viename nustatytame atstume nuo hmC arba ghmC, skėlimo saitas gali būti stebimas 11-13 nukleotidų atstumu nuo hmC arba ghmC vienoje grandinėje ir 9-10 nukleotidų atstumu nuo hmC arba ghmC priešingoje grandinėje (ghmCN<sub>11-13</sub>/N<sub>9-10</sub>). Apie 20-23 bp skėlimo fragmentas atsiranda iš skėlimo simetriniame hmC saite (hmCG/GhmC) dviejų grandinių DNR.

Įvairovės įvertinimas dėl skėlimo atstumo ZZYX fermentų šeimos pagalba yra svarbus hmC bazių lokacijos kartografavimui genome. Fig. 8C taip pat atskleista, kad ghmC buvimas hmC vietoje neturi pastebimo efekto skėlimo atstumui. Atkreiptas dėmesys, vis dėlto, kad skėlimo aktyvumas gali būti padidintas, jeigu G yra fiksuotas prie 21-22 nukleotidų iš ZZYX atpažinimo saito priešingoje vijoje, net jei čia nėra absoliutaus reikalavimo, skirto G šioje lokacijoje, net jei čia yra bet koks žymus arba net aptinkamas nukleotido nuokrypis tuoju pat flankuojantis ghmC arba hmC.

hmC arba ghmC lokacija didelėje DNR gali būti nustatyta klonuojant skėlimo produktus ir/arba sekvenuojant, pavyzdžiui, ultra didelio našumo sekvenavimo platformomis (žiūr. 7 pavyzdį). Naudojant ZZYX šeimos narius, yra galimybė nustatyti ir kartografuoti hmC arba ghmC ir juos diferencijuoti iš mCs vienoje fermento reakcijoje. Kur du hemi-metilinti CpG saitai yra arti vienas kito, yra galimybė nustatyti, kuri iš 4 galimų pozicijų turi hmC (pavyzdžiui, Fig. 9-11).

ZZYX fermentų šeimos narių panaudojimas generuoti fragmentų rinkinius hidroksimetilinimo genominei analizei arba kitiems tikslams, galima pasikliauti pavieniu fermentu arba galima įjungti fermentų gausybę, kur keli arba visi iš fermentų yra: ZZYX šeimos nariai; kildinti iš ZZYX šeimos narių; MspJI fermentų šeimos įjungti nariai (PCT publikacija WO2010/075375) ir/arba MmeI-kaip endonukleazų restrikcijos šeima; ir/arba endonukleazų restrikcijos kiti tipai.

Vienas iš požymių, kuris charakterizuoja ZZYX baltymų šeimą yra surišimo domeno (C-galinio) ir katalitinio domeno buvimas. ZZYX panašių fermentų surišimo domenas gali būti panaudotas abiemis *in vivo* ir *in vitro* pritaikymams. Surišimo domeno panaudojimo pavyzdžiai apima: hmC turinčio DNR *in vitro* praturtinimą, pavyzdžiui, kaip reagentą afininiam gryninimui arba hmC *in vitro* taiknį. Pastarasis gali būti pasiektas pažymint surišimo domeną arba suliejant surišimo domeną su kitais domenais įvedant jį į hmC saitus genome. Surišimo domenas taip pat gali būti sulietas su nukleazės domenu, kad iššauktų dvigubos grandinės trūkius gretimuose hmC

saituose, kad aktyvuotų DNR reparacijos kelią taip, kad šių saitų epigenetinis statusas galėtų būti pakeistas.

Epigenetinėms studijoms yra svarbus DNR pavyzdžių hidroksimetilinimo lygio nustatymas. Genomo epigenetinė reguliacija apima chromatino remodeliavimą, kuris dalinai liečia mC konvertavimą į hmC. Skirtumai hidroksimetilinimo modeliuose gali būti lemiamas indikatorius netinkamų vystimosi procesų, pavyzdžiui, embrioninėms kamieninėms ląstelėms. Dabar šie skirtumai gali būti tiriami įprastu būdu panaudojant fermentus iš ZZZY baltymų šeimos, kuri selektyviai nukreipia taikinį į hmC ir ghmC. Šios nukleotido modifikacijos tada gali būti kartografuotos ant metilomos arba genomo.

PCT paraiška PCT/US2010/046632, kuri čia įtraukta visa apimtimi, toliau detaliau aiškina apie kartografavimo reikšmę hidroksimetilintų nukleotidų genome, kad būtų suprastas ląstelės šeimininkės ir organizmo fenotipas. ZZZY šeima viena arba kartu su kitais fermentais (pavyzdžiui, MspJI šeimos nariu arba MspI arba fermentų, funkciškai giminingų MspI) gali būti panaudota pateikti metilomą arba hidroksimetilomą kartu panaudojant sekvenavimo techniką, aprašytą čia arba kitas sekvenavimo technikas, žinomas iš technikos lygio.

Fragmentų rinkinys(iai), gautas(ti) iš fermentų skaldymo su vienu arba daugiau fermentų iš ZZZY šeimos ir pasirinktinai kitų DNR skėlimo fermentų, kurie skelia modifikuotus nukleotidus, gali būti sekvenuojami panaudojant didelio našumo įvairių rūšių sekvenavimo būdus, kurie šiuo metu prieinami naudojant, pavyzdžiui, NextGen sekvenavimo metodus, kad identifikuotų ir kartografuotų hmCs arba ghmCs DNR-je. Specifinių skaldymo produktų parinkimas, kurie genome hibridizuoja tam tikrus regionus, gali būti panaudoti greitiems diagnostikos būdams, atskleidžiantiems nenormalų hmCs arba ghmCs buvimą arba nebuvimą, koreliuojantį su liga, tokia kaip vėžys. Individo konkretaus fenotipo nustatymui gali būti panaudoti specifiniai oligonukleotidai. Pavyzdžiui, fragmentų rinkinio hibridizacija apibrėžtai sekai arba sekų rinkiniui, esančioms ant kieto paviršiaus (hibridizacijos nustatymas) arba pažymėtoms tirpale (arba *visa versa*) gali atskleisti skirtumus tarp fragmentų rinkinyje ir standartinių fragmentų rinkinių, kurie charakterizuoja metilomą. qPGR arba hibridizacijos nustatymas taip pat gali būti panaudoti detaliam iširti vieną arba daugiau žinomų tiriamų lokacijų kopijavimui. hmCs arba ghmC arba surišimo domenai gali būti pažymėti fluorescentine arba chemiliuminescentine žyme arba kitais žymėjimo būdais, žinomais technikos lygyje, kad palengvintų aptikimą.

ZZZY fermentų šeimos narių gryninimas ir jų skaldymo specifiškumo charakterizavimas ir jų gebėjimas atskirti hemi-hmC nuo simetrinio hmC pirmiausiai pateikia duomenų generacijos

tipą, kuris reikalingas tyrime, medicininėje diagnostikoje ir gydyme, kad būtų suprasta hidroksimetilintos DNR rolė ir prasmė bet kuriam specialiam fenotipui.

Medicininį sąlygų detalus aprašymas pateiktas PCT publikacijoje WO2010/037001, kuri čia įtraukta kaip nuoroda visa apimtimi. Kiekvienu atveju, kur TET baltymai yra įtraukti į genų ekspresiją, bus būtina nustatyti, kur hmC, kuri yra sukurta TET baltymų, yra lokalizuota, ir tokiu būdu bus pageidautina panaudoti ZYZZ šeimą, čia aprašytą.

1 Lentelė: ZYZZ baltymų šeimos narių pavyzdžiai

| Pavadinimas | Alternatyvus pavadinimas | hmC  | ghmC | mCG | C | Ilgis, panašumas ir tapatumas su PvuRtsII | Šaltinis                                             |
|-------------|--------------------------|------|------|-----|---|-------------------------------------------|------------------------------------------------------|
| PvuRtsII    | PvuRtsII                 | 2000 | 2000 | 8   | 1 | 293, 100 %, 100 %                         | <i>Proteus vulgaris</i> (RtsI)                       |
| PpeHI       | PROPEN_01738             | 128  | 256  | 2   | 1 | 294, 72.5 %, 64.6 %                       | <i>Proteus penneri</i> ATCC 35198                    |
| PatTI       | Pat_1499                 | 2000 | 8000 | 128 | 1 | 306, 51.4 %, 36.5 %                       | <i>Pseudoalteromonas Atlantica</i> T6c ATCC BAA-1087 |
| Ykri        | Ykris0001_38440          | 250  | 1000 | 64  | 1 | 299, 49.6 %, 36.1 %                       | <i>Yersinia kristensenii</i> ATCC33638               |
| EsaNI       | GOS_7810815              | 128  | 64   | 8   | 1 | 299, 47.8 %, 39.6 %                       | Jūrinis metagenomas                                  |
| AbaSDFI     | ABSDF3356                | 500  | 8000 | 1   | 0 | 317, 47.8 %, 35.4 %                       | <i>Acinetobacter baumannii</i> SDF                   |
| SpeAI       | Spea_2956                |      |      |     |   | 285, 47.1 %, 35.6 %                       | <i>Shewanella pealeana</i> ATCC 7000345              |
| BbiDI       | BbifN4_02244             |      |      |     |   | 311, 43.3 %, 32.0 %                       | <i>Bifidobacterium bifidum</i> NCIMB 41171           |
| BmeDI       |                          |      |      |     |   | 300, 40.0 %, 30.5 %                       | <i>Bacillus megaterium</i> kamienas DSM319           |

Pateikti eksperimentų protokolai nenumatyti, kad būtų apribojami. Bet kuris, patyręs šioje srityje, galėtų panaudoti eksperimentinį planą kaip pasiūlymą bet kuriam naujai nustatytos šeimos papildomam nariui.

Visos čia cituojamos nuorodos, įskaitant JAV atskirtas paraiškas 61/376932, pateiktą 2010 metų rugpjūčio 25 dieną ir 61/296630, pateiktą 2010 metų sausio 20 dieną, yra įtrauktos kaip nuorodos.

## **PAVYZDŽIAI**

### **1 pavyzdys: ZZZY baltymų gryninimas ir aktyvumo bei specifiškumo tyrimas**

Pasirinktos sekos atpažįstamos BLAST paieškos sistema (Zhang et al., *J. Comput. Biol.* 7(1-2): 203-214 (2000)) ir tiriamasis yra ZZZY šeimos narys, kuris gali būti ekspresuotas dabar žinomomis technikomis, pavyzdžiui *in vitro* transkripcija-transliacija (PURExpress®, NEB). Genas gali būti pasirinktinai kodono optimizuotas ir po to klonuotas kaip aprašyta žemiau PvuRts1I baltymui.

PvuRts1I buvo išgrynintas sekančiu būdu: A genas, koduojantis PvuRts1I, buvo įterptas į pTXB1 vektorių (NEB), klonuotas į ER2566 (T7 ekspresijos šeimininkas) ir augintas LB kultūroje su Ampicilinu (100 µg/ml) 30°C temperatūroje 4×10 ml, kad gautų naktinę kultūrą. 4×10 ml naktinės kultūros buvo pasėtos į 2×1L LB su Ampicilinu (100 µg/ml). Po to 1L kultūros buvo augintas 30°C temperatūroje 6 valandas, į kultūrą pridėtas IPTG iki galutinės 500 µM koncentracijos. Kultūra dar buvo inkubuota 16°C temperatūroje per naktį. PvuRts1I buvo išgrynintas naudojant chitino granulių gravitacinę kolonėlę. Po to PvuRts1I buvo patalpintas į kolonėlę druskingame buferyje (10 mM tris-acetatas, 500 mM, KOAc, pH 8.0), kolonėlė buvo plaunama vėl druskingu buferiu. Druskingo buferio, turinčio 30 mM DTT, trys kolonėlės turiai buvo panaudoti perpilti per kolonėlę. Kolonėlė buvo inkubuota 16 valandų 4°C temperatūroje druskingame buferyje su DDT. Baltymas buvo eliuotas ir sukonzentruotas. Bendrai buvo gauta apie 10 mg PvuRts1I. Baltymas buvo pridėtas į 50 % glicerolį ir laikytas -20°C temperatūroje tolimesniam gelio charakterizavimui. Kai 1 µg eliuanto nubėgo ant nuo 10 % iki 20 % poliakrilamido gelio, buvo aptiktos išgryninto fermento juostos, patvirtinančios fermentų grynumą (žiūr. Fig. 1). Šis būdas buvo taip pat pritaikytas išgryninti kitus ZZZY šeimos narius, tokius kaip PpeHI, AbaSDFI, PatTI, EsaNI ir YkrI.

Išgryninto fermento substrato aktyvumui ir specifiškumui ištirti, sintetinis substratas buvo amplifikuotas panaudojant PGR kaip aprašyta Fig. 2. Citozinas, mC arba hmC buvo panaudoti

PGR reakcijoje, kad gautų produktą su arba nemodifikuotu C, arba pilnai modifikuotu C. PGR produktas su hmC vėliau įvestas į reakciją su BGT ir UDP-gliukoze, kad konvertuotų hmC į ghmC.

Substrato specifiškumo ir aktyvumo analizės buvo atliktos sekančiai: buvo sudarytas reakcijos mišinys, kuris turėjo fermentą (3  $\mu$ l, turintį 33-123  $\mu$ g fermento), 3  $\mu$ l kalio acetato (galutinė koncentracija 250 mM) ir NEB4 buferio (NEB), 50 ng DNR substrato, kuriame citozinas yra nepakeistas arba pakeistas į hmC, ghmC arba mC, ir pridėta 30  $\mu$ l vandens iki galutinio reakcijos tūrio.

Fermentas buvo nuosekliai praskiestas ir pridėtas į reakcijos mišinį, kuris buvo inkubuotas 20 minučių kambario temperatūroje (23°C) ir po to patalpintas į 1,8 % agarozės gelį. Reakcija buvo sustabdyta stop-reagentu, o rezultatai parodyti Fig. 3-5 ir susumuoti 1 lentelėje.

Skėlimo sritis 1 lentelėje parodo, kad hmC buvo atpažintas ir suskaldytas mažiausiai 50 kartų aktyviau negu citozinas ir daugiau negu 8 kartus už mC. Kas dėl AbaSDFI, tai jo santykinis skėlimo aktyvumas buvo 8000 kartų aktyvesnis ghmC atžvilgiu, negu hmC ir 500 kartų aktyvesnis hmC atžvilgiu, negu mC ir nebuvo aptiktas aktyvumas C atžvilgiu.

Priedui prie aktyvumo ir santykinio specifiškumo (greičių) tyrimo, ZZZY šeimos nariams gali būti nustatyta disociacijos pusiausvyros konstanta ( $K_d$ ). Atitinkamai, fermentų reakcijos greitis su skirtingai modifikuotais substrato oligonukleotidais gali būti nustatytas panaudojant gelio mobiliojo poslinkio tyrimus panaudojant mikrogrynas-EZ kolonėles (Millipore, Billerica, MA). Pavyzdžiui, žymėtų oligonukleotidų pastovūs kiekiai gali būti sumaišyti su fermento didėjančiais kiekiais neskėlimo sąlygose, pvz. be  $Mg^{2+}$ , ir fermento-DNR mišinys analizuotas ant gamtinio poliakrilamido gelio, arba panaudojant sukinio kolonėles, kurios tik atpalaiduoja DNR oligonukleotidus (skėlimo produktus), kai fermentas yra prijungtas.  $K_d$  duomenys gali būti panaudoti įmirkymo eksperimentuose, kad įgytų kokristalus. Priedo, C-galinio domeno  $K_d$ , kuris tarnauja kaip surišimo modulis hidroksimetilintai bazei ir jo flankuojančios sekos substrate gali būti paprastai nustatytos.

## 2 pavyzdys: Skėlimo lokacijos nustatymas substrate ZZZY šeimos nariais

Patvirtinimui, kad ZZZY baltymų šeima skelia dsDNR, pažinimo nukleotido 3' transkripcijos kryptimi fiksuotoje padėtyje, buvo sukurti sintetiniai nukleotidai, kurie buvo pažymėti vienoje grandinėje arba abiejose grandinėse, ir kurie turėjo simetrinius hmC arba ghmC nukleotidus arba buvo hemi-hidroksimetilinti kaip parodyta Fig. 7. Visi nukleotidai su fluorescentinėmis žymėmis buvo pagaminti panaudojant organinės sintezės standartines metodikas panaudojant hmC fosforamiditą (Glen Research, sterling, VA). dsDNRs buvo gautos

leidžiant dviems oligonukleotidams aneliuoti į 10  $\mu$ M tirpalą. Visi skaldymai buvo atlikti NEB4 buferyje (NEB). Kiekviename skaldyme 10 pmol substrato buvo inkubuota su 1 PpeHI vienetu 37°C temperatūroje 1 valandai. Vienas vienetas yra apibrėžtas kaip fermento kiekis, reikalingas suskaldyti T4gt DNR 1  $\mu$ g į stabilų mažų fragmentų vaizdą 37°C temperatūroje 1 valandai. Tada 2  $\mu$ l kiekvieno reakcijos mišinio buvo patalpinta į 20 % poliakrilamido gelį su 7M karbamido. Rezultatai yra parodyti Fig. 7. Strėlės rodos juostas, atitinkančias suskilusiems fragmentams, parodytiems gelio dešinėje. Rezultatai rodo, kad fermentai gali skelti atpažinimo sekos abiejose pusėse apytiksliai iš esmės fiksuotame atstume nuo atpažinimo saito.

### 3 pavyzdys: Skėlimo atstumo nustatymas kiekvienoje dviguboje DNR grandinėje

Taip pat buvo nustatytas skėlimo atstumo specifiskumas. Pažymėti ir nepažymėti nukleotidai, turintys hmC, buvo susintetinti panaudojant organinės sintezės standartines metodikas kaip aprašyta 2 pavyzdyje. Oligonukleotido dydžio žymuo buvo taip susintetintas, kad nustatytų skėlimo saitus kaip parodyta Fig. 8A. 3 ir 4 juostos parodo, kad skėlimo saitai hmC turinčios grandinės komplementarioje grandinėje buvo 9-10 nt ant 3' pusės nuo hmC.

Kad būtų nustatyti skėlimo saitai hmC toje pačioje grandinėje, dsDNRs buvo sudaryti leidžiant oligonukleotidams, turintiems hmC, aneliuoti į oligonukleotidus, turinčius nemodifikuotą citoziną į 20  $\mu$ M tirpalą. 40  $\mu$ l aneliuotos DNR buvo paveikta su T4 BGT (NEB) pagal gamintojo protokolą. Po reakcijos DNR buvo išgryninta per QIAquick® nukleotido pašalinimo rinkinį (Quigen, Germantown, MD) ir eliuuota 30  $\mu$ l tūryje. 3' žymėjimo reakcija atlikta panaudojant Taq polimerazę (NEB) dalyvaujant alfa-<sup>33</sup>P dATP pagal standartinius protokolus. Apie 10 pmol substrato buvo inkubuota su AbaSDFI 1 vienetu 4 buferyje 37°C temperatūroje 1 valandai. 2  $\mu$ l reakcijos turinio, kartu su žymenimis, buvo patalpinti į 20 % poliakrilamido gelį su 7M karbamido. Žymenys buvo pagaminti sintetiškai ir pažymėti su gama-<sup>33</sup>P ATP naudojant T4 polinukleotido kinazę (NEB). Gelis buvo išdžiovintas, eksponuotas į fosforo ekraną ir nuskenuotas. Rezultatai yra parodyti Fig. 8B. Fig. 8B skėlimo vietos gali būti išvestos, kad yra 12-13 nukleotidai nuo hmC toje pačioje grandinėje.

Nors specifiniai dydžio skėlimo fragmentai vyrauja Fig. 8A ir 8B, rasta, kad sekos kontekstas ir fermento koncentracija gali sukelti abejones dėl fermento skėlimo saito, tokias kaip skėlimo saitai gali kisti vienu arba dviem nukleotidais.

### 4 Pavyzdys: Nustatymas, ar ghmC turi skirtingą ar tokį pat poveikį į hmC skėlimo vaizdus

Naudojant metodus, aprašytus aukščiau, buvo palygintas skėlimo specifiskumas saitams, modifikuotiems su hmC ir ghmC.

Visi nukleotidai su fluorescentiniais žymenimis ir hmC buvo gauti organinės sintezės būdu, kaip aprašyta aukščiau. hmC turintys substratai buvo konvertuoti į ghmC turinčius substratus naudojant T4 BGT (NEB) pagal gamintojo instrukcijas. Rezultatai Fig. 8C parodė, kad hmC turinti DNR 1 ir 3 juostose turi tokį pat skėlimo atstumo specifiškumą kaip ir ghmC 2 ir 4 juostose.

#### 5 Pavyzdys: hmC fragmentų klonavimas iš žiurkės smegenų genominės DNR

Buvo ištirtas ZZZY fermentų naudingumas atskleidžiant hmC saitus žiurkės smegenų genominėje DNR. 1 µg žiurkės smegenų genominės DNR pradžioje buvo paveiktas su T4 BGT, kad hmC konvertuotų į ghmC. DNR po reakcijos buvo išsodinta ir buvo pajungta AbaSDFI skaldymui. Suskaldyti DNR produktai buvo vėl išgryninti ir pateikti ligavimo reakcijai su sintetiniu dsDNR adapteriu. Vienas adapterio galas turėjo 3' lipnų galą su dviem degeneruotomis bazėmis taip, kad ji galėjo liguoti skeltų produktų galus su 2-bazės 3'-lipniu galu. Tada liguoti produktai nupilti ant žemai besilydančios agarozės gelio ir parinktas dydis buvo nuo 1 kb iki 3 kb. Tada parinkto dydžio DNR PGR amplifikuota naudojant specifinį pradžmenį liguotam adapteriui. PGR pradmuo gali būti suplanuotas turėti daug restrikcijos saitų. Šiame pavyzdyje XbaI buvo naudojamas klonavimo palengvinimui. PGR produktai buvo išgryninti ir suskaldyti su XbaI ir liguoti iki atitinkamo vektoriaus, skirto transformacijai. Kolonijos buvo augintos ir sekvenuotos, kad nustatytų klonuotus įterpimus.

Vėlesnėje kompiuterinėje analizėje klonuotų įterpimų sekose viskas atitiko nurodytam žiurkės genomui. Šių sekų galai, pažymėti 2 bp 3'-lipniu galu, kuriuos generavo AbaSDFI. Tada mes išekstrahavome 40 bp sekos langą abiemis - kairiam ir dešiniam aplink skėlimo saitą ir išdėstėm sekas atskleisdami potencialų kosensą, kaip parodyta Fig. 10. Buvo stebimi du matomi CG klasteriai prie abiejų skėlimo saitų dešinės ir kairės pusės. Atstumas tarp šių CG klasterių ir skėlimo saito atitiko AbaSDFI fermento skėlimo atstumui. Kol nenorima, kad būtų apribota teorijos, siūloma, kad AbaSDFI dimerizacija gali realizuotis apie 22 nukleotidų skėlimų atstume modifikuoto citozino vienoje pusėje. Taigi, gali būti padaryta išvada, kad mažiausiai vienas iš CG saitų yra hidroksimetilintas.

#### 6 Pavyzdys: Hidroksimetilintų bazių genominėje DNR aptikimas

E14 pelės embrioninės kamieninės ląstelės ir NIH3T3 pelės genominės DNR ~ 500 ng buvo pajungta skaldymui su padidintos koncentracijos PvuRts1I 25°C temperatūroje valandai NEB4 buferyje. Molekulinės masės žymenys buvo pBR322 DNR skaldymo su MspI, su kelių juostų dydžiu, išvardintu dešinėje. Suskaldytas produktas buvo paskleistas ant 20 % TBE gelio 2

valandoms prie 140 V, nudažytas su SYBER®Gold nukleorūgšties gelio dažikliu (Invitrogen, dabar Life Technologies, Carlsbad, CA) ir kiekybiškai nustatytas naudojant Typhoon skanerį (GE Healthcare Life Sciences, Piscataway, NJ) naudojant 488 nM lazerį (žiūr. Fig. 15).

#### 7 Pavyzdys: Sekų, turinčių hmCs, kartografavimas

Kai hmC pasitaiko genome, atrodo, kad tai dažniausiai pasitaiko CpG saite. Vis dėl to taip pat tikėtina, kad DNR yra hemi-hidroksimetilinta CpG saite. Jeigu čia DNR sekoje yra du kaimyniniai CpG saitai ir skėlimo vaizdas rodo hemi-hidroksimetilinimą, tai būtina nustatyti, kuris iš dviejų galimų CpGs yra hidroksimetilintas. Sekantis eksperimentas iliustruoja kaip hemi-hidroksimetilinto saito lokacija net gali būti nustatyta kai yra du kaimyniniai CpGs.

DNR fragmentai yra sukurti po specifinio arba nespecifinio genomo skėlimo. Egzistuojantys trys galai ant šių fragmentų yra blokuoti, pavyzdžiui, „užbukinimu“ arba defosforiliniu. Skeliami fragmentai yra liguojami prie oligonukleotidų adapterio, tada fragmentai pajungiami skėlimui fermentu iš ZZZY šeimos (pavyzdžiui, AbaSDFI). Kad nustatytų, ar DNR skėlimo saito ant abiejų pusių turi hmC ar tik viena pusė turi hmC, skaldyta genomine DNR yra „užbukinama“ ir liguojama prie X-kontrolės-koduoto SOLID™ pradmens ir P1 pradmens naudojant NEBNext® greito ligavimo modulį (NEB M2200). Po ligavimo, DNR yra vėl skaldoma ZZZY fermentų šeima, ir fragmentai, kurie negali būti nuskelti nuo X-kontrolės-koduoto SOLID™ pradmens (Applied Biosystems, dabar Life Technologies, Carlsbad, CA) turės seką hmC transkripcijos kryptimi. Fragmentai, kurie gali būti suskaldyti, gamins produktą su specifiniu 2 nt 3' galu, kuris yra iš X-kontrolės-koduoto pradmens. Tada antras Y-kontrolės-koduotas SOLID™ pradmuo yra liguotas prie laisvojo galo; šis SOLID™ judantis ratu sekvenavimas atskleis seką, turinčią hmC. Nesvarbu, jeigu šio metodo 5-hmC saitai yra sugrupuoti arba atsitiktinai paskirstyti (Fig. 11).

hmC pozicijos nustatymas DNR sekoje ir sekos kartografavimas ant genomo žemėlapiu įgalina hidroksimetilomos sukūrimą. Tai taip pat įgalina diagnostikos testą nustatyti hmC buvimą ant taikinio DNR ir po to tai palyginti su etalonine hidroksimetiloma, kad nustatytų koreliaciją su fenotipu. Šalia specifinio fermentinio skaldymo su arba be gliukozilavimo, surišto su qPCR, gali būti nustatytas kiekybinis hidroksimetilinimas specifiniuose saituose.

Identifikuojant poziciją ir hmC kiekį iš skirtingų šaltinių, apimant skirtingus audinius ir ląstelės kultūrą prie skirtingu laiko taškų, galima pateikti svarbių įžvalgų į genų ekspresiją ir reguliaciją.

#### 8 Pavyzdys: hmC aptikimo būdas DNR pavyzdyje

ZZYZ baltymų šeimos charakterizavimas atidaro naujas galimybes didelio skaičiaus mėginių greitai analizei arba diagnostiniams testams, skiriamais gydytojo priimamajame.

Kadaise hmCs yra kartografuoti ant hidroksimetilomos ir parinkta hmC pozicija koreliavo su fenotipu, todėl norėsis nustatyti specifinių hmCs buvimą genomo taikinio regionuose. Tai gali būti lengvai pasiekama naudojant ZZZY baltymų šeimą. Jeigu DNR yra simetriškai hidroksimetilinta, tai atsiras specifinio dydžio fragmentas (20-30 nt). Jeigu DNR yra hemi-hidroksimetilinta, tai hmC buvimas gali būti nusisytas kaip nurodyta žemiau.

Genomini DNR yra pasirinktinai gliukozilinama ir po to paveikiama fermentu iš ZZZY baltymų šeimos atitinkamame buferyje, tokiam kaip 250 mM kalio acetate, kad leistų skėlimą į fragmentus prie nukleotidų 11-13 saitų 3' transkripcijos kryptimi nuo hmC toje pačioje grandinėje ir 9-10 nukleotidų 3' transkripcijos kryptimi nuo hmC komplementarioje grandinėje. Yra parinkti pradmenys, kurie yra komplementarūs sekoms bet kurioje ZZZY baltymo skėlimo sito pusėje, tokia kaip DNR, kuri nėra skelta yra amplifikuota. Tada aptikti fragmentai atitiks hmC nebuvimui. Kur DNR yra skelta ZZZY baltymu, nurodančiu hmC buvimą, jokio amplifikacijos produkto nebus aptikta. Sulyginant amplifikacijos produktą genominiame bandinyje paveiktus ir nepaveiktus ZZZY baltymu, pastarasis gali potencialiai įvertinti hidroksimetilinimo procentinį dydį. Alternatyviai, adapteriai, turintys pradmens seką gali būti liguoti prie pakopinių galų prie 3' skilimo sito arba liguoti prie DNR „užbukintų galų“ ir tik tokie fragmentai, turintys hmC, bus amplifikuoti nuo pradmens priklausančia amplifikacija. Taigi, hmC buvimas bus aptiktas kaip amplifikacijos produktas. Šiuo atveju amplifikuoti fragmentai gali būti sekvenuoti.

Šio metodo pranašumas yra tas, kad visos reakcijos gali būti įvykdytos paprastame reakcijos inde arba mikrofluidiniame įtaise arba luste.

#### 9 Pavyzdys: Rinkinys, skirtas aptikti hmC DNR mėginyje

Pateikiamas rinkinys, kuris turi vieną arba daugiau išgrynintų rekombinantinių baltymų iš ZZZY baltymų šeimos, jų funkcinių darinių arba jų katalitinių fragmentų, pavyzdžiui AbaSDF1, kartu su atitinkamu reakcijos buferiu ir papildomai BGT ir UDP-gliukoze. Be to, rinkinys turi oligonukleotidų adapterius, kad palengvintų didelio našumo sekvenavimą. Be to, rinkinys dar gali apimti fermentus, tinkančius „galų užbukinimui“ ir ligavimui, ir papildomuose įgyvendinimuose gali apimti specifinius pradmenis. Išradimo įgyvendinime, rinkinys dar apima įpakavimo medžiagas ir rinkinių naudojimo instrukcijas.

## SEKŲ SĄRAŠAS

<110> New England Biolabs, Inc.

Zhu, Zhenyu

Guan, Shengxi

Zheng, Yu

Wang, Hua

Quimby, Aine

Zhang, Penghua

Apone, Lynne

<120> Kompozicijos, būdai ir susiję panaudojimai modifikuotos DNR skėlimui

<130> NEB-328-PCT

<150> 61/296,630

<151> 2010-01-20

<150> 61/376,932

<151> 2010-08-25

<160> 16

<170> Patentas 3.5 versijoje

<210> 1

<211> 948

<212> DNR

<213> dirbtinė

<220>

<223> oligonukleotidas

<400> 1

ttttaaatta aaaatgaagt tttaaataca tctaaagtat atatgagtaa acttggtctg 60

acagttacca atgcttaac agtgaggcac ctatctcagc gatctgtcta ttctgtcat 120

ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg 180

gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa 240  
 taaaccagcc agccggaagg gccgagcgca gaagtgggcc tgcaacttta tccgcctcca 300  
 tccagtctat caattgttgc cgggaagcta gagtaagtag ttgccagtt aatagtttgc 360  
 gcaacgttgc tgccattgct acaggcatcg tgggtgcacg ctcgtcgttt ggtatggctt 420  
 cattcagctc cggttcccaa cgatcaaggc gagttacatg atccccatg ttgtgcaaaa 480  
 aagcgggttag ctcttcgggt cctccgatcg ttgtcagaag taagtggcc gcagtgttat 540  
 cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 600  
 ttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 660  
 gttgctcttg cccggcgta atacgggata ataccgccc acatagcaga actttaaag 720  
 tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 780  
 gatccagttc gatgtaacct actcgtgcac ccaactgac ttacgatct tttactttca 840  
 ccagcgttgc tgggtgagca aaaacaggaa ggcaaatgc cgcaaaaag ggaataaggg 900  
 cgacacggaa atgttgaata ctcatctct tctttttca atattatt 948

<210> 2  
 <211> 56  
 <212> DNR  
 <213> dirbtiné  
 <220>  
 <223> oligonukleotidas

<400> 2  
 cggcgtttcc ggggtccata ggctccgcca cgtactctga tgaccagggc atcaca 56

<210> 3  
 <211> 55  
 <212> DNR  
 <213> dirbtiné  
 <220>  
 <223> oligonukleotidas

<400> 3  
 ccatacatat cccttacttc tcctaactg gatgataaag gtagtttatg tggaa 55

<210> 4

<211> 55  
 <212> DNR  
 <213> dirbtinė

<220>  
 <223> oligonukleotidas

<400> 4  
 tccacataaa ctacctttat catccacgtt aggagaagta agggatatgt atgga 55

<210> 5  
 <211> 334  
 <212> PRT  
 <213> *Bacillus megaterium*

<400> 5

Met Asn Lys Tyr Asp Tyr Ile Lys Arg Gln Leu Ala Lys Thr Asn Lys  
 1 5 10 15

Lys Asn Asp Glu Asn Tyr Ile Val Thr Arg Ile Trp His Leu Leu Asp  
 20 25 30

Asn Tyr Asp Ile Lys Ile Asn Thr Gln Gln Tyr Val Val Arg Ser Asn  
 35 40 45

Lys Asn Gln Lys Ala Glu Tyr Gly Leu Ile Asp Leu Tyr Phe Pro Gln  
 50 55 60

Phe Asn Leu Ala Val Glu Ile Asp Glu Ala His His Lys Asn Asp Ile  
 65 70 75 80

Asn Gln Thr Leu Asp Glu Ile Arg Lys Asn Asp Ile Val Asn Ala Leu  
 85 90 95

Asp Cys Glu Phe Ile Arg Ile Asp Ala Thr Gln Ser Phe Glu Lys Ile  
 100 105 110

His Glu Lys Ile Asp Gln Val Val Glu Lys Ile Asn Leu Leu Thr Lys  
 115 120 125

Glu Lys Trp Phe Ile Pro Trp Asp Leu Glu Lys Glu Tyr Asp Pro Asn  
 130 135 140

Thr Tyr Ile Glu Gln Gly Tyr Ile Asp Ala Asp Asp Asn Val Ser Leu  
 145 150 155 160

Arg Leu Val Ala Asp Cys Cys Asn Val Phe Gly Ala Gly Tyr Ala His  
 165 170 175

Gly Ile Gln Lys Ser Gly Ala Pro His Lys Phe Glu Glu Asp Thr Asp  
 180 185 190

Ile Lys Arg Leu Lys Phe Phe Pro Asn Glu Thr Trp Asn Asn Gln Leu  
 195 200 205

Leu Glu Asn Glu Glu Ile Phe Ile Glu Tyr Asn Thr Ile Pro Glu Glu  
 210 215 220

Asn Glu Ala Tyr Phe Gln Lys Arg Met Tyr Gln Leu Asn Gln Lys Ile  
 225 230 235 240

Ala Leu Phe Ala Tyr Ala Lys Thr Ser Ser Gly Arg Phe Glu Ala Ile  
 245 250 255

Phe Lys Gly Leu Tyr Leu Leu Asn Arg Glu Lys Ser Lys Asn Thr Gly  
 260 265 270

Val Leu Thr Tyr Asn Arg Ile Ser Thr Ile Met Pro Thr Tyr Tyr Pro  
 275 280 285

Lys Asp Val Lys Gln Pro Leu Arg Ile Ala Glu Ala Tyr Asn Asn Asp  
 290 295 300

Glu Tyr Lys Val Ala His Phe Tyr Thr Glu Asn Gln Val Arg Lys Phe  
 305 310 315 320

Glu Gly Lys Tyr Lys Lys Arg Tyr Lys Ile Ile Ser Tyr Ser  
 325 330

&lt;210&gt; 6

&lt;211&gt; 309

&lt;212&gt; PRT

<213> *Bifidobacterium bifidum*

&lt;400&gt; 6

Met Ser Gly Ala Asp Lys Leu Gly Tyr Leu Ile Arg Ala Leu Ser Arg  
 1            5            10            15

Thr Lys Arg Lys Asp Tyr Glu Asn Tyr Val Val Asn Ala Ile Trp Asn  
          20            25            30

Arg Leu Ala Met Asp Asp Val Lys Pro Val Thr Gln Gln Leu Val Leu  
      35            40            45

Trp Pro Asp Gly Arg Arg Ser Phe Val Asp Leu Tyr Phe Pro Gln Ala  
      50            55            60

Met Ile Gly Val Glu Cys Asp Glu Ala Tyr His Gln Arg Gln Arg Glu  
 65            70            75            80

Arg Asp Arg Glu Arg Glu Ile Thr Ile Thr Asp Val Leu Arg Gln Ile  
      85            90            95

Arg Gly Glu Asp Tyr Arg Ala Leu His Val Asp Val Ser Gly Ser Tyr  
      100            105            110

Glu Gln Val Glu Arg Ser Ile Asp Asp Cys Val Arg Arg Ile Arg Ala  
      115            120            125

Glu Ile Glu Arg Arg Arg Gln Ala Asn Glu Phe Thr Pro Trp Thr Glu  
      130            135            140

Ala Tyr Val Asp Tyr Lys Glu Phe Tyr Lys Thr Arg Asp Ala Val Ser  
      145            150            155            160

Val Ala Asp Asp Val Gly Phe Pro Arg Ile Ala Asp Ala Val Asn Thr

165                      170                      175

Leu Cys Gly Ser Glu Tyr Lys Arg Phe Gln Glu Ser Trp Phe Val Pro  
180                      185                      190

Ser Val Met Arg Gln Trp Tyr Gly Asp Arg Tyr Arg Val Trp Phe Pro  
195                      200                      205

Lys Leu Ala Ile Gly Gly Lys Ala Val Ala Asn Gly Trp Asn Asn Arg  
210                      215                      220

Leu Ser Asp Asp Gly Thr Tyr Ile Tyr Glu Tyr Asn Glu Asp Ala Asp  
225                      230                      235                      240

Leu Val Asp Pro Val Gly Asp Gly Asp Pro Asn Asp Ile Arg Ile Thr  
245                      250                      255

Phe Ala Lys Ser Ala Asp Pro Val Thr Arg Ile Gln Ala Tyr Arg Phe  
260                      265                      270

Val Gly Val Phe Arg Arg Ile Ser Asn Ser Glu Asp Gly Thr Arg Lys  
275                      280                      285

Arg Tyr Gln Arg Ile Glu Thr Val Phe Pro Ile His Arg Thr Pro Cys  
290                      295                      300

Leu Pro Ile His Arg  
305

<210> 7

<211> 321

<212> PRT

<213> *Acinetobacter baumannii*

<400> 7

Met Cys Asn Lys Ala Ser Ser Asp Leu Thr Asp Tyr Val Ile Arg Gln  
1                      5                      10                      15

Leu Gly Arg Thr Lys Asn Lys Arg Tyr Glu Ala Tyr Val Val Ser Arg  
 20 25 30

Ile Ile His Leu Leu Asn Asp Phe Thr Leu Lys Phe Val Thr Gln Gln  
 35 40 45

Phe Val Arg Leu Ser Asn Lys Lys Ile Ala Leu Thr Asp Leu Tyr Phe  
 50 55 60

Pro Gln Leu Gly Ile His Ile Glu Val Asp Glu Gly His His Phe Leu  
 65 70 75 80

Arg Asn Ser Lys Met Glu Tyr Ser Leu Asn Gln Ile Asp Glu Pro Leu  
 85 90 95

Tyr Ser Ile Ser Gln Thr Glu Ser Asp Ala Met Arg Glu Glu Asp Ile  
 100 105 110

Ile Ser Ile Thr Gly His Lys Ile Phe Arg Val Asn Val Phe Lys Asn  
 115 120 125

Gln Glu Gly Gln Pro Gln Asn Leu Glu Asn Ile His Gln Gln Ile Asp  
 130 135 140

Lys Ile Ile Glu Glu Ile Lys Thr Ala Lys Asn Lys Leu Ile Glu Ala  
 145 150 155 160

Ser Thr Phe Lys Glu Trp Asn Ile Glu Thr Glu Tyr Asn Pro Gln Thr  
 165 170 175

Tyr Ile Asp Leu Gly Arg Ile Ser Leu Ala Asp Asn Val Val Leu Lys  
 180 185 190

Thr Thr Lys Asp Val Cys Asn Cys Phe Gly Tyr Ser Tyr Lys Asn Tyr  
 195 200 205

Gln Arg Gly Gly Ala Leu His Pro Tyr Lys Lys Asp Thr Leu Ile Trp  
 210 215 220

Phe Pro Arg Leu Tyr Glu Asn Lys Asp Trp Ile Asn Thr Ile Ser Pro  
 225                    230                    235                    240

Asp Gly Leu Thr Ile Thr Glu Lys Ser Thr Asp Glu Thr Ile Thr Leu  
                   245                    250                    255

Lys Lys Leu Glu Glu Trp Lys Asn Gly Pro Gln Lys Arg Ile Val Phe  
                   260                    265                    270

Ala Arg Val Lys Asp Asn Leu Ser Ser Arg Ala Met Tyr Arg Phe Met  
                   275                    280                    285

Gly Leu Tyr Glu Phe Gln Lys Ala Asp Leu Lys Asp Gly Ala Val Trp  
                   290                    295                    300

Lys Arg Val Lys Cys Glu Val Gln Thr Tyr Ser Pro Lys Glu Thr Lys  
 305                    310                    315                    320

Cys

<210> 8

<211> 285

<212> PRT

<213> *Pseudoalteromonas atlantica*

<400> 8

Met Asp Lys Lys Glu Tyr Ile Ile Arg Gln Leu Gly Arg Thr Lys Asn  
 1                    5                    10                    15

Lys Lys Tyr Glu Ala Tyr Val Val Thr Arg Ile Ile His Leu Leu Asn  
                   20                    25                    30

Asp Phe Ser Ile Lys Phe Ile Thr Gln Gln Tyr Val Thr Arg Pro Lys  
                   35                    40                    45

Gly Arg Ala Leu Thr Asp Leu Tyr Phe Pro Gln Phe Ala Phe His Ile  
                   50                    55                    60

Glu Val Asp Glu Gly Gln His Phe Asn Gln Ala Asn Ile Glu Ala Asp  
65                      70                      75                      80

Lys Ile Arg Glu Ala Asp Ile Ile Asn Ala Thr Gly His Glu Ile Leu  
                    85                      90                      95

Arg Ile Asp Val Thr Lys Ser Phe Asp Asp Ile Asn Thr Gln Ile Asp  
                    100                      105                      110

Ala Ala Val Asn Lys Ile Lys Ser Met Arg Gln Glu Ile Ser Phe Ile  
                    115                      120                      125

Pro Trp Asp Ile Asp Ser Glu Phe Asp Ser Ala Thr Tyr Ile Lys Arg  
                    130                      135                      140

Gly Tyr Ile Asp Ile Lys Asp Asn Val Ala Phe Lys Thr Ile Lys Asp  
145                      150                      155                      160

Ala Cys Asn Cys Phe Gly His Asn Tyr Thr Gly Tyr Gln Arg Ala Gly  
                    165                      170                      175

Ala Ala His Pro Asp His Tyr Ile Met Leu Trp Phe Pro Lys Leu Phe  
                    180                      185                      190

Pro Asn Gly Glu Trp Asp Asn Gln Ile Ser Ser Asp Glu Glu Val Ile  
                    195                      200                      205

Thr Glu Arg Asn Glu Asp Asp Lys Lys Ala Lys Gln His Val Ser Ser  
                    210                      215                      220

His Val Asn Asn Lys Glu Lys His Lys His Gln Arg Ile Val Phe Ala  
225                      230                      235                      240

Lys Val Arg Gly Asn Leu Gly Asp Val Leu Tyr Arg Phe Arg Gly Gln  
                    245                      250                      255

Tyr Gln Leu Asp Val Lys Asp Ser Asn Glu Lys Thr Gly Leu Ile Trp

260                      265                      270

Arg Arg Ile Lys Thr Arg Val Lys Thr Tyr Glu Gln Asn  
           275                      280                      285

<210> 9

<211> 289

<212> PRT

<213> *Yersinia kristensenii*

<400> 9

Met Ile Ser Gln Tyr Glu Tyr Val Val Arg Gln Leu Ala Arg Thr Lys  
   1                      5                      10                      15

Asn Lys Lys His Glu Gln Tyr Val Val Thr Gly Ile Val His Lys Leu  
           20                      25                      30

Asn Arg Asp Asp Ile Lys Phe Val Thr Gln Gln Tyr Val Lys Arg Glu  
           35                      40                      45

Ser Gly Arg Ala Leu Thr Asp Leu Tyr Phe Pro Ala Ile Asn Leu His  
           50                      55                      60

Ile Glu Ile Asp Glu Pro Phe His Leu Lys Gln Ala Glu His Asp Asn  
   65                      70                      75                      80

Leu Arg Glu Ala Asp Ile Ile Asp Ala Thr Gly His Glu Val Ile Arg  
           85                      90                      95

Ile Ser Val Asp Gly Ser Leu Arg Gln Met Asn Glu Arg Ile Asp Asp  
           100                      105                      110

Cys Val Ala Ala Ile Lys Ser Lys Ile Ser Ala Leu Gly Asp Cys Phe  
           115                      120                      125

Glu Pro Trp Asp Met Asp Lys Glu Leu Ser Ile Glu Pro His Ile Arg  
           130                      135                      140

Arg Gly Tyr Ile Asp Val Lys Asp Asn Val Ala Phe Arg Arg Ile Thr  
 145                    150                    155                    160

Asp Ala Cys Asn Cys Phe Gly His Asn Tyr Lys Phe Leu Gln Lys Ala  
                   165                    170                    175

Gly Ala Lys His Pro Tyr His Asp Asp Ile Leu Ile Trp Leu Pro Lys  
                   180                    185                    190

Leu Phe Asp Asn Glu His Trp Ser Asn Gln Ile Ser Asn Asp Glu Asn  
                   195                    200                    205

Val Ile Thr Glu Ile Pro Lys Ser Glu Asp Ala Gln Ala Ala His Phe  
                   210                    215                    220

Asp Lys Trp Met Ala Glu Thr Arg Asn Lys Arg Leu Val Phe Ala Lys  
 225                    230                    235                    240

Ala Lys Asp Asn Leu Gly Met Thr Leu Tyr Arg Phe Lys Gly Leu Tyr  
                   245                    250                    255

Glu Leu Asn Pro Lys Lys Ser Asn Arg Thr Ile Gly Leu Tyr Trp Gln  
                   260                    265                    270

Arg Ile Ser Thr Arg Val Lys Thr Tyr Pro Ser Pro Ala Arg Asn Pro  
                   275                    280                    285

Asp

<210> 10

<211> 283

<212> PRT

<213> *Shewanella pealeana*

<400> 10

Met Lys Thr Leu His Met Thr Arg Gln Leu Gln Arg Cys Lys Asn Lys  
 1                    5                    10                    15

Arg Phe Glu Leu Tyr Ala Ile Thr Arg Ile Ile His Lys Val Asp Asp  
 20 25 30

Leu Asp Ile Lys Phe Ile Thr Gln Gln Tyr Val Ala Arg Pro Asp Gly  
 35 40 45

Phe Ala Leu Thr Asp Leu Tyr Leu Pro Gln Leu Lys Leu His Ile Glu  
 50 55 60

Ile Asp Glu Gly Phe His Lys Gln Gln Val Asp Ala Asp Lys Val Arg  
 65 70 75 80

Glu Leu Asp Ile Ile Thr Ala Thr Asp His Gln Val Lys Arg Ile Asp  
 85 90 95

Ala Ser Val Ala Ile Glu Gln Ile Asn Leu Gln Val Asp Gln Ile Val  
 100 105 110

Ala Glu Ile Leu Gln Ser Val Glu Val Gln Lys Ala Ala Gly Asn Phe  
 115 120 125

Gln Ala Trp Asp Pro Glu Thr Glu His Ser Val Ser Tyr Leu Gln Ser  
 130 135 140

Arg Gly Val Ile Arg Ala Ser Glu Asn Val Ala Leu Arg Thr Ser Ala  
 145 150 155 160

Glu Val Cys Asn Leu Leu Gly His Asn Tyr Lys Gly Trp Gln Arg Ser  
 165 170 175

Ser Ala Ser Val Pro His Tyr Pro Asn Ile Arg Leu Trp Phe Pro Lys  
 180 185 190

Leu Tyr Pro Asn Glu Gln Trp Phe Asn His Ile Ser Tyr Asp Gly Cys  
 195 200 205

Glu Ile His Glu Tyr Cys Ile Glu Ser Glu Thr Lys Lys Arg Gln Phe  
 210 215 220

Ile Asp Lys Asn Leu Ser Glu Asn Ile Gln Gln Met Val Phe Ala Arg  
 225                230                235                240

Val Lys Asp Glu Leu Gly Gln Thr Met Tyr Arg Phe Lys Gly Leu Phe  
                  245                250                255

Ile Leu Asp Arg Asp Lys Thr Cys His Glu Ser Gly Val Tyr Trp Lys  
                  260                265                270

Arg Val Ala Thr Glu Phe Glu Leu Gly Leu Thr  
                  275                280

<210> 11

<211> 296

<212> PRT

<213> Aplinkos mėginys, Newport Harbor, Rhode Island

<400> 11

Met Ser Lys Phe Thr Lys Glu Thr Tyr Val Thr Arg Asn Phe Gln Lys  
 1                5                10                15

Ile Ser Gly Lys Arg Trp Glu Leu Tyr Val Ile Thr Arg Val Ile His  
                  20                25                30

Leu Leu Asn Asp Pro Asp Ile Glu Tyr Val Cys Gln Gln Tyr Ile Asn  
                  35                40                45

Pro Pro Gln Asn Lys Asp Tyr Tyr Leu Ala Asp Leu Ala Phe Pro Ser  
                  50                55                60

Leu Lys Leu Tyr Leu Glu Ile Asp Glu Gly Gln His Gly Ser Glu Met  
 65                70                75                80

His Gln Thr Ser Asp Leu Lys Arg Asp Ala Glu Ile Leu Glu Ala Thr  
                  85                90                95

Asp Trp Thr Cys Lys Arg Ile Pro Val Phe Val Lys Lys Gly Ser Ser  
                  100                105                110

Lys Ile Asp Lys Ser Leu Glu Ala Leu Asn Lys Glu Ile Asp Asp Phe  
 115                      120                      125

Val Ser Tyr Val Glu Gln Lys Lys Gln Lys Met Val Ser Ala Gly His  
 130                      135                      140

Lys Ile Ser Trp Asn Tyr Glu Gln Lys Phe Ser Pro Asp Thr Phe Ile  
 145                      150                      155                      160

Ala Lys Glu Lys Ile Lys Val Ser Asp Asn Val Ala Leu Leu Asn His  
 165                      170                      175

Arg Asp Val Leu Arg Leu Phe Gly Tyr Lys Lys Gly His Tyr Gln Arg  
 180                      185                      190

Ala Val Trp Thr Ile Lys Lys Thr Asn Gln Met Val Trp Phe Pro Lys  
 195                      200                      205

Leu Tyr Pro Asn Ser Asp Trp Val Asn Ser Phe Asp Asp Lys Ser Gly  
 210                      215                      220

Tyr Ile His Gln Phe Arg Lys Asp Asn Gln Pro His Pro Met Pro Lys  
 225                      230                      235                      240

Glu Gly Asp Pro Asp Arg Ile Val Phe Ala His Gln Lys Asn Ile Phe  
 245                      250                      255

Gly Gln Thr Val Tyr Lys Phe Phe Gly Ile Phe Arg Ala Asp Leu Asn  
 260                      265                      270

Lys Thr Asp Pro Val His His Tyr Phe Lys Arg Ile Asn Thr Cys Leu  
 275                      280                      285

Asp Leu Ser Arg Tyr Ser Ala Asn  
 290                      295

<210> 12

&lt;211&gt; 293

&lt;212&gt; PRT

<213> *Proteus penneri*

&lt;400&gt; 12

Met Ser Lys Thr Asp Tyr Ile Leu Arg Ser Leu Ser Lys Ile Thr Lys  
 1            5            10            15

Lys Arg Trp Glu His Tyr Val Ile Asn Arg Ile Phe His Lys Leu Asp  
           20            25            30

Asp Pro Glu Ile Glu Phe Val Cys Gln Gln Cys Ile Arg Lys Ala Asn  
           35            40            45

Ser Pro Asp Lys Ile Tyr Leu Ala Asp Leu Phe Phe Pro Gln Leu Ala  
           50            55            60

Leu Tyr Leu Glu Ile Asp Glu Glu His His Asp Ser Asp Glu Ala Lys  
 65            70            75            80

Lys Lys Asp Ala Lys Arg Arg Leu Asp Ile Ile Glu Ala Thr Gly Phe  
           85            90            95

Ile Glu Lys Arg Ile Pro Ala Ser Asn Val Thr Ile Glu Gln Leu Asn  
           100            105            110

Thr Ser Ile Asp Glu Phe Val Lys Leu Leu Ile Asp Thr Lys Glu Lys  
           115            120            125

Gln Lys Ala Gln Lys Lys Phe Ile Pro Trp Asp Tyr Ser Ala Gln Tyr  
           130            135            140

Thr Ala Lys Arg His Ile Asp Ala Gly Phe Ile Glu Val Gly Pro His  
 145            150            155            160

Ala Ile Phe Arg Tyr His Arg Asp Ala Leu Glu Cys Phe Gly Tyr Ile  
           165            170            175

Asn Lys Gly His His Gln Ser Gly Ser Trp Lys Leu Pro Glu Asn Ile  
 180 185 190

Val Arg Glu Ile Gly Leu Ser Gly Arg Ile Met Val Trp Phe Pro Arg  
 195 200 205

Leu Tyr Asn Ala Gly Val Trp Asn Asn Glu Leu Ser Pro Asp Gly Glu  
 210 215 220

Trp Ile Thr Glu Glu Ser Leu Glu Val Asp Asn Asn Tyr Ile Glu Asp  
 225 230 235 240

Trp Asp Tyr Arg Ile Val Met Ala His Ser Arg Asp Glu Leu Asn Arg  
 245 250 255

Val Leu Tyr Arg Phe Leu Gly Val Phe Gln Ile Asp Lys Asn Lys Ser  
 260 265 270

Val Glu Gly Lys Asn Ile Phe Lys Arg Ile Asn Thr Lys Val Lys Val  
 275 280 285

Phe Asn Ser Tyr Asn  
 290

<210> 13

<211> 293

<212> PRT

<213> *Proteus vulgaris*

<400> 13

Met Ser Lys Thr Asp Tyr Ile Leu Arg Ala Leu Ser Lys Ile Ser His  
 1 5 10 15

Lys Arg Trp Glu His Tyr Ile Ile Asn Arg Val Val His Thr Leu Asp  
 20 25 30

Asp Pro Asp Ile Glu Phe Val Cys Gln Gln Cys Ile Arg Lys Glu Gly  
 35 40 45

His Leu Gly Lys Ile Tyr Leu Ala Asp Leu Leu Phe Pro Gln Leu Asn  
 50                      55                      60

Leu Tyr Leu Glu Ile Asp Glu Ala His His Asp Ser Asn Asp Ala Arg  
 65                      70                      75                      80

Lys Ala Asp Ala Val Arg Arg Leu Asp Ile Val Glu Ala Thr Gly Phe  
                     85                      90                      95

Gln Glu Glu Arg Ile Pro Ala Ser Asn Ile Thr Leu Ser Glu Val Asn  
                     100                      105                      110

Lys Leu Val Asp Glu Phe Val Arg Leu Val Lys Asp Lys Lys Glu Glu  
                     115                      120                      125

Leu Glu Asn Gln Gly Leu Phe Phe Arg Trp Asp Tyr Asp Glu Arg Tyr  
                     130                      135                      140

Ser Ala Lys Lys His Ile Asn Thr Gly Tyr Met Ala Val Gly Pro Asn  
                     145                      150                      155                      160

Ser Val Phe Arg Tyr His Arg Asp Ala Leu Gln Cys Phe Gly Tyr Arg  
                     165                      170                      175

Arg Glu Gly His His Gln Ser Gly Gly Trp Ala Leu Pro Ala Glu Val  
                     180                      185                      190

Ala Gln Ser Ile Gly Leu Thr Gly Arg Val Met Val Trp Phe Pro Arg  
                     195                      200                      205

Leu Tyr Glu Ala Gly Glu Trp Lys Asn Ala Leu Ser Ala Asp Gly Asn  
                     210                      215                      220

Lys Ile Thr Glu Gln Ser Leu Asn Ala Thr Arg Asn Tyr Gln Glu Thr  
                     225                      230                      235                      240

Trp Asp Tyr Arg Ile Val Met Ala His Ser Arg Asp Glu Leu Asn Arg  
                     245                      250                      255

Thr Leu Tyr Arg Phe Leu Gly Val Phe Ala Ile Asp Val Asp Lys Ser  
 260 265 270

Ser Asp Glu Val Lys Val Phe Ser Arg Val Tyr Ser Arg Val Asn Val  
 275 280 285

Tyr Arg Ser Gln Asn  
 290

<210> 14  
 <211> 102  
 <212> PRT  
 <213> dirbtinė  
 <220>  
 <223> konservatyvi seka N-galiniame domene

<220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (2)..(8)  
 <223> X = bet kuri aminorūgštis

<220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (10)..(11)  
 <223> X= bet kuri aminorūgštis

<220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (13)..(13)  
 <223> X= bet kuri aminorūgštis

<220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (15)..(32)  
 <223> X= bet kuri aminorūgštis

<220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (35)..(50)  
 <223> X gali būti pakartota minimaliai 11 kartų ir maksimaliai 16 kartų  
 <220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (53)..(54)  
 <223> X= bet kuri aminorūgštis

<220>

<221> ĪVAIRŪS POŽYMAI

<222> (56)..(61)

<223> X= bet kuri aminorūgštis

<220>

<221> ĪVAIRŪS POŽYMAI

<222> (63)..(63)

<223> X= bet kuri aminorūgštis

<220>

<221> ĪVAIRŪS POŽYMAI

<222> (66)..(67)

<223> X= bet kuri aminorūgštis

<220>

<221> ĪVAIRŪS POŽYMAI

<222> (69)..(94)

<223> X gali būti pakartota minimaliai 6 kartus ir maksimaliai 26 kartus

<220>

<221> ĪVAIRŪS POŽYMAI

<222> (96)..(97)

<223> X= bet kuri aminorūgštis

<220>

<221> ĪVAIRŪS POŽYMAI

<222> (99)..(101)

<223> X= bet kuri aminorūgštis

<400> 14

Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Lys Xaa Xaa Glu Xaa Tyr Xaa Xaa  
1                    5                    10                    15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
                  20                    25                    30

Gln Gln Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
          35                    40                    45

Xaa Xaa Asp Leu Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Glu Xaa Asp  
          50                    55                    60

Glu Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
65                    70                    75                    80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asp Xaa  
 85 90 95

Xaa Arg Xaa Xaa Xaa Ile  
 100

<210> 15  
 <211> 79  
 <212> PRT  
 <213> dirbtinė

<220>  
 <223> konservatyvi seka C-galiniame domene

<220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (2)..(2)  
 <223> X= bet kuri aminorūgštis

<220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (4)..(43)  
 <223> X= bet kuri aminorūgštis; X gali būti pakartota minimaliai 30 kartų ir maksimaliai 40 kartų

<220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (45)..(57)  
 <223> X= bet kuri aminorūgštis; X gali būti pakartota minimaliai 12 kartų ir maksimaliai 13 kartų

<220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (59)..(59)  
 <223> X= bet kuri aminorūgštis

<220>  
 <221> ĮVAIRŪS POŽYMAI  
 <222> (61)..(78)  
 <223> X= bet kuri aminorūgštis; X gali būti pakartota minimaliai 16 kartų ir maksimaliai 18 kartų

<400> 15

Trp Xaa Asn Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 1            5            10            15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
           20            25            30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ala Xaa Xaa Xaa Xaa  
           35            40            45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Xaa Gly Xaa Xaa Xaa Xaa  
           50            55            60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg  
 65            70            75

<210> 16

<211> 102

<212> PRT

<213> dirbtinė

<220>

<223> konservatyvi seka N-galiniame domene

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (2)..(8)

<223> X= bet kuri aminorūgštis

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (10)..(11)

<223> X= bet kuri aminorūgštis

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (13)..(13)

<223> X= bet kuri aminorūgštis

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (15)..(32)

<223> X= bet kuri aminorūgštis

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (35)..(50)

<223> X= bet kuri aminorūgštis, X gali būti pakartota minimaliai 11 kartų ir maksimaliai 16tų

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (53)..(54)

<223> X= bet kuri aminorūgštis

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (56)..(61)

<223> X= bet kuri aminorūgštis

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (63)..(63)

<223> X= bet kuri aminorūgštis

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (66)..(67)

<223> X= bet kuri aminorūgštis

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (69)..(94)

<223> X= bet kuri aminorūgštis

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (96)..(97)

<223> X= bet kuri aminorūgštis

<220>

<221> ĮVAIRŪS POŽYMAI

<222> (99)..(101)

<223> X= bet kuri aminorūgštis

<400> 16

Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Lys Xaa Xaa Glu Xaa Tyr Xaa Xaa  
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 20 25 30

Gln Gln Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 35 40 45

Xaa Xaa Asp Leu Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Glu Xaa Asp  
 50 55 60

Glu Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asp Xaa  
 85 90 95

Xaa Arg Xaa Xaa Xaa Ile  
 100

## Išradimo apibrėžtis

1. Preparatas b e s i s k i r i a n t i s tuo, kad apima vieną arba daugiau išgrynintų rekombinantinių baltymų, kuriame baltymai yra ZZZY baltymų šeimos nariai, kai kiekvienas ZZZY baltymų šeimos narys turi N-galinį domeną, kur N-galinis domenas turi mažiausiai 95 % homologiją su SEQ ID Nr. 14; ir reakcijos buferį.

2. Preparatas pagal 1 punktą, b e s i s k i r i a n t i s tuo, kad vienas arba daugiau išgrynintų baltymų pirmenybiškai rišasi prie hidroksimetilinto citozino (hmC) ir/arba prie glikozilinto hidroksimetilinto citozino (ghmC), palyginus su rišimusi prie DNR metilinto citozino (mC).

3. Preparatas pagal bet kurį iš nuo 1 iki 2 punktą, b e s i s k i r i a n t i s tuo, kad vienas arba daugiau išgrynintų baltymų pirmenybiškai skelia DNR hmC arba ghmC nustatytu atstumu nuo 3'.

4. Preparatas pagal bet kurį iš nuo 1 iki 3 punktų, b e s i s k i r i a n t i s tuo, kad nustatytas atstumas yra 11-13 nukleotidų grandinėje, turinčioje hmC arba ghmC, arba 9-10 nukleotidų komplementarioje grandinėje.

5. Preparatas pagal bet kurį iš nuo 1 iki 4 punktų, b e s i s k i r i a n t i s tuo, kad vienas arba daugiau išgrynintų baltymų skelia DNR, turinčią ghmC, hmC ir/arba mC santykiu mažiausiai 8:1 pagal ghmC arba hmC su mC.

6. Preparatas pagal bet kurį iš nuo 1 iki 5 punktų b e s i s k i r i a n t i s tuo, kad buferis apima druską, besiskiriančią anijonu, parinktu iš grupės, susidedančios iš sulfato, fosfato, acetato ir citrato.

7. Preparatas pagal bet kurį nuo 1 punkto iki 6 punktų, b e s i s k i r i a n t i s tuo, kad buferis neturi chlorido, nitrato, karbonato arba imidazolo druskų.

8. Preparatas pagal 6 punktą, b e s i s k i r i a n t i s tuo, kad druskos koncentracija yra 50-500 mM.

9. Preparatas pagal bet kurį iš nuo 1 iki 8 punktą, b e s i s k i r i a n t i s tuo, kad bent vienas narys apima SEQ ID Nr. 14.

10. Preparatas pagal bet kurį iš nuo 1 iki 9 punktą, b e s i s k i r i a n t i s tuo, kad bent vienas išgrynintas baltymas turi aminorūgščių seką su mažiausiai 95 % sekos homologija su SEQ ID Nr. 7.

11. Preparatas pagal bet kurį iš nuo 1 iki 10 punktą, b e s i s k i r i a n t i s tuo, kad bent vienas išgrynintas baltymas turi aminorūgščių seką, apimančią mažiausiai 95 % sekos homologiją su fermentu, parinktu iš PvuRts1I, PpeHI, EsaSS310P, EsaRBORFBP, PatTI, YkrI, EsaNI, SpeAI, BbiDI, PfrCORF1I80P, PcoORF314P, BmeD1, AbaSDF1, AbaCI, AbaAI, AbaSI, AbaUMB3ORFAP ir Asp6ORFAP.

12. Preparatas pagal bet kurį iš nuo 1 iki 11 punktą, b e s i s k i r i a n t i s tuo, kad bent vienas iš išgrynintų baltymų yra parinktas iš grupės, susidedančios iš: PvuRts1I, PpeHI, EsaSS310P, EsaRBORFBP, PatTI, YkrI, EsaNI, SpeAI, BbiDI, PfrCORF1I80P, PcoORF314P, BmeD1, AbaSDF1, AbaCI, AbaAI, AbaSI, AbaUMB3ORFAP ir Asp6ORFAP ir jų katalitiškai aktyvių mutantų ir jų darinių.

13. hmC aptikimo būdas genominės DNR mėginyje, apimantis:

- a) preparato pagal bet kurį iš nuo 1 iki 12 punktų pridėjimą į genominės DNR mėginį;
- b) sudarymą galimybės baltymui skelti genominę DNR skėlimo vietoje;
- c) skėlimo vietos DNR sekos bent vienos pusės nustatymą;
- d) pasirinktinai, DNR sekos kartografavimą ant referencinės genominės DNR sekos; ir
- e) hmC aptikimą.

14. Būdas pagal 13 punktą, b e s i s k i r i a n t i s tuo, kad preparatas a) pakopoje dar apima: vieną arba daugiau DNR polimerazių, pradmenų ir adapterių.

15. Būdas pagal bet kurį iš nuo 13 iki 14 punktų, b e s i s k i r i a n t i s tuo, kad dar apima: prieš nustatant DNR seką c) pakopoje, suskaldytos genominės DNR amplifikaciją b) pakopoje.

16. Būdas pagal bet kurį iš nuo 13 iki 15 punktų, b e s i s k i r i a n t i s tuo, kad prieš a) pakopą genominė DNR reaguoja su  $\beta$ -gliukoziltransferaze (BGT).

17. Būdas pagal bet kurį iš nuo 13 iki 16 punktų imtinai, dar apimantis: a)-c) pakopų vykdymą viename reakcijos inde arba mikrofluidiniame įtaise.

18. Rekombinantinio ZZZY šeimų baltymų gryninimo būdas, apimantis:  
klonavimą ir ekspresavimą sulietų baltymų, apimančių ZZZY baltymus, inteiną ir afininio surišimo baltymą;  
sukėlimą sulieto baltymo rišimosi prie matricos afininio surišimo baltymo priemonėmis;  
ZZZY baltymo atskėlimą nuo inteino; ir  
išgryninto baltymo regeneravimą eliuote.

19. Būdas pagal 18 punktą, b e s i s k i r i a n t i s tuo, kad afininio surišimo baltymas yra chitiną surišantis domenas.

20. Fragmentų rinkinys, b e s i s k i r i a n t i s tuo, kad apima mažiausiai vieną iš a) oligonukleotido fragmentą, turintį 20-23 nukleotidų dydį su patalpintu centre hmC arba ghmC; arba b) didelę DNR arba oligonukleotidų fragmentus, turinčius hmC arba ghmC, patalpintus DNR duplekso vienguboje grandinėje prie 11-13 nukleotidų padėtimi nuo grandinės 3' galo.

21. Rinkinys, b e s i s k i r i a n t i s tuo, kad apima: vieną arba daugiau išgrynintų rekombinantinių ZZZY baltymų šeimų baltymų, jų funkcinių darinių arba jų katalitinių fragmentų veiksmingame buferyje fermento aktyvumui; ir naudojimo instrukciją.

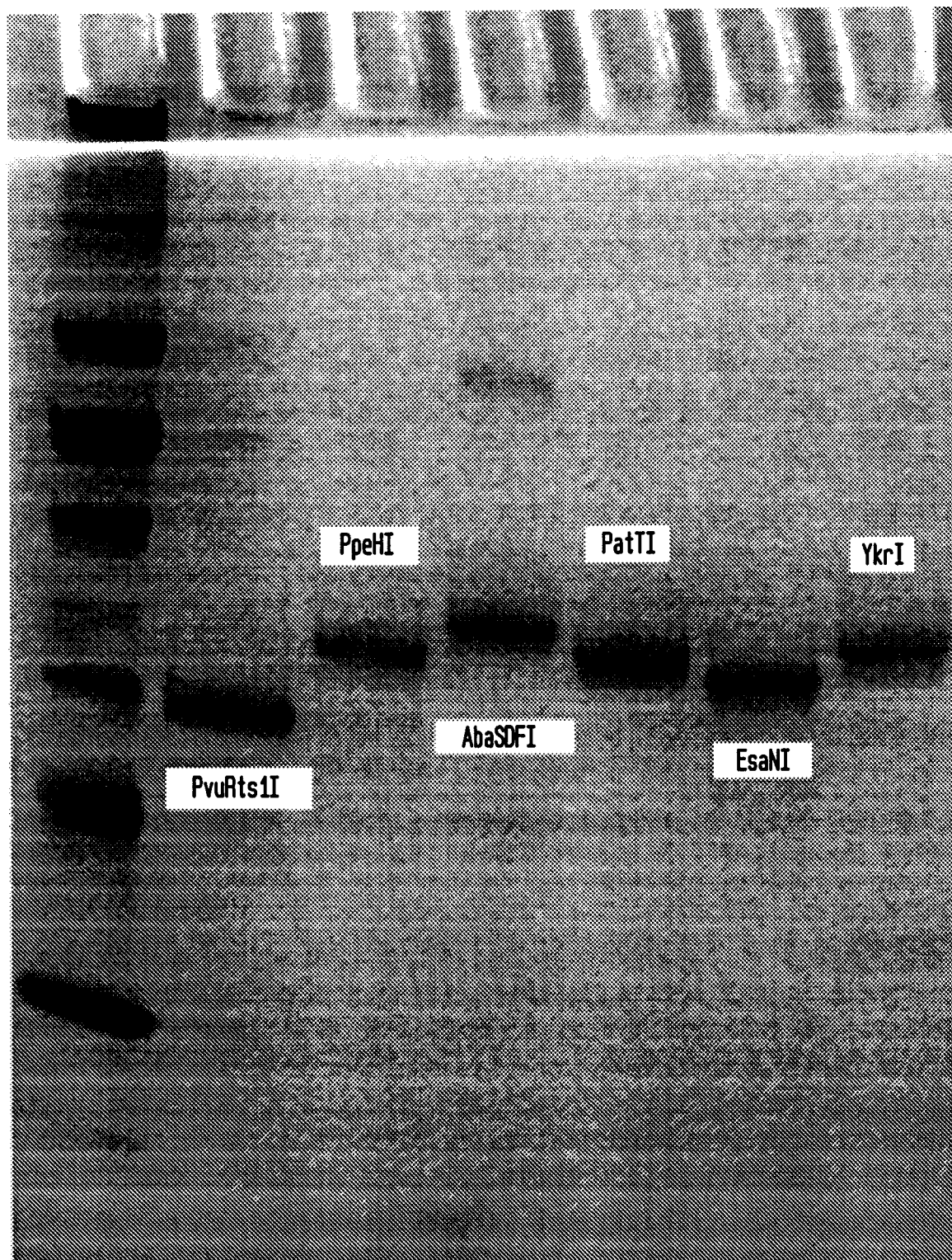
22. Rinkinys pagal 21 punktą, b e s i s k i r i a n t i s tuo, kad dar apima BGT ir UDP-gliukozę.

23. Rinkinys pagal 21 arba 22 punktą, b e s i s k i r i a n t i s tuo, kad dar apima pradmenis.

24. Reakcijos indas, apimantis: preparatą pagal bet kurį iš 1-12 punktų, pradmenis ir/arba adapterius ir termostabilią polimerazę.

25. Preparatas, b e s i s k i r i a n t i s tuo, kad apima: ZZZY baltymų šeimų nario baltymo funkcinius darinius arba katalitinius fragmentus; ir reakcijos buferį.

**FIG. 1**

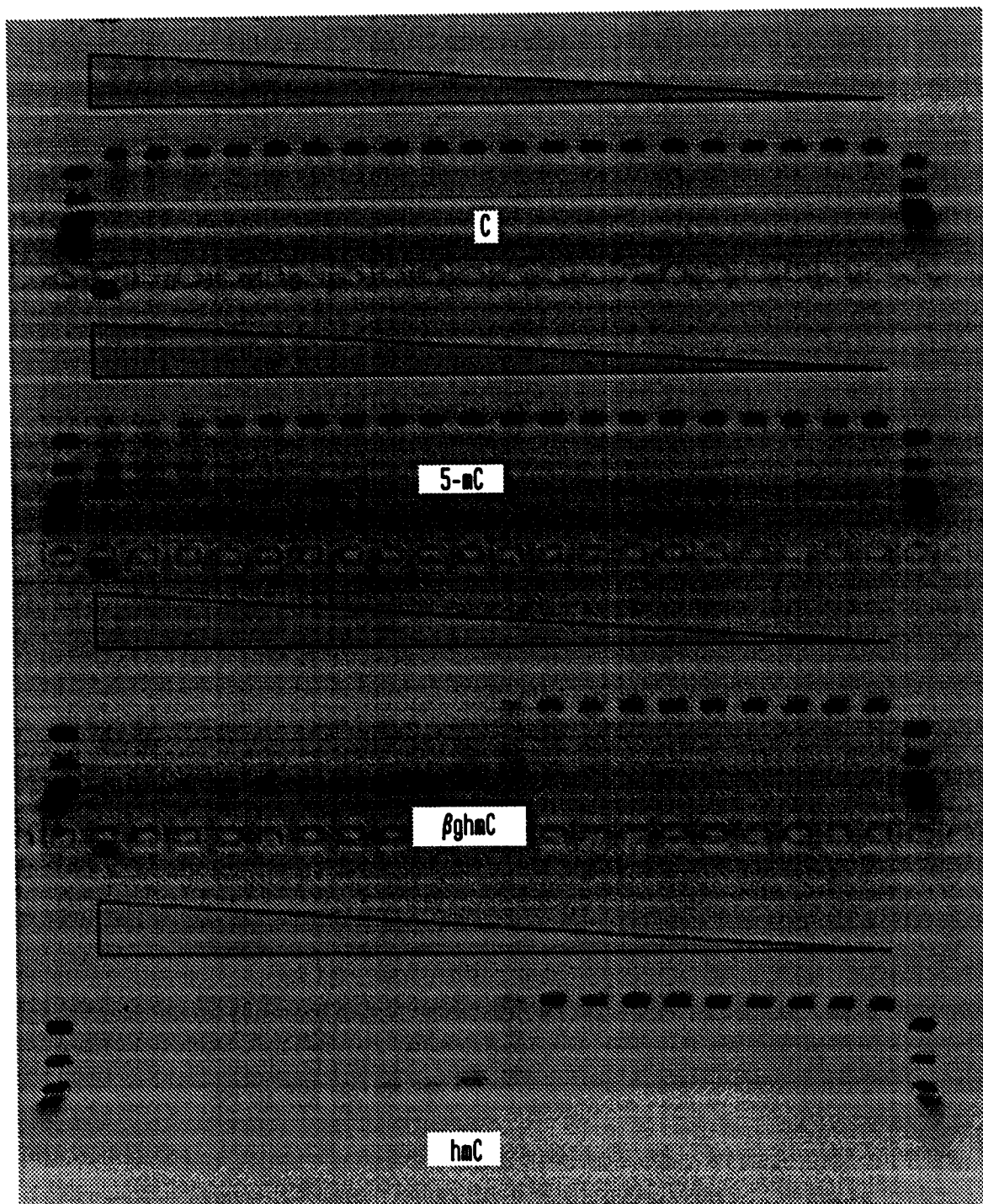


**FIG. 2**  
(SEQ ID NO: 1)

1 TTTTAAATTA AAAATGAAGT TTTAAATCAA TCTAAAGTAT ATATGAGTAA  
51 ACTTGGTCTG ACAGTTACCA ATGCTTAATC AGTGAGGCAC CTATCTCAGC  
101 GATCTGTCTA TTTCGTTTAT CCATAGTTGC CTGACTCCCC GTCGTGTAGA  
151 TAACTACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA  
201 CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC  
251 AGCCGGAAGG GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA  
301 TCCAGTCTAT CAATTGTTGC CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT  
351 AATAGTTTGC GCAACGTTGT TGCCATTGCT ACAGGCATCG TGGTGTACAG  
401 CTCGTGCTTT GGTATGGCTT CATTGAGCTC CGGTTCCCAA CGATCAAGGC  
451 GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT  
501 CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTAT CACTCATGGT  
551 TATGGCAGCA CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT  
601 TTTCTGTGAC TGGTGAGTAC TCAACCAAGT CATTCTGAGA ATAGTGTATG  
651 CGGCGACCGA GTTGCTCTTG CCCGCGTCA ATACGGGATA ATACCGCGCC  
701 ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT TCTTCGGGGC  
751 GAAAACTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC  
801 ACTCGTGAC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC  
851 TGGGTGAGCA AAAACAGGAA GGCAAAATGC CGCAAAAAG GGAATAAGGG  
901 CGACACGGAA ATGTTGAATA CTCATACTCT TCCTTTTTC AATATTAT

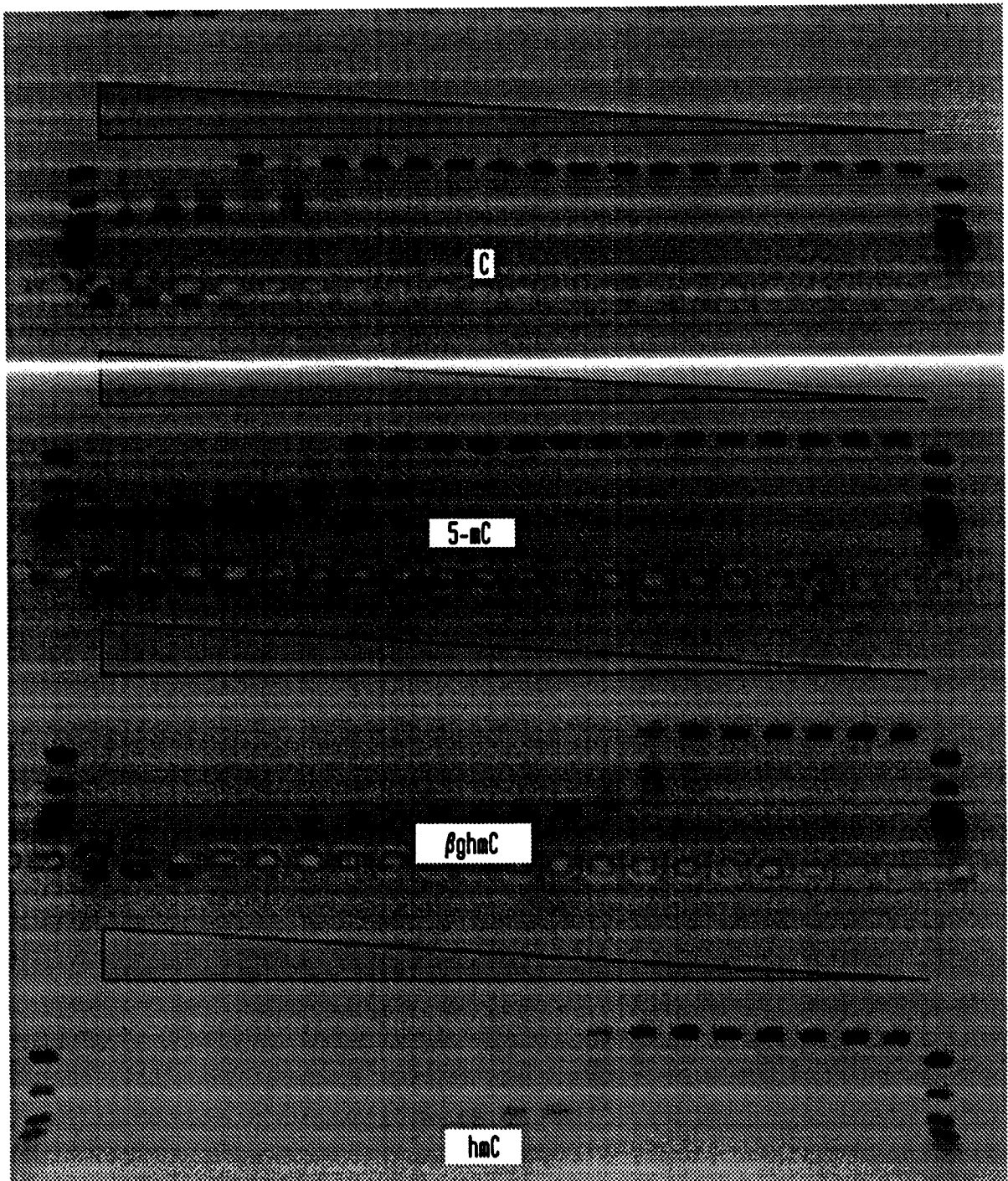
3/18

FIG. 3



4/18

FIG. 4



5/18

**FIG. 5**

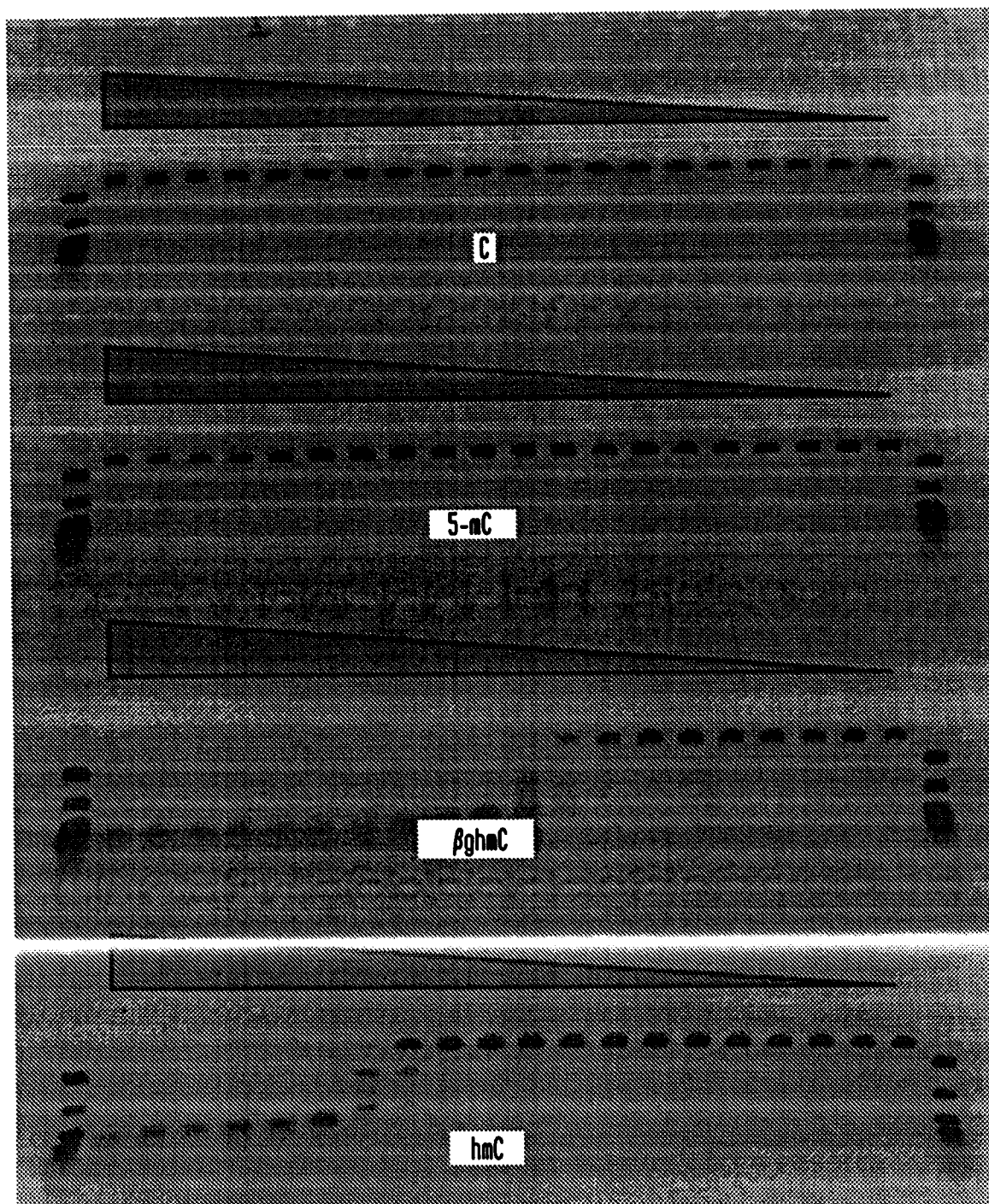


FIG. 6B

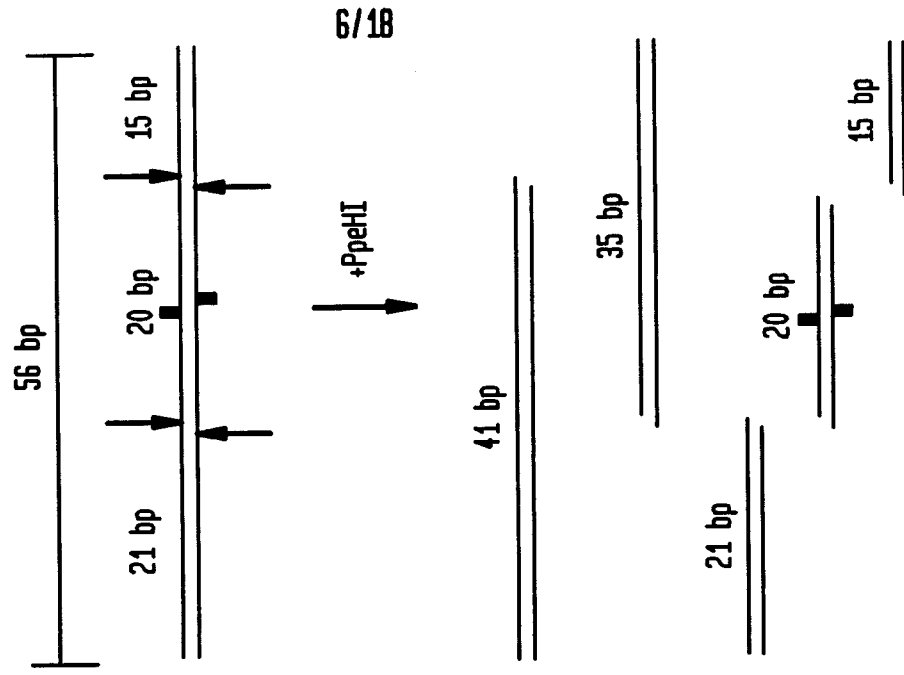
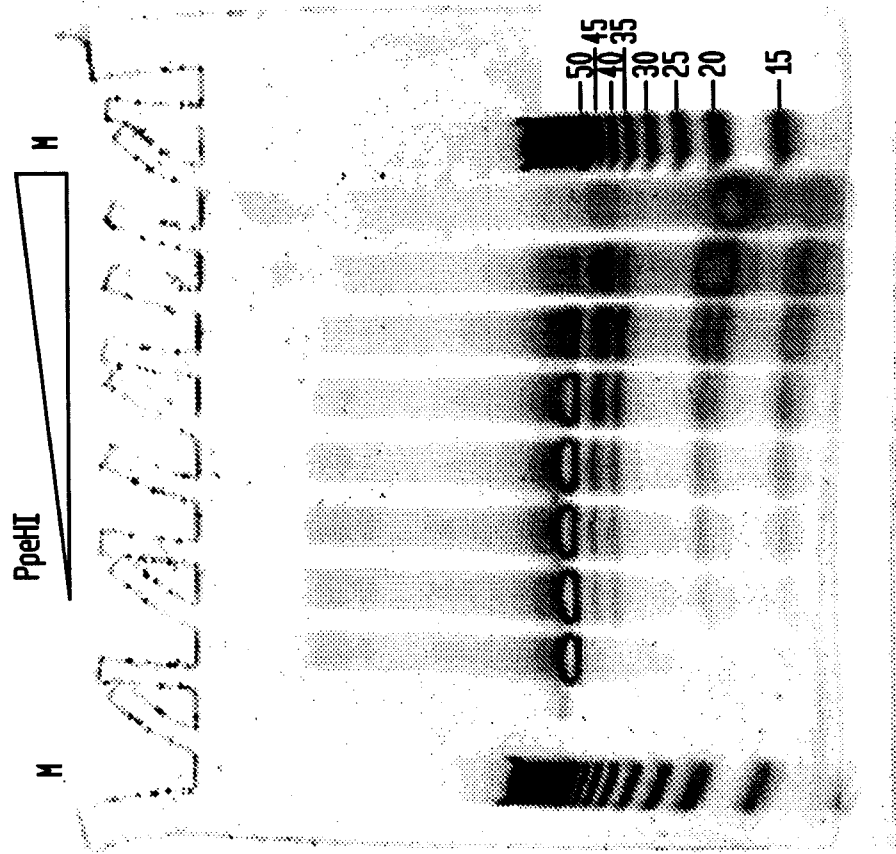


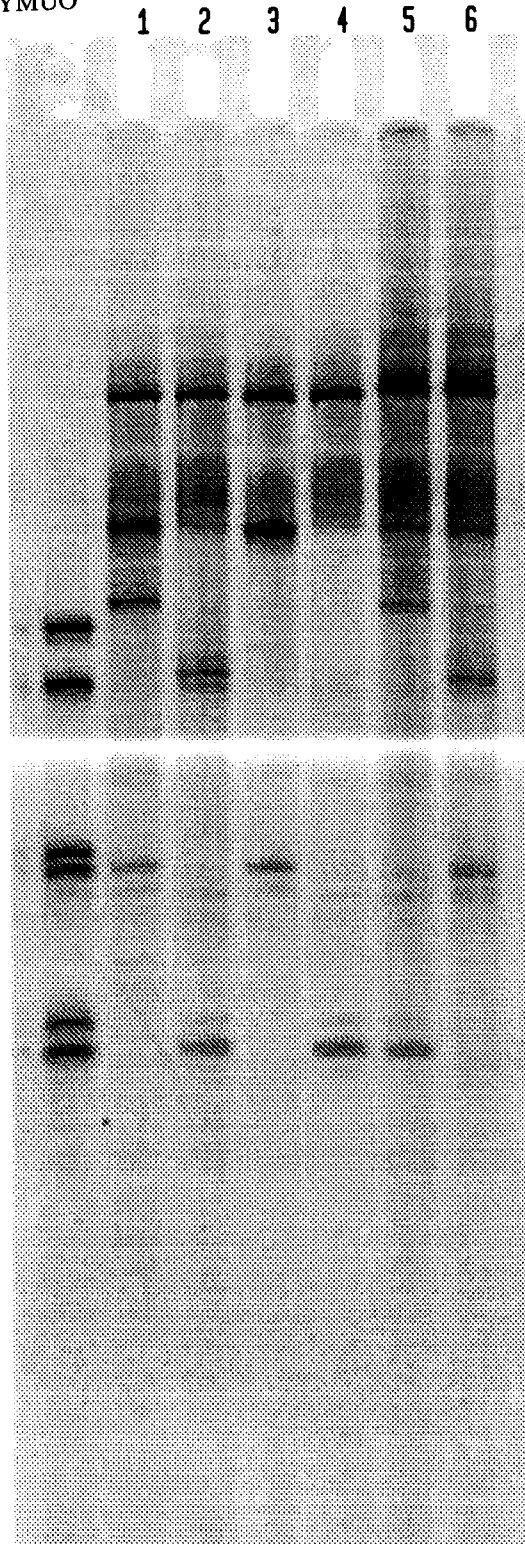
FIG. 6A



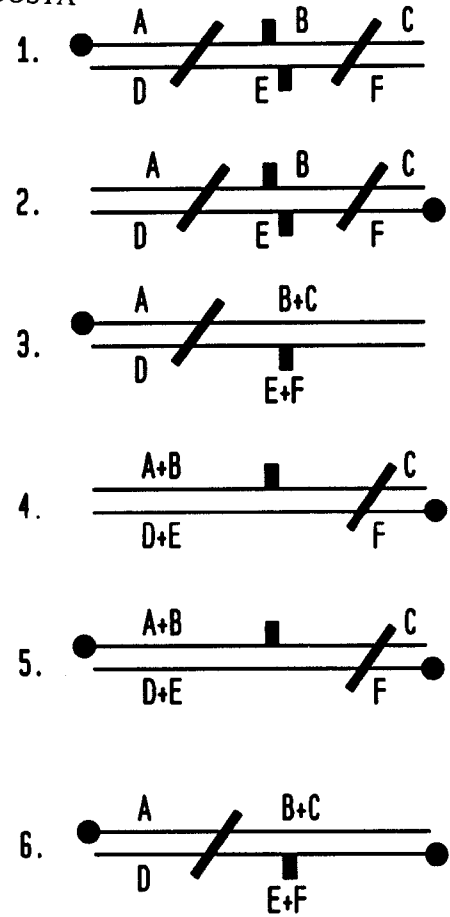
7/18

FIG. 7

ŽYMUO



JUOSTA



ŽYMUO

1 2 3 4 5 6

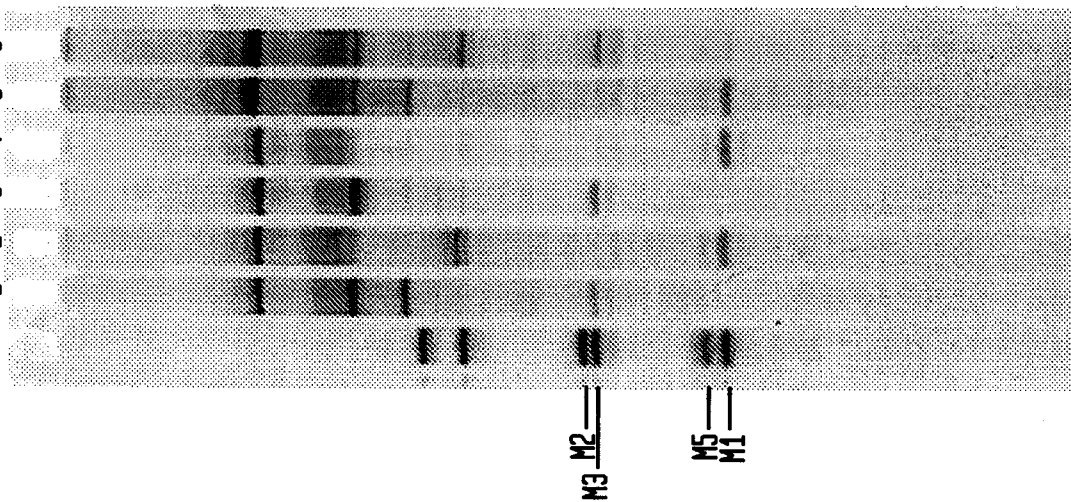
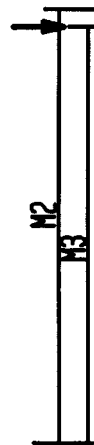


FIG. 8A

JUOSTA

3.



(SEQ ID NO: 2)

(FAM) CCGCGTTTCCGGGTTCCATAGGCTCCGCCACGTACTCTGATGACCAAGGCATCACA  
GCCGCAAGGCCCAAGGTATCCGAGGCGGTGCATGAGACTACTGGTCCCGTAGTGT

JUOSTA

4.

hm

(SEQ ID NO: 2)

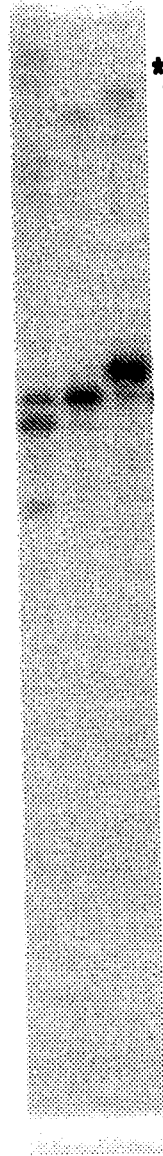
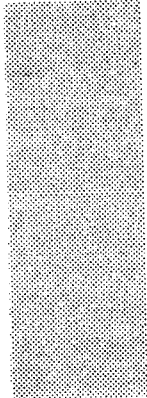
CGGCGTTTCCGGGTTCCATAGGCTCCGCCACGTACTCTGATGACCAAGGCATCACA  
GCCGCAAGGCCCAAGGTATCCGAGGCGGTGCATGAGACTACTGGTCCCGTAGTGT (FAM)



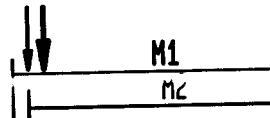
9/18

FIG. 8B

R2 M2 M1



sh



CCATACATATCCCTTACTTCTCCTAACGTGGATGATAAAGGTAGTTTATGTGGAA\* (SEQ ID NO: 3)  
\*AGGTATGTATAGGGAATGAAGAGGATTGCACCTACTATTCCATCAAATACACCT (SEQ ID NO: 4)

10/18

FIG. 8C

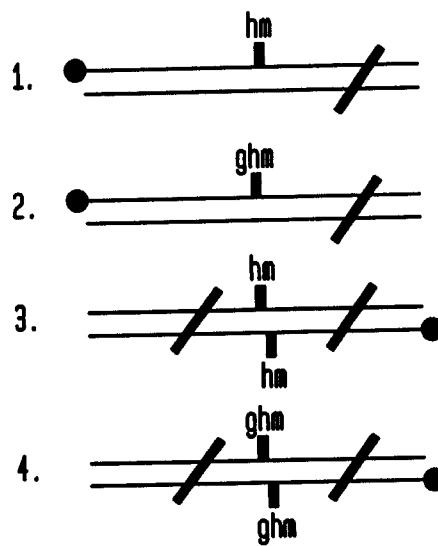
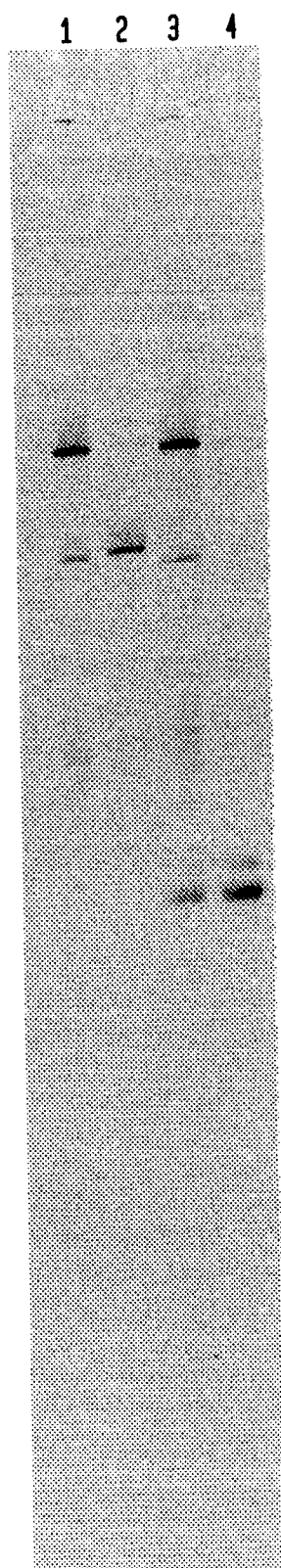


FIG. 9

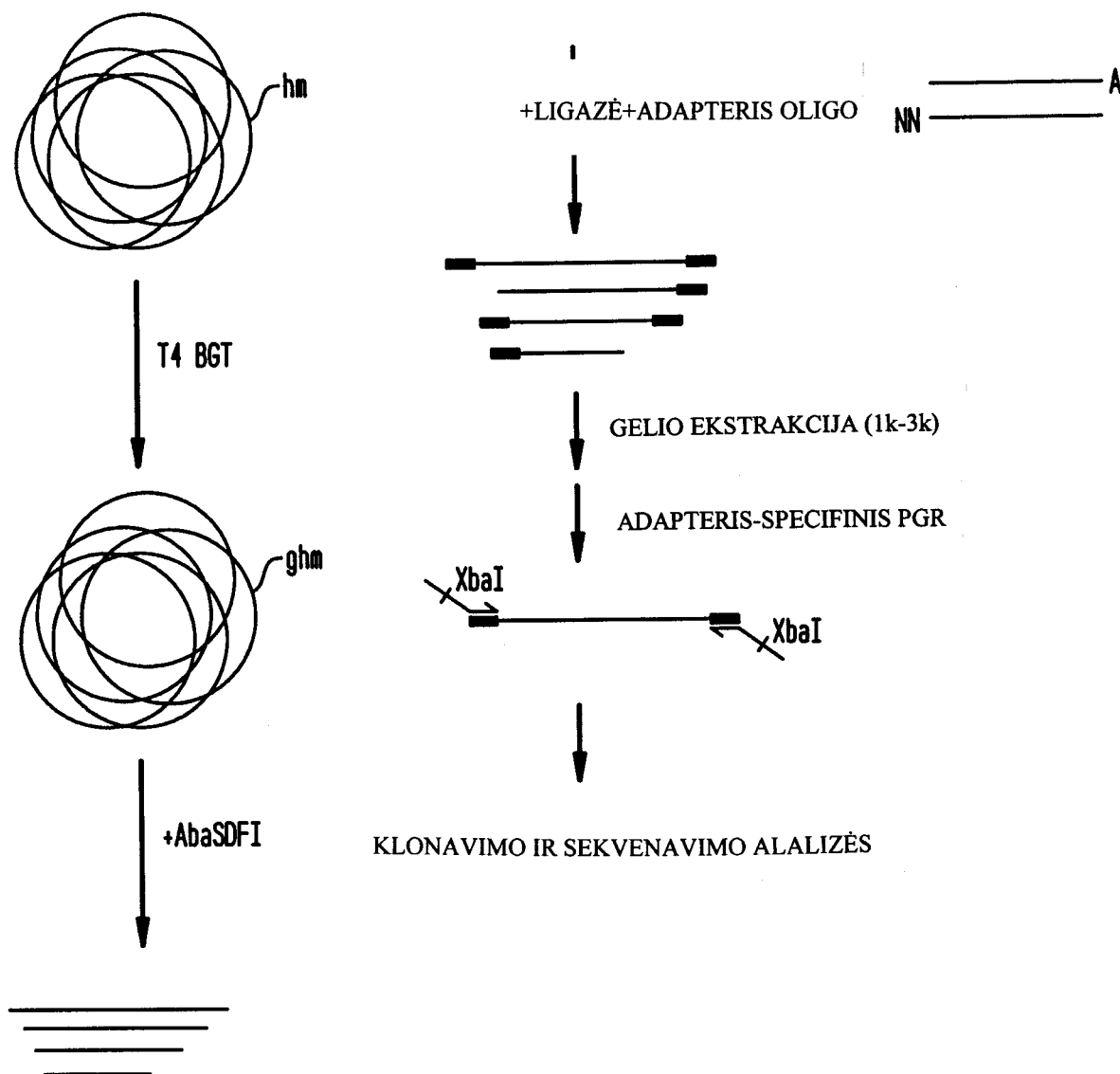


FIG. 10

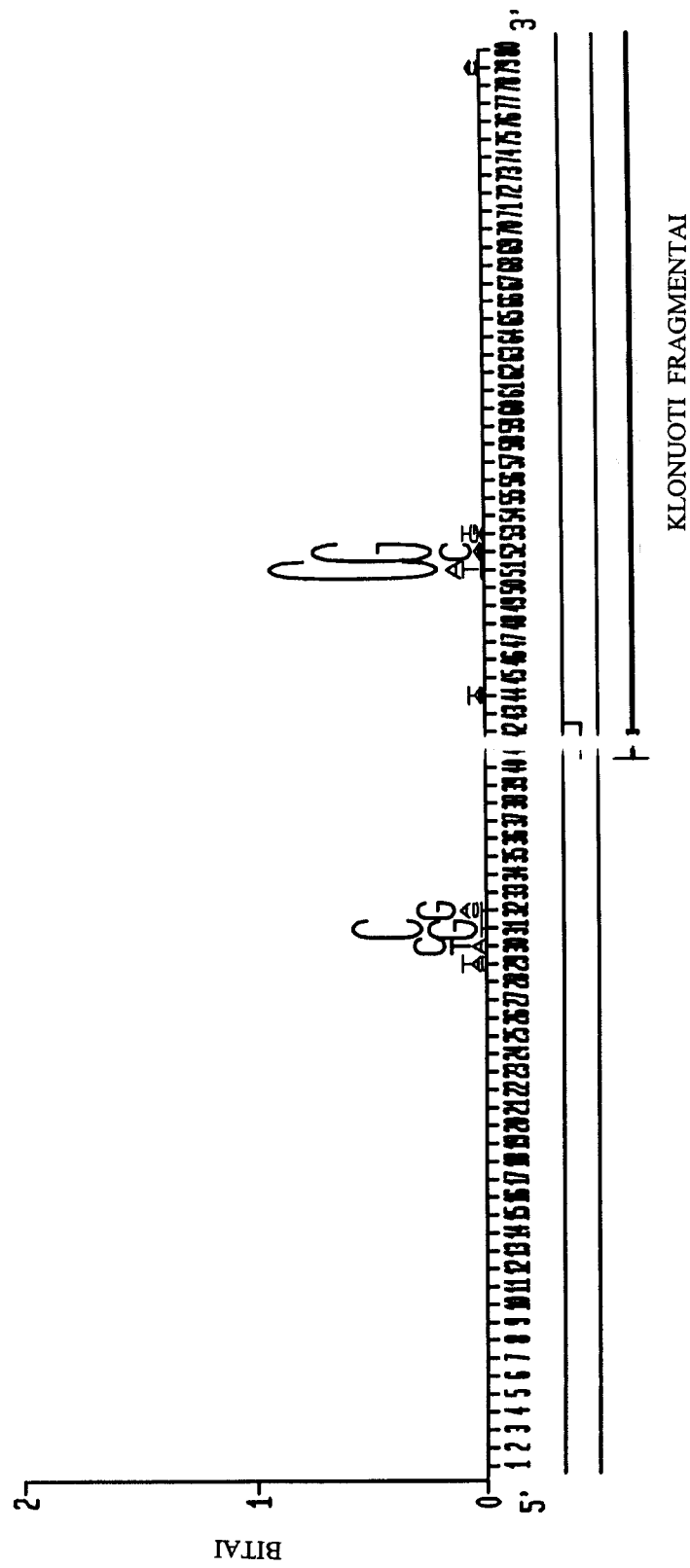


FIG. 11

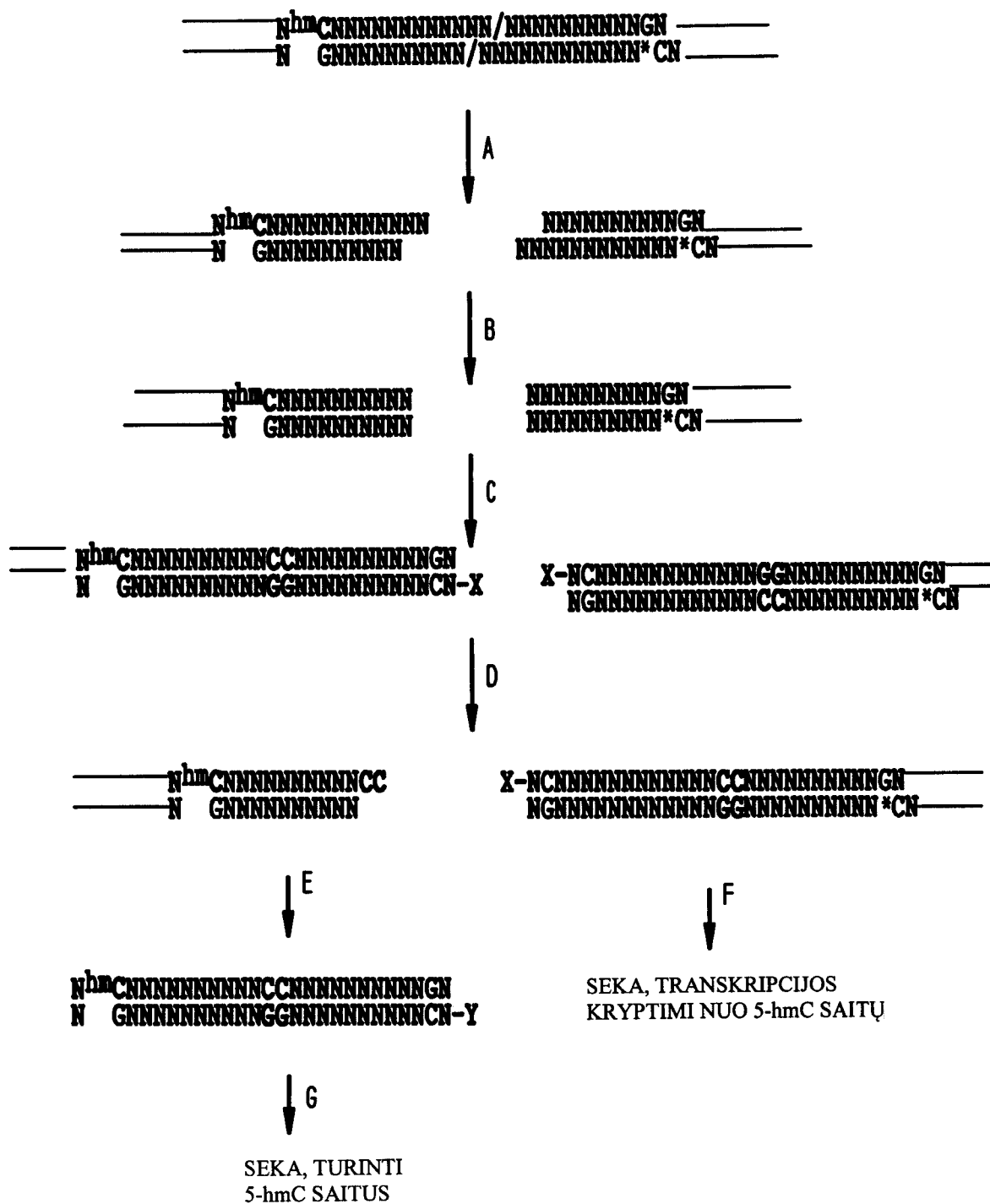
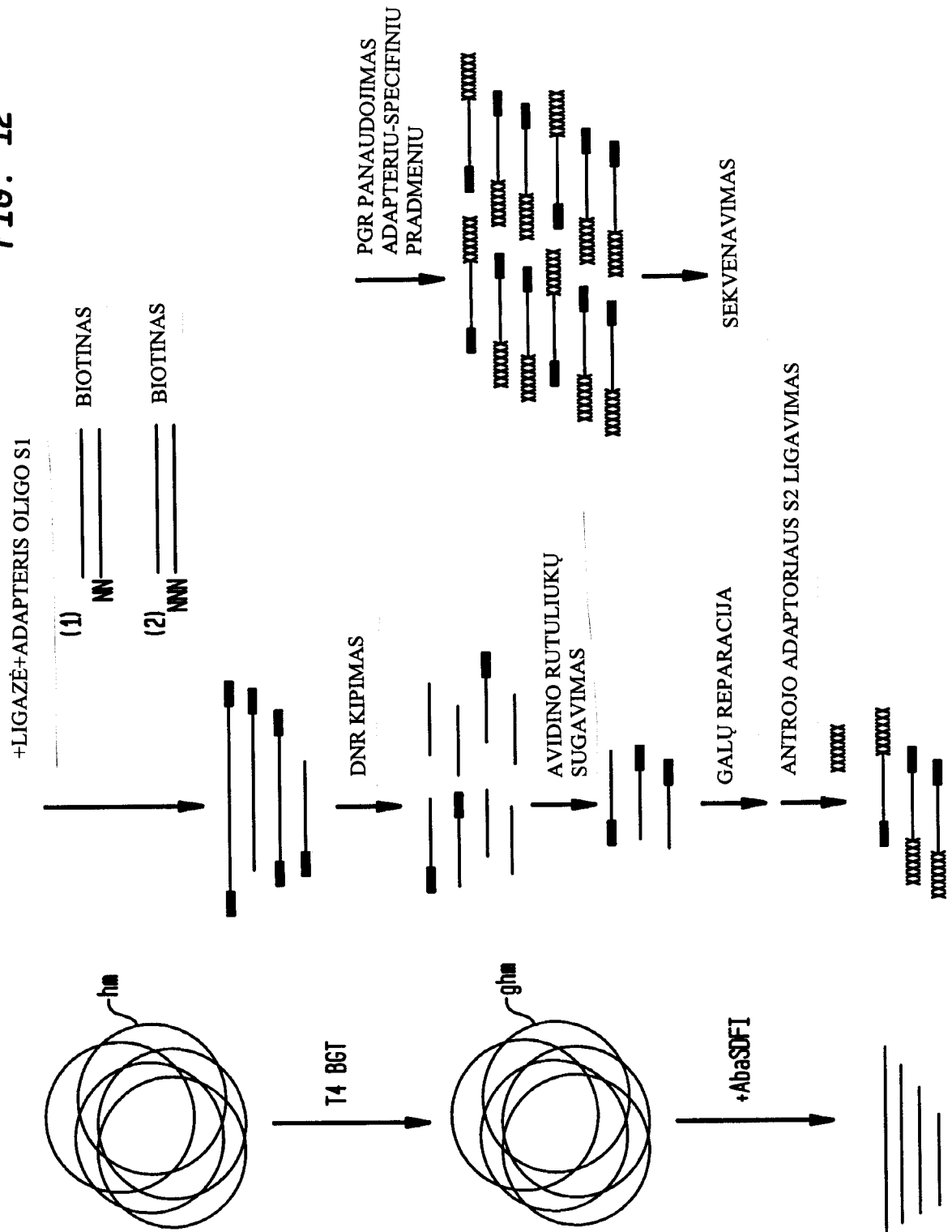


FIG. 12



**Konservatyvumas:**

## Konservatyvios sekos

Konservatyvumas:

|          |    |                           |                          |                  |                  |                |       |     |   |   |   |   |   |  |
|----------|----|---------------------------|--------------------------|------------------|------------------|----------------|-------|-----|---|---|---|---|---|--|
|          | 5  | 5                         | 5                        | 9                | 9                | 9              | 9     | 5   | 5 | 6 | 7 | 7 | 6 |  |
| BaeDI    | 65 | FNLAVEIDEAHHKNDIL         | -----                    | QTLDEIRKNDIVNA   | -----            | LDCEPIRIDATQ   | ----- | 107 |   |   |   |   |   |  |
| BbIDI    | 64 | AMIGVCEDEAYHQOQ           | -----                    | RERDRERETITDVLRO | IRGEDYRALHVDVSG  | -----          | 110   |     |   |   |   |   |   |  |
| AbaSDFI  | 67 | LGIIHIEVDEGHHFLRHSMEYSLNQ | IDEPLYSISQTESDAMREEDILSI | -----            | TGHEKIFRVNVFKNOE | -----          | 130   |     |   |   |   |   |   |  |
| PatTI    | 60 | FAPHIEVDEGQHFTQAM         | -----                    | TEADKIREADIVNA   | -----            | TGHEILRIDVTK   | ----- | 102 |   |   |   |   |   |  |
| YlrI     | 61 | INLHIEIDEFPHLKQ           | -----                    | AEDNMLREADIIDA   | -----            | TGHEVIRISVDG   | ----- | 101 |   |   |   |   |   |  |
| SpeAI    | 59 | LKLHIEIDEFGFHKQO          | -----                    | VDADKVRELDIITA   | -----            | TDHQVKRIDASV   | ----- | 99  |   |   |   |   |   |  |
| EsaNI    | 65 | LKLYLEIDEQGJGSEMI         | -----                    | QTSDLKRADEITLEA  | -----            | TDWTCKRIPVFKAG | ----- | 110 |   |   |   |   |   |  |
| PpeHI    | 63 | LALYLEIDEKHHDSDEA         | -----                    | KKQDAKRRLDIIEA   | -----            | TGFIEKRIPASH   | ----- | 105 |   |   |   |   |   |  |
| PvUrts1I | 63 | LALYLEIDEAHHDSNDA         | -----                    | RKADAVRRLDIVEA   | -----            | TGFQZERIPASH   | ----- | 105 |   |   |   |   |   |  |
|          |    | eeeeee                    | h                        | hhhh             | hhhhhhhhhhhhhhhh | eeeeee         |       |     |   |   |   |   |   |  |

Konservatyvijos sekos

## Konservatyvios sekos

FIG. 13-2

Konservatyvumas:

|                       |     |                                                   |                                                |                                                  |                                 |                              |                              |     |   |   |    |   |    |
|-----------------------|-----|---------------------------------------------------|------------------------------------------------|--------------------------------------------------|---------------------------------|------------------------------|------------------------------|-----|---|---|----|---|----|
| Konservatyvumas:      | 5   | 79                                                | 7                                              | 5                                                | 5                               | 7                            | 97                           | 5   | 5 | 5 | 55 | 6 | 75 |
| BaeDI                 | 108 | -----                                             | SFEKHKKIDQVVEKINL                              | -----                                            | LTKKMFIPNDLEKEYDPTTYIEQ         | -----                        | GYTDADQNVSLRLVADC            | 106 |   |   |    |   |    |
| RbiDI                 | 111 | -----                                             | SYEQVERSIDDVRRIRAEIERRRQAMEFTPH                | -----                                            | TEAYVDYKEFYKTRDAVSVADDVGEFRIADA | -----                        |                              | 103 |   |   |    |   |    |
| AbaSDFI               | 131 | GQ                                                | -----                                          | PQMLENIHQIDKILKEIKTANKKLIRASTFKAMNILETEYNPQTYIDL | -----                           | GRISLADNVVLAKTKOV            | 107                          |     |   |   |    |   |    |
| PatFI                 | 103 | -----                                             | SFDDINTQIDAAVVKIKSRQEI                         | -----                                            | SFIPMDIDSEFDSATYIKR             | -----                        | GYIDIKONVAFKTIKA             | 101 |   |   |    |   |    |
| YkrI                  | 102 | -----                                             | SLRQMERIDDCVAAIKSKISALGO                       | -----                                            | CTEPMDNDKELSLIEPHIR             | -----                        | GYIDVKNVAFRRITDA             | 102 |   |   |    |   |    |
| SpeAI                 | 100 | -----                                             | ALTEQINLQVDQIVAEILQSVVEVQKAGHFQANDPETHSVSYLQSR | -----                                            | GVIRASENVALRTSAEV               | -----                        |                              | 102 |   |   |    |   |    |
| EsaNI                 | 111 | SSKIDKSLEALKEIDDFVSVEQKQKQNVSAGHKISNNYEQKSPDFFIAK | -----                                          | EKIKVSDNVALLNHRDV                                | -----                           |                              |                              | 103 |   |   |    |   |    |
| PpeHI                 | 106 | -----                                             | VTIEQLTTSIDEFVKLLIDTKKEQKAKKFI                 | -----                                            | PNDYSAQYTAKRHIDA                | -----                        | GFIEVGPHALFRYHRDA            | 103 |   |   |    |   |    |
| Pvurt.sII             | 106 | -----                                             | YTLSEVNKLVDSEFVRLVADKKEELEHQGLF                | -----                                            | RNDYDERYSAKKHINT                | -----                        | GTMAVGPMSVFRYHRDA            | 103 |   |   |    |   |    |
| Konservatyvijos sekos |     | hhhhhhhhhhhhhhhhhhhhhhhhhhhh                      | hhhhhhhhhhhhhhhhhhhhhhhhhhhh                   | hhhhhhhhhhhhhhhhhhhhhhhhhhhh                     | hhhhhhhhhhhhhhhhhhhhhhhhhhhh    | hhhhhhhhhhhhhhhhhhhhhhhhhhhh | hhhhhhhhhhhhhhhhhhhhhhhhhhhh |     |   |   |    |   |    |

Konservatyvijos sekos

Konservatyvumas:

|                         |     |                                                 |       |                                        |       |                      |        |      |      |      |      |      |      |      |
|-------------------------|-----|-------------------------------------------------|-------|----------------------------------------|-------|----------------------|--------|------|------|------|------|------|------|------|
| Konservatyvumas:        | 5   | 77                                              | 7     | 55                                     | 6     | 7                    | 575775 | 5    | 9    | 9    | 55   | 55   | 7    | 7    |
| BaeDI                   | 167 | CHVFGAGYANGIQKSG-APHKFE                         | ----- | EDTDIKRIAKFFPMEI                       | ----- | NNQJLENEEIEFYNTI-    |        |      |      |      |      |      |      | 201  |
| RbiDI                   | 174 | VNTLCSEYKRPQES--NFVPSVMRWYIG                    | ----- | DRYRVWFPKLAIGGKAVANGNHRISLDDGTYIYEYNEO | ----- |                      |        |      |      |      |      |      |      | 203  |
| AbaSDFI                 | 198 | CNCFGYSY-KNYQCG-ALHPYK                          | ----- | KDTLLNFPRLYENKO                        | ----- | NITITSPDGLTYTEKSTD   |        |      |      |      |      |      |      | 201  |
| PatTI                   | 162 | CNCFGHHY-TGYQAG-AAHPD                           | ----- | HYTNLWFPKLPNGE                         | ----- | NDSQISSDEEVITERNEO   |        |      |      |      |      |      |      | 201  |
| YkrI                    | 163 | CNCFGHHY-KFLOKAG-AKHPYH                         | ----- | DDILLNLPLKLPMEH                        | ----- | NMQISNDENVITETPKS-   |        |      |      |      |      |      |      | 205  |
| SpeAI                   | 163 | CHLLGHHY-KGNQSS-ASVPHY                          | ----- | PHIRLWFPKLPMEQ                         | ----- | NFHLISYDGCIEIHEYCTE- |        |      |      |      |      |      |      | 205  |
| EsaNI                   | 180 | LRLEFYAK-CHYORAV-NTIKKT                         | ----- | NQAVNFPKLPYPSD                         | ----- | NVNSFDQKSGYIEQFRKON  |        |      |      |      |      |      |      | 203  |
| PpeHI                   | 170 | LECFTGYIN-KGHHQSGSNKLPENIVREIGLSGRILAVFPRLYNAGV | ----- | NNHLSPDGEMITERSLE-                     | ----- |                      |        |      |      |      |      |      |      | 202  |
| PvurTsII                | 170 | LQCFGTNR-KGHHQSGGNALPAEVAQSIGLTRGVNWFPRLYEAGE   | ----- | NKHALSADGKHITQSLN-                     | ----- |                      |        |      |      |      |      |      |      | 202  |
|                         |     | hhhh                                            | hhhh  | hhhh                                   | hhhh  | hhhh                 | hhhh   | hhhh | hhhh | hhhh | hhhh | hhhh | hhhh | hhhh |
| Konservatyvumas salton: |     |                                                 |       |                                        |       |                      |        |      |      |      |      |      |      |      |

Konservatyvijos sekos

FIG. 13-3

Konservatyvumas:

|           |     |                       |                                   |                      |     |
|-----------|-----|-----------------------|-----------------------------------|----------------------|-----|
| BaeDI     | 222 | PEEN--EAYFQKRN---     | YQIN-QKIALFAVAKTSSG-RFAIFKGLYLN   | VEKSNITGLTYHRISTIMP  | 284 |
| Ebidi     | 239 | ADLVD-PVEDG-----      | DEPD-IRI-TFAKSADPVTRIQAIRFVGVPRI  | INSEDC-TRKRYQRIETVFP | 298 |
| AbaSD: I  | 252 | ETITL-AKLEENK-----    | NGPQ-KRI-VFAVKDMLSSRAVRFNGLYEFC   | ADLAD--CAVMKRVACEVQ  | 312 |
| PatTI     | 215 | DKKAK-QVSSSHVNKKHKKH- | QRI-VFAVKGHLG-DVLYRFRGQYQLL       | KDSNEKKTGLIMRIKTRVK  | 280 |
| YkrI      | 217 | EDAQA-AHFDKMA-----    | ETRN-KRL-VFAKAKDHLG-MLYRFKGLYELN  | KESNRTIGLYWQRISTRVK  | 279 |
| SpeAI     | 217 | SETKKRQFDKML-----     | SENI-QQK-VFAVKDELG-QVMYRFKGLFILL  | DKTCHESGVYMKRVATEFE  | 279 |
| EsaNI     | 234 | QPEPM-----P-----      | KEGDPRI-VFAKQKNIFC-QTVYKPFGLFRAL  | NKTDPVHH-YFKRINTCLD  | 289 |
| PpeHI     | 233 | VDMNY-----I-----      | EDWD-YRI-VMAHSRDELM-RVLYRFLGVFQIC | NKSVEGKN-IFKRINTKVK  | 287 |
| PvURt: 1I | 233 | ATRN Y-----Q-----     | ETWD-YRI-VMAHSRDELM-RVLYRFLGVFALC | DKSSDEVK-VFSRVYSKVN  | 287 |
|           |     | hhhhh hhhhhh          | eee eeeee hhh hhhhhh eeeee        | eeeeeeee ee          |     |

Konservatyvijos sekos

Konservatyvumas:

|           |     |                                               |     |                 |
|-----------|-----|-----------------------------------------------|-----|-----------------|
| BaeDI     | 285 | TYTP-KDVQPLRIAKAYNDEYKVAHYFTEHQVRKTEGKRYKIISY | 334 | (SEQ ID NO: 5)  |
| Ebidi     | 299 | IHRTPC-----LP IHR-----                        | 309 | (SEQ ID NO: 6)  |
| AbaSD: I  | 313 | TYSPK-----ETKC-----                           | 321 | (SEQ ID NO: 7)  |
| PatTI     | 281 | TYEQN-----                                    | 285 | (SEQ ID NO: 8)  |
| YkrI      | 280 | TYPSPA-----RNPD-----                          | 289 | (SEQ ID NO: 9)  |
| SpeAI     | 280 | LGLT-----                                     | 283 | (SEQ ID NO: 10) |
| EsaNI     | 290 | LSRYSA-----N-----                             | 296 | (SEQ ID NO: 11) |
| PpeHI     | 288 | VFNSTH-----                                   | 293 | (SEQ ID NO: 12) |
| PvURt: 1I | 288 | VYRSON-----                                   | 293 | (SEQ ID NO: 13) |
|           |     | ee                                            |     |                 |

Konservatyvijos sekos

**FIG. 14**