

Opinnäytetyö (AMK)

Bio- ja elintarviketekniikka

Biotekniikka

2012

Anna-Kaisa Kartesalo

# INFLUENSSA B-VIRUKSEN NUKLEOPROTEIINIA TUOTTAVAN *PICHIA PASTORIS* -KANNAN LUOMINEN



TURUN AMMATTIKORKEAKOULU  
TURKU UNIVERSITY OF APPLIED SCIENCES

OPINNÄYTETYÖ (AMK) | TIIVISTELMÄ

TURUN AMMATTIKORKEAKOULU

Bio- ja elintarviketekniikka | Biotekniikka

2012 | 52

Kaj Sjöblom FM, Jarno Pusa Ins., Petri Susi FT

Ilari Suominen, yliopettaja, FT

Anna-Kaisa Kartesalo

## ***INFLUENSSA B-VIRUKSEN NUKLEOPROTEIINIA TUOTTAVAN PICHIA PASTORIS-KANNAN LUOMINEN***

Tämän opinnäytetyön tarkoituksena oli kehittää geeniteknisin menetelmin *Pichia pastoris*-hiivakanta, joka tuottaa influenssa B-viruksen nukleoproteiinia (InBNP) fuusioituna merkkiproteiiniin. Erilaisia geenikonstrukteja tehtiin yhteensä kuusi kappaletta, joiden erona oli erilainen promoottori tai merkkiproteiini.

Työ tehtiin HydroBody-projektissa, joka on Turun AMK:n ja VTT:n yhteistyöprojekti ja jonka tarkoituksena on kehitellä uusia tuotokonsepteja teollisesti arvokkaille proteiineille (<http://hydrobody.turkuamk.fi/>). Projektin nimi HydroBody tulee sanasta hydrofobiini, joka on alun perin rihmasienten tuottama monipuolinen pieni proteiini. Hydrofobiinin ainutlaatuisen ominaisuuksien vuoksi se voidaan konsentroida surfaktantti-pohjaisessa kaksifaasiuutossa faasista toiseen. Tällöin myös kohdeproteiinin puhdistus helpottuu käytettäessä hydrofobiinia merkkiproteiinina proteiinintuotannossa.

Erilaisia *Pichia pastoris*-kannan transformoitavia vektoreita on kaksi kappaletta, PICZ A ja GAPZ A. PICZ A on metanoli-indusoitava proteiinituotannon aloittamiseksi ja GAPZ A:ssa proteiinin tuotto on aina päällä johtuen erilaisista promoottoreista. Erilaisia proteiinia tuottavia geenejä valmistettiin kolme kappaletta, joista kaikki tuottavat InBNP:tä, vain merkkiproteiini muuttuu. Kaikki geenit ilmentettiin molempien promoottorien yhteydessä, jolloin lopulta päästiin kuuteen erilaiseen InBNP:tä tuottavaan kantaan.

Lopuksi *Pichia pastoris*-kannat testattiin InBNP-tuottotasojen selvittämiseksi yhteistyössä ArcDian kanssa käyttäen MariPOC-testimenetelmää. Mittauksissa todettiin tuotokantojen tuottavan nukleoproteiinia muodossa, jonka laitteisto tunnistaa sekä pystyttiin alustavasti seulomaan paremmin tuottavat kannat huonommista.

ASIASANAT:

*Pichia pastoris, influenssa b viruksen nukleoproteiini, hydrofobiini, geenitekniikka*

BACHELOR'S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Bio- and food technology | Biotechnology

2012 | 52

Kaj Sjöblom M. Sc., Jarmo Pusa B. Eng., Petri Susi Ph. D.

Ilari Suominen, Senior Lecturer, Ph. D.

Anna-Kaisa Kartesalo

## ***CREATION OF INFLUENZA B VIRUS NUCLEOPROTEIN PRODUCING PICHIA PASTORIS STRAIN***

The aim of this Bachelor Thesis was to create a *Pichia pastoris* yeast strain that produces the nucleoprotein of Influenza B virus (InBNP) combined to a tag protein. Six different versions of the InBNP producing construct were created, each differing from the others by a different tag protein or vector in which the construct was introduced to *Pichia*.

The practical part of the thesis was performed as part of the HydroBody project, which is a joint project between Turku University of Applied Sciences and VTT, in the protein production laboratories at TUAS. The project aims to develop new production concepts for industrially valuable proteins. The project name HydroBody, comes from the word hydrophobin which is a small protein produced by filamentous fungi. As a result of the versatile and exceptional properties of the protein it can be concentrated from one phase to another by surfactant-based two-phase extraction thus facilitating the purification of the target fusion protein, InBNP.

*Pichia pastoris* was transformed by two different vectors, PICZ A and GAPZ A. PICZ A has to be induced by methanol to initiate protein production whereas the production in GAPZ A is always on resulting from different promoters integrated into the vector. Three different protein-coding genes were prepared all of them producing InBNP but varying in tag protein. All genes were expressed combined to both vectors resulting in six different InBNP-producing vectors.

Finally the *Pichia* strains created were tested for their level of InBNP production in cooperation with ArcDia International using their MariPOC testing method. In the measurements strains were proven to produce InBNP in a form recognized by the testing device and initial screening of strains could be performed based on the gathered data.

KEYWORDS:

*Pichia pastoris, Influenza B virus nucleoprotein, hydrophobin, gene technology*

# SISÄLTÖ

<b>LYHENTEET JA SANASTO</b>	<b>7</b>
<b>1 JOHDANTO</b>	<b>9</b>
<b>2 <i>PICHA PASTORIS</i> JA TUOTTOSYSTEEMI</b>	<b>11</b>
<b>3 INFLUENSSA B-VIRUS</b>	<b>15</b>
<b>4 HYDROFOBIINIT MERKKIPROTEIININA</b>	<b>17</b>
4.1 Hydrofobiinien käyttö geenitekniikan sovelluksissa	18
<b>5 SEKVENSSEIEN LUOMINEN JA OPTIMOINTI</b>	<b>20</b>
5.1 PlasmaDNA	20
5.2 Influenssa B viruksen nukleoproteiinia koodaava geenisekvenssi	21
5.3 Hydrofobiini 3:n, hydrofobiini 4:n ja Tag6:n sekvenssien luominen	23
5.4 Sekvenssien optimointi	26
<b>6 VEKTORIKONSTRUKTIEN LUOMINEN <i>E. COLISSA</i></b>	<b>29</b>
6.1 <i>Escherichia colin</i> transformaatio plasmidistokista, kloonaus ja seulonta	29
6.2 Kasvatus Midiprep-plasmidin geelieristystä varten	32
6.3 Ligaatiot	35
6.4 <i>Escherichia colin</i> transformaatio vektorilla	39
<b>7 <i>PICHA PASTORIS</i> X33-KANNAN TRANSFORMOINTI</b>	<b>41</b>
<b>8 <i>PICHA PASTORIS</i>-TUOTTOKANTOJEN SEULONTA</b>	<b>44</b>
<b>9 TYÖN ARVIOINTI, TULOKSET JA PÄÄTELMÄT</b>	<b>46</b>
9.1 Työ	46
9.2 ArcDian MariPOC influenssa B-virus-testin tulokset	48
9.3 Päätelmät	49
9.4 Lopuksi	50
<b>LÄHTEET</b>	<b>51</b>

## KUVAT

Kuva 1. Työssä käytettyjen vektorien rakenne. ....	13
Kuva 2. Hydrofobiinin rakenne. Vihreä alue kuvaa hydrofobisen saarekkeen sijaintia molekyylissä. Molekyylin N-pää ja C-pää on merkattu muuhun rakenteeseen. (Linder <i>et al.</i> , 2004) .....	17
Kuva 3. Kozakin sijainti aloituskodoniin nähden. Kunkin kirjaimen koko viittaa kyseisen emäksen tärkeyteen voimakkaan Kozak-sekvenssin muodostumisessa. (Amiruddin <i>et al.</i> , 2012) .....	22
Kuva 4. HFBI_3-hydrofobiinin sekvenssin rakenne ja siihen liitetyt osiot. Geenin koko on 315 bp.....	24
Kuva 5. HFBI_4:n b(vasemmalla) ja Tag6:n (oikealla) sekvenssien rakenteet. HFBI_4-geenin koko on 360 bp ja Tag6:n 332 bp. ....	24
Kuva 6. PlasmaDNA:ssa virtuaalisesti digestoitu Tag6-geeni. Kohesiiviset päät näkyvät värillisinä geenin päissä.....	25
Kuva 7. Valittujen pesäkkeiden (kloonien) maljauskuvio. ....	30
Kuva 8. Digestoimaton InBNP pUC57-plasmidissa. ....	31
Kuva 9. HFBI_3:n digestiot kaivoissa 6-7 geelillä 1 sekä geelillä 2 kaivoissa 2 ja 3. HFBI_4:n digestiot geelillä 2 kaivoissa 4-7.....	32
Kuva 10. pUC57-Tag6:n digestiot, kloonit 1.-4. kaivoissa 2.-5. ....	32
Kuva 11. InBNP-geeni digestoituna. Kaivoissa 1. ja 2. on yhteensä jaettuna digestionäyte. 3. ja 6. kaivo on jätetty tyhjäksi, 5. kaivossa 1 µl digestionäytettä, 7.:ssa katkaisematon InBNP+pUC57 ja 4. sekä 8. kaivossa GeneRuler 1 kb.....	33
Kuva 12. 1. & 2. kaivossa on HFBI_3:n digestiot, 3. katkaisematon näyte, 4. GeneRuler 1 kb, 5. HFBI_4:n katkaisematon näyte, 6. & 7. kaivossa HFBI_4:n digestiot sekä 8. kaivossa GeneRuler 1 kb.....	34
Kuva 13. 1. ja 4. kaivossa GeneRuler 1 kb, 2. ja 3. Tag6:n digestiot, 5. katkaisematon näyte ja 6. 1,5 µl:n näyte digestiosta. ....	34
Kuva 14. PICZ 1. ja PICZ 2. plasmidivektorit. ....	37
Kuva 15. GAPZ 1. ja GAPZ 2. plasmidivektorit. ....	37
Kuva 16. GAPZ 3. ja PICZ 3. plasmidivektorit. ....	38
Kuva 17. PICZ 1., PICZ 2., GAPZ 1. ja GAPZ 2. nimettyjen vektorien seulontadigestiot AGE-geelillä.....	39
Kuva 18. PICZ 3. ja GAPZ 3. seulontadigestiot.....	40
Kuva 19. pGAPZ A-vektorien linearisointi tehtiin AvrII-entsyymillä. ....	41
Kuva 20. pPICZ A-vektorien linearsiointissa käytettiin SacI-entsyymiä. ....	42
Kuva 21. Jcat sekvenssinmuokkaussivuston aloitussivu. ....	58
Kuva 22. Optimoitavan sekvenssin soveltuvuus <i>Saccharomyces cerevisiaen</i> kodoninkäyttöön ennen kodonioptimointia. ....	59
Kuva 23. Sekvenssin soveltuvuus <i>S. cerevisiaen</i> kodoninkäyttöön optimoinnin jälkeen. ....	59
Kuva 24. NEBcutter.sekvenssinanalysointi .....	60
Kuva 25. NEBcutterista saatava analysoidun sekvenssin restriktioentsyymien tunnistuskohta-kartta. ....	61
Kuva 26. NEBcutterissa entsyymejä voi seuloa sekvenssistä löytyvien restriktioentsyymien tunnistus-/katkaisukohtien määrän mukaan. ....	62
Kuva 27. Ligaatiolaskuri.....	63
Kuva 28. MariPOC-InB- määritysten data. ....	65

## TAULUKOT

Taulukko 1. <i>Pichia pastoris</i> -hiivassa tuotettujen rekombinanttiproteiinien tuottotasoja. (Macauley-Patrick <i>et al.</i> , 2005) .....	14
Taulukko 2. Geenien saannot ja liuosten konsentraatiot geelieristyksen jälkeen.....	35
Taulukko 3. Vektorit. ....	36
Taulukko 4. <i>E. colista</i> eristettyjen vektorien saanto ja pitoisuus. ....	40
Taulukko 5. Linearisoitujen vektorien pitoisuudet ja saannot.....	42

## KUVAAJAT

Kuvaaja 1. InBNP+HFBI_3-fuusioproteiinia tuottavien vektorien tuottovertailu.....	66
Kuvaaja 2. InBNP+HFBI_4-fuusioproteiinia tuottavien vektorien tuottovertailu.....	66

## LIITTEET

Liite 1. Elektrokompenttien solujen valmistus ja mediumit	
Liite 2. Työohjeet	
Liite 3. Projektin sekvenssipohja uusille inserteille	
Liite 4 Sekvenssien optimointi ja analysointi	
Liite 5. Ligaatiolaskuri	
Liite 6. Näytteiden purkukoodi	
Liite 7. MariPOC Influenssa B-virus-testin tulokset	

# LYHENTEET JA SANASTO

## Lyhenteet

InB	Influenssa B-virus
InBNP	Influenssa B-viruksen nukleoproteiinia koodaava geeni
LS-LB-medium	low salt LB-medium, ks liitteet; mediumien valmistus

## Sanasto

amfifiili/amfipaattinen	molekyyli, jolla on sekä hydrofobinen että -fiilinen alue
aggregoitua	proteiinien kyky muodostaa järjestelmällinen rakenne tai pinta liittymällä yhteen
elektroporaatio	transformointikeino, jännitteen avulla soluseinämä saatetaan läpäisevään muotoon, jolloin vierasperäinen geeni saadaan solun sisään osaksi sen genomia
geenikasetti	geenitekniikan keinoja käyttäen luotu DNA-jakso, joka voidaan integroida osaksi kohdeorganismien genomia
konstitutiivinen	proteiinin tuotanto riippuu promoottorista, jota ei induoida vaan se on aina aktiivinen
metanoli-indusoitava	proteiinin tuotanto aloitetaan lisäämällä kasvatusmediumiin metanolia
metylotrofi	metanolia hiililähteenä käyttävä organismi
plasmidivektori	ympyränmuotoinen DNA-molekyyli, joka voi monistua vieraassa organismissa
post-translacionaalinen	proteiinisynteesin jälkeinen
promoottori	DNA:n alue, johon RNA-polymeraasi kiinnittyy ja aloittaa transkription

resistenssigeeni	geeni, jonka avulla organismi saa vastustuskyvyn antibiootille (mm. ampisilliini, zeosiini)
restriktioentsyymi	endonukleaasi, DNA:ta katkova entsyymi
retentiosignaali	aminohapposekvenssi, joka toimii signaalina solulle pidättää proteiini solun sisällä sen ulos erittämisen sijaan
surfaktantti	pinta-aktiivinen aine
transformointi	vieraan genomin integrointi osaksi organismin genomia



# 1 JOHDANTO

Rekombinanttiproteiinien tuottoa hiivassa on tutkittu useita vuosia ja vuosien saatossa metylootrofisten hiivojen käyttöön on keskitytty niiden suotuisten ominaisuuksien vuoksi. Rekombinanttiproteiinien tuotto metylootrofisissa eli metanolia hiililähteenä käyttävissä hiivakannoissa linkittyy alkoholioksidaasia koodaavaan promoottoriin (AOX1), jonka tuotto hiivassa on korkealla tasolla kyseisen entsyymin hapettaessa metanolia formaldehydiksi. Kyseisen promoottorin aktiivisuuden avulla hiivassa voidaan tuottaa rekombinanttiproteiineja säädellysti. (Potvin *et al.*, 2010), (Macauley-Patrick *et al.*, 2005)

HydroBody-projekti on Turun AMK:n ja Espoon VTT:n yhteistyöprojekti, jonka tavoitteena on kehittää uusia tuottokonsepteja erilaisille proteiineille *Pichia pastoris*-hiivassa ja *Trichoderma reesei*-rihmasienessä. Turun AMK:n projektiryhmässä käytettiin hyödyksi kahta erilaista *Pichia pastoris*-proteiinintuottosysteemiä (PICZ A ja GAPZ A, Invitrogen). (<http://hydrobody.turkuamk.fi/>)

Biotekniikan alalla on yleistietoa, että rekombinanttiproteiinin tuotannossa aikaa ja rahaa vievin osuus on proteiinin puhdistus lopputuotteen vaaditulle puhtaustasolle. Tästä syystä, projektin nimen mukaan, tuottoproteiiniin liitetään hydrofobiini-proteiini, jolloin proteiinin puhdistukseen saadaan uusia mahdollisuuksia surfaktantti-vesi kaksifaasiuutolla, jolla puhdistus helpottuisi ja sen kuluja saataisiin alennettua. (Linder *et al.*, 2004)

Tässä opinnäytetyössä tuottoproteiininä on influenssa B-viruksen nukleoproteiini (InBNP), jota voidaan puhdistettuna käyttää virusdiagnoosissa positiivisena kontrollina. Erilaisia *Pichia*-tuottokantoja InBNP:lle rakennettiin yhteensä kuusi kappaletta.

Työ aloitettiin geenien sekvenssien optimoinnilla ja synteettiset geenit tilattiin Genscriptiltä. Geenien kokoaminen vektoriin tehtiin *Escherichia coli*-bakteerissa, jonka jälkeen valmiit tuottovektorit siirrettiin elektroporaatiolla *Pichia pastoris*-kannoille. Kymmenillä *Pichia pastoris*-kannoilla tehtiin testikasvatuk-

set, joiden tuottotasoja mitattiin yhteistyössä ArcDia International-yrityksen kanssa MariPOC-mittauslaitteella siihen kehitetyllä Influenssa B-virus-testillä.

## 2 *PICHIA PASTORIS* JA TUOTTOSYSTEEMI

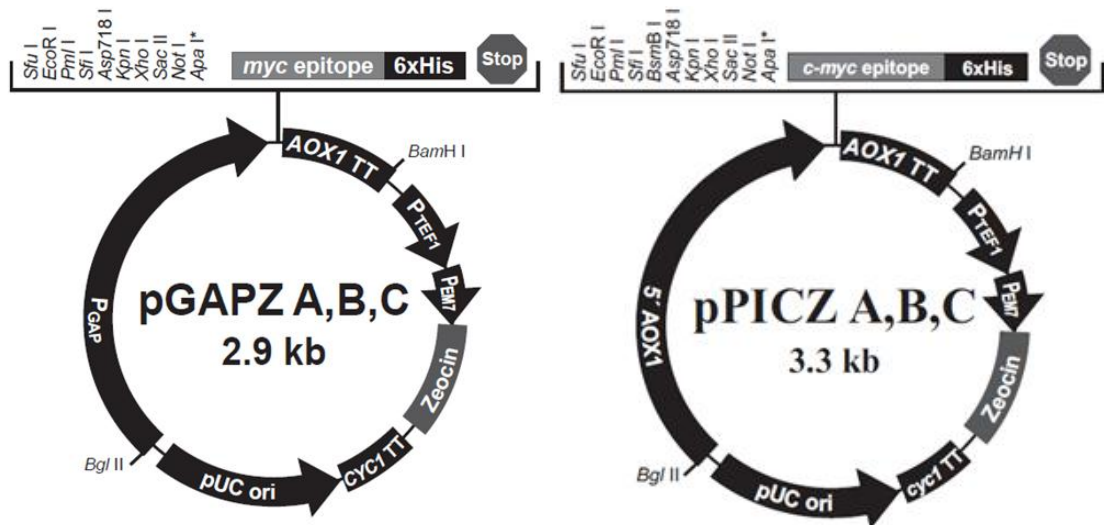
*Pichia pastoris*-hiivan käyttöä rekombinanttiproteiinien tuottokantana on tutkittu 1970-luvulta lähtien. Paljon käytetyllä ja raportoidulla metyloτροφisella eukaryootilla on runsaasti etuja verrattuna esimerkiksi yleisesti käytettyyn prokaryoottiin *Escherichia coliin* tuottokantana. Näitä ovat muun muassa *Pichian* rekombinanttiproteiinien korkea tuottotaso sekä post-translacionaaliset proteiinin muokkaukset, kuten glykolysaatio tavalla, jota yleensä tavataan korkeammilla eukaryooteilla. (Macauley-Patrick *et al.*, 2005) Lisäksi etuna *Pichia pastoris*-tuotossa on rekombinanttigeenin integrointi osaksi hiivan genomia, jolloin tuottoorganismi on stabiilimpi ja geenin integrointikohta genomissa voidaan määrittää (Paus *et al.*, 2002, Skreekrishna *et al.* 1997). Verrattuna eläinsoluilla tai bakulovirussysteemillä tuotettuihin proteiineihin, *P. pastoris*-systeemi on halvempi, nopeampi ja helpommin käsiteltävissä geenitekniikan menetelmin parhaan mahdollisen tuoton saavuttamiseksi (Macauley-Patrick *et al.*, 2005).

Rekombinanttiproteiinin tehokas tuotto *Pichia pastoriksessa* on saatu aikaan voimakkaan promoottorin avulla. Yleensä tällä viitataan *Pichia*-hiivan kykyyn käyttää metanolia ainoana hiililähteenä (Macauley-Patrick *et al.*, 2005). AOX1 on pääasiallinen alkoholioksidaasia (AOX) koodaava geeni *P. pastoriksessa*. Geeni tuottaa 95 % kaikesta solussa tuotetusta entsyymistä, joka toimii metanolin hyödyntämisketjussa silloin, kun metanolia on käytettävissä. Tämän lisäksi samaan tarkoitukseen solussa esiintyy AOX2-geeni, joka siis ei ole yhtä tehokas, mutta pystyy tuottamaan pieniä määriä entsyymiä. (Potvin *et al.*, 2010) Metanoli-indusoitavia fenotyyppisiä on yhteensä olemassa 3, joita nimitetään Mut<sup>+</sup>, Mut<sup>s</sup> ja Mut<sup>-</sup>. Mut<sup>+</sup> käyttää metanolia samoin kuin villityypin *P. pastoris*-kanta eli se tarvitsee metanolia suuria määriä, jonka kaiken se pystyy myös hyödyntämään kasvussaan AOX1:n avulla. Mut<sup>s</sup>-kannasta on poistettu AOX1-geeni, vain AOX2-geeni on olemassa eli se voi käyttää edelleen metanolia, mutta se kasvaa ja käyttää metanolia hitaammin. Mut<sup>-</sup>-kannasta puuttuu sekä AOX1, että AOX2, jolloin solut eivät kasva metanolissa. Kaikkia fenotyyppisiä

käytetään proteiinituotannossa, mutta AOX1 mut<sup>+</sup>- ja mut<sup>s</sup>-systeemit ovat yleisimpiä. (Macauley-Patrick *et al.*, 2005)

GAP on geeni, joka koodaa glyseraldehydi3-fosfaattidehydrogenaasia (GAPDH). GAP-tuottoa ei tarvitse induoida, joka ei johda haasteelliseen metanolin käyttöön fermentoinnin aikana. Tämän lisäksi proteiinisynteesi alkaa heti kasvatuksen alussa, joten mitä suurempi biomassa, sitä suurempi proteiinituotto. (Potvin *et al.*, 2010). Joissain tapauksissa GAP-promoottorin aikaansaama proteiinin tuottotaso on ollut korkeampi kuin AOX1-promoottorin (Macauley-Patrick S. *et al.*, 2005). Esimerkiksi hepatiitti B:n pinta-antigeenin tuotannossa GAP-promoottorin tehokkuus oli saatu aikaiseksi integroimalla *Pichian* genomiin useita geenikasetteja. Tällöin proteiinin tuottotasot nousivat suhteessa geenikasetin kopiolukuun genomissa. (Vassileva *et al.*, 2000)

Invitrogenilla on tarjolla edellä mainittujen lisäksi lukuisia muita induoitavia tai konstitutiivisia promoottorivaihtoehtoja, mutta tässä opinnäytetyössä rekombinanttiproteiinin tuottoon valittiin AOX1- ja GAP-promoottorit sisältävät plasmidivektorit, joita kuvaavat myös vektorien nimet; pPICZ A ja pGAPZ A (Kuva 1). Vektoreista on saatavilla myös pPICZ B ja C sekä pGAPZ B ja C, kirjaimet viittaavat vektoreiden eri lukukehyksiin. Invitrogenin kitin (pPICZ A, B, and C; Cat. no. V195-20) suosituksen mukaan käytetty *Pichia pastoris*-kanta oli X-33, jonka fenotyyppi on Mut<sup>+</sup>. Sama kanta toimii myös käytettäessä pGAPZ A-vektoria (Kitti: pGAPZ A, B, and C pGAPZ $\alpha$  A, B, and C, Cat. No. V205-20). Molemmat vektorit sisältävät zeosiini-resistenssigeenin. (pPICZ A, B, and C User Manual, 2010; pGAPZ A, B, and C pGAPZ $\alpha$  A, B, and C User Manual, 2010)



Kuva 1. Työssä käytettyjen vektorien rakenne.

Rekombinanttiproteiinien tuottotasojen muutoksia ja niiden ongelmakohtia on tutkittu paljon (Macauley-Patrickin *et al.*, 2005). Rekombinanttiproteiinin tuottoon vaikuttavat tuottokanta, tuottokannan luonnollinen proteolyysi, promoottori, itse tuotettava proteiini sekä kasvatusolosuhteet, joihin lukeutuu mm. pH, lämpötila sekä ravinteiden määrä, indusoitavilla kannoilla metanolin määrä. Lisäksi on todettu, että tuottotasoissa on valtava ero fermentorien ja ravistelupullojen välillä johtuen tilavuuden erosta sekä erilaisista kasvatuksen aikaisista olosuhteiden kontrolloinnista tai sen puutteesta.

Edellä mainitussa artikkelissa (Mavauley-Patrick *et al.*, 2005) on listattu vuoteen 2002 mennessä tuotettujen proteiinien tuottotasoja (Taulukko 1.), joista vain yksi oli viruksen proteiini, tuhkarokon nukleoproteiini, joka oli käsittänyt 29 % kokonaisproteiinista. Proteiinit, joita oli *P. pastoriksessa* vuoteen 2002 mennessä tuotettu, olivat peräisin bakteereista, sienistä, selkärangattomista, nisäkkäistä ym. selkärangallisista eliöistä ja ihmisistä sekä kasveista. Tuottotasot vaihtelivat milligrammoista muutamiin grammoihin per litra. Proteiinituotanto oli toteutettu työkohtaisesti eri proteiinien ilmentämiskeinoilla eli käytettyjä promoottoreita olivat AOX1:n ja GAP:n lisäksi muut saatavilla olevat, *P. pastoriksen* promoottorit.

Taulukko 1. *Pichia pastoris*-hiivassa tuotettujen rekombinanttiproteiinien tuotto-  
tasoja. (Macauley-Patrick *et al.*, 2005)

Protein expressed	Function	Expression level	Reference
<b>Bacteria</b>			
<i>Clostridium botulinum</i> serotype F [BoNTF(Hc)] heavy chain fragment C	Antigen	205 mg/kg of cells	55
<i>Escherichia coli</i> AppA	Use in the animal feed industry to release inorganic phosphate	117 ± 1 U/mL	105
<i>Escherichia coli</i> phytase	Use in animal feed industry to improve phosphorus utilization	6.4 g/L	23
<i>Glucanacetobacter diazotrophicus</i> exo-levanase	Fructose-releasing potential for high-fructose syrup production	0.46 g/L	73
<i>Thermus aquaticus</i> YT-1 aqualysin I	Heat-stable subtilisin-type serine protease	1 g/L	81
<b>Fungi</b>			
<i>Aspergillus oryzae</i> tannase	Hydrolyses the ester and depside bonds of gallotannins and gallic acid esters	72 mg/L	132
<i>Candida antarctica</i> CBM–CALB fusion protein	Lypase	1.3 g/L	54
<i>Candida parapsilosis</i> lipase/acyltransferase	Catalyses alcoholysis of esters	5.8 g/L	15
<i>Rhizopus oryzae</i> lipase	Catalyses breakdown of triglycerides	19 U/mL	89
<i>Trametes versicolor</i> cellobiose dehydrogenase (CDH)	Oxidation of reducing-end groups of cellobiose	225 U/mL	106
<i>Trametes versicolor</i> laccase (lcc1)	Phenoloxidase	23.9 U/mL	79
<i>Trigonopsis variabilis</i> D-amino acid oxidase (DAAO)	Flavoprotein, oxidoreductase	1.46 U/mg	129
<b>Invertebrates</b>			
<i>Boophilus microplus</i> BM95 antigen	Protects cattle from <i>B. microplus</i> infestation	550.0 mg/L	7
<i>Pandalus borealis</i> (shrimp) cathepsin L	Cysteine proteinase	60.0 mg/L	1
<b>Vertebrates (non-human)</b>			
<i>Bothrops atrox</i> moojeni venom batroxobin	Cleaves fibrinogen $\alpha$ -chain	7 mg/L	128
Bovine enterokinase light chain (EK <sub>L</sub> )	Duodenal serine proteinase	350.0 mg/L	85
Equine lactoferrin (ELF)	Iron-binding protein, family of transferrins	40.0 mg/L	83
Hen egg H5-lysozyme	Bactericidal effects	2.1 mg/L	66
Hen egg white ovotransferrin	Iron-binding protein, family of transferrins	57.0 mg/L	63
Mouse endostatin	Tumour growth suppressor	133.33 mg/L	110
Ovine interferon- $\tau$ ( $\tau$ -oIFN- $\tau$ )	Antiviral and antiproliferative activity	391.7 mg/L	101
Porcine lactoferrin (PLF)	Iron-binding glycoprotein, family of transferrins	12.0 mg/L	116
<b>Humans</b>			
$\mu$ -Opioid receptor (HuMOR)	G-protein coupled receptor (GPCR)	4.0 pmol/mg	94
$\mu$ -Opioid receptor fused to green fluorescent protein	G-protein coupled receptor (GPCR)	1250 pmol/l	86
Angiostatin	Used in long-term therapy in suppression of metastases	108.0 mg/L	124
Anti-HBs Fab fragment	Prevention and treatment of hepatitis B virus	50.0 mg/L	77
DNA topoisomerase I (hTopo I)	Role in DNA replication, transcription, recombination, chromosome assembly	11–58 mg/L	126
Dopamine D2S receptor	G-protein	3–13 pmol/mg	33
Granulocyte-macrophage colony-stimulating factor (hGM-CSF)	Regulation of activity of mature, differentiated granulocytes and macrophages	180.0 mg/L	123
Serine protease inhibitor SERpin C1-INhibitor (C1-Inh)	Plasma glycoprotein, inhibitor of inflammatory reactions	180.0 mg/L	6
Serum albumin	Binding and transport, colloid osmotic pressure	13.4 mg/g	16
Serum amyloid P (SAP)	G-protein, inhibition of DNA auto-antibody formation	12.0 mg/L	8
Transforming growth factor- $\beta$ type II receptor extracellular domain (SolRII)	TGF- $\beta$ antagonist, regulates cell proliferation and differentiation	20.0 mg/L	95
<b>Plants</b>			
Amaryllidaceae snowdrop agglutinin	Binding of $\alpha$ -D-mannose groups	80.0 mg/L	3
<i>Arabidopsis thaliana</i> core $\alpha$ 1,3-fucosyltransferase	Biosynthesis of N- and O-linked oligosaccharides	1.0 U/L	4
Gsp (Panax ginseng C, medicinal peptide)	Potential use as drug against diabetes	800.0 mg/L	125
Maize cysteine proteinase (Mir1)	Seed germination, senescence	Insoluble/inactive	84
Rice $\alpha$ -amylase (Amy 1A/3D)	Hydrolysis of starch	340.0 mg/L	63
Rice $\beta$ -D-fructofuranosidase (Os $\beta$ fruct3)	Hydrolysis of sucrose	0.35 U/L	37
<i>Solanum tuberosum</i> (potato) apyrase	Phosphohydrolase, platelet-aggregation inhibition	1.0 mg/L	78
<i>Thaumatococcus daniellii</i> thaumatin II	Sweet-tasting protein	25 mg/L	69
<b>Viruses</b>			
Measles virus nucleoprotein (MeN)	Responsible for viral genome packaging	29% of total protein	102

### 3 INFLUENSSA B-VIRUS

Influenssa B-virus (InB-virus) on ortomyksovirusiin kuuluva, lähinnä ihmistä infektoiva, rajua hengitysteiden infektiota aiheuttava virus. Tauti leviää kausittaisina epidemioina ja voi aiheuttaa riskiryhmiin kuuluville hengenvaaran sekä yhteiskunnalle taloudellisia menetyksiä. Muita influenssaviruksia ovat A- ja C-virukset, joista influenssa A-virus aiheuttaa yleensä rankempia epidemioita kuin B-virus, kun taas influenssa C-virus on heikompi taudinaiheuttaja. (Pyhälä, Ziegler, 2005)

Virukset on jaoteltu viruksen genomirakenteen ja replikaatiomekanismin mukaan, influenssa B kuuluu myksovirusiin. Influenssa B-viruksen rakenne käsittää membraanimaisen kuoriosan, jossa on kaksi proteiinia, hemagglutiniini (HA) ja neuraminidaasi (NA). Tämän sisällä genomi on jakautunut kahdeksaksi nukleoproteiinin ympäröimäksi jaokkeeksi negatiivisäikeistä RNA:ta, jonka virus infektoidessaan solun vapauttaa sytoplasmaan. Jaokkeisiin on liittyneenä RNA-polymeraasi, jota tarvitaan genomien replikaatiossa. RNA-jaokkeiden limittäiset lukukehykset sekä silmukointi johtavat kahdeksan jaokkeen koodaamiin 11 eri polypeptidiin. Sytoplastasta genomijaokkeet kulkeutuvat nukleoproteiinin peittäminä tumaan. Tumassa viruksen genomi on ikään kuin osa solun omaa genomia ja solu tuottaa viruspartikkeleita, jotka myöhemmin muodostavat uusia viruksia. (Pyhälä, Ziegler, 2005)

Virusten muuntelukyky perustuu kahteen tapaan. Replikaation aikana RNA-polymeraasin tekemät virheet voivat johtaa HA- tai NA-proteiinien muunteluun. Uusi virustyyppi voi muodostua solussa jos kaksi eri virustyyppiä infektoivat solua samaan aikaan. Tällöin nämä kaksi virustyyppiä voivat yhdistyä eli voi tapahtua reassortatio, jolloin muodostuu uusi virustyyppi. Esimerkiksi H1N1 ja H3N2 muodostavat H3N1-tyyppin viruksen. Uuden tyyppiset virukset ovat vaarallisia, koska ihmisillä ei ole olemassa immuniteettia muuntunutta virusta vastaan, jolloin viruksen leviäminen on helppoa ja nopeaa. (Pyhälä, Ziegler, 2005)

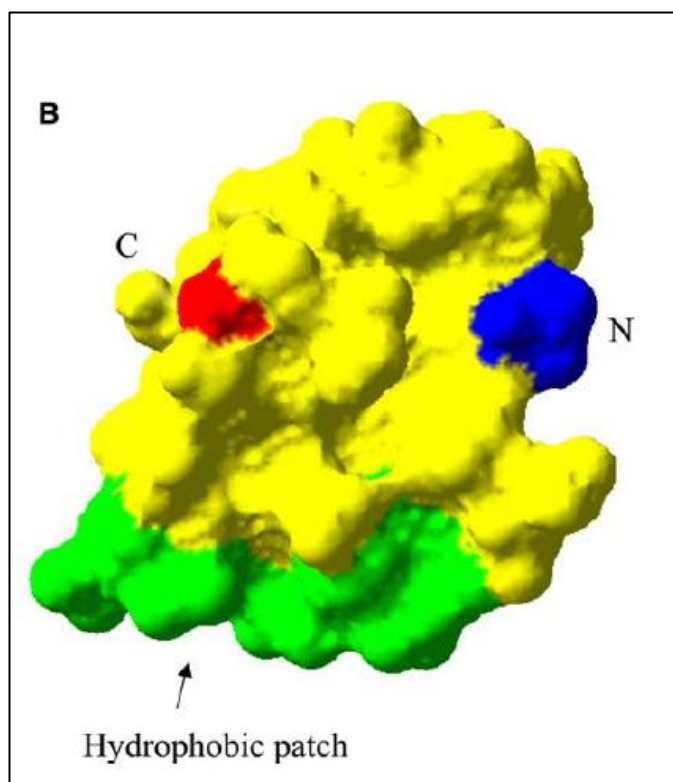
Viruksen nukleoproteiini on konservoitunut eli geneettisiä muutoksia siinä ei juuri tapahdu. Tästä syystä nukleoproteiini antigeeninä on erinomainen diagnostiikan kohde influenssa B-viruksen tunnistamisessa spesifisesti. (Glikmann *et al.*, 1995)

Tässä työssä käytetty InB-virus kanta oli WHO:n jaotteluperusteiden mukaan B/Virginia/05/2011. Kyseinen jaottelu ilmaisee influenssa-tyypin (B), joka on peräisin tietystä paikasta (Virginia), jolla on tietty kannan numero (05) ja viruksen eristysvuosi on 2011. (Assaad *et al.*, 1979)



## 4 HYDROFOBIINIT MERKKIPROTEIININA

Hydrofobiinit ovat alun perin joukko rihmasienten tuottamia proteiineja, joiden yhteisiä piirteitä ovat suhteellisen pieni koko sekä sekvenssin kahdeksan kysteiniä, jotka stabiloivat proteiinin erityisen rakenteen. Kysteiniin lisäksi suuri osa proteiinin aminohappoketjusta muodostuu hydrofobisista aminohapoista, jotka muodostavat proteiinin tertiäärirakenteeseen niin kutsutun hydrofobisen saarekkeen (Linder *et al.*, 2005) (Kuva 2.).



Kuva 2. Hydrofobiinin rakenne. Vihreä alue kuvaa hydrofobisen saarekkeen sijaintia molekyylissä. Molekyylin N-pää ja C-pää on merkattu muuhun rakenteeseen. (Linder *et al.*, 2004)

Hydrofobiinit on tutkittujen ominaisuuksiensa mukaan jaoteltu kahteen ryhmään; luokan I ja luokan II hydrofobiinit. Jaottelu perustuu mm. proteiinin hydrofobisten ja hydrofiilisten alueiden ryhmittymiskuvioon. Hydrofobiinit ovat amfiifiilejä eli niillä on sekä hydrofobinen että hydrofiilinen osa eli ne ovat pinta-aktiivisia molekyylejä. (Wösten *et al.*, 2000) Pinta-aktiivisten ominaisuuksiensa ansiosta hyd-

rofobiineilla on taipumus muodostaa yksikerroksinen filmimäinen kerros veden ja ilman rajapinnalle (Szilvay *et al.*, 2007). Luokan I hydrofobiinien agglomeraatit ovat erittäin vaikealiukoisia ja luokan II hieman epästabiileja (Linder *et al.*, 2004).

Näiden ominaisuuksien vuoksi hydrofobiineilla on useita rooleja sienen rakenteessa sekä suuri merkitys sienten selviytymisessä ja sopeutumisessa erilaisiin ympäristöihin. Hydrofobisen ja hydrofiilisen osan omaava proteiini on pinta-aktiivinen, millä on molekyyllitasolla suuri merkitys. Maan vetovoimalla ei ole juurikaan merkitystä näin pienessä mittakaavassa, sen sijaan veden pintajännitys on valtava voima, johon hydrofobiinit voivat erityispiirteidensä vuoksi vaikuttaa. (Linder *et al.*, 2005) Hydrofobiineilla on osansa esimerkiksi sienen maanpäällisen osan ja itiöiden muodostumisessa ja kasvussa eri vaiheissa, sekä sienen kiinnittymisessä hydrofobisiin kiinteisiin pintoihin, kuten kasvien eri osiin. (Wösten *et al.*, 2000)

#### **4.1 Hydrofobiinien käyttö geenitekniikan sovelluksissa**

Hydrofobiinien käyttö rekombinanttiproteiinin merkki-/puhdistusproteiinina, niin kutsuttuna tagina, perustuu sen amfifiiliseen olemukseen, jolloin hydrofobiinin ja proteiinin fuusio voidaan eristää kasvatusmediumista tai solulysaatista surfaktanttipohjaisella kaksifaasiuutolla. Tämä perustuu siis siihen, että hydrofobiineilla on kyky muodostaa erilaisia järjestäytyneitä rakenteita nesteiden rajapinnoilla, joten ne voivat siirtyä faasista toiseen. (Linder *et al.*, 2005) Hydrofobiinien käyttöä biomateraleina tai pinnoitteina on tutkittu, jolloin hydrofobiinien avulla voidaan muokata partikkelin pinnan hydrofobisuutta. (Bimbo L. M. *et al.*, 2011)

Linder *et al.* (2004) ovat kuvanneet surfaktantti-pohjaisen kaksifaasiuuton periaatteen. Aluksi kohdeproteiinin sisältävä näyte vesifaasissa sekoitetaan surfaktanttiin eli pinta-aktiiviseen aineeseen. Nämä kaksi erottuvat kahdeksi faasiksi, muodostuu surfaktantti- ja vesifaasi, jolloin kohdeproteiini on siirtynyt hydrofobiinin mukana surfaktanttiin. Vesifaasi erotetaan ja proteiini otetaan talteen surfaktanttifaasista sekoittamalla siihen isobutyylialkoholia. Tällöin surfaktantti ja

isobutyylialkoholi muodostavat yhden faasin ja surfaktanttiin aiemmin sitoutunut vesi toisen faasin, johon myös hydrofobiinifuusioproteiini siirtyy. Näin fuusioproteiini on konsentroitunut vesifaasiin.

Kuten muutkin tagit, hydrofobiini voidaan katkaista irti fuusiosta, mikäli sekvenssi on alun perin optimoitu tätä toimenpidettä varten. Erilaisten proteaasien ja restriktioentsyymien käytöstä on kerrottu enemmän kappaleessa 5 Sekvenssien luominen ja optimointi.

## 5 SEKVENSIIEN LUOMINEN JA OPTIMOINTI

Geenien tilaaminen synteettisesti valmistettuna vektorissa on nykypäivää. Tätä työtä varten kunkin geenin sekvenssi luotiin, siihen lisättiin tarvittavat palat ja se optimoitiin. Optimoitu geenisekvenssi lähetettiin sekvenssointipalvelun tarjoajalle, joka syntetisoi geenin ja lähetti sen sopivaan vektoriin kloonattuna tilaajalle. Tätä työtä varten käytettiin GenScript-yhtiön palveluja.

Restriktioentsyymit lukevat sekvenssiä samassa suunnassa kuin proteiinisyn-  
teesikin eli 5'-päästä 3'-päähen ja ne katkaisevat DNA:n jokaisesta tunnista-  
mastaan katkaisukohtasta. Tunnistuskohta on yleensä neljä-kuusi emäsparia  
pitkä ja entsyymi voi katkaista metyloimattoman DNA:n joko tunnistuskohdan  
sisältä tai sen ulkopuolelta (Jeltsch, 2002).

Tässä työssä genomien muokkauksessa käytettiin luokan II restriktioentsyymejä,  
jotka katkaisevat DNA:n fosfodiesterisidokset tunnistussekvenssinsä ulkopuolel-  
ta luoden kohessiiviset päät DNA:han. Huolellisesti suunniteltuna digestoituja  
DNA-pätkiä voidaan liittää yhteen esimerkiksi luoden uusia restriktioentsyymien  
katkaisukohtia. (Engler *et al.*, 2008)

### 5.1 PlasmaDNA

PlasmaDNA on Helsingin yliopiston kehittämä ja ylläpitämä DNA:n analysointi-  
ja manipulointiohjelmisto, joka on ilmaiseksi saatavilla verkossa. Ohjelmiston  
hyvä puoli on se, ettei sitä tarvitse asentaa tietokoneelle, vaan se muodostuu  
tekstikansioista, jotka voi tallentaa esimerkiksi muistitikulle tai koneelle ja se on  
saatavilla sekä PC- että Mac-versioina. (PlasmaDNA-verkkosivusto, viitattu  
30.11.2012)

Kuten tässäkin opinnäytetyössä on käytetty erilaisia internetin ilmaisohjelmia  
sekvenssien analysoimiseen, myös PlasmaDNA:n avulla voi toteuttaa samoja  
toimenpiteitä kuten etsiä restriktioentsyymien katkaisukohtia. Tämän lisäksi  
PlasmaDNA etsii automaattisesti esimerkiksi resistenssigeenejä, promoottereita

ym. sekvenssistä löytyviä piirteitä. PlasmaDNA myös kuvaa sekvenssin jo valmiiksi DNA:n kaksoisjuosteena (suorana viivana tosin), joten sekvenssin hahmottaminen on helppoa ja jollei ohjelma tunnista työssä käytettyä geeniä, voi sen vaivattomasti lisätä osaksi meneillään olevaa projektia. (PlasmaDNA-verkkosivusto, viitattu 30.11.2012)

Ohjelmassa on valmiina runsaasti tallennettua dataa, kuten entsyymejä, joista voi valita oman projektin mukaisen yhdistelmän, DNA:ta voi digestoida ja ligoida ja kun työ on PlasmaDNA:ssa tehty, tiedot voi tallentaa omaan kansioonsa, josta myöhemmin voi nähdä koko projektin alusta loppuun. Ligaation osalta PlasmaDNA tunnistaa myös luokan II restriktioentsyymien muodostamat kohessiiviset päät ja näin liittää palat toisiinsa oikein.

Tämän opinnäytetyön geeniteknisessä osuudessa on käytetty PlasmaDNA:ta varmistuskeinona ennen varsinaisen työn suoritusta, joista kuvia löytyy useassa eri työvaiheessa.

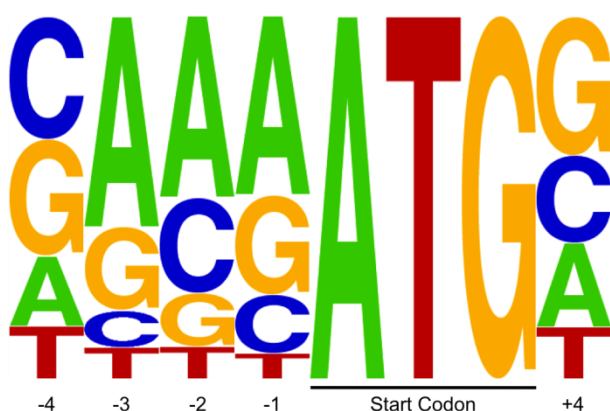
## **5.2 Influenssa B viruksen nukleoproteiinia koodaava geenisekvenssi**

Tuotettavan influenssa B viruksen nukleoproteiinia koodaava geenisekvenssi eli emäsjärjestys etsittiin GISAID-tietokannasta (Global Initiative on Sharing Avian Influenza Data). GISAID on kansainvälinen tutkijoiden perustama yhteisö influenssadatan jakamiseen, analysointiin ja julkaisuun. Tietokannasta voi valita erikseen koko genomien tai osia siitä. Kanta, joka valittiin, oli influenssa B viruksen kanta B/Virginia/05/2011 ja tästä valittiin siis nukleoproteiinin (NP) sekvenssi. (GISAID-verkkosivusto, viitattu 1.11.2012)

Geenisekvenssi sijoitettiin projektin käyttämään sekvenssipohjaan, johon on määriteltä valmiiksi Bsal:n tunnistuskohta, translaation aloituskohta (Kozak),  $\alpha$ -faktori, tuottoproteiinin sekvenssin paikka sekä loppuun lisättävät emäkset sekä Bsal:n katkaisukohta. Sekvenssipohja on opinnäytetyön liitteenä 4.

Tässä työssä käytettiin translaation aloituskohtana Kozak-sekvenssiä, jonka voidaan ajatella olevan 21 emäsparia pitkä, joista kuitenkin muutamalla on suu-

rempi merkitys aloitussignaalin kannalta. Tässä työssä kiinnitettiin erityistä huomiota seitsemän emäsparin optimoimiseksi, jotka on esitetty kuvassa 3. Kozakiin sisältyy varsinainen proteiinin aloitusmetioniini, josta lähtien tuotettavan proteiinin sekvenssin tulee olla oikeassa 3 nukleotidin tahdissa eli lukukehyksessä. Yleensä mRNA:n translaation aloituskohta on ensimmäinen AUG eli metioniini. Proteiinin tehokkaan translaation takaamiseksi tulee tarkistaa, että sekvenssissä on voimakas translaation aloitussignaali aloitusmetioniin yhteydessä. Tällöin ribosomin 40S-osa tunnistaa metioniinin ja sitoutunut 60S-osa yhdessä 40S-osan kanssa muodostaa translaatiokompleksin. (Kozak, 1999)



Kuva 3. Kozakin sijainti aloituskodoniin nähden. Kunkin kirjaimen koko viittaa kyseisen emäksen tärkeyteen voimakkaan Kozak-sekvenssin muodostumisessa. (Amiruddin *et al.*, 2012)

Kozakin jälkeen, ennen proteiinin geenisekvenssiä on *Saccharomyces cerevisiae* prepro- $\alpha$ -faktoria koodaava sekvenssi.  $\alpha$ -faktori on *Saccharomyces cerevisiae* feromonin koodaava n. 200 emäsparin sekvenssi. Feromoni käynnistää erilaiset biokemialliset ja morfologiset muutokset hiivasolussa, jotta se voi lisääntyä suvullisesti.  $\alpha$ -faktoriin tuotto hiivasolussa (*S. cerevisiae*) on tehokasta ja feromonin tuotto suunnataan ulos solusta. (Bitter *et al.*, 1984)

Tässä työssä  $\alpha$ -faktoriin prepro-osan sisällyttämisellä kohdeproteiiniin (InBNP) sekvenssiin on pyritty nimenomaan suuntaamaan tuotto kasvatusmediumiin. Prepro- $\alpha$ -faktoriin geenisekvenssin olisi myös voinut tilata suoraan Pichiaan transformoitavassa vektorissa (pPICZ $\alpha$ /pGAPZ $\alpha$ ) tai sisällyttää syntetisoituihin merkkiproteiinien geenisekvensseihin. Sijoittamalla faktori kohdeproteiiniin gee-

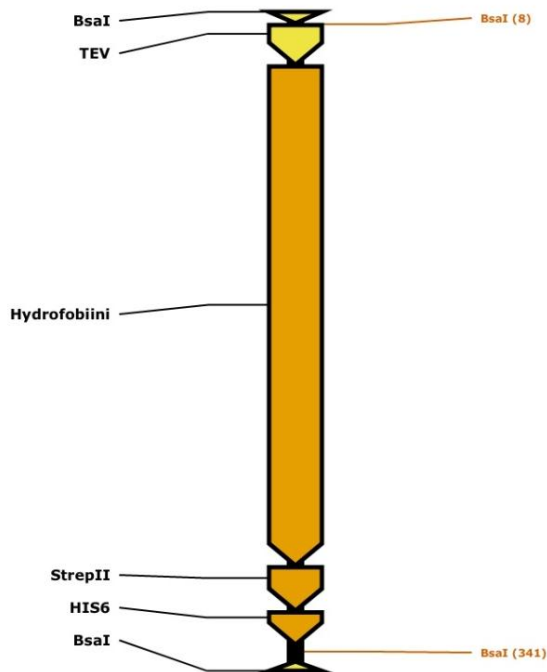
niin pyrittiin varmistumaan kohdeproteiinin  $\alpha$ -faktorin mukaisesta käsittelystä solussa, eli siirtymisestä endoplasmiseen kalvostoon ja oikeaa erityisreittiä ulos solusta, jolloin post-translacionaaliset muokkaukset solussa tapahtuvat.

Lisäksi geenin molempiin päihin oli suunniteltu sopivat emäkset siten, että geenin leikkaaminen irti vektorista luokan II restriktioentsyymillä ja seuraavan vaiheen eli ligationin toteuttaminen spesifisti kohessiivisten päiden avulla onnistuivat. Tätä havainnollistavat PlasmaDNA:n kuvat 4.-6. kappaleessa 5.3.

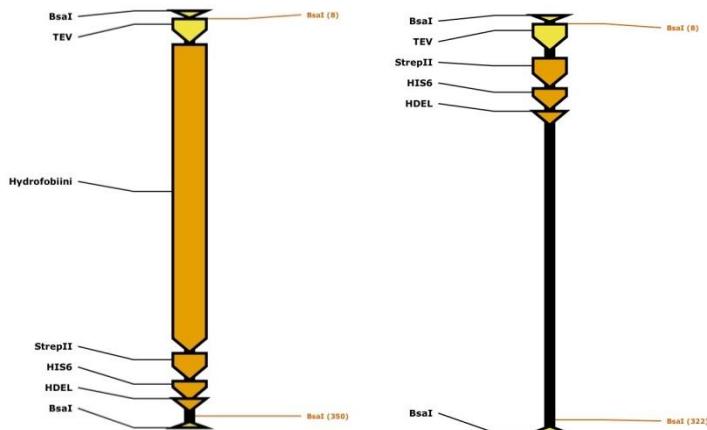
GenScriptiltä tilatut geenien (InBNP, HFBI\_3, HFBI\_4 ja Tag6) sekvenssit toimitettiin pUC57-vektorissa, josta oli poistettu *Bsa* I:n katkaisukohtat, sillä *Bsa* I:tä käytettiin restriktioentsyyminä irrotettaessa geeniä pUC57-vektorista. Toisin sanoen, tilatun sekvenssin molempiin päihin sijoitettiin *Bsa* I:n katkaisukohta, ja restriktioentsyymiä käytettäessä plasmidi katkesi kahteen osaan, kohdegeeniin ja pUC57-plasmidiin, jolloin geenin tunnistaminen ja eristäminen AGE-geeliltä onnistui.

### **5.3 Hydrofobiini 3:n, hydrofobiini 4:n ja Tag6:n sekvenssien luominen**

Koska itse InBNP on varsinainen tuotettava proteiini ja siihen liitetään merkki-proteiini, on järkevää lisätä ylimääräiset tarvittavat geenit näihin merkki-proteiineihin; hydrofobiini 3 (HFBI\_3), hydrofobiini (HFBI\_4) ja histidiini-tag (Tag6). Oheisesta näkyvät sisällytettävät osiot itse proteiinien geenien lisäksi (Kuva 4.).



Kuva 4. HFBI\_3-hydrofobiinin sekvenssin rakenne ja siihen liitetyt osiot. Geenin koko on 315 bp.

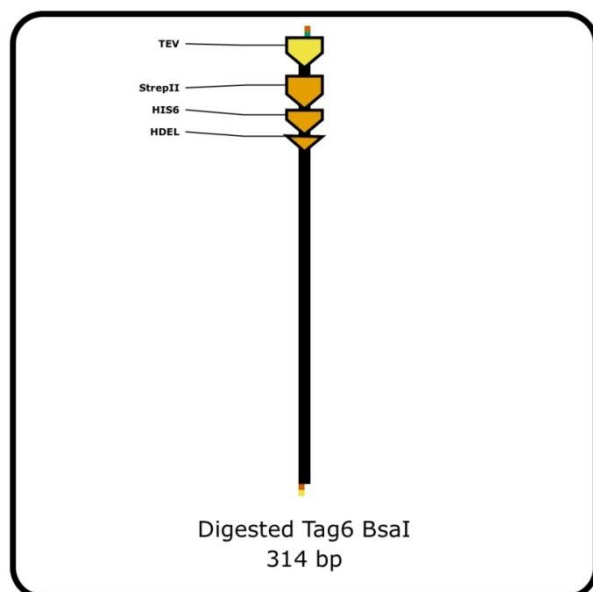


Kuva 5. HFBI\_4:n b(vasemmalla) ja Tag6:n (oikealla) sekvenssien rakenteet. HFBI\_4-geenin koko on 360 bp ja Tag6:n 332 bp.

Ensimmäisenä oleva BsaI-entsyymin tunnistuskohta jättää katkaistessaan sekvenssiin EcoRI-entsyymin katkaisukohtaan, jolla merkkiproteiini liitetään InBNP-geeniin lugaatiossa (Kuva 6.). Tämä on siis mahdollista, koska BsaI on luokan II



restriktioentsyymi, joka katkaisee DNA:n tunnistuskohtansa ulkopuolelta. (ks Liite 3)



Kuva 6. PlasmaDNA:ssa virtuaalisesti digestoitu Tag6-geeni. Kohessiiviset päät näkyvät värillisinä geenin päissä.

Merkkiproteiini voidaan leikata irti tuottoproteiinista ja tätä varten merkkiproteiinin alkupäähän, jonka eteen siis InBNP-geeni tulee, liitettiin TEV-proteaasin leikkauskohta. TEV-proteaasi (Tobacco Etch Virus) on viruksesta eristetty proteaasi jonka tunnistuskohta on glutamiinihappo-asparagiini-leusiini-tyrosiini-fenyylialaniini-glutamiini-(glysiini/seriini) ja katkaisukohta on glutamiinin ja sitä seuraavan glysiinin tai seriinin välissä. (Fang *et al.*, 2007)

StrepII-tagin on kahdeksasta peptidistä muodostuva proteiini, joka voi sitoutua streptavidiniin. Tällä tagilla varustettu fuusioproteiini voidaan puhdistaa myös käyttämällä affiniteettikromatografiaa ja streptavidinimatriksia proteiinien sitouttamiseen. Tämän jälkeen proteiini voidaan eluoida käyttämällä biotiinianalogia. (Arnau *et al.*, 2005) StrepII-tagin liitettiin tässä työssä käytettyihin proteiineihin varmuuden vuoksi, jotta käytettävissä olisi erilaisia puhdistuskeinoja.

Kohdeproteiinin haluttu tuottotapa eli eritetäänkö se solusta ulos vai jääkö proteiini solun sisään tulee olla selvillä, sillä retentiosignaali tulee sisällyttää merkkiproteiinin sekvenssiin. Ilman retentiosignaalia proteiini todennäköisesti erittyy

merkkiproteiinissa olevan prepro- $\alpha$ -faktorin ansiosta endoplasmisen kalvoston ja Golgin laitteen läpi ulos solusta eritettäväksi. Eritysnopeus riippuu proteiinin konsentraatiosta kussakin soluorganelissa. Tässä työssä käytettiin retentiosignaali-HDEL:iä, jonka on todettu riittävän pidättämään proteiini endoplasmisen kalvoston lumenissa. HDEL-nimitys tulee kyseisen signaalin sekvenssistä ja nimenomaan aminohappojen yksikirjaimisesta nimeämisestä eli histidiini – asparagiinihappo - glutamiinihappo – leusiini, joka sijoitetaan proteiinin 3'-päähän eli translaation lukusuunnassa geenin loppupäähän. (Gomord *et al.*, 1998)

HFBI\_3:n ja HFBI\_4:n ero on HDEL-signaalissa, kolmosessa sitä ei ole, nelosessa on. Tag6-proteiini sen sijaan on rakennettu histidiini-tagiksi proteiinin loppupäähän ja siitä puuttuu varsinainen hydrofobiini-proteiini kokonaan (Kuvat 4. ja 5.)

Työssä käytettävien restriktioentsyymien tunnistussekvenssien osalta proteiinin sekvenssi tulee tarkistaa, jottei käytettävän restriktioentsyymin katkaisukohtia ole proteiinin geenissä enempää kuin mahdollisten työvaiheiden kannalta on tarpeellista. Jos esimerkiksi eristetyistä plasmidista halutaan leikata tietty geeni irti liitettäväksi myöhemmin tuotettavaan vektoriin, käytettävän restriktioentsyymin katkaisukohdat täytyy poistaa geenin sekvenssistä.

#### **5.4 Sekvenssien optimointi**

Geenin optimointi tarkoittaa geenin sekvenssin muokkaamista siten, että emästen käyttö tuotto-organismissa olisi mahdollisimman optimaalista ja tuotto helppoa eli mahdollisimman tehokasta. ”Kodonien optimointi on niiden kodonien sekvenssien valintaa, jotka määrittävät tiettyä proteiinia siten, että valituilla kodoneilla on suurin mahdollinen esiintymistiheys organismissa.” (Condon, Thachuk, 2012)

Internetistä löytyvien kodonin optimointiohjelmien tärkein informaatio on CAI (codon adaptation index) eli kodonin sovitusindeksi, joka kertoo kunkin aminohapon kodonin käytön sopivuudesta tuotto-organismien kodoninkäytössä.

Työ aloitettiin yksinkertaisesti syöttämällä geenin sekvenssi tekstinkäsittelyohjelmaan. Kaikkien sekvenssien, InBNP, hydrofobiini 3 (HFBI\_3) tai hydrofobiini 4 (HFBI\_4) kohdalla työ suoritettiin samalla tavalla.

Aluksi sekvenssin optimoinnissa hyödynnettiin JCat (JAVA Codon Adaptation Tool)-ohjelmaa. Ohjelman on luonut Andreas Grote Mikrobiologian instituutissa Braunschweigin teknillisessä yliopistossa. Andreas Grote *et al.* ovat kirjoittaneet JCat-ohjelmasta artikkelin (2005). Viimeisin päivitys ohjelmasta on 9.2.2008.

NEBCutter on New England Biolabsin ohjelma, joka löytää restriktioentsyymien katkaisukohtat. Ohjelmasta voi valita eri restriktioentsyymejä, seuloa niiden tunnistuskohtia sekvenssistä ja etsiä mitkä entsyymit katkaisevat geenin kerran tai useammin sekä mitkä eivät katkaise kertaakaan. Jos tietää mitä entsyymejä tulee geenin muokkauksessa käyttämään, voi seuloa katkaisukohtat läpi ohjelman avulla. (Vincze *et al.*, 2003)

Expasy on SIB:n (Swiss Institute of Bioinformatics) bioinformatiikkaportaali, josta on pääsy tieteellisiin tietokantoihin sekä biotieteiden alan ohjelmiin, jotka voivat liittyä esimerkiksi proteomiikkaan, genomiikkaan tai systeemibiologiaan. Tässä työssä ExPASya käytettiin optimoidun sekvenssin translaation kuvantamiseen, vaikka ExPASysta olisi löytynyt työkalut useaan muuhunkin optimoinnin työvaiheeseen. ExPASysta saa translaation tekstimuodossa, josta näkee sekä aloitusmetioniinit että mahdolliset stop-kodonit eri lukukehyksissä.

Edellä mainituista sekvenssinkäsittelyohjelmista on kuvia liittyen InBNP-geenin optimointiprojektiin liitteessä 5.

Lisäksi geenin tarkastamisessa sekä translaation että restriktioentsyymien varalta käytettiin Sequence Manipulation Suitea (SMS), josta myös löytyy useita työkaluja genomien optimointiin. SMS:n tekijänoikeudet omistaa Paul Stothard ja sivustoa päivitetään säännöllisesti.

Työssä käytettiin siis useita eri ohjelmia sekvenssin tarkastamiseen, jotta voitiin olla varmoja oikeasta kokoonpanosta. Sekvenssintalvot tarjoava GenScript ei tarkasta sekvenssiä, vaikka se vielä varmistaa asiakkaan halukkuuden tilauk-

sen viemiseen eteenpäin ennen sekvensoinnin aloitusta sekä antaa asiakkaalle mahdollisuuden sekvenssin tarkistamiseen.

## 6 VEKTORIKONSTRUKTIEN LUOMINEN *E. COLISSA*

Vieraan proteiinin ekspressio *Pichia pastoris*-systeemissä noudattaa kolmea pääperiaatetta. Geeni sisällytetään vektoriin, joka transformoidaan *Pichia pastoris*-kantaan ja transformoituneista soluista tehdään omat solulinjansa, soluklooninsa, joiden joukosta seulotaan parhaat proteiinia tuottavat kannat. (*Pichia Expression Kit User Manual*, 2010)

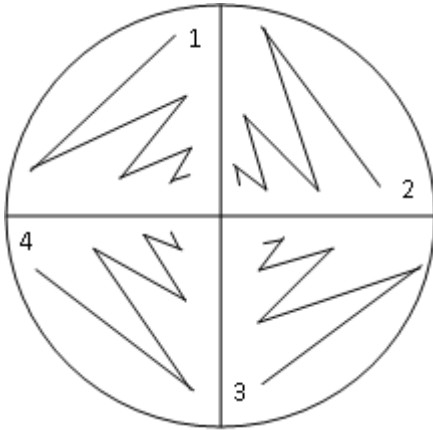
Hiivan kasvuvauhti on hidasta verrattuna *E. coliin*, joten aluksi geenin muokkaus tehtiin *E. colissa*, jolloin saatiin nopeasti seulottua tarvittava määrä *Pichiaan* transformoitavaa vektoria.

Genscriptiltä tilattu geeni toimitettiin *BsaI*-katkaisukohta-vapaassa pUC57-plasmidissa, joka on *E. coliin* korkean kopioluvun plasmidi. pUC57-plasmidiin sisältyy ampicilliiniresistenssigeeni, jonka vuoksi *E. coli*a kasvatetaan elektroporaation jälkeen selektiivisesti ampicilliinia sisältävillä maljoilla tai mediumissa, johon on lisätty antibioottia. Itse PICZ A ja GAPZ A vektorit sisältävät zeosiiniresistenssi-geenin, eli kun *E. coli* tai *Pichia* on transformoitu valmiilla vektorilla, käytetään zeosiinia seulontamenetelmänä.

### 6.1 *Escherichia coli*n transformaatio plasmidistokista, kloonauksen ja seulonta

pUC57-plasmidistokit laimennettiin ennen elektroporaatiota. 50 µg Genscriptiltä tilattua plasmidia transformoitiin elektrokompetentteihin *E. coli* soluihin XL-1 (50 µL) elektroporaatiolla (Liite 1). *E. coli*n elektroporaatiossa käytettiin 1 mm:n elektroporaatiokyvettejä (BTX, electroporation cuvettes plus™). Lisäksi tehtiin negatiivinen kontrolli, jossa DNA korvattiin vedellä. Elektroporaatiolaitteen asetukset olivat low range 200 Ω, kapasitanssi 25 µF ja jännite 1,5 kV. Elektroporaatioseokseen lisättiin viipymättä elektroporaation jälkeen huoneenlämpöistä, steriiliä SOC-mediumia, jonka jälkeen suspensiota inkuboitiin +37 °C:ssa tunnin ajan. Tämän jälkeen 50 µl ja 5 µl solususpensiota maljattiin 75 µg/ml ampicilliinia (AMP75) sisältäville low salt-LB-maljoille ja inkuboitiin +37 °C:ssa O/N.

Seuraavana päivänä maljoilta valittiin neljä pesäkettä jatkotutkimuksia varten, jotka maljattiin uudelle AMP75 LS-LB-maljoille ja inkuboitiin +37 °C:ssa O/N (Kuva 7.) Valituista pesäkkeistä tehtiin esikasvatusten jälkeen tutkimukset AGE:lla sekä DNA-pitoisuusmittauksin, joiden perusteella valittiin parhaimman saannon antava solulinja. Tutkimusten perusteella valituilla solulinjoilla tuotettiin riittävä määrä DNA:ta ligaatioita sekä Pichian transformaatiota varten.



Kuva 7. Valittujen pesäkkeiden (kloonien) maljauskuvio.

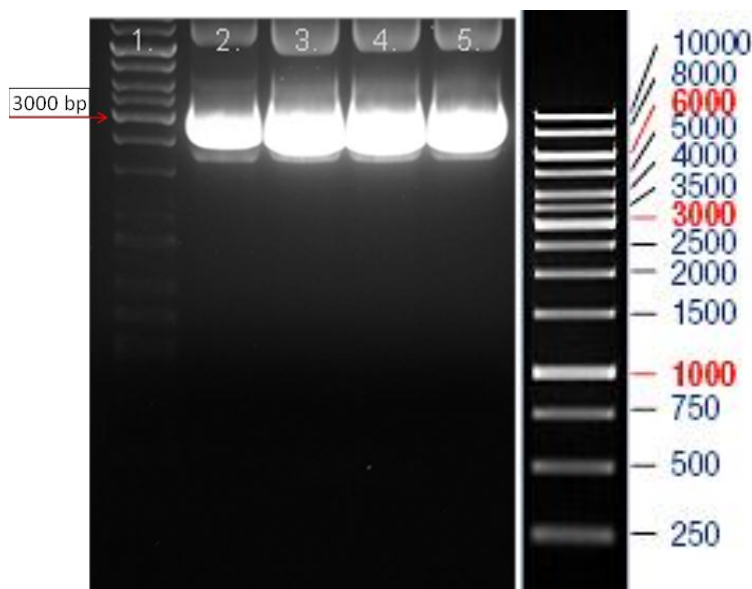
Klooneista tehtiin 5 ml:n kasvatukset LS-LB-mediumissa, johon oli lisätty 100 mg/ml ampicilliinia. Kasvatusparametrit olivat + 37 °C ja 250 rpm. pUC57-plasmidi eristettiin kasvatusmediumista käyttäen NucleoSpin® Plasmid-kittiä (Macherey-Nagel). DNA-pitoisuudet mitattiin käyttäen Nanodrop-laitetta. Näytteistä osa digestoitiin Bsal-entsyymillä (Fermentas), jonka jälkeen digestioseos ajettiin AGE:lla, 100 V noin 35 minuuttia (Kuvat 8.-10.). Geenien DNA-sekvenssien koko geelillä määritettiin käyttäen Fermentaksen GeneRuler 1 kb-kokomarkkeria.

Elektrokompetenttien solujen valmistus ja mediumien reseptit löytyvät liitteestä 1.

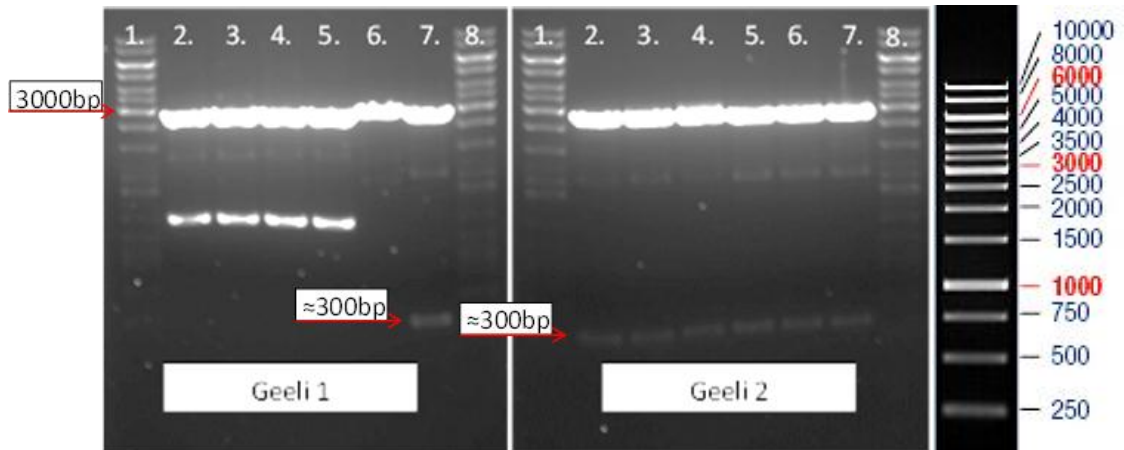
AGE-geelien kuvien (Kuvat 8.-10.) ja Nanodropin DNA-pitoisuuksien (perusteella valittiin parhaiten oikeaa plasmidia tuottava kanta, jota käytettiin seuraavassa vaiheessa. Tässä vaiheessa DNA-pitoisuuden tuli olla riittävän suuri, jotta

geelieristyksen saannosta tulisi riittävä. Dataa ei ole esitetty, koska sillä ei ole työn etenemisen kannalta merkitystä.

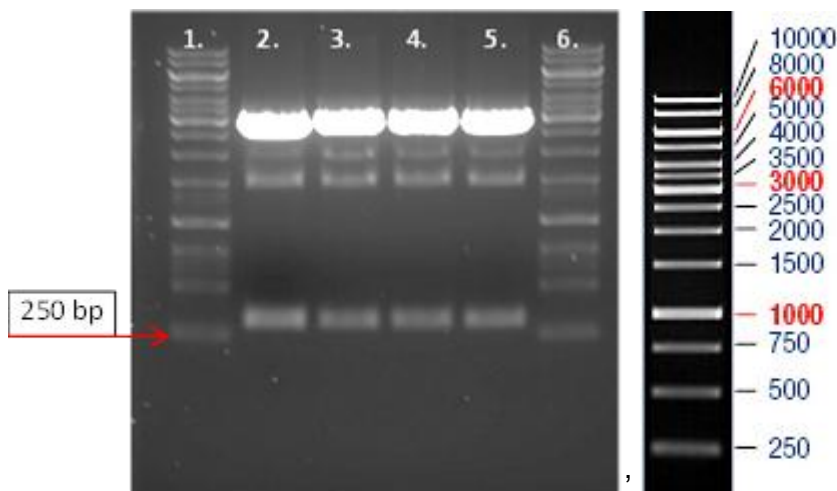
Kuvista etsittiin toisistaan eronneet pUC57-plasmidi sekä kohdegeeni, lisävyöhykkeet merkitsevät vääränlaista geeniä. Koska InBNP-pUC57-plasmidia ei digestoitu, itse kohdegeeniä ei voitu nähdä geelillä, joten Midiprepiin valittiin kloni numero 3. kaivossa 4 (Kuva 8.). Kuvien 9. perusteella parhaiten HFBI\_3:a tuotti kasvatus 2. kaivossa 7 geelillä 1 ja HFBI\_4:a kasvatus 2. kaivossa 5 geelillä 2 sekä Tag6-geeniä kasvatus 1 kaivossa 2 (Kuva 10.).



Kuva 8. Digestoimaton InBNP pUC57-plasmidissa.



Kuva 9. HFBI\_3:n digestiot kaivoissa 6-7 geelillä 1 sekä geelillä 2 kaivoissa 2 ja 3. HFBI\_4:n digestiot geelillä 2 kaivoissa 4-7.



Kuva 10. pUC57-Tag6:n digestiot, kloonit 1.-4. kaivoissa 2.-5.

## 6.2 Kasvatus Midiprep-plasmidin geelieristystä varten

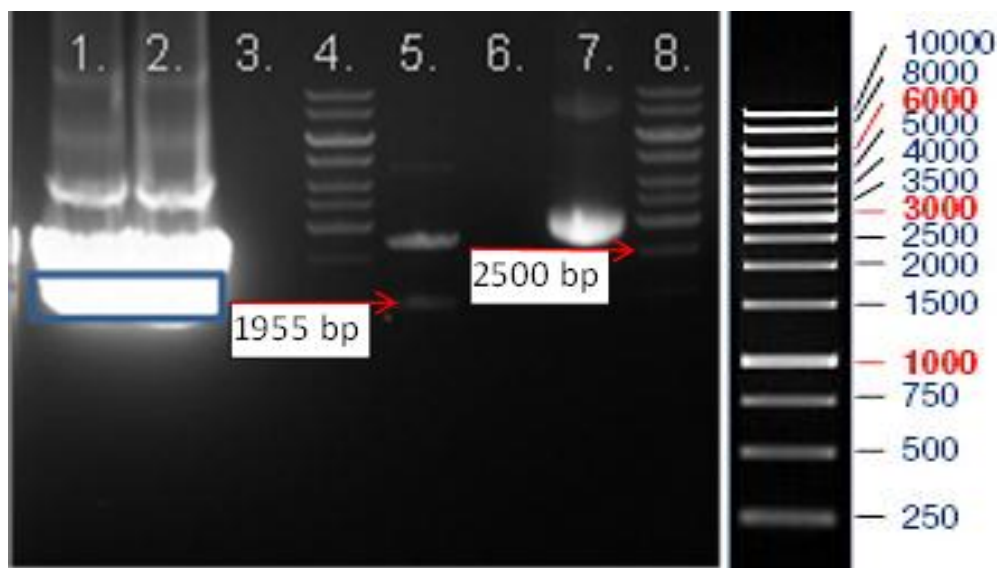
Plasmidin eristystä varten valittiin seuraavat kannat; InBNP kloni 3., HFBI\_3, kloni 2., HFBI\_4 kloni 2. ja Tag6 kloni 1. Kloonit kasvatettiin 100 µg/ml ampisilliini-antibioottiselektion avulla LS-LB-mediumissa (Liite 1). Kasvatusaika oli 20-22 tuntia, lämpötila +37 °C ja sekoitusnopeus 250 rpm.

150 – 200 ml:n *E. coli*-kasvatus käytettiin kokonaisuudessaan pUC57-plasmidin eristykseen. Käytettävä kitti oli NucleoBond® Xtra Midi (Macherey-Nagel). Eristys suoritettiin ohjeen mukaan, viimeisenä DNA:lle tehtiin isopropanolisaostus.



Saostuksen jälkeen DNA-pelletti liuotettiin 100–200 µl:an steriiliä milli-Q-vettä. Näytteet digestoititiin Bsal-entsyymillä kohdegeenin irrottamiseksi plasmidista ja näytteet ajettiin AGE:lla, josta tutkittiin geenin koko. DNA-pitoisuudet mitattiin Nanodropilla. Saantojen viitteelliset arvot on laskettu NanoDrop-tuloksista, jolloin saatiin InBNP:n saannoksi 423 µg ja konsentraatioksi 4230 ng/µl, HFBI\_3:n 950 µg ja 4770 ng/µl, HFBI\_4:n 670 µg ja 3370 ng/µl ja Tag6:n saanto oli 480 µg plasmidia, liuoksen konsentraatio ≈ 900 ng/µl. Seuraavan vaiheen DNA-määrä laskettiin näistä konsentraatioista.

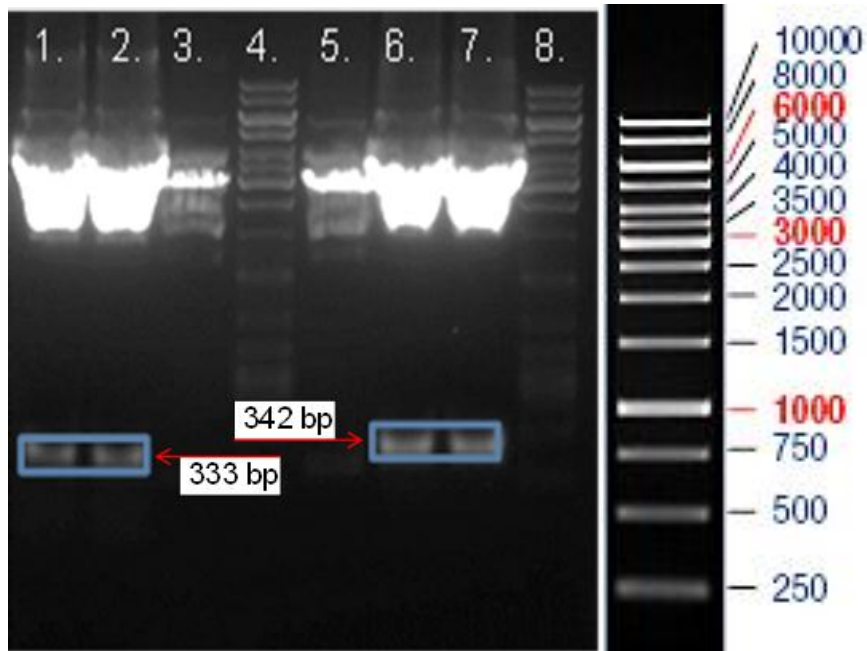
Geelieristystä varten Midi-tuotoksen saannosta käsiteltiin n. 30 µg DNA:ta digestioimalla. Plasmidista digestoititiin irti haluttu geeni Bsal-restriktioentsyymillä (Fermentas) ja koko digestio pipetoitiin AGE-geelille. Geeli ajettiin 100 V, noin 35 minuuttia ja geelille ajautunut kohdegeenipala leikattiin skalpellilla irti UV-valon avulla (Kuvat 11.–13.).



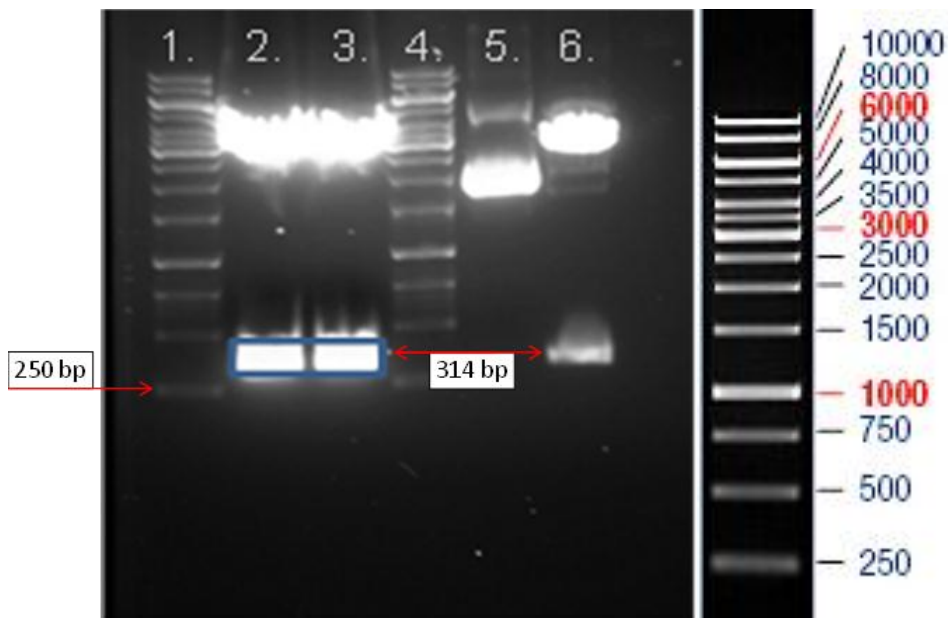
Kuva 11. InBNP-geeni digestoituna. Kaivoissa 1. ja 2. on yhteensä jaettuna digestionäyte. 3. ja 6. kaivo on jätetty tyhjäksi, 5. kaivossa 1 µl digestionäytettä, 7.:ssa katkaisematon InBNP+pUC57 ja 4. sekä 8. kaivossa GeneRuler 1 kb.

Koska pUC57-plasmidi on kooltaan noin 2700 bp ja InBNP 1955 bp, ne ajautuvat geelillä lähes yhtä matkaa. Lisäksi, kun geelille oli ladattu mahdollisimman paljon geeniä yhteen kaivoon, jotta geelieristykseen tulisi mahdollisimman vähän epäpuhtauksia leikattaessa, kuvat ovat huonoja. Silmin nähden InBNP-

geeni ja pUC57-plasmidi erottuivat UV-valon alla paremmin, jolloin laatikoidun kohdegeenin talteenotto onnistui kohtalaisen hyvin (Kuva 11.).



Kuva 12. 1. & 2. kaivossa on HFBI\_3:n digestiot, 3. katkaisematon näyte, 4. GeneRuler 1 kb, 5. HFBI\_4:n katkaisematon näyte, 6. & 7. kaivossa HFBI\_4:n digestiot sekä 8. kaivossa GeneRuler 1 kb.



Kuva 13. 1. ja 4. kaivossa GeneRuler 1 kb, 2. ja 3. Tag6:n digestiot, 5. katkaisematon näyte ja 6. 1,5 µl:n näyte digestiosta.

Laatikoidut alueet leikattiin geelieristykseen. Geeni eristettiin geeliltä NucleoSpin® Gel and PCR Clean-Up.-kitillä (Macherey-Nagel) ja eluoiitiin 20-30 µL:an kitin NE-puskuria, jonka jälkeen DNA-pitoisuudet mitattiin Nanodropilla (Taulukko 2.).

Taulukko 2. Geenien saannot ja liuosten konsentraatiot geelieristuksen jälkeen.

<b>Näyte</b>	<b>DNA-pitoisuus (ng/µl)</b>	<b>Saanto (µg)</b>
<b><i>InBNP</i></b>	360,7	9,0175
<b><i>HFBI_3</i></b>	11,86	0,356
<b><i>HFBI_4</i></b>	11,71	0,350
<b><i>Tag6</i></b>	145	2,9

### 6.3 Ligaatiot

Erilaisia vektorikonstrukteja tehtiin 6 kappaletta. Jokaisen pääasiallinen tuotto-proteiini on influenssa B viruksen nukleoproteiini (InBNP), vektori ja merkkiproteiini eli tagi vaihtuvat. Tägeja työssä valmisteltiin 3 kappaletta, HFBI\_3, HFBI\_4 ja Tag6 ja valittavia vektoreita oli 2 kpl, PICZ A ja GAPZ A.

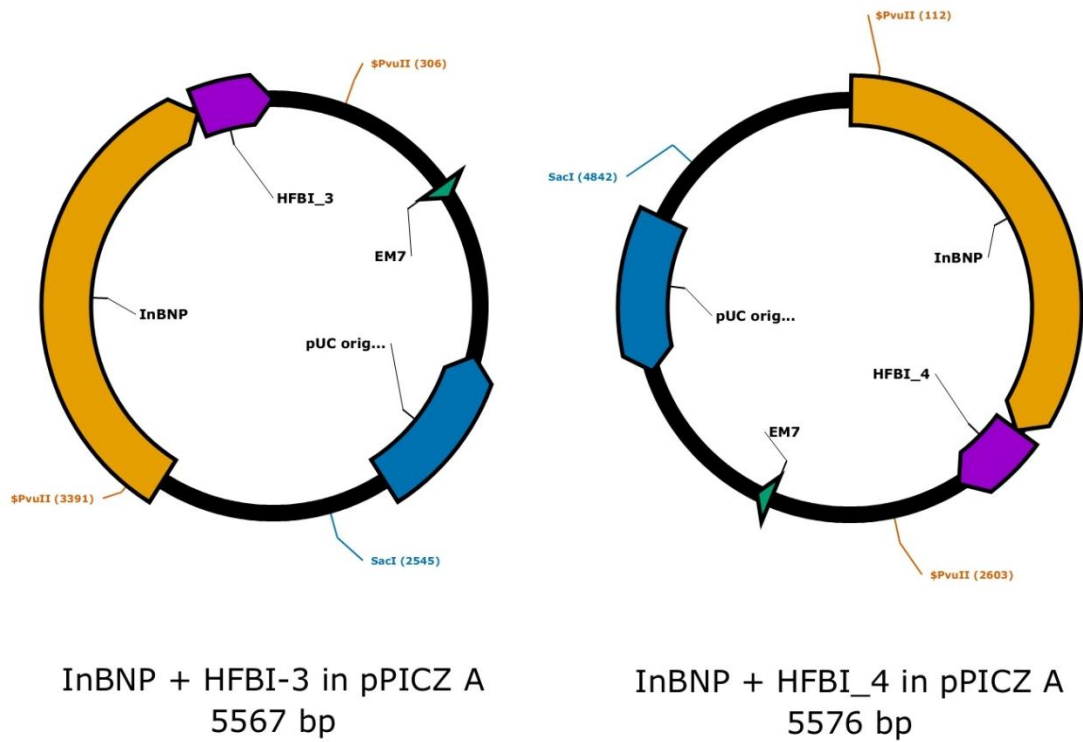
Vektorit oli esikäsitelty luokan II restriktioentsyymeillä EcoRI ja NotI (Fermetas), jolloin suunnitellut insertit ja vektori liittyivät toisiinsa oikeassa järjestyksessä (ks. Liite 4).

Taulukko 3. Vektorit.

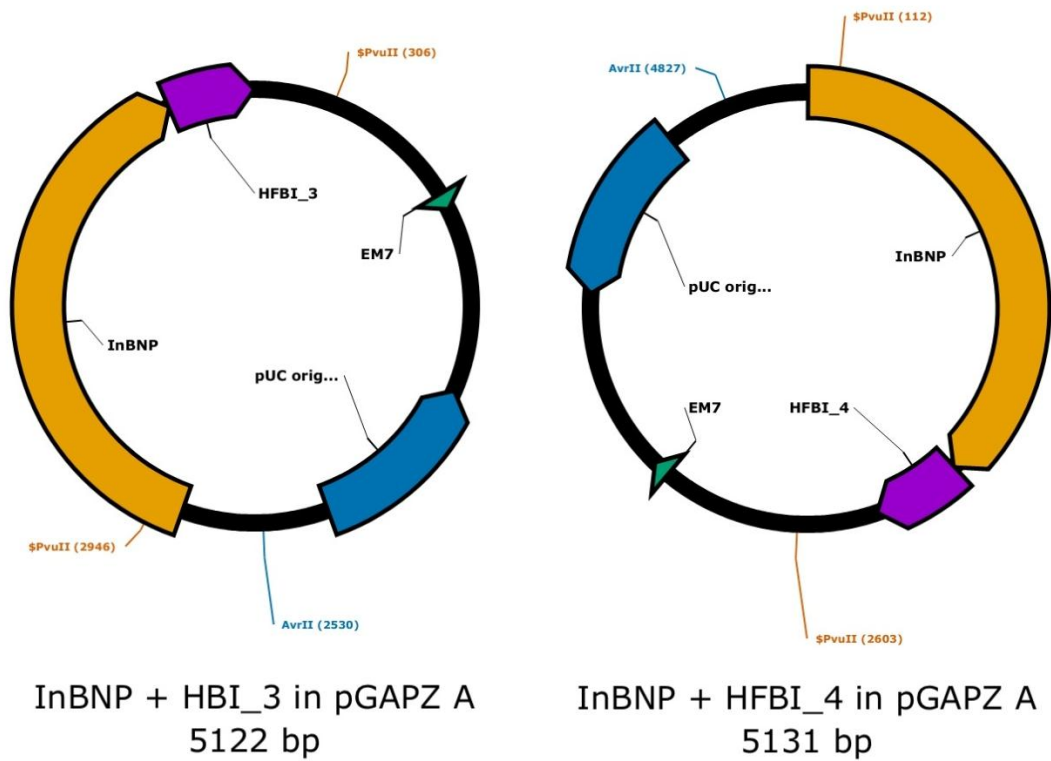
<b><i>Uuden vektorin nimi:</i></b>	<b>Insertti:</b>	<b>Insertti:</b>	<b>Vektori:</b>	<b>Numerokoodi (Liite 3)</b>
<b><i>PICZ 1.</i></b>	HFBI_3	InBNP	PICZ A	631X0 / 631X1
<b><i>PICZ 2.</i></b>	HFBI_4	InBNP	PICZ A	641X0 / 641X1
<b><i>PICZ 3.</i></b>	Tag6	InBNP	PICZ A	651X0 / 651X1
<b><i>GAPZ 1.</i></b>	HFBI_3	InBNP	GAPZ A	632X1
<b><i>GAPZ 2.</i></b>	HFBI_4	InBNP	GAPZ A	642X1
<b><i>GAPZ 3.</i></b>	Tag6	InBNP	GAPZ A	652X1

Käytännön työssä valmiiden vektorien kanssa käytettiin HydroBody-projektin omaa numerointikoodijärjestelmää (Liite 3), mutta tässä opinnäytetyössä vektoreita on selvyiden vuoksi kuvattu taulukon 3. mukaan PICZ 1., PICZ 2 jne. Myös tulosten käsittelyn kirjallisessa osuudessa on käytetty Taulukon 3. mukaista nimeämistä, kun taas varsinainen data on esitetty numerokoodien mukaan.

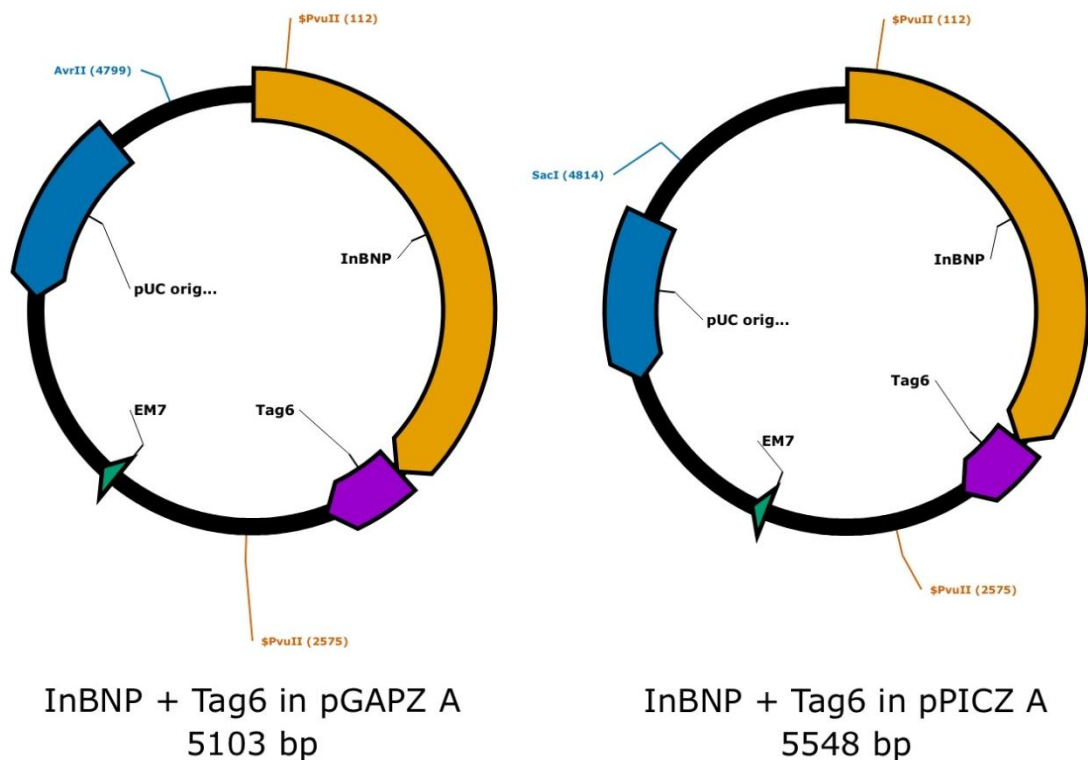
Ennen työn suoritusta ligaatioita testattiin PlasmaDNA:n avulla, jossa myös muodostettiin kuvat tulevista plasmidivektoreista (Kuvat 14.–16.).



Kuva 14. PICZ 1. ja PICZ 2. plasmidivektorit.



Kuva 15. GAPZ 1. ja GAPZ 2. plasmidivektorit.



Kuva 16. GAPZ 3. ja PICZ 3. plasmidivektorit.

Kuvista on jätetty selkeyden vuoksi pois edellisissä kuvissa (Kuvat )olleet domainit. Kuvissa on esitetty InBNP- ja merkkiproteiini-geenit sekä käytettävien restriktioentsyymien tunnistuskohdat. Lisäksi EM7-promoottori sekä pUC-ori-alue on jätetty kuviin, mutta niillä ole merkitystä tämän työn kannalta.

Ligaatiot tehtiin moolisuhteessa 3:1, jossa vektorin määrä on vakio 150 ng. Kumpaakin inserttiä, InB- ja tagi-geeniä, lisätään siis 1/4 osa vektoriin nähden. Määrien laskemiseen käytettiin internetissä saatavilla olevaa laskuria (Liite 6.). Jos vektorin ja insertin/inserttein koot emäspareina tiedetään, laskuri laskee tarvittavat määrät inserttiä (ng). Näiden massojen perusteella laskettiin tarvittavat tilavuudet kunkin geenin geelieristysaannosta.

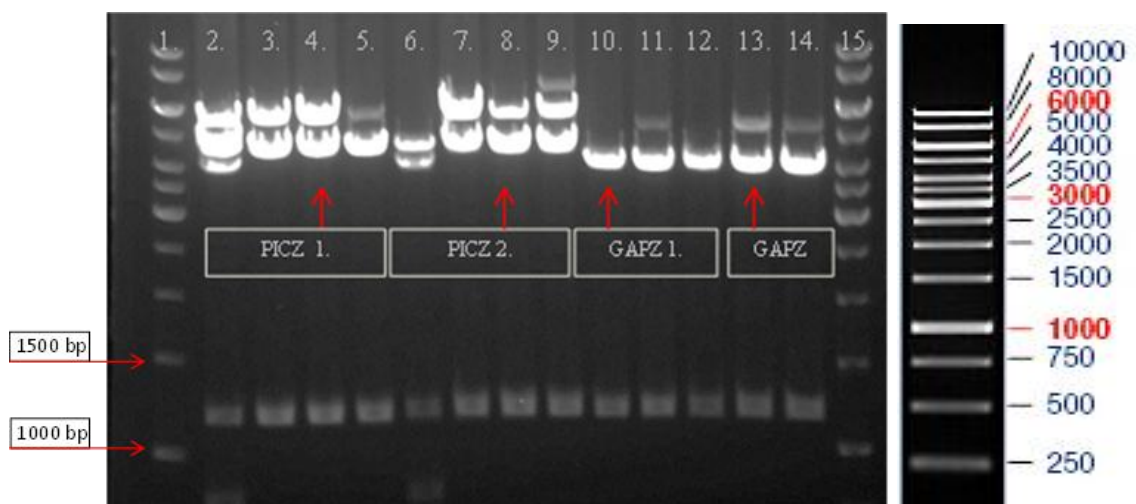
Ligaatiossa käytettiin Fermentaksen T4 DNA Ligase-ensyymiä. Inkubointiaika oli kolme tuntia huoneenlämmössä, jonka jälkeen entsyymi lämpöinaktivoitiin + 70 °C:ssa .

## 6.4 *Escherichia coli* transformaatio vektorilla

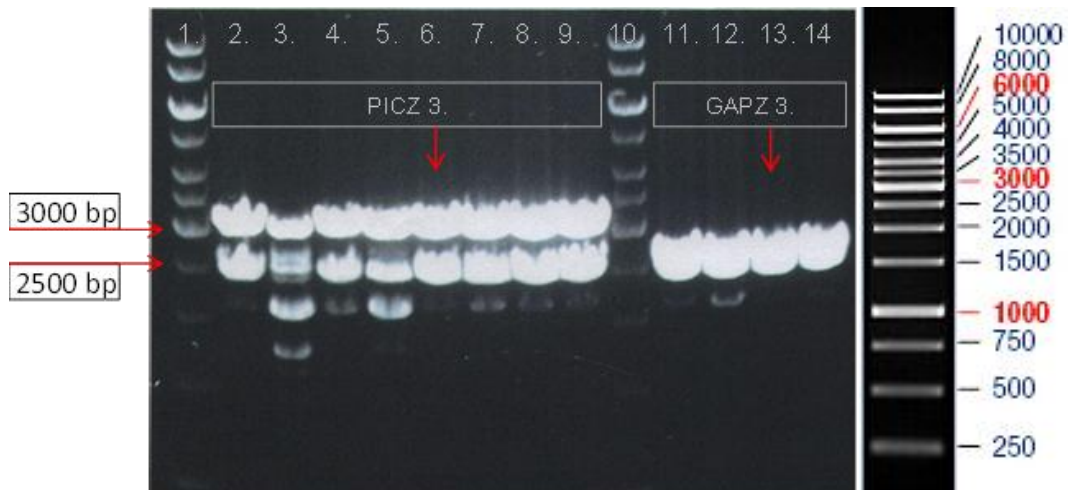
Jotta vektoria olisi tarvittava määrä *Pichian* transformointia varten, vektorikonstrukteja monistettiin *E. coli*ssa.

Elektrokompetentteihin *E. coli*-soluihin (Liite 1) (50 µl) elektroporoitiin 1,5 µl:ta ligaatioseosta. Elektroporaatio ja maljaus suoritettiin kuten *E. coli* transformointi plasmidistokista.

Kloonit seulottiin samaan tapaan kuin aiemmin, käyttäen aluksi NucleoSpin® Plasmid-kittiä ja digestiota ennen agaroosigeelielektroforeesia. Plasmidit digesioitiin restriktioentsyymeillä Sall ja Xbal (PICZ 1. ja PICZ 2. sekä GAPZ 1. ja GAPZ 2.) ja PvuII (PICZ3. ja GAPZ 3.), jolloin kustakin plasmidista muodostui kaksi helposti toisistaan erotettavaa palaa. Sall+Xbal-digestioissa muodostuneiden palasten koot olivat 1200 ja 4400 bp ja Pvu-digestiossa muodostui 2400 ja 3000 bp:n kokoiset palat. Kaikki käytetyt restriktioentsyymit olivat Fermentaksen. AGE:n ja NanoDrop-DNA-pitoisuusmääritysten perusteella päätettiin, mikä kloni otettiin jatkoon eli millä kannalla tuotettiin vektoria *Pichian* transformointia varten. Kuvissa on merkitty kunkin konstruktin valitut kloonit punaisiin nuolin (Kuvat 17.-18.).



Kuva 17. PICZ 1., PICZ 2., GAPZ 1. ja GAPZ 2. nimettyjen vektorien seulontadigestiot AGE-geelillä.



Kuva 18. PICZ 3. ja GAPZ 3. seulontadigestiot

Kuvien perusteella tehdyt solulinjojen valinnat perustuivat paitsi etsittyjen vyöhykkeiden kokoon, myös siihen, löytyykö ajautuneesta näytteestä ylimääräisiä vyöhykkeitä, jotka indikoivat vääränlaista DNA:ta. Kuvassa 17. todettiin kuitenkin ylimmän, suurimman vyöhykkeen olevan katkeamatonta plasmidia.

Tämän jälkeen käytettiin jälleen NucleoBond® Xtra Midi-kittiä suuremmista esikasvatuksista. Eristykseen käytettiin klooneja PICZ 1; 3., PICZ 2.; 3., GAPZ 1.; 1., GAPZ 2.; 1 sekä PICZ 3.; 5. ja GAPZ 3.; 3.

Vektorit eristettiin siis suoraan NucleoBond® Xtra Midi-kitillä ja isopropanolisostuksella kitin ohjeen mukaan. Saostuksesta saatu pelletti liuotettiin 200 µl:an milli-Q-vettä ja DNA-pitoisuus mitattiin NanoDropilla.

Taulukko 4. *E. colista* eristettyjen vektorien saanto ja pitoisuus.

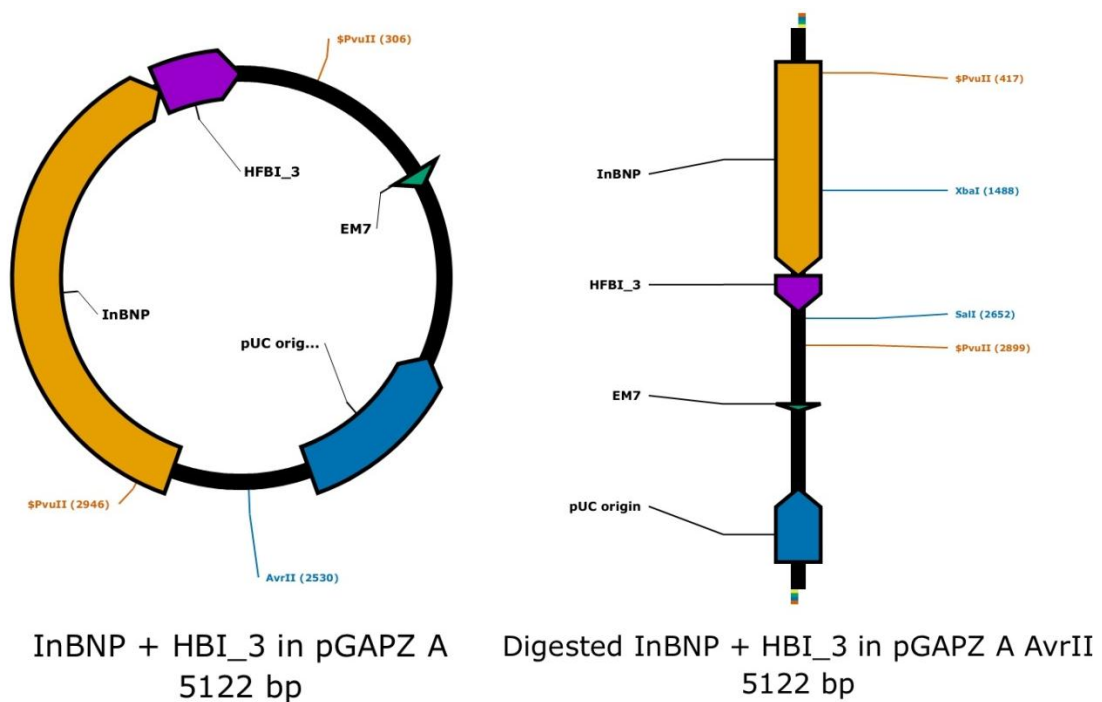
<b>Vektori</b>	<b>Pitoisuus (ng/µl)</b>	<b>Saanto (µg)</b>
<b>PICZ 1.</b>	6400	1280
<b>PICZ 2.</b>	9700	1940
<b>PICZ 3.</b>	11300	2260
<b>GAPZ 1.</b>	4600	920
<b>GAPZ 2.</b>	2500	500
<b>GAPZ 3.</b>	4700	940



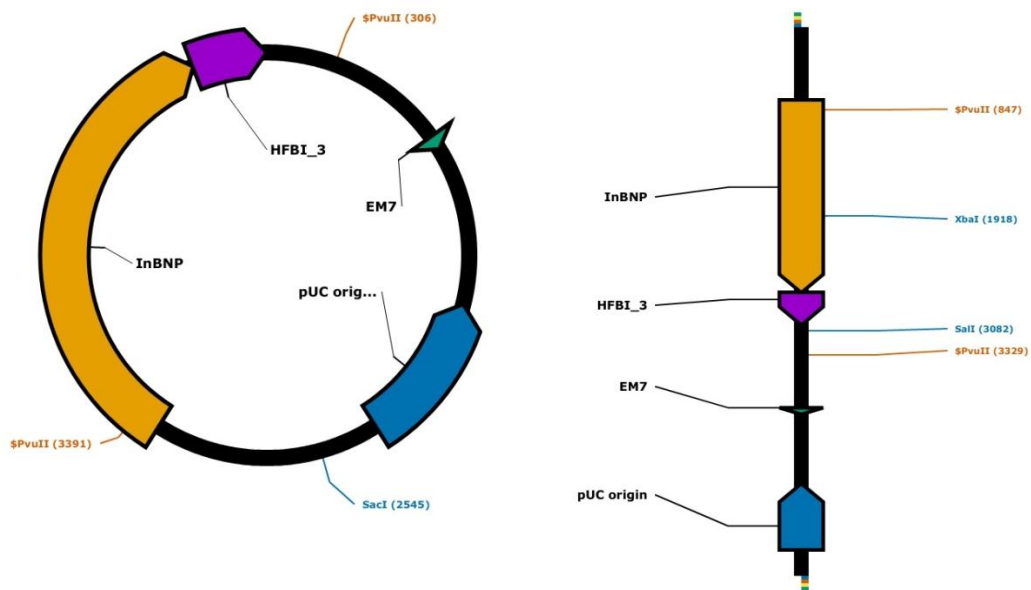
## 7 *PICHA PASTORIS* X33-KANNAN TRANSFORMOINTI

Ennen *Pichia pastoris*-kannan transformointia pitää vektori linearisoida, jotta geeni integroituisi tehokkaasti hiivan genomiin. Lisäksi elektrokompetentteja soluja tulee olla valmistettuna (Liite 1)

Linearisointi suoritettiin vektorista riippuen eri restriktioentsyymillä. PICZ A-vektorien katkaisemiseen käytettiin *SacI*-entsyymiä ja GAPZ A:lle *AvrII*-entsyymiä. *SacI* olisi katkaissut GAPZ A-vektorin translaation kannalta väärästä kohdasta. Restriktioentsyymien etsimiseen käytettiin avuksi sekä Invitrogenin *Pichian* transformointi-kitin manuaalia, että PlasmaDNA:ta, jossa voi virtuaalisesti digestoida DNA:ta (Kuvat 19.-20.).



Kuva 19. pGAPZ A-vektorien linearisointi tehtiin *AvrII*-entsyymillä.



InBNP + HFBI-3 in pPICZ A  
5567 bp

Digested InBNP + HFBI-3 in pPICZ A SacI  
5567 bp

Kuva 20. pPICZ A-vektorien linearsiointissa käytettiin SacI-entsyymiä.

Linearisoinnissa käytettiin Fermentaxen entsyymejä, digestioaika oli kolme tuntia ja lisäksi entsyymit lämpöinaktioitiin Fermentaxen suositusten mukaan 20 minuutin ajan, SacI +60 °C:ssa ja AvrII +80 °C:ssa.

Digestion jälkeen plasmidit puhdistettiin NucleoSpin® Gel and PCR Clean-Up-kitillä (Macherey-Nagel), ohjeen 5.1 mukaan. Saanto eluoiitiin 18–20 µl:an kitin NE-puskuria (5mM Tris/HCl, ph 8.5). Eristettyjen plasmidikolloidien DNA-pitoisuudet mitattiin NanoDropilla. Saannot liikkuvat välillä 10–14,6 µg, konsentraatiot kullekin linearisoidulle vektorille on esitetty taulukossa 5.

Taulukko 5. Linearisoitujen vektorien pitoisuudet ja saannot.

<b>Vektori</b>	<b>Pitoisuus (ng/µl)</b>	<b>Saanto (µg)</b>
<b>PICZ 1.</b>	730	14,6
<b>PICZ 2.</b>	730	14,6
<b>PICZ 3.</b>	570	11,4
<b>GAPZ 1.</b>	500	10
<b>GAPZ 2.</b>	590	11,8
<b>GAPZ 3.</b>	570	11,4

Transformointiin käytettiin 10 µl:a puhdistettua linearisoitua vektoria/kyvetti. Kyvetteinä *Pichian* transformoinnissa käytettiin *E. coli*n transformaatiosta poiketen BTX:n 2 mm:n kyvettejä (Electroporation Cuvettes Plus™ Model No 620, 2mm Gap Cuvette, BTX). Kyvettiin lisättiin suoraan vektoriliuoksen päälle 80 µl *Pichia pastoris* X-33 elektrokompetentteja soluja (Liite 1), jonka jälkeen seosta inkuboitiiin jäällä 5 minuuttia ennen elektroporaatiota. Laitteen asetukset *Pichian* transformoinnissa olivat low range 200 Ω, 25 µF kapasitanssi ja jännite 1.5 kV. Välittömästi sähköpulssein jälkeen solususpensioon lisättiin 1 ml steriiliä 1M D-sorbitoli-liuosta ja solususpensiota inkuboitiiin 1-2 tuntia + 30 °C:ssa.

Solususpensiosta maljattiin 50 µl/malja, joita inkuboitiiin + 30 °C:ssa kunnes pesäkkeet olivat sopivan kokoisia, käytännössä 2-4 päivää. Pesäkkeistä valittiin ulkonäön perusteella ja poimittiin 6 kpl/konstrukti uusille maljoille, joiden kanssa jatkettiin seuraavaan vaiheeseen. Maljauskuvio oli sama kuin kuvassa 7.

## 8 *PICHIA PASTORIS*-TUOTTOKANTOJEN SEULONTA

Viimeiset kaksi konstruktia, PICZ 3 ja GAPZ 3., jäivät ajan loppumisen vuoksi tähän vaiheeseen eli niitä ei ehditty seulomaan tämän opinnäytetyön sisällä, jolloin vain neljä konstruktia eli PICZ 1.- ja 2.- sekä GAPZ 1.- ja 2.-kannat seulottiin.

Kaikista transformanteista otettiin kuusi erillistä pesäkettä seulontavaiheeseen. Kannoista tehtiin 25 ml:n testikasvatukset, joista otettiin solulysaattinäytteet sekä kasvatusmediumin supernatanttinäytteet InB -nukleoproteiinin määrittämiseen.

PICZ A-kannoista tehtiin esikasvatus BMGY-mediumissa (Liite 1), jotka indusoitiin vuorokauden kasvatuksen jälkeen metanolilla käyttäen BMMY-mediumia (Liite 1). Metanolia lisättiin tämän jälkeen päivittäin niin, että kasvatuksen metanolin loppukonsentraatio oli 0,5 %. Kasvatusta jatkettiin 192 tuntia. GAPZ A-kannan kasvatus aloitettiin suoraan YPDS-mediumissa (Liite 1) ja kasvatusta jatkettiin suoraan 192 tuntia. Kasvatukset suoritettiin huoneen lämpötilassa, 250 rpm.

Transformoiduista *Pichia pastoris*-kannoista haluttiin selvittää ne kannat, jotka tuottavat influenssa B-viruksen nukleoproteiinia. Nukleoproteiinia voidaan käyttää positiivisena kontrollina diagnostiikassa, joten seulonta suoritettiin suoraan ArcDian MariPOC-laitteella, jota käytetään terveydenhuollossa taudinaiheuttajan määrittämisessä.

Konstruktien nimeämisessä siirryttiin HydroBody-projektin käyttämään numerointijärjestelmään, jota käytettiin näytteiden nimeämiseen ArcDian MariPOC-systeemiä käytettäessä. Numerokoodien avain on liitteenä 6.

MariPOC-mittauslaitteen toiminta perustuu ArcDia™ TPX (two-photon excitation) -teknologiaan. Teknologia perustuu antigeenin muodostamaan immuunkompleksiin mikropartikkelin pinnan vasta-aineiden ja vapaiden leimattujen vasta-aineiden välille. Mikropartikkelin pintaan on kiinnitetty useita (tuhansia) primäärivasta-aineita, jotka voivat sitoutua InB-viruksen nukleoproteiiniin, jolloin

leimatut sekundäärivasta-aineet kiinnittyvät nukleoproteiiniin konsentroituen mikropartikkelin pintaan. (Koskinen et al, 2005)

Käytännössä edellä kuvattu sandwich-reaktio tapahtui sekoittamalla reagensiseos, johon oli sekoitettu vasta-aineella päällystetyt mikropartikkelit sekä leimattu vasta-aine. Näyte ja edellä mainittu seos pipetoitiin 384-kuoppalevyille 1:1-suhteessa. Tässä tapauksessa näytteistä oli tehty 10- ja 1000-kertaiset laimennokset, jolloin lopullisen TPX-laitteella mitattavan näytteen pitoisuus oli 20- tai 2000-kertainen. Mikropartikkelista laser-säteen avulla saatu siroava signaali sekä fluoresenssi mitattiin. Mittauslaitteen data on liitteenä 7.

## 9 TYÖN ARVIOINTI, TULOKSET JA PÄÄTELMÄT

Työn lähtökohdat olivat selkeät ja HydroBody-projektin tavoitteet sisältyivät työn suoritukseen kiinteästi. Tavoitteena oli luoda influenssa B-viruksen nukleoproteiinia tuottava *Pichia pastoris*-kanta, jolloin itse kohdeproteiinin tuli tuottua oikeassa muodossa ja riittävässä määrin.

Työn onnistumisen kannalta kriittisiä vaiheita olivat sekvenssien optimointi, oikeanlaisten vektorien aikaansaaminen sekä onnistunut *Pichia pastoriksen* transformointi, jolloin rekombinantit kannat saatiin tuottamaan oikeaa proteiinia.

Nukleoproteiinin käytettävyyden kannalta oli tärkeää, että proteiini tuottuu muodossa, jota voidaan tulevaisuudessa käyttää hyväksi diagnostiikassa ja muissa sovelluksissa virusviljelmällä tuotetun proteiinin sijaan. Proteiinin tuottuminen sekä sen käyttökelpoisuus testattiin ArcDian MariPOC-InB-mittausmenetelmällä.

### 9.1 Työ

Sekvenssien optimoinnissa käytettiin hyväksi useita eri sekvenssinmanipulointiohjelmia ja sekvenssit tarkastettiin useampaan otteeseen ennen varsinaisten sekvenssien tilaamista, jolloin varmistuttiin niiden oikeellisuudesta.

Työn kulku oli aina riippuvaista edellisen vaiheen onnistumisesta, geenin eristyksestä pUC57-plasmidista aina vektorien luomiseen ja *Pichia pastoriksen* transformaatioon tapahtui ketjussa.

Oikeanlaisten geenien monistaminen *E. coli*ssa pUC57-plasmidissa ja geenin eristäminen plasmidista onnistui jokaisen geenin kohdalla oletetusti hyvin tuloksin. Jokaisen geenin saanto oli kohtalainen. Digestiota seuraavan AGE:n ja NanoDrop-mittausten perusteella pääteltiin ne kannat, joilla saatiin suurimmat saannot oikeanlaista DNA:ta.

Ennen geenien ligaatiota PICZ A- ja GAPZ A-vektoreihin, rekombinantti-DNA:n tuoton tasolla *E. colissa* oli merkitystä lähinnä työn tehokkuuden, ei niinkään työn onnistumisen kannalta. Koska jokaista geeniä tuotettiin useammalla kloonilla, joista seulottiin paras kloonin jatkoon, päästiin lopputulokseen eli vektoriin, jonka jokainen osa oli läpäissyt seulonnan. Vaikka joissain tapauksissa *E. colista* eristettyä geeniä ei digestoitu ja näin voitu todistaa sen oikeellisuutta, pääteltiin AGE:lla tutkitun geenin koosta, että se oli oikeanlainen.

Valmiita vektorikonstrukteja monistettiin *E. colissa* ennen *Pichian* transformointia ja vektoreille suoritettiin samanlaiset varmistukset digestion ja AGE:n sekä tuottotasojen perusteella, jolloin myös *Pichian* elektroporaatiossa käytetty vektori oli paras mahdollinen valmistetuista.

Suurin työn tehokkuuteen vaikuttava työvaihe oli geelieristys, jossa oletetun DNA:n määrä lopullisessa eluoidussa näytteessä jäi alle odotusten. Tämä johtuu työvaiheen luonteesta, jossa aluksi geeni leikattiin geeliltä skalpellilla. Leikkaus vaatii tarkkuutta, jottei näytteeseen tule ylimäärin geeliä huonontamaan puhdistusvaiheen tehokkuutta, mutta samalla koko vyöhyke, jolla geeni on, tulee saada mukaan, jottei DNA:ta menisi hukkaan.

Myös geelieristyksen saannot vaihtelivat suuresti eri geenien välillä, mikä johtuu geenien koosta. Influenssa B-viruksen NP:n koko (1955 bp) on moninkertainen verrattuna hydrofobiineihin (~300 bp), joten on ymmärrettävää, että myös eristetyn NP-DNA:n määrä on moninkertainen verrattuna pienemmän molekyylin DNA:n määrään, kun plasmidin määrä on suunnilleen sama. Tietenkin myös esikasvatuksella, sen kestolla sekä eluointilavuudella on merkitystä eristetyn DNA:n määrään.

Lopullisen vektorin toimivuuteen vaikuttavat kaikkien työvaiheiden lisäksi valittu *Pichia*-kanta, johon vektori transformoidaan sekä transformoinnin eli tässä tapauksessa elektroporaation onnistuminen.

Lopuksi todettakoon, että ArcDian MariPOC:iin perustuva seulontamenetelmä ei takaa, että fuusioproteiini tuottuu surfaktanttipohjaisen kaksi-faasiuuton kannalta oikeanlaisena, nukleoproteiinin ja hydrofobiinin kompleksina, sillä Ma-

riPOCin InB-testimenetelmä mittaa vain influenssa B-viruksen nukleoproteiinin määrää näytteessä. Ennen työn jatkamista kaksi-faasiuuttojen suunnitteluun, tulisi *Pichia*-kantojen tuottamien proteiinien rakenne määrittää.

## 9.2 ArcDian MariPOC influenssa B-virus-testin tulokset

Testissä käytetyn positiivisen kontrollin signaali on tasoa, joka voitiin maksimissaan käytetyllä valmisteella tällä laitteella saavuttaa. Positiivisen kontrollin signaali 2 tunnin aikapisteessä tasapainotilanteessa oli 117, josta detektorajaksi kompensoiduille näytesignaaleille määrytyi 1,17 tasapainotilanteessa. Signaali oli verrannollinen näytteessä olleen influenssa B-viruksen nukleoproteiinin määrään.

Koska jokaisen kannan kohdeproteiini on influenssa B-viruksen nukleoproteiini ja kannat eroavat toisistaan merkkiproteiinin suhteen, on vertailu ja työn tulokset esitetty tämän merkkiproteiinin mukaan, vaikka todellisuudessa mittauksissa on mitattu nimenomaan nukleoproteiinin määrää kussakin näytteessä. Mittausten data ja siitä muotoillut kuvaajat ovat liitteenä 7.

Fuusioproteiini tuottui kaikilla kannoilla pääasiassa solunsisäisesti, lymfaattinäytteistä saatu signaali oli aina suurempi kuin supernatantinäytteiden signaali. Kuvaajat (Liite 7) on tehty tästä syystä vain lymfaattinäytteistä. Eroja voi kuitenkin vähäisessä määrin huomata joillain InB+HFBI\_4-kannoilla, joissa supernatantista saatu signaali on korkeampi kuin muiden supernatantinäytteiden. Merkkiproteiinien ero oli nimenomaan retentiosignaali HDEL. HFBI\_3:n geenin HDEL oli lisätty ja HFBI\_4:n ei. Näin ollen HFBI\_4:n olisi pitänyt erittyä kasvatusmediumiin, mitä se ei ainakaan suuressa määrin ole näyttänyt tekevän. Tämä havainto viittaa mahdollisesti influenssa B viruksen nukleoproteiinin ja hydrofobiihin erkanemiseen toisistaan eritysvaiheessa.

Vertailtaessa kahta eri vektoria, PICZ A:ta ja GAPZ A:ta, voi nopeasti huomata PICZ A:ssa tuotettujen tasojen olevan noin 5-kertaisia verrattuna GAPZ A:n tuottotasoihin (Kuvaajat 1. ja 2. liitteessä 7)). Koska näytteet ovat testikasvusten loppunäytteitä, kumulatiivista tietoa tuottomääristä ajan suhteen ei ole saa-



tavilla. PICZ-kantojen korkeammat tuottotasot eivät kuitenkaan takaa parasta tuottokantaa, sillä kuten edellä on mainittu, proteiinien rakenne ei selviä tällä testimenetelmällä.

Vertailtaessa GAPZ A-vektorin molempien fuusioproteiinien tuottotasoja 1/20-laimennoksilla tasapainotilanteessa, voidaan päätellä, mitkä kannat ovat edullisimpia tuotannon kannalta. Kuitenkin on todettava, että PICZ A-vektorien tuottotasojen ollessa jälleen moninkertaisia verrattuna GAPZ A-vektoreihin, ei niiden keskinäistä paremmuusjärjestystä ole järkevää arvioida. Kaikilla PICZ A kannoilla saatiin sekä 1/20- että 1/2000-laimennoksillakin positiiviset tulokset InB-nukleoproteiinille. GAPZ A:lla tuotettujen lyaattinäytteiden 1/20-laimennoksilla positiivisen tuloksen antoivat molemmilla fuusioilla viisi kloonia ja 1/2000-laimennoksilla vain yksi HFBI\_3-fuusiokanta ja 4 HFBI\_4-fuusiokantaa.

Laitteiston herkkyuden vuoksi 1/20-laimennosten positiiviset tulokset eivät ole vertailukelpoisia keskenään. 1/2000-laimennosten vertailussa voidaan tässä tapauksessa päätellä parhaat tuottokannat kyseiselle rekombinanttiproteiinille. Yksi GAPZ 1.-kanta tuotti huomattavasti enemmän kuin muut GAPZ A-vektorit, mutta senkin signaali on vain hieman parempi kuin huonoimman PICZ 2.-kannan tuotto ja miltei kaksi kertaa huonompi kuin huonoimman PICZ 1.-kannan tuotto.

PICZ A-kannoista voitiin päätellä parhaat tuottokannat kyseisille kahdelle fuusioproteiinille. HFBI\_3-fuusioproteiinin tuotto oli kaikilla klooneilla suhteellisen tasaista, kuitenkin klooni numero 4. osoittaa vahvinta signaalia. HFBI\_4-fuusioproteiinille paras tuottokanta on klooni 6., jonka jälkeen klooni 5. antaa parhaan signaalin 1/2000-laimennoksella.

### 9.3 Päätelmät

Tulokset olivat sekä influenssa B-viruksen nukleoproteiinin tuoton että opinnäytetyön kannalta hyviä ja lupaavia, sillä seulotuilla kannoilla saatiin signaalia herkillä InB-testimenetelmällä. Näin ollen ainakin InB:n nukleoproteiinin tuotto

*Pichia pastoris*-hiivassa onnistui. Tuottuuko proteiini oikeassa muodossa eli InBNP-HFBI-fuusiona vai kahtena erillisenä proteiinina vaatii lisätutkimuksia.

Seulonnan tulosten perusteella voitiin päätellä, että PICZ A-vektori influenssa B viruksen nukleoproteiinin tuottovektorina *Pichia pastoris* X-33-kannassa on sekä InB+HFBI\_3- että InB+HFBI\_4- fuusioproteiinin kohdalla tehokkaampi kuin GAPZ A.

PICZ A-kannoista parhaaksi InB+HFBI\_3-proteiinin tuoton kannalta pääteltiin kloni numero 4. ja InB+HFBI\_4-proteiinille kloni numero 6.

#### **9.4 Lopuksi**

Haluan kiittää HydroBody-projektin henkilökuntaa kärsivällisyydestä, erityisesti Joonas Terävälle kiitos avusta AGE-geelien puitteissa, joita en raskauteni vuoksi voinut käsitellä.

Juustolle ja litulle kiitos tuesta ja rakkaudesta.

Valmistumisen loppukiri-kiitokset Henna Anderssonille ja Ilmari Äijölle.

## LÄHTEET

- 1 Assaad F. A., Brés P., Chi-Ming Chu, Dowdle W. R., Kendal A. P., Kilbourne E. D., Kit Lam Sai, Milouchine V., Nejmi S., Noble G., Oya A., Palese P., Schild G. C., Scholtissek G., Skehel J. J., Webster R. G., Yakhno M., 1979, A revision of the system of nomenclature for influenza viruses: a WHO Memorandum, Bulletin of the World Health Organization, **58**, 585-591
- 2 Amiruddin N., Lee X.-W., Blake D. P., Suzuki Y., Tay Y.-L., Lim L.-S., Tomley F. M., Watanabe J., Sugimoto C., Wan K.-L., 2012, Characterisation of full-length cDNA sequences provides insights into the *Eimeria tenella* transcriptome, BMC Genomics 2012, **13**:21, saatavilla verkossa: <http://www.biomedcentral.com/1471-2164/13/21>
- 3 Arnau J., Lauritzen C., Petersen G. E., Pedersen J., 2005, Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins, Protein expression & purification, **48**, 1-13
- 4 Bimbo L. M., Mäkilä E., Raula J., Laaksonen T., Laaksonen P., Strommer K., Kauppinen E. I., Salonen J., Linder M., Hirvonen J., Santos H. A., 2011, Functional hydrophobin-coating of thermally hydrocarbonized porous silicon microparticles, Biomaterials, **32**, 9089-9099
- 5 Bitter G, Chen K., Banks A., Lai P.-H., 1984, Secretion of foreign proteins from *Saccharomyces cerevisiae* directed by  $\alpha$ -factor gene fusions, Biochemistry, **81**, ss. 5330-5334
- 6 Condon A., Thachuk C., 2012, Efficient codon optimization with motif engineering, Journal of Discrete Algorithms, **16**, ss 104-112
- 7 Engler C, Kandzia R., Marillonnet S., 2008, A One Pot, One Step, Precision Cloning Method with High Throughput Capability, PLoS ONE **3**; e3647
- 8 Fang L., Jia K.-Z., Tang Y.-L., Ma D.-Y., Yu M., Hua Z.-C., 2007, An improved strategy for high-level production of TEV protease in *Escherichia coli* and its purification and characterization, Protein Expression and Purification, **51**, 102-109
- 9 GISAID-influenssatietokanta verkossa, [www.gisaid.org](http://www.gisaid.org), – About GISAID, viitattu 1.11.2012
- 10 Gomord V., Wee E., Faye L., 1998, Protein retention and localization in the endoplasmic reticulum and the Golgi apparatus, Biocimica, **81**, 607-618
- 11 Grote A., Hiller K., Scheer M., Münch R., Nörtemann B., Hempel D. C., Jahn D., 2005, JCat: a novel tool to adapt codon usage of a target gene to its potential expression host, Nucleic Acids Research, **33** (Web Server issue) W526-W531
- 12 Jeltsch A., 2002, Maintenance of species identity and controlling speciation of bacteria: a new function for restriction/modification systems?, Gene, **317**, 13-16
- 13 Koskinen J., Vaarno J., Vainionpää R., Meltola N. J., Soini A. E., 2005, A novel separation-free assay technique for serum antibodies using antibody bridging assay principle and two-photon excitation fluorometry, Journal of Immunological Methods, **309**, ss 11-24
- 14 Kozak M., 1999, Initiation of translation in prokaryotes and eukaryotes, Gene, **234** (1999) 187-208

- 15 Linder M., Qiao M., Laumen F., Selber K., Hyytiä T., Nakari-Setälä T., Penttilä M. E., 2004, Efficient Purification of Recombinant Proteins Using Hydrophobins as Tags in Surfactant-Based Two-Phase Systems, *Biochemistry*, **43**, ss. 11873-11882
- 16 Linder M., Szilvay G., Nakari-Setälä T., Penttilä M., 2005, Hydrophobins: the protein-amphiphiles of filamentous fungi, *FEMS Microbiology reviews*, **29**, 877-896
- 17 Macauley-Patrick S., Fazenda M. L., McNeil B., Harvey L. M., 2005, Heterologous protein production using the *Pichia pastoris* expression system, *Yeast*, **22**, ss. 249-270
- 18 Paus E. J., Willey J., Ridge R. J., Legg C. R., Finkelman M. A., Novitsky T. J., Ketchum P. A., 2002, Production of recombinant endotoxin neutralizing protein in *Pichia pastoris* and methods for its purification, *Protein Expression and Purification*, **26**, 202-210
- 19 Pichia Expression Kit For Expression of Recombinant Proteins in *Pichia pastoris* User Manual, 2010, Invitrogen, Catalog no. K1710-01, Revision date: 07 September 2010, published online in Wiley InterScience
- 20 PlasmaDNA sivusto, <http://research.med.helsinki.fi/plasmadna/>, viitattu 30.11.2012
- 21 Potvin G., Ahmad A., Zhang Z., 2010, Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: A review, *Biochemical Engineering Journal* **64** (2012) 91-105
- 22 Pyhälä R., Ziegler T., 2005, Influenssavirukset. Kirjassa: Huovinen P., Meri S., Peltola H., Vaara M., Vaheri A., Valtonen V., *Mikrobiologia ja infektiosairaudet Kirja I*, 413- 428, Duodecim, Gummerus Jyväskylä
- 23 Szilvay G. R., Kisko K., Serimaa R., Linder M., 2007, The relation between solution association and surface activity of the hydrophobin HFBI from *Trichoderma reesei*, *FEBS Letters*, **581**, 2721-2726
- 24 Sreekrishna Koti, Brankamp R. G., Kropp K. E., Blankenship D. T., Tsay Jiu-Tsiar, Smith P. L., Wierschke J. D., Subramaniam A., Birkenberger L. A., 1996, Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*, *Gene* **190** (1997) 55-62
- 25 Vassileva A., Chugh D. A., Swaminathan S., Khanna N., 2000, Expression of hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris* using the GAP promoter, *Journal of biotechnology* **88** (2001) 21-35
- 26 Vincze, T., Posfai, J., Roberts, R.J., 2003, NEBcutter: a program to cleave DNA with restriction enzymes, *Nucleic Acids Research*, **31**, 3688-3691
- 27 Wösten H., de Vocht M., 2000, Hydrophobins, the fungal coat unravelled, *Biochimica et Biophysica Acta* **1469** (2000) 79-86

## Elektrokompetenttien solujen valmistus ja mediumit

### *Escherichia coli*

Siirrostetaan silmukallinen *Escherichia coli*-kanta XL-1 Blue 5 ml:n LS-LB-mediumia. Inkubointi O/N + 37 °C:ssa, 250 rpm.

Seuraavana aamuna 5 ml:n esikasvatus siirrostetaan 200 ml:n LS-LB-mediumia. Inkubointi +37 °C:ssa, 250 rpm, kunnes OD<sub>600</sub> ~0,7. Kun OD<sub>600</sub> saavuttaa halutun arvon siirretään kasvatus jäille. Tästä eteenpäin työ suoritetaan esiviilennetyin tarvikkein jäillä.

1. Sentrifugointi 2000 g, + 4 °C, 15 min. Kaada supernatantti pois.
2. Suspensointi 200 ml:n kylmää steriiliä vettä. Sentrifugointi 2000 g, + 4 °C, 15 min. Kaada supernatantti pois.
3. Liuotetaan pelletti jäljelle jäävään nesteeseen ja lisätään noin 25 ml kylmää steriiliä 20 % glyserolia. Sentrifugoi 2000 g, 15 min, +4°C. Kaada supernatantti pois.
4. Suspensoi pelletti 500 µl:aan kylmää steriiliä 20% glyserolia.
5. Jaetaan suspensio 110 µl:n eriin eppareihin.
6. Jäädytetään putket sisältöineen nestetyypellä. Upota epparit muutaman sekunnin ajaksi nestetyyppeen, kunnes suspensio on täysin jäätynyt. Siirrä putket -80 °C:n pakkaseen odottamaan käyttöä.

### *Pichia pastoris*

Käytetty *P. pastoris*-kanta X-33.

Tehdään 50 ml:n esikasvatus YPD-mediumiin (Liite 1), lisätään noin silmukallinen X-33:sta, inkubointi 250 rpm, + 30 °C, n. 24 h.

Mitataan esikasvatuksen absorbanssi. Siirrostetaan esikasvatus 1000 ml:an RT YPD-mediumia ja lasketaan kasvatuksen teoreettinen absorbanssi. Jatketaan

inkubointia 250 rpm, +30 °C, mitataan absorbanssia ja kun se on noin 1, siirrytään seuraavaan vaiheeseen.

Kerätään solut sentrifugoimalla 2000 g, 5 min, +4 °C. Kaadetaan supernatantti pois.

Suspensoidaan solut 200 ml:aan YPD -medium + HEPES (RT).

Lisätään 5 ml STERIILIÄ 1 M DTT:tä.

Inkubointi +30 °C, 15 min.

→ Tästä eteenpäin työ suoritettiin jäällä.

Lisätään noin 1000 ml kylmään steriiliä milli-Q H<sub>2</sub>O:ta.

Sentrifugointi 2000 g, 5 min, +4 °C. Kaadetaan supernatantti pois.

Suspensoidaan solut noin 250 ml:aan kylmään steriiliä milli-Q H<sub>2</sub>O:ta

Sentrifugointi 2000 g, 5 min, +4 °C. Kaadetaan supernatantti pois.

Suspensoidaan solut 20 ml:aan kylmää steriiliä 1 M sorbitolia.

Sentrifugointi 2000 g, 5 min, +4 °C. Kaadetaan supernatantti pois.

Suspensoidaan solut 0,5 ml:aan kylmää steriiliä 1 M sorbitolia.

Jaetaan kompetentit solut 170 µl:n eriin eppareihin ja pakastetaan suoraan -80 °C pakkasessa tai käytetään välittömästi.

## **Mediumien valmistusohjeet**

Seuraavien mediumien reseptit löytyvät Invitrogenin EasySelect™ Pichia Expression Kitin käyttäjämanuaalista, sivunumerot on mainittu mediumkohtaisesti (Kit: Cat. no. K1740-01, Rev Date 18 June 2010, Manual part no. 25-0172).

**Low salt LB-medium ja agar**

Sivu 53.

**YPD(S) –medium**

Sivut 55 – 56.

YPD -medium + HEPES on valmistettu lisäämällä YPD-mediumiin 20 til- % 1 M HEPES-puskuria (pH 8).

**BMMY/BMGY –medium**

Sivut 55 ja 58.

**SOC –medium**

Tryptoni	20 g
Hiivauute	5 g
NaCl	0.5 g
1M KCl	2.5 ml
KO vesi	1000 ml

Jaa sopivan suuruisiin eriin (esim. 50 ml)

Steriloi 121 ° C, 20 min → jäähditys RT

Lisätään 20 ml steriiliä 1 M glukoosia / 1000 ml mediumia juuri ennen käyttöä. Heitä käyttämättä jäänyt SOC -medium, mihin on lisätty glukoosi pois.

## Työohjeet

### ***E. coli*-esikasvatukset NucleoSpin® Plasmid ja NucleoBond® Xtra Midi-kittejä varten**

Miniprepin esikasvatus:

LS-LB-medium	5 ml
Zeosiini/Ampisilliini	25/100 µg/ml
1 pesäke seulottavaa bakteeria	

Midiprepin esikasvatus (jatketaan miniprepin esikasvatuksesta)

LS-LB-medium	150 – 200 ml
Zeosiini/Ampisilliini	25/100 µg/ml

siirrostetaan 5 ml:n esikasvatus mediumiin

Esikasvatukset voi käyttää plasmidineristykseen kun OD<sub>600</sub> on noin 1.

### **Genscriptiltä saapuneiden plasmidistokkien käsittely**

Plasmidi toimitetaan kylmäkuivattuna pellettinä mikrosentrifugiputkessa.

1. Sentrifugointi 6000g, +4 °C, 1 min
2. Lisää 20 µl steriiliä vettä, vorteksoi. Pitoisuus 200 ng/ml.

Elektroporaatioon tarvittava DNA:n määrä 50 ng/ml.

Pipetoidaan 1 µl liuotettua plasmidia ja 3 µl steriiliä vettä, jolloin 1 µl vastaa 50 ng:a DNA:ta.



## Projektin sekvenssipohja uusille inserteille

```
GGTCTCGAATTCAACCATGAGATTCCCATCTATCTTCACTGCTGTTTTGTT 50
CGCTGCTTCTTCTGCTTTGGCTGCTCCAGTTAACACTACTACTGAAGACG 100
AAACTGCTCAAATCCCAGCTGAAGCTGTTATCGGTTACTCTGACTTGAA 150
GGTGACTTCGACGTTGCTGTTTTGCCATTCTCTAACTCTACTAACAACGG 200
TTTGTTGTTTCATCAACACTACTATCGCTTCTATCGCTGCTAAGGAAGAAG 250
GTGTTTCTTTGGAAAAGAGAGAAGCTGAAGCT Tähän kohdeproteiinin
```

optimoitu DNA-sekvenssi TTGGAGAGACC

### Selitykset:

Alleiviivattu KOZAK

*S. cerevisiae*lle optimoitu prepro- $\alpha$ -faktori sinisellä

**BsaI:n tunnistuskohta**

BsaI-entsyymien katkaisukohta on kuvattu alla:

```
5' ---GGTCTCN  NNNN--- 3'
3' ---CCAGAGNNNNN  --- 5'
```

EcoRI:n tunnistuskohta:

```
5' GAATTC
3' CTTAAG
```

ja katkaisukohta:

```
5' ---G °  AATTC--- 3'
3' ---CTTAA °  G--- 5'
```

NotI:n tunnistus- ja katkaisukohta:

```
5' GC °  GGCCGC3'
3' CGCCGG °  CG5'
```

Entsyymi siis katkaisee geenin sekvenssipohjan alussa olevan punaisella merkityn alueen jälkeen. Sekvenssin lopussa entsyymien tunnistuskohta on "väärinpäin", jolloin DNA katkeaa tunnistuskohtaa ennen. Näin geenin päihin jäävät DNA-jaksot -AATT (alussa) ja TTGG- (lopussa), jotka ligaatiossa liittyvät EcoRI:lla ja NotI:lla linearisoituun vektoriin.

# Sekvenssien optimointi ja analysointi

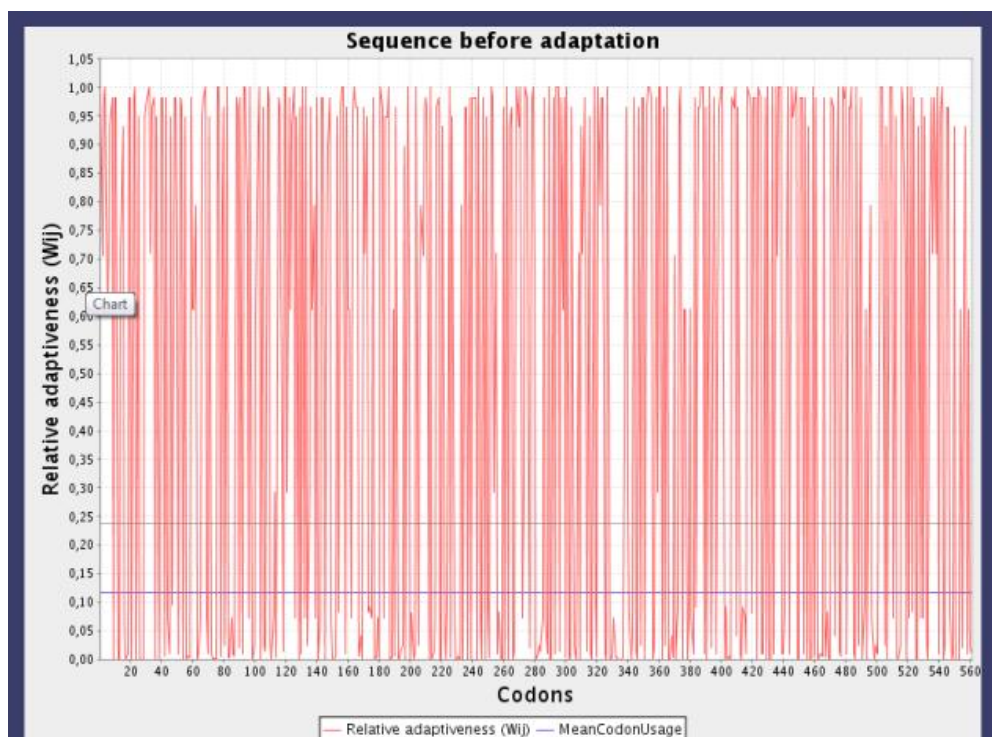
## JCat –sekvenssinoptimointiohjelma kuvina

The image shows the 'Codon-Adaptation' web interface. It is divided into four main sections:

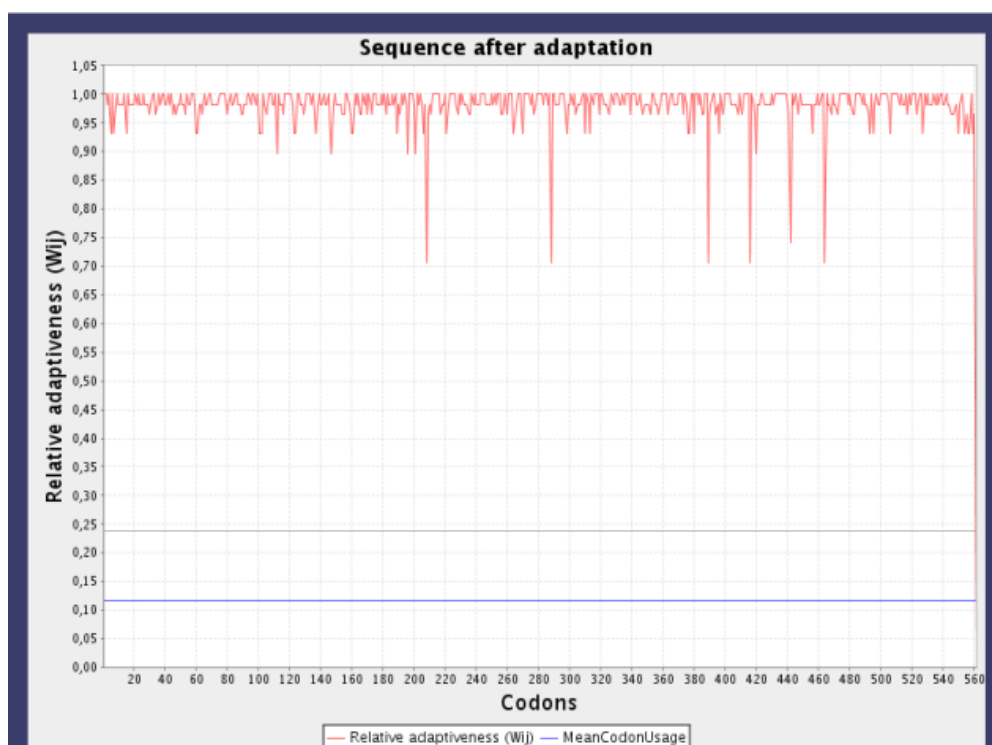
- 1. Type/paste sequences below:** A large empty text input area for pasting a sequence.
- 2. Specify the pasted Sequence:** Two radio buttons: 'DNA/RNA Sequence' (selected) and 'Protein Sequence'.
- 3. Select organism:** A dropdown menu with 'Saccharomyces cerevisiae' selected.
- 4. Additional Options:** A list of checkboxes:
  - Avoid rho-independent transcription terminators.
  - Avoid prokaryotic ribosome binding sites.
  - Avoid Cleavage Sites of Restriction Enzymes: A dropdown menu is open showing 'SacI', 'SacII', 'Sall', 'SapI', and 'Sau96I'.
  - Only partly optimization in order to apply site directed mutagenesis.

At the bottom right, there are two buttons: 'Submit' and 'Reset'.

Kuva 21. Jcat sekvenssinmuokkaussivuston aloitussivu.



Kuva 22. Optimoitavan sekvenssin soveltuvuus *Saccharomyces cerevisiae* kodoninkäyttöön ennen kodonioptimointia.



Kuva 23. Sekvenssin soveltuvuus *S. cerevisiae* kodoninkäyttöön optimoinnin jälkeen.

## NEBcutter V2.0 kuvina



## NEBcutter V2.0

[Program Guide](#)
[Help](#)
[Comments](#)
[\[NEB homepage\]](#)

This tool will take a DNA sequence and find the large, non-overlapping open reading frames using the E. coli genetic code and the sites for all Type II and commercially available Type III restriction enzymes that cut the sequence just once. By default, only enzymes available from NEB are used, but other sets may be chosen. Just enter your sequence and "submit". Further options will appear with the output. **The maximum size of the input file is 1 MByte, and the maximum sequence length is 300 KBases.**

[What's new in V2.0](#)   [Citing NEBcutter](#)

Local sequence file:  [Browse...](#)

GenBank number:  [\[Browse GenBank\]](#)

or paste in your DNA sequence: *(plain or FASTA format)*

Standard sequences:

# Plasmid vectors ▾

# Viral + phage ▾

[Submit](#)

[More options](#)

[Set colors](#)

The sequence is:  Linear    Circular

Enzymes to use:

- NEB enzymes
- All commercially available specificities
- All specificities
- All + defined oligonucleotide sequences
- Only defined oligonucleotide sequences

[\[define oligos\]](#)

Minimum ORF length to display:  a.a.

Name of sequence:  *(optional)*

**Earlier projects:**

*Note: Your earlier projects will be deleted 2 days after they were last accessed. You need to have cookies enabled in your browser for this feature to work.*

Disable NEBcutter cookies

[Delete projects](#)

## Kuva 24. NEBcutter.sekvenssinalysointi

**NEW ENGLAND Biolabs NEBcutter**

**Linear Sequence: unnamed sequence**

[Help](#) [Comments](#)

**Display:** - All single cutter restriction enzymes  
- Main non-overlapping, min. 100 aa ORFs

GC=43%, AT=57%

1 174 aa 1973

Cleavage code	Enzyme name code
▬   blunt end cut	Available from NEB
▬   5' extension	Has other supplier
▬   3' extension	Not commercially available
▬   cuts 1 strand	*: cleavage affected by CpG meth.
	#: cleavage affected by other meth.
	(enz.name): ambiguous site

**Main options**

- New DNA
- Custom digest
- View sequence
- ORF summary
- Save project
- Print

**Availability**

- NEB
- All commercial

**Display**

- 2 cutters
- 3 cutters

**Zoom**

- Zoom in
- More...

**List**

- 0 cutters
- 1 cutters
- All sites
- Save all sites
- Flanking enzymes

Minimum ORF length to display: 100 a.a.

Kuva 25. NEBcutterista saatava analysoidun sekvenssin restriktioentsyymien tunnistuskohta-kartta.

**2 cutters**  
*unnamed sequence*

Number of cuts =

#	Enzyme	Specificity	Sites & flanks	Cut positions (blunt - 5' ext. - 3' ext.)
1	<a href="#">AlwI</a>	GGATCNNNN <sup>*</sup> N <sub>1</sub>	<a href="#">list</a>	#597/598, #1299/1300
2	<a href="#">BaeI</a>	<sub>1</sub> (N) <sub>5</sub> <sup>*</sup> (N) <sub>10</sub> ACNNNNGTAYC(N) <sub>7</sub> <sub>1</sub> (N) <sub>5</sub> <sup>*</sup>	<a href="#">list</a>	1016/1011+1049/1044, 1185/1180+1218/1213
3	<a href="#">BpuEI</a>	CTTGAG(N) <sub>14</sub> NN <sup>*</sup>	<a href="#">list</a>	1416/1414, 1491/1489
4	<a href="#">BsaBI</a>	GATNN <sub>1</sub> NNATC	<a href="#">list</a>	24, 1588
5	<a href="#">BsaI</a>	GGTCICH <sup>*</sup> NNNN <sub>1</sub>	<a href="#">list</a>	7/11, 1962/1966
6	<a href="#">BsrGI</a>	T <sup>*</sup> GTAC <sub>1</sub> A	<a href="#">list</a>	824/828, 1367/1371
7	<a href="#">BtsI</a>	GCAGTG <sub>1</sub> NN <sup>*</sup>	<a href="#">list</a>	35/33, 1931/1929
8	<a href="#">CjeI</a>	<sub>1</sub> (N) <sub>6</sub> <sup>*</sup> (N) <sub>5</sub> CCA(N) <sub>6</sub> GT(N) <sub>9</sub> <sub>1</sub> (N) <sub>6</sub> <sup>*</sup>	<a href="#">list</a>	951/945+985/979, 1608/1602+1642/1636
9	<a href="#">CjuII</a>	CAY(N) <sub>5</sub> CTC	<a href="#">list</a>	1374, 1407
10	<a href="#">CspCI</a>	<sub>1</sub> NN <sup>*</sup> (N) <sub>11</sub> CAA(N) <sub>5</sub> GTGG(N) <sub>10</sub> <sub>1</sub> NN <sup>*</sup>	<a href="#">list</a>	714/712+749/747, 1002/1000+1037/1035
11	<a href="#">EcoP15I</a>	CAGCAG(N) <sub>25</sub> <sup>*</sup> NN <sub>1</sub>	<a href="#">list</a>	10/12, 1906/1908
12	<a href="#">FaeI</a>	<sub>1</sub> (N) <sub>5</sub> <sup>*</sup> (N) <sub>8</sub> AAG(N) <sub>5</sub> CTT(N) <sub>8</sub> <sub>1</sub> (N) <sub>5</sub> <sup>*</sup>	<a href="#">list</a>	1503/1498+1535/1530, 1762/1757+1794/1789
13	<a href="#">HgaI</a>	GACGC(N) <sub>5</sub> <sup>*</sup> (N) <sub>5</sub> <sub>1</sub>	<a href="#">list</a>	*1420/1425, *1768/1773
14	<a href="#">Hin4I</a>	<sub>1</sub> (N) <sub>5</sub> <sup>*</sup> (N) <sub>5</sub> GAY(N) <sub>5</sub> VTC(N) <sub>8</sub> <sub>1</sub> (N) <sub>5</sub> <sup>*</sup>	<a href="#">list</a>	362/357+394/389, *838/833+870/865
15	<a href="#">HincII</a>	GTY <sub>1</sub> RAC	<a href="#">list</a>	81, 1173
16	<a href="#">MfeI</a>	C <sup>*</sup> AATT <sub>1</sub> G	<a href="#">list</a>	1312/1316, 1429/1433
17	<a href="#">PcsI</a>	WCGNNN <sub>1</sub> N <sup>*</sup> MNCGW	<a href="#">list</a>	*689/688, *1211/1210
18	<a href="#">RdeGBIII</a>	<sub>1</sub> NN <sup>*</sup> (N) <sub>5</sub> TGRYCA(N) <sub>8</sub> <sub>1</sub> NN <sup>*</sup>	<a href="#">list</a>	1163/1161+1189/1187, 1649/1647+1675/1673
19	<a href="#">SmlI</a>	C <sup>*</sup> TYRA <sub>1</sub> G	<a href="#">list</a>	1395/1399, 1470/1474
20	<a href="#">SstE37I</a>	CGAAGAC(N) <sub>18</sub> <sub>1</sub> NN <sup>*</sup>	<a href="#">list</a>	*1241/1239, *1490/1488
21	<a href="#">TatI</a>	N <sup>*</sup> GTAC <sub>1</sub> N	<a href="#">list</a>	824/828, 1367/1371
22	<a href="#">UbaF9I</a>	TAC(N) <sub>4</sub> RIGT	<a href="#">list</a>	306, 1755

Kuva 26. NEBcutterissa entsyymejä voi seuloa sekvenssistä löytyvien restriktioentsyymien tunnistus-/katkaisukohtien määrän mukaan.

## Ligaatiolaskuri

([http://www.insilico.uni-duesseldorf.de/Lig\\_Input.html](http://www.insilico.uni-duesseldorf.de/Lig_Input.html))

**LIGATION CALCULATOR**

Please provide the following information:

.....vector size (in bp):

vector amount (in ng):

.....insert size (in bp):

Please enter the molar vector :  insert ratio:

(normally a vector to insert ratio of 1 to 3 is used of cohesive end ligations. higher molar ratios can be used for blunt end ligations)

When pressing the "do calculation" button the tool calculates the required amount of insert DNA (in ng) resulting in the given molar ratio

Kuva 27. Ligaatiolaskuri.

## Näytteiden purkukoodi

HydroBody-projektissa käsitellään useita *Pichia pastoris*-kantoja, joten jokaiselle eri tuottoproteiinille, merkkiproteiinille sekä tuottovektorille on oma merkkinsä näytekohtaisessa koodissa joka koostuu kuudesta-seitsemästä numerosta. Tässä opinnäytetyössä on käsitelty 6 erilaista konstruktia, joiden numerokoodi purkautuu seuraavasti:

1. numero: 6 = Influenssa B viruksen nukleoproteiini
2. numero: 3 = HFBI\_3            4 = HFBI\_4            5 = Tag6
3. numero: 1 = PICZ A            2 = GAPZ A
4. numero: Lähtöpesäkkeen juokseva numero; 1, 2, 3, 4, 5...  
(Kuva 7.)
5. numero: 0 =indusoimaton    1 = indusoitu
6. numero: L = lyaatti            (P = pelletti            S = supernatantti)



## MariPOC Influenssa B virus-testin tulokset

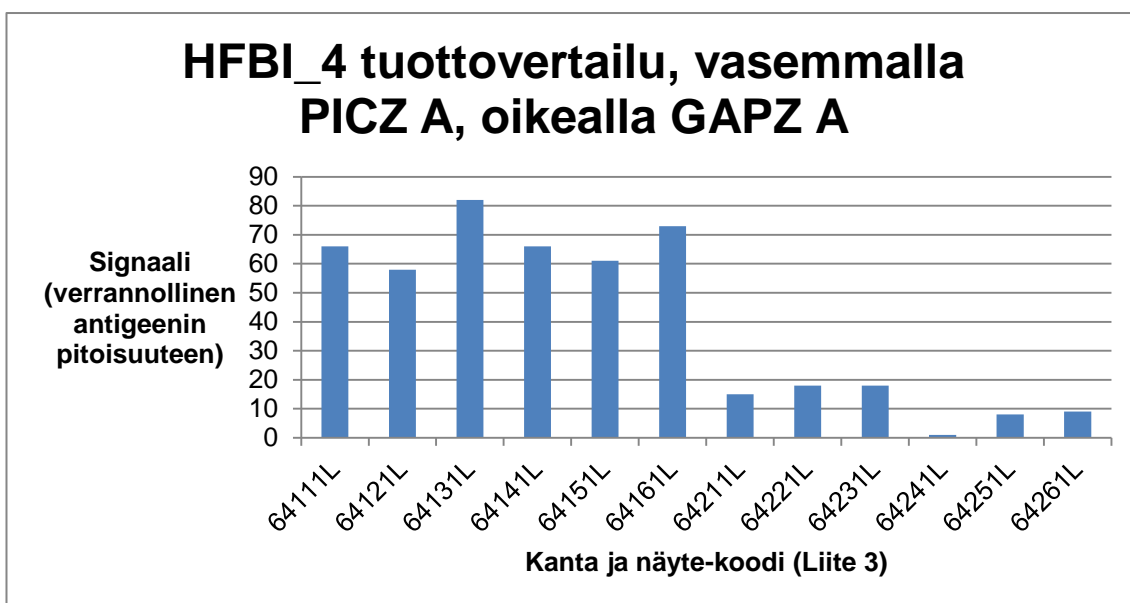
Ks. näytteiden purkukoodi Liite 3.

Näytteiden sijainti kuoppalevyllä												
	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos
	1	2	3	4	5	6	7	8	9	10	11	12
A	63111L	63111L	63111S	63111S	63121L	63121L	63121S	63121S	63131L	63131L	63131S	63131S
B	63141L	63141L	63141S	63141S	63151L	63151L	63151S	63151S	63161L	63161L	63161S	63161S
C	64111L	64111L	64111S	64111S	64121L	64121L	64121S	64121S	64131L	64131L	64131S	64131S
D	64141L	64141L	64141S	64141S	64151L	64151L	64151S	64151S	64161L	64161L	64161S	64161S
E	63211L	63211L	63211S	63211S	63221L	63221L	63221S	63221S	63231L	63231L	63231S	63231S
F	63241L	63241L	63241S	63241S	63251L	63251L	63251S	63251S	63261L	63261L	63261S	63261S
G	64211L	64211L	64211S	64211S	64221L	64221L	64221S	64221S	64231L	64231L	64231S	64231S
H	64241L	64241L	64241S	64241S	64251L	64251L	64251S	64251S	64261L	64261L	64261S	64261S
Suhteutetut mikropartikkeli-liuossignaalit (kompensoitu signaali) tasapainoilanteessa (2 tunnin aikapiste)												
"Mikropartikkelisignaali jaettuna liuossignaalilla"												
	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos
	1	2	3	4	5	6	7	8	9	10	11	12
A	72,78	2,10	1,61	1,01	97,70	2,11	1,24	1,06	74,91	2,13	1,62	0,92
B	78,67	2,67	1,74	0,83	71,67	2,39	1,23	1,04	84,37	2,40	2,83	1,01
C	66,06	1,76	1,31	1,01	57,59	1,30	1,22	1,02	81,91	1,48	4,50	1,12
D	66,41	1,97	1,42	1,00	61,44	2,19	1,55	0,92	73,03	2,48	1,58	0,99
E	10,57	1,01	1,12	0,98	11,39	1,07	1,30	0,94	19,37	1,33	1,34	1,04
F	18,25	1,12	1,36	1,04	5,88	1,16	1,09	1,04	1,10	1,01	1,06	0,99
G	15,13	1,18	1,55	1,00	18,25	1,20	1,61	1,12	18,11	1,17	4,52	1,05
H	0,98	0,90	0,99	0,99	7,54	0,98	1,10	0,99	9,38	1,18	1,15	1,11
Signaali on verrannollinen määrittelyssä olleen nukleoproteiinin pitoisuuteen.												
Kuopat, joissa nukleoproteiinin pitoisuus ylittää detektorirajan												
	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos	1/10 laimennos	1/1000 laimennos
	1	2	3	4	5	6	7	8	9	10	11	12
A	+	+	+	-	+	+	+	-	+	+	+	-
B	+	+	+	-	+	+	+	-	+	+	+	-
C	+	+	+	-	+	+	+	-	+	+	+	-
D	+	+	+	-	+	+	+	-	+	+	+	-
E	+	-	-	-	+	-	+	-	+	+	+	-
F	+	-	+	-	+	-	-	-	-	-	-	-
G	+	+	+	-	+	+	+	-	+	+	+	-
H	-	-	-	-	+	-	-	-	+	+	-	-
+	positiivinen näyte tässä määrittelyssä ->						Nukleoproteiinin pitoisuus yli detektorirajan					
-	negatiivinen näyte tässä määrittelyssä ->						Nukleoproteiinin pitoisuus alle detektorirajan					

Kuva 28. MariPOC-InB- määrittysten data.



Kuvaaja 1. InBNP+HFBI\_3-fuusioproteiinia tuottavien vektorien tuottovertailu.



Kuvaaja 2. InBNP+HFBI\_4-fuusioproteiinia tuottavien vektorien tuottovertailu.

