



Andy Bui

Fragment Analyzer -mikrobiyhteisöanalyysimenetelmän soveltuvuus teollisuuden prosessiolosuhteiden seurantaan

Metropolia Ammattikorkeakoulu

Laboratorioanalytiikka (AMK)

Laboratorioanalytiikka

Opinnäytetyö

21.11.2022

Tiivistelmä

| | |
|-------------------|--|
| Tekijä: | Andy Bui |
| Otsikko: | Fragment Analyzer -mikrobiyhteisöanalyysimenetelmän soveltuvuus teollisuuden prosessiolosuhteiden seurantaan |
| Sivumäärä: | 65 sivua |
| Aika: | 21.11.2022 |
| Tutkinto: | Laboratorioanalyttikko (AMK) |
| Tutkinto-ohjelma: | Laboratorioanalytiikka |
| Ohjaajat: | Lehtori Jarmo Palm Vanhempi tutkija Jaakko Ekman Teknikko Tinja Kanerva |

Opinnäytetyö suoritettiin Kemira Oyj:n Espoo R&D and Technologies -toimipisteessä Mikrobiologian ja Bioteknologian osastolla osana Metropolian Ammattikorkeakoulun laboratorioanalyttikon koulutusta. Tavoitteena oli selvittää, kuinka hyvin Fragment Analyzer -laite soveltuu mikrobiyhteisön seurantaan teollisuuden prosessiolosuhteissa. Menetelmä olisi nopeampi ja kustannustehokkaampi verrattuna sekvensointiin.

Projektissa valitut DNA-näytteet monistettiin PCR:llä ja analysoitiin Fragment Analyzer -laitteella. DNA-näytteitä oli 50. Monistettava geeni näytteissä oli 16S-rRNA:n V1–V2-hypervariaabelialue. Kyseinen geeni on kaikissa bakteereissa, ja menetelmä perustuu geenissä esiintyviin emäsparipituuden vaihteluihin eri bakteerien välillä. Fragment Analyzer -laite tunnistaa nämä erikokoiset monistustuotteet ja luo sormenjäljen, niin sanotun piikkiprofilin. Näitä piikkiprofileja analysoimalla ja seuraamalla voidaan määrittää, onko bakteeriyhteisö pysynyt samana tai onko se muuttunut. Menetelmä on siis yhteisösormenjälkimenetelmä, tarkemmin Length Heterogeneity PCR (LH-PCR).

Tuloksien perusteella todettiin, että menetelmä soveltuu yksinkertaisten mikrobiyhteisöjen seurantaan. Piikkiprofileista voidaan tunnistaa vallitsevia bakteereita sekvenssitietojen sekä edellisten tulosten avulla. Menetelmää ei suositella sellaisenaan laajempaan käyttöön, tai määrittämään kvantitatiivisesti bakteerien osuuksia mikrobiyhteisössä. Tuloksista saatuja osuuksia voidaan kuitenkin katsoa suuntaa antavana.

Pääongelmana on monistetun geenin pieni osuus PCR-tuotteissa, mikä näkyy tuloksia tarkastellessa. Menetelmän parantamiseksi suositellaan tutkimaan, jos eri alukkeet toimisivat paremmin, PCR-reaktioseoksen säätöä, mahdollisesti PCR-ajo-ohjelman optimoimista, jotta saataisiin monistetun DNA:n osuus korkeammaksi. Juuri nyt vallitseva osuus PCR-tuotteesta on alukedimeereitä. Näytteistä suositellaan myös tekemään rinnakkaisnäytteet, mutta sen tarpeellisuus voi muuttua, jos menetelmää saadaan parannettua luotettavammaksi.

Avainsanat: mikrobiyhteisö, Fragment Analyzer, geeni, piikkiprofiili, LH-PCR, sekvenssi

Abstract

Author: Andy Bui
Title: Suitability of Fragment Analyzer Application to Monitor Microbial Community Changes in Industrial Processes
Number of Pages: 65 pages
Date: 21 November 2022

Degree: Bachelor of Laboratory Services
Degree Programme: Laboratory Sciences
Supervisors: Jarmo Palm, Senior Lecturer
Jaakko Ekman, Senior Research Scientist
Tinja Kanerva, Technician

This thesis work was carried out for Kemira Oyj. The work was implemented in Microbiology and Biotechnology laboratories at Espoo R&D and Technologies center. The goal of the project was to assess if Fragment Analyzer could be used to monitor microbial communities in industrial processes. Procedure in question would be faster and more cost efficient compared to sequencing done for these industrial process samples.

50 DNA samples were chosen for the project. Samples were duplicated in PCR and the amplicons, PCR products, were analyzed with the Fragment Analyzer instrument. The amplified gene was the 16S-rRNA V1–V2 hypervariable regions. Every bacterium has this gene, and the procedure is based on the differences of base pair length of this gene among different bacteria. The Fragment Analyzer can detect these differences from the amplicons and creates an image that is called a peak profile. The peak profile is a fingerprint of that sample's microbiome and by analyzing and keeping track of these fingerprints, it is possible to determine if the microbiome has changed or not. Procedure is a Length Heterogeneity PCR (LH-PCR).

Based on the results, procedure is fit for analyzing and keeping track of simple microbiomes. It's possible to recognize the dominant bacteria from microbiome with previous results and information gained from sequencing. Procedure is not recommended for wider use or to quantify certain bacteria from the microbiome. However, the data gained regarding the quantities of the bacteria can be used as a directive information.

Main problem in the procedure is to get the targeted gene duplicating better. To improve the procedure, it is suggested to test other primers, possible tweaking of the mastermix and optimizing the PCR program for duplication. The goal of these actions is to increase the share of the targeted gene in the amplicons, as now the primer dimers are the prevalent product in the amplicon. It is recommended to do duplicates when analyzing samples, but the need for them can be reassessed after improving the procedure.

Keywords: microbial community, Fragment Analyzer, gene, fingerprint, LH-PCR, sequence

Sisällys

Lyhenteet

| | | |
|-------|---|----|
| 1 | Johdanto | 1 |
| 2 | Mikrobien vaikutus paperintuottoprosesseissa | 2 |
| 3 | Kvantitatiivinen reaaliaikainen polymeerasiketjureaktio | 3 |
| 4 | Yhteisösormenjälkimenetelmät | 5 |
| 5 | Fragment Analyzer | 6 |
| 5.1 | Yleistiedot | 6 |
| 5.2 | Toimintaperiaate | 6 |
| 5.3 | Parametrit | 8 |
| 6 | Työn toteutus | 10 |
| 6.1 | Työn pohjustus | 10 |
| 6.1.1 | Näytteiden valinta | 10 |
| 6.1.2 | Silva ja sekvenssit | 12 |
| 6.2 | Testiajot | 14 |
| 6.3 | Varsinaiset ajot | 14 |
| 7 | Tulokset ja niiden tulkinta | 17 |
| 7.1 | Testiajojen tulokset | 19 |
| 7.2 | Varsinaisten ajojen tulokset | 26 |
| 7.2.1 | Tehdas 1 | 26 |
| 7.2.2 | Tehdas 2 | 32 |
| 7.2.3 | Tehdas 3 | 40 |
| 7.2.4 | Tehdas 4 | 43 |
| 7.2.5 | Tehdas 5 | 45 |
| 7.2.6 | Tehdas 6 ja 7 | 48 |
| 7.2.7 | Tehdas 8 | 52 |
| 7.3 | Projektin haasteet | 57 |
| 8 | Yhteenveto | 59 |
| | Lähteet | 63 |

Lyhenteet

| | |
|--------|---|
| Bp: | Base pair, emäspari |
| DNA: | Deoxyribonucleic acid, deoksiribonukleiinihappo |
| FA: | Fragment Analyzer, Fragment Analyzer -laite |
| GMO: | Genetically modified organism, geenimuunneltu organismi |
| LH-PCR | Length Heterogeneity Polymerase Chain Reaction (ei suomenkielistä vastinetta) |
| LM: | Lower marker, alamarkkeri |
| NGS: | Next generation sequencing, uuden sukupolven sekvensointi |
| NTC: | No Template Control, negatiivinen kontrolli |
| OTU: | Operational Taxonomic Unit, operatiivinen taksonominen yksikkö |
| PCR: | Polymerase chain reaction, polymeerasiketjureaktio |
| qPCR: | quantative real-time polymerase chain reaction, kvantitatiivinen reaaliaikainen polymeerasiketjureaktio |
| RCF: | Recycled Cellulose Fiber, kiertokuitu |
| RFU: | Relative fluorescence units, suhteelliset fluoresenssiyksiköt |
| rRNA: | Ribosomal ribonucleic acid, ribosomaalinen ribonukleiinihappo |
| UM: | Upper marker, ylämarkkeri |
| UV: | Ultraviolet, ultravioletti |

1 Johdanto

Opinnäytetyö suoritettiin Kemira Oyj:n Espoo R&D and Technologies toimipisteessä Mikrobiologian ja Bioteknologian osastolla. Opinnäytetyön tarkoituksena oli tutkia Fragment Analyzer (FA) -laitteen soveltuvuutta seurata mikrobiyhteisöjä ja niiden muutoksia erilaisissa teollisuuden prosessiolosuhteissa. Menetelmä olisi hyvä lisä sekvensoinnin rinnalle. Nopeampi ja kustannustehokkaampi menetelmä voisi sopia paremmin joillekin asiakkaille. Asiakkaat eivät aina tarvitse kaikkea tietoa, mitä tulee sekvensoinnista, joten FA-laitteen käyttäminen Length Heterogeneity Polymerase Chain Reaction (LH-PCR) -yhteisösormenjälkimenetelmässä olisi hyvä ja nopeampi vaihtoehto nykyiselle sekvensoinnille. Mikrobiyhteisön sormenjäljistä on myös mahdollista tunnistaa tiettyjä bakteerisukuja.

Alustavia tutkimuksia ja analyysyjä on Kemiralla kuvattu sisäisissä tutkimusraporteissa, joiden perusteella Agilent, Fragment Analyzer 5200 -laite hankittiin. Tässä työssä hyödynnetään jo kehitettyjä kvantitatiivisen polymeerasiketjureaktion (qPCR) ja FA:n menetelmiä ja ohjeita.

Näytteitä valittiin 50 kappaletta kahdeksalta eri tehtaalta. Valintakriteerit näytteille olivat mikrobiyhteisö, sijainti, näytteenottoaika prosessissa ja näytteenottoaika. Osa näytteistä oli sekvensoitu, eli näytteistä oli tiedossa eri bakteerien osuudet mikrobiyhteisössä. Näitä tietoja vertailtiin tuloksiin FA:lta, ja selvitettiin, olisiko mahdollista yhdistää bakteereita niitä vastaaviin piikkeihin näytteen piikki-profiilissa. Työssä verrattiin myös piikki-profiilissa olevien piikkien osuuksia sekvenssien osuuksiin, mikä kertoisi kuinka tarkka menetelmä mahdollisesti olisi.

Näytteiden sekvensoinneista saadaan operational taxonomic units (OTU), jotka edustavat tässä kontekstissa käytännössä bakteeriryhmiä ja ovat lyhyitä sekvensoituja 16S-rRNA-geenin alueita [1]. Tässä työssä käytetyistä näytteistä on sekvensoitu 16S-rRNA-geenin alue V4 tai alueet V4–V5. Näiden lyhyiden 16S-

rRNA-geenien sekvenssien avulla voidaan sekvenssitietokannoista hakea kokopitkä 16S-rRNA-geenin sekvenssi, jota tarvitaan, kun ennakoidaan polymeraasi-ketjureaktion (PCR) tuotteiden pituuksia. PCR-tuotteiden (amplikonien) pituuksien odotettiin olevan 300–400 emäsparin (base pair, bp) välillä perustuen käytettäviin alukkeisiin ja aikaisempiin tutkimuksiin. Ennustettavien amplikonien pituudet *in silico* käytettiin enemmän suuntaan antavana eikä täysin lähtökohtana, koska aikaisemmin mainituissa tutkimusraporteissa niiden pituudet eivät koskaan täsmänneet.

Kemira Oyj on globaali kemikaaliyhtiö, joka tuottaa kestäviä kemikaaliratkaisuja vettä käyttäville ja hyödyntäville teollisuudenaloille. Sen pääkonttori on Helsingissä ja sillä on 62 tuotantolaitosta sekä 3 päätutkimuskeskusta, jotka työllistävät työntekijöitä yli 100:ssa eri maassa. Yhtiö keskittyy massa- ja paperi-, öljy- ja kaasuteollisuuteen sekä veden käsittelyyn. Liiketoiminta on jakautunut kahteen asiakassegmenttiin, josta Pulp & Paper -segmentin osuus konsernin liikevaihdosta on noin 60 % ja Industry & Water -segmentin osuus noin 40 %. [2.]

Aikaisemmin mainitut tutkimuskeskukset, eli Research & Development centers, sijaitsevat Yhdysvalloissa, Kiinassa ja Suomessa. Tutkimuskeskukset tekevät yhteistyötä lukuisien tahojen kanssa ja ne keskittyvät kehittämään ratkaisuja asiakkaiden ja yhteiskunnan haasteisiin. R&D asiantuntijoita on noin 250 ja patentteja löytyy lähemmäs 2000 kappaletta. [2.]

2 Mikrobin vaikutus paperintuottoprosesseissa

Paperitehtaat ovat avoimia, vetisiä ja optimaalisen lämpötilan systeemejä mikrobikasvustoille. Mikrobit tulevat esimerkiksi veden tai raaka-aineiden mukana systeemiin. Niiden liiallinen kasvu aiheuttaa taloudellista vahinkoa ja ongelmia monessa vaiheessa prosessia, kuten lopputuotteen laadun huononemista ja ongelmia jäteveden käsittelyssä. Vahingot voivat syntyä esimerkiksi happamoitumisesta, haisevien tuotteiden muodostumisesta tai pigmenttien värivirheistä. Erytisesti tehtaissa, joissa hyödynnetään valkaisua ilman klooria, käsittelyä neutraalissa pH:ssa, suljettuja systeemejä ja kierrätettyä paperia,

mikrobikasvuston ja biofilmin muodostuminen on runsaampaa. [3.] Biofilmi on kokoelma mikro-organismeja, paperitehtailla pääosin bakteereja, ja sen lisääntyneen muodostuminen laitteen kiertokulussa voi aiheuttaa vikoja esim. paperissa (läikkiä ja reikiä) [4, s. 10].

Mikrobeja löytyy jokaisesta paperitehtaasta, ja niiden läsnäoloa prosesseissa ei pysty välttämään [4, s. 9]. Vastatoimenpiteitä voidaan kuitenkin tehdä, jos pystytään ennakoimaan esimerkiksi biofilmin muodostumista. On lukuisia tekniikoita, joilla voidaan analysoida biologista likaantumista. Tekniikat jakautuvat fysikaalisiin (ATP- ja sulfidimittaus), kemiallisiin, geneettisiin (PCR) ja biologisiin (bakteerin identifiointi) metodeihin. [4, s. 22–24.]

3 Kvantitatiivinen reaaliaikainen polymeerasiketjureaktio

Kvantitatiivinen reaaliaikainen polymeerasiketjureaktio, lyhenteenä qPCR, on reaktio, jossa monistetaan tietty DNA-jakso ja seurataan sen monistumista reaaliaikaisesti. Reaktiot ajetaan qPCR-laitteella ja monistustulokset käsitellään instrumentin omalla ohjelmalla. [5.]

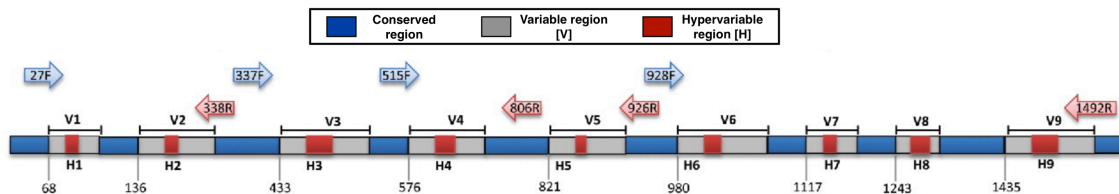
DNA-tuotteiden reaaliaikainen seuranta qPCR:llä perustuu DNA:han sitoutuvan fluoroforin olemassaoloon jokaisessa reaktiokaivossa, jolloin DNA-tuotteen lisääntyminen lisää fluoresenssia. Fluoroforit sitoutuvat DNA-juosteeseen, jolloin DNA:n monistuessa instrumentissa olevat fluoresenssidetektorit detektoivat fluoresenssisignaalit. Mitattu fluoresenssi on suoraan verrannollinen DNA:n määrään. [5.]

qPCR-analyysiä käytetään, kun halutaan kvantifioida nopeasti ja tarkasti tunnettuja DNA-jaksoja erilaisissa biologisissa näytteissä. Sillä on monia applikaatioita, kuten geeniekspressioanalyysi, geenimuunneltujen organismien havaitseminen ruoassa ja syövän fenotyypin selvitys. [5.]

16S ribosomaalinen ribonukleiinihappo, 16S-rRNA, on geeni, jota esiintyy kaikissa bakteereissa [6; 7]. Se on prokaryoottiribosomin 30S-alayksikön

komponentti ja välttämätön geeni proteiinien synteesiä varten. Geeni on pituudeltaan noin 1500 bp ja se muodostuu variaabelialueista ja muuttumattomista alueista. Muuttumattomat alueet geenissä ovat kaikilla bakteereilla samat, mutta variaabelialueilla on taas vaihtelua bakteeriryhmistä tai -lajeista riippuen. [7.]

16S-RNA-geeni sisältää 9 hypervariaabelialuetta (V1–V9) [7; 8]. Kuvassa 1 näkyy 16S-rRNAgeenin rakenne.



Kuva 1. 16S-rRNA-geenin rakenne. Muuttumaton alue on merkitty sinisenä, variaabelialueet ovat harmaita ja hypervariaabelialueet punaisia [9.]

Kyseiset hypervariaabelialueet havainnollistavat bakteerien välillä merkittäviä sekvenssieroja, toisin sanoen emäsparipituuksien vaihtelua. Nämä alueet on mahdollista monistaa PCR:ssä käyttäen yleisalukkeita, mutta ne tuottavat kuitenkin eri tason sekvenssieroja, eikä pelkästään yhtä aluetta tutkimalla voida erottaa bakteereita toisistaan. Hypervariaabelialueita tutkittaessa on todettu, että alueet V2 ja V3 ovat parhaita kaikkien bakteereiden tunnistamiseen lajitasolle saakka. Poikkeuksena tässä vain *Enterobacteriaceae*-bakteeriheimoa lähellä olevat suvut. [8.]

16S-rRNA-geenin monistustuotteiden sekvensointi on ensisijainen vaihtoehto tutkittaessa mikrobiomeja. Sekvensointiin perustuvat 16S-rDNA-tekniikat ovat käytössä laajasti bakteerityypityksessä ja mikrobiyhteisöjen karakterisoinnissa, mutta valitettavasti tutkimuksia ei ole tehty paljon menetelmien eroavaisuuksien vertailuun. Analyysien tarkkuus on vahvasti riippuvainen käytetyistä alukkeista ja on tehty tutkimuksia, jossa testataan eri alukkeita erikokoisille amplikoneille. Mahdollisia tutkittavia hypervariaabelialueita ovat V1–V2, V1–V3, V3–V4, V4, V4–V5, V6–V8 ja V7–V9. [10; 11; 12.]. Tässä projektissa tutkittiin hypervariaabelialueita V1–V2.

4 Yhteisösormenjälkimenetelmät

Bakteerin yhteisösormenjälkimenetelmiä käytetään biodiversiteetin mittaamiseen tai profiloimaan mikrobiyhteisön rakennetta esimerkiksi kokeellisissa olosuhteissa. Analysointi alkaa halutun näytteen DNA:n eristämällä, minkä jälkeen haluttu kohdegeeni monistetaan PCR:llä. Sieltä tulevat monistustuotteet voidaan erottaa lukuisilla menetelmillä ja tuotos voi olla esimerkiksi 2D-geelikuva tai elektroferogrammi [13.] Yksi profiloititapa perustuu 16S-rRNA-geenin tai pienien rRNA-alayksikön geenien monistamiseen ja niistä syntyvien amplokonien pituuserojen tunnistamiseen. Jokainen geenivariantti (vaikka raita elektroferogrammissa tai piikki piikkiprofiilissa) vastaa tiettyä mikrobitalsonia. Usein sillä viitataan OTUihin, joita voidaan vertailla sekvenssitietokantoihin ja yhdistää siten tiettyihin taksonomisiin sekvensseihin. [13; 14; 15]. Yksi hyvä esimerkki yhteisösormenjälkimenetelmistä on LH-PCR (Length Heterogeneity Polymerase Chain Reaction), joka hyödyntää aikaisemmin mainittuja geneeissä esiintyviä pituuseroja mikrobiyhteisöjen seurannassa.

LH-PCR on molekulaarinen ja biologinen matalan suoritustehon menetelmä mikrobiyhteisöjen seurantaan eri ympäristöissä. Matalan suoritustehon menetelmät ovat yksinkertaisempia, laajemmassa käytössä ja kustannustehokkaampia. [14.] Metodin käyttö mikrobiyhteisön seurannassa on tehostanut ja laajentanut ymmärrystä mikrobiyhteisöjen dynamiikasta tietyissä ympäristöissä. [15.]

On olemassa myös muita yhteisösormenjälkimenetelmiä, jotka tutkivat variantteja tietyistä geeneistä. Esimerkiksi denaturoiva gradienttielektroforeesi (DGGE) ja DNA:n katkoskirjo (RFLP). [13.]

5 Fragment Analyzer

Fragment Analyzer -laite käsitellään yleisesti. Käydään läpi sen toiminnot ja tärkeimmät parametrit tulosten ymmärtämiseksi. Käytettävä laite on 5200 Fragment Analyzer System (Agilent). Käytössä olevat ohjelmistot FA-laitteella ovat Fragment Analyzer controller software ja ProSize data analysis software. Ensimmäinen on laitteen kontrolliohjelma, jossa on ajo-ohjelmat ja muut laitteen käyttöä varten vaativat ominaisuudet. Jälkimmäinen on tulosten analyysiä varten, jossa ohjelma luo piikkiprofilit näytteille sekä luo myös teoreettisen kuvan elektroferogrammeista. [16.] Pääosin raakadata käsiteltiin manuaalisesti, kuten yhdistämällä vallitsevat piikit piikkiprofilissa vallitseviin bakteerisukuihin sekvenssitiedoissa ja vertailemalla niiden pituuksia ennustettuihin amplikonien pituuksiin.

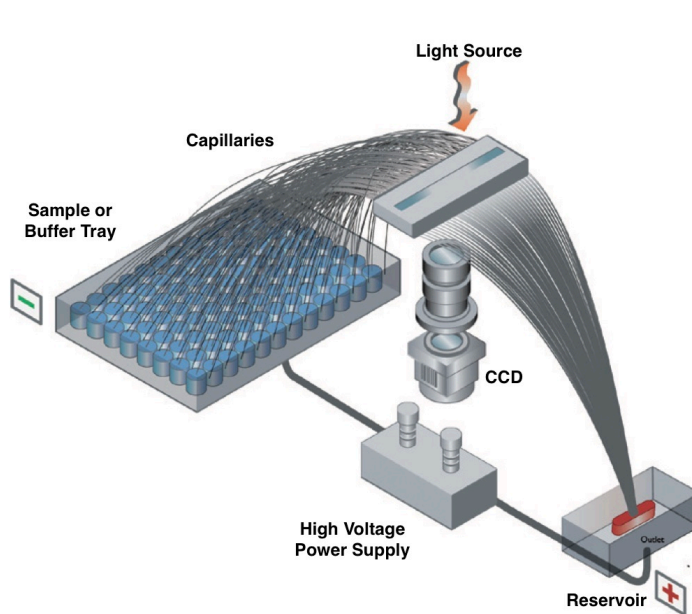
5.1 Yleistiedot

Fragment Analyzer -laite on paralleeli-kapillaarielektroforeesi-instrumentti. Se sisältää 12 kapillaaria, eli se voi ajaa yhdessä ajossa sen verran näytteitä, sekä kolme 96-kaivon levyä, joissa valmistellut näytteet voivat odottaa seuraavia ajoja. Laitteella on myös kaksi eri geelimatriisia, eli se voi ajaa monenlaisia näytteitä saman ajon aikana. Laite voi esimerkiksi analysoida genomista DNA:ta, pieniä RNA-fragmentteja, soluvapaata DNA:ta, suuria DNA-fragmentteja ja totaali-RNA:ta. [16.] Riippuen mikä aikaisemmin mainituista vaihtoehdoista on tutkittavana, valitaan sopivat reagenssit. On olemassa kvantitatiivisia ja kvalitatiivisia kittejä. Kvantitatiiviset kitit mittaavat näytteistä näytekoon (bp tai nt) ja konsentraation. Kvalitatiiviset kitit mittaavat myös tarkan koon, mutta ero kvantitatiiviseen kittiin on, että näytteen kvantifointi mitataan suhteellisesti eikä numeerisesti [17].

5.2 Toimintaperiaate

Fragment Analyzer -laitteessa on kapillaareja, jotka täytetään erotusgeelimatriisilla, jossa on fluoresoivaa väriainetta. Täytettyihin kapillaareihin injektoidaan

näytteet ja kokostandardi elektrokineettisesti (jännitteen avulla injektoidaan varautuneet DNA-näytteet kapillaarien läpi), jolloin ne erottuvat erikokoisiksi fragmenteiksi. Niiden edetessä, omissa kapillaareissaan, ne keräävät fluoresoivaa väriä mukaan ja ohittaessaan detektori-ikkunan fluoresoiva väriaine viritetään jatkuvalla valon lähteellä. Valon emission rekisteröi herkkä CCD-kenno (charged coupled devise), joka kaappaa näytteen koon ja konsentraatiotason. Näytteen läpikulkuun mennyt aika detektori-ikkunasta antaa tietoa fragmenttien koosta ja valon emission määrä konsentraatiosta verrattuna kalibroituun kokostandardiin. [18.] Kuvassa 2 näkyy kuva FA-laitteen sisäisestä rakenteesta.

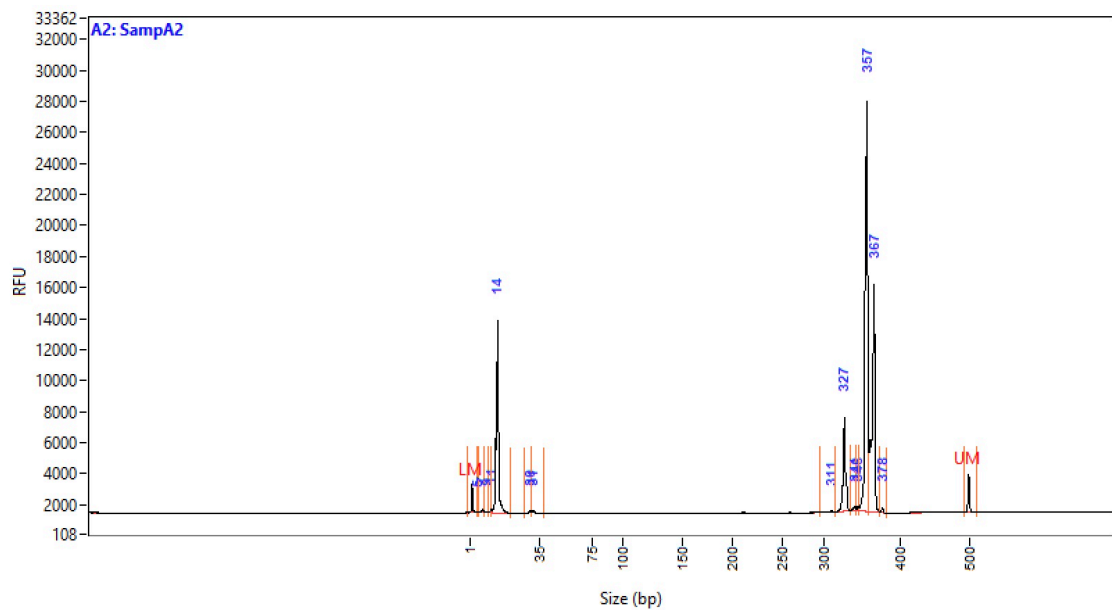


Kuva 2. Kapillaarit ovat näyte- tai puskurilevyssä, josta näytteet/puskuri liikkuvat jännitteen avulla kapillaarien läpi. CCD-kenno havaitsee detektori-ikkunan läpi kulkevat näytteet [18].

Kuvassa 2 näkyy 96 kapillaarin systeemi, jota käytetään 5300- ja 5400 Fragment Analyzer -laitteissa [18].

5.3 Parametrit

FA-laitteella käytetään kahta ohjelmaa: Fragment Analyzer controller software ja ProSize data analysis software. ProSize-data-analyysiohjelmalla käsiteltiin kaikki tulokset ja tässä luvussa 5.3 käydään tarkemmin läpi ohjelman/tulosten relevantit parametrit tulosten ymmärtämiseksi. Kuvassa 3 on käsittelemätön piikkiprofiili.



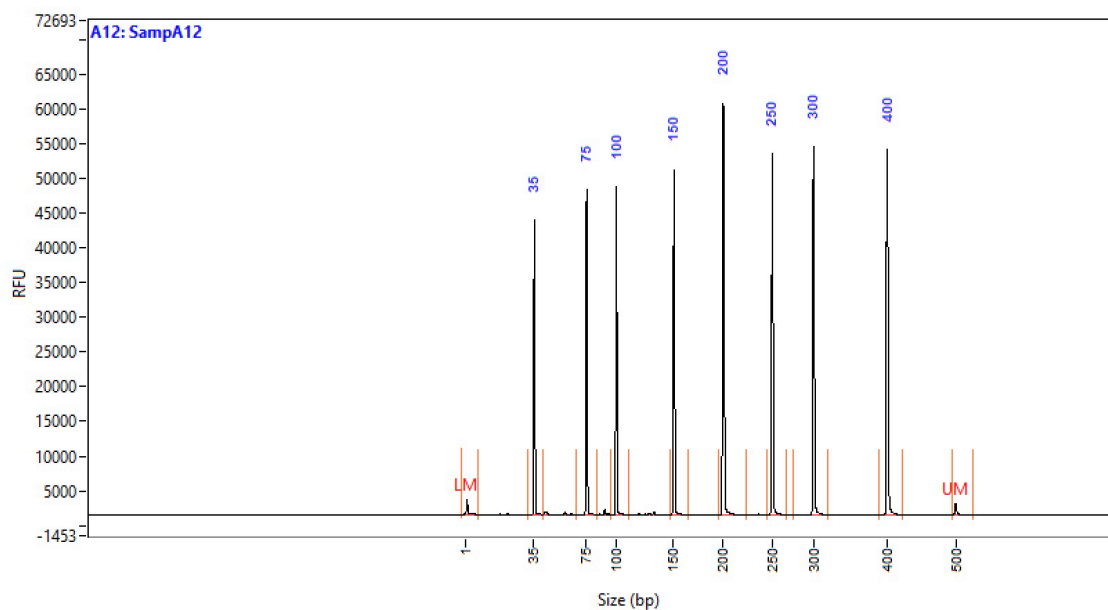
Kuva 3. Esimerkki käsittelemättömästä piikkiprofiilista ProSize-data-analyysiohjelmassa.

Kuvassa 3 Y-akselilla on parametri RFU (relative fluorescence units), joka kertoo, kuinka paljon mitäkin DNA-fragmenttia on näytteessä. Se toimii aikaisemmin mainitun CCD-kennon avulla mittaamalla fluoresenssin näytteistä, jotka on erotettu fragmentteihinsa kapillaarielektroforeesin avulla. [19.] Tällä tavalla saadaan myös X-akselilla esiintyvä parametri bp, joka kertoo näytteen DNA-fragmenttien koon emäspareina. DNA-fragmentit voi ilmoittaa ”peak size”- tai ”smear size”-muodossa. Peak size ilmoittaa runsaimman fragmentin pituuden määritellyllä alueella, kun smear size ilmoittaa määritellyn alueen fragmenttien keskiarvon [20.] Y-akselissa parametri RFU esiintyy tuloksissa Amplitude-

nimellä. Tämä seuraa siitä, kun vertaillaan lukuisia piikkiä toisiinsa ja FA-laitteen ohjelma muuttaa RFU:t amplitudiksi vertailtavuuden vuoksi.

Kuvassa 3 piikit kooltaan alle 100 bp ovat alukedimeerejä, jotka ovat PCR-reaktion sivutuotteita. Alukedimeerit syntyvät, kun alukkeet tarttuvat toisiinsa templaatin sijasta reaktiokaivossa PCR:ssä ja monistuvat. Tulosten tulkinnessa alukedimeerit saattavat häiritä ja ne ovat epärelevantteja, joten tuloksissa ne on poistettu.

Kuvissa 3 ja 4 näkyy piikit 1 bp ja 500 bp, jotka on merkitty nimillä LM ja UM (lower marker ja upper marker). Markkerit ovat kaikissa kaivoissa, ja luotettavien tulosten saamiseksi markkerien pitää olla onnistuneita. Data-analyysiohjelma käyttää markkereita näytteiden koon ja pitoisuuden määrittämiseen kokostandardin (ladder) kanssa. Kuvassa 4 esimerkki onnistuneesta kokostandardista.



Kuva 4. Ladder eli kokostandardi.

Kuvassa 4 nähdään, että piikkiä on LM ja UM piikit ja 8 kokostandardin piikkiä. Data-analyysi -ohjelma vertailee markkereiden pitoisuuden ja koot tiedettyihin kokostandardin piikkeihin ja kokoihin. Näin ohjelma määrittää markkereiden avulla näytteistä pitoisuudet ja koot.

6 Työn toteutus

Työn toteutuksessa käydään läpi tehdyt esivalmistelut projektia varten ja tyypillinen työn toteutus, kun esivalmistelut oli tehty. Työssä käytettiin FA-laitteella kvalitatiivista kittiä (cat. 5191–6583; DNF-905-FR; 5191–6614 Agilent) näytteiden analysoinnissa ja DNA-koot (piikkiprofiilien piikit) ilmoitettiin ”peak size” -muodossa ”smear size” -muodon sijaan.

6.1 Työn pohjustus

Työn pohjustukseen kuului näytteiden valinta ja odotettujen amplikonien pituuksien selvittäminen sekvensoitujen näytteiden mikrobiyhteisöistä. qPCR:llä monistetaan käytännössä 16S-rRNA-geeniä, tarkemmin V1–V2-hypervariaabelialuetta käytössä olevilla yleisalukkeilla. Amplikonien pituuksien ennakoitiin, jotta voitaisiin potentiaalisesti selvittää ja yhdistää piikit FA-laitteelta niitä täsmäviin bakteereihin sekvenssitiedoista. Suurimmasta osasta näytteistä DNA oli eristetty valmiiksi, joten vain osalle näytteistä piti suorittaa totaali-DNA:n eristys.

6.1.1 Näytteiden valinta

Näytteitä valittiin 50 kappaletta. Ne koodattiin tehdaskohtaisesti, esimerkiksi T3.3 on tehtaan 3 kolmas näyte. Testinäytteet nimettiin testinäyte 1–4 ja ovat tehtaan 1 peräisin. Valintakriteerit olivat mikrobiyhteisö (kiinnostavat bakteerit), näytteenottoaika prosessissa (tutkitaan lukuisia näytteitä samasta näytteenottoaikasta mikrobiyhteisön vertailua varten) ja näytteenottoaika (eri ajoilta näytteet, että voidaan tutkia mahdollisia muutoksia/muuttumattomuutta mikrobiyhteisöstä).

Tehtaan 1 näytteitä valittiin 11 kappaletta, josta neljä olivat testiajoja varten. Tehtaan 1 näytteet olivat mikrobiyhteisöllisesti yksinkertaisia, ja siksi näytteet sopivat testiajoihin. Loput seitsemän näytettä valittiin eri ajankohdilta mikrobiyhteisön pysyvyyden/muutoksen seuraamisen arvioimisen vuoksi. Myös

kiinnostava *Dysgonomonas*-sukuinen haitallinen bakteeri esiintyy osassa näytteissä suurissa määrin, minkä vuoksi kyseiset näytteet valittiin. Tehtaassa otetaan näytteitä kahdesta eri näytteenottoapaikasta, kiertokuitu ja 0-vesi. Odotettavasti näytteenottoa ei vaikuttaisi suuresti nähtävään mikrobiyhteisöön perustuen aikaisempiin tutkimuksiin.

Tehtaasta 2 valittiin yhdeksän näytettä. Näytteet on otettu viideltä eri ajankohdalta kolmesta eri näytteenottoapaikasta. Kiinnostuksen kohteena oli mikrobiyhteisön mahdollisen seurannan johdonmukaisuus. Kyseisissä näytteissä on myös enemmän vallitsevia bakteereita kuin muissa, joten haluttiin selvittää, kuinka monta eri bakteeria voidaan identifioida luotettavasti menetelmän avulla.

Tehtaasta 3 valittiin neljä näytettä. Näytteet oli otettu kaikki samana ajankohdalla, mutta eri paikoista. Näytteiden sekvensoinnissa näkyi *Tepidimonas*- ja *Pseudoxanthomonas*-sukuisten bakteerien osuuksien tasaisesti laskevan, mitä pitemmälle prosessia päästiin. Tavoitteena oli selvittää, kuinka luotettavasti pysyisi menetelmä määrittämään bakteerien osuuksia mikrobiyhteisössä. Mikrobiyhteisöt myös erosivat keskenään näytteissä, mikä helpottaisi niiden tunnistamista piikkiprofiilista.

Tehtaasta 4 valittiin kuusi näytettä. Kaksi näytettä oli tehtaan prosessista ja 4 näytettä oli lopputuotteesta. Tehtaan näytteet valittiin erilaisen mikrobiyhteisön vuoksi, ja näytteissä oli myös *Meiothermus*-sukuinen haittabakteeri, jonka tunnistaminen olisi hyvä lisä.

Tehtaasta 5 valittiin kolme näytettä. Niillä oli sama näytteenottoaika, mutta eri näytteenottoapaikat. Mikrobiyhteisö näissä sisälsi sekvenssien mukaan lukuisia samanlaisia bakteereja kuin muissakin tehtaista, joten bakteerien tunnistaminen piikkiprofiilista näistä tuloksista helpottaisi muiden tehtaiden tulosten tarkasteluja, joissa on enemmän vallitsevia bakteereja mikrobiyhteisössä.

Tehtaasta 6 valittiin viisi näytettä ja tehtaasta 7 yksi näyte. Näytteistä haluttiin tarkastella ja tunnistaa *Thermus*- ja *Thermicanus*-sukuun kuuluvat

haittabakteerit. Näytteet ovat kaikki eri näytteenottoaikoista, joten ei keskitytty mikrobiyhteisön muutoksen seurantaan näissä näytteissä.

Tehtaasta 8 valittiin 11 näytettä, josta yksi oli lopputuotenäyte ja muut prosessista viidestä eri näytteenottoaikaista kahdelta eri näytteenottoapäiviltä. Näitä näytteitä ei ole ennen analysoitu, joten mikrobiyhteisö oli tuntematon. Haluttiin simuloida, kuinka luotettavasti voidaan soveltaa menetelmää, kun ei ole ennestään informaatiota ja tukia, voitaisiinko tunnistaa joitakin bakteereita piikkiprofilista hyödyntäen muita tuloksia.

6.1.2 Silva ja sekvenssit

Ennen varsinaisia laboratoriotöitä tiedossa oli jo joidenkin näytteiden sekvenssitetiedot. Sekvensoinnin kohteena oli lyhyt pätkä 16S-rRNA-geeniä (V4 tai V4–V5-geenin alue), ja sen perusteella oli selvitetty bakteeriryhmät ja niiden osuudet mikrobiyhteisöstä. Näytteiden lyhyiden sekvenssien avulla haettiin kokopitkät sekvenssit Silvan SSU r138.1-sekvenssitetokannasta käyttämällä Silvan ACT Toolia (<https://www.arb-silva.de/aligner/>). Sekvenssit ladattiin *FASTA with gaps*-muodossa. Kun sekvenssit oli haettu, niiden pituudet ennakoitiin *in silico* käytettävien alukkeiden kanssa Sequence Extractorissa (<https://www.bioinformatics.org/seqext/>). Jos sitä ei ollut sekvenssissä valmiiksi, niin etualukkeen emäsparijärjestys piti lisätä sekvenssin alkuun, jotta saataisiin oikea odotettava pituus amplikoneille. Taulukossa 1 on esitetty tunnistetut bakteerisuvut ja niitä täsmäävät valitut kokosekvenssien inventaarionumerot (accession number) Silvan SSU r138.1-sekvenssitetokannasta sekä Sequence Extractorissa ennakoitujen ampikonien pituudet.

Taulukko 1. Taulukossa on näytteistä tunnistettujen bakteerisukujen nimet, inventaarionumerot Silvan SSU r138.1-sekvenssietokannasta ja ennakoitujen ampikonien pituudet Sequence Extractorista.

| Bakteerisuku | Inventaarionumero | <i>In silico</i> V1-V2 pituus (bp) |
|--------------------------------|--------------------------------|---|
| Cloacibacterium | GQ097896 & GQ098213 | 343 |
| Dysgonomonas | AB447648 | 356 |
| Prevotella_7 | JN245803 & JQ085649 | 354 |
| Prevotellaceae_unclassified | JN245803 & JQ085650 | 354 |
| Weeksellaceae_unclassified | AM159183 & HM625714 | 343 & 365 |
| Niabella | KT950742 | 348 |
| Anoxybacillus | AJ586360 & KX505021 | 357 |
| Ethanoligenens | HM309996 | 348 |
| Thermicanus | AJ242495 & AZNU01000004 | 368 & 358 |
| Enterococcus | CP010059 & AF061009 | 365 |
| Clostridium_sensu_stricto_N | Y13594 & EU247890 & KY608080 | 333 & 336 & 354 |
| Bacillus | JX232168 & DQ374637 | 327 & 364 |
| Bacillales_unclassified | CP023704 & AJ586361 & AB618491 | 357 |
| Pseudoxanthomonas | FR774574 & AB681369 | 349 & 351 |
| Tepidimonas | MF509266 & MH684019 & AM042693 | 344 & 351 & 355 |
| Tepidiphilus | HM543264 & AJ504663 & KX767837 | 313 & 344 & 362 |
| Sphingomonadaceae_unclassified | LC349997 | 317 |
| Pseudomonadaceae_unclassified | HM989889 | 345 |
| Meiothermus | Y13594 & EU247890 | 332 & 335 |
| Thermus | AY550101 | 341 |

Tietyistä bakteerisuvuista on tutkittu monta eri bakteerilajia, minkä vuoksi niissä näkyy monta eri inventaarionumeroa. Tutkittaessa *in silico* emäsparipituudet saattoivat vaihdella suvun sisällä, kuten *Tepidimonaksella* ja *Meiothermuksella*, kun taas *Enterococcusella* eroa ei ollut.

Tietyistä bakteereista ei ollut tiedossa niitä täsmäviä OTUja, joten ne haettiin nimellä (bakteerisuku) käyttäen Silvan Search-hakutoimintoa ja valitsemalla paras tai sopivin täsmävyys. Kriteerinä käytettiin sekvenssin laatua ja bakteerilajia.

6.2 Testiajot

Ennen varsinaisia ajoja FA-laitteen kapillaarit huollettiin ja testiajettiin. Kapillaarit oli juuri vaihdettu, joten valittiin muutama näyte tehtaasta 1 kapillaarien testausta varten. Samalla tehtiin myös rinnakkaisnäytteitä, joiden avulla päätettiin niiden tarve varsinaisissa ajoissa. Toteamisrajaksi päätettiin markkerien korkeudet. Jos piikki oli sitä pienempi, sitä ei pidetty luotettavana tuloksena ja se jätettiin huomioimatta, vaikka piikit näkyisivät tuloskuviissa.

6.3 Varsinaiset ajot

Laboratoriotyöt jakautuivat kahteen osaan: pre-PCR:ään ja post-PCR:ään. Ennen varsinaisia laboratoriotöitä laskettiin tarvittavat qPCR-reagenssien määrät ja tarvittavat laimennokset analysoitavista näytteistä. Näytteistä luotiin pipetointikartta, jota seurattiin qPCR:n ja FA-laitteen näytelevyillä.

Yhdessä FA-ajossa analysoitiin siis enimmillään 11 näytettä, vaikka kaivoja on 96. Analyysien välissä oli hyvä tarkastella tuloksia ja päättää vasta sen jälkeen seuraavat näytteet ajoon. PCR-ajossa kaivo 12 oli aina negatiivinen kontrolli (NTC) ja FA-laitteella siellä on kokostandardi. PCR-ajossa ajettiin myös näytteiden rinnakkaisnäytteitä, jotta nähtiin, ovatko näytteet luotettavasti monistuneet. Jos rinnakkaisnäytteiden sykliarvot erosivat tietyn määrän, näyte ei ollut

luotettava FA-ajoa varten. Näytteettömät kaivot olivat vain TE-puskuria FA-laitteella, jos ei ollut näytteelle tarvetta. Ne toimivat myös qPCR:ssä NTC:nä.

Fragment Analyzer -laite valmisteltiin näyteajoja varten suorittamalla full condition -ohjelma. Ohjelma valmistelee kapillaarit näyteajoja varten ja sen pituus vaihteli riippuen edellisten ajojen tuloksista tai siitä, milloin viimeisin ajo oli suoritettu. Ajoliuksia valmistettiin full condition -ohjelman ja näyteajon verran ohjeiden mukaisesti. Full condition-ohjelman suoriutuessa siirryttiin toiseen laboratorioon, jossa oli qPCR-laite ja DNA-näytteet. Valitut näytteet siirrettiin pakastimesta qPCR-reagenssien kanssa jääkaappiin sulamaan. Näytteiden ja qPCR-reagenssien sulaessa siirryttiin takaisin laboratorioon, jossa FA-laite oli suorittanut full condition-ohjelman. Jos ohjelma oli onnistunut moitteettomasti, valmisteltiin Marker- ja Buffer-levyt ja tyhjennettiin Waste-levy. Tämän jälkeen FA-laite oli valmis näyteajoja varten. Jos kapillaareista olisi tullut Waste-levylle epätasainen määrä tai ei ollenkaan nesteitä ohjelman jälkeen, se toteutettiin uudestaan. Tämä viittasi siihen, että kapillaarit vaativat lisää huoltamista. Nämä vaiheet kuuluivat pre-PCR-työosuuteen.

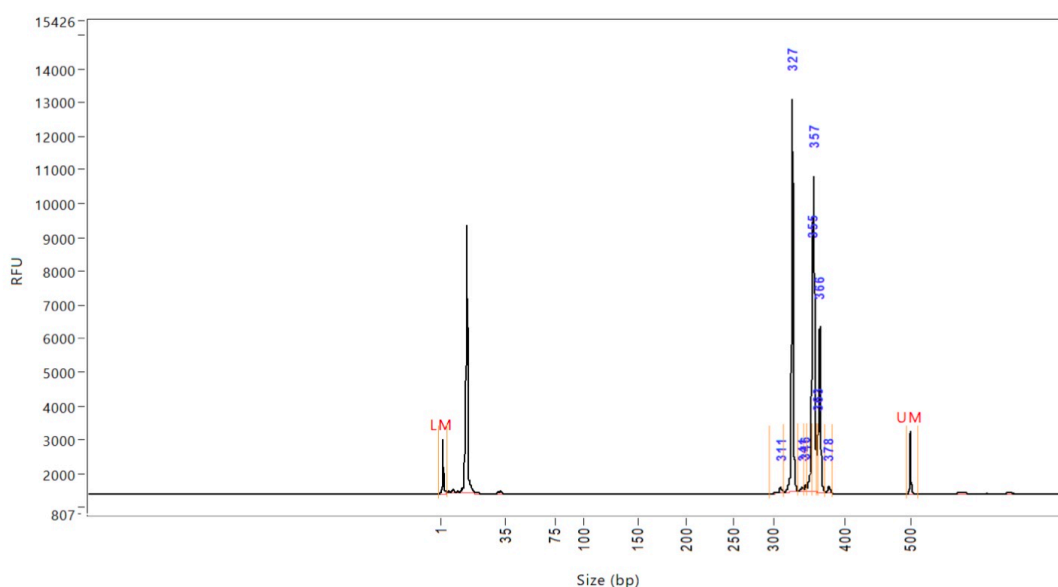
Kun DNA-näytteet ja qPCR-reagenssit olivat sulaneet, näytteet laimennettiin saman vahvuiseksi, normaalisti laimeimman näytteen mukaan. Osa näytteistä sisälsi liian vähän DNA:ta, jolloin reaktiivilavuutta nostettiin ja lisättiin templaatin osuutta, jotta DNA:n kokonaismäärä olisi suurempi PCR-monistumisen onnistumiseksi. Nämä kyseiset näytteet ajettiin omissa ajoissaan, jotta näytteiden pitoisuudet eivät eroaisi liian paljon. Jos DNA-pitoisuudet eroaisivat liikaa, näytteet monistuisivat eri vauhtia, mikä olisi ongelmallista. PCR-monistumisen onnistumisen todettiin reaaliajassa SYBR Green -väriaineen avulla, ja monistuminen pysäytettiin ennen kuin näytteet saavuttivat plateau-vaiheen. Plateau-vaiheessa eri amplikonien osuudet näytteissä vääristyisivät, mikä vaikeuttaisi tulosten tulkintaan ja aiheuttaisi niiden epäluotettavuutta. Kun näytteet olivat monistuneet tarpeeksi, monistus keskeytettiin ja tarkistettiin näytteiden ja niiden rinnakkaisnäytteiden sykliarvot. Jos ne olivat kunnossa, näytteet oli tarpeeksi luotettavasti valmistettu ja monistuneet ja valmiina Fragment Analyzer -näyteajoja varten. Jos jonkin näytteen ja sen rinnakkaisnäytteen sykliarvot erosivat toisistaan liikaa,

niiden tulokset analysoitiin tilannekohtaisesti FA-näyteajon jälkeen ja todettiin vasta myöhemmin, ovatko ne luotettavat. Riippuen tilanteesta luotettavuus arvioidtiin vertailemalla näytteen piikkiprofileja saman tehtaan näytteisiin tai näytteen omiin rinnakkaisnäytteisiin.

PCR-monistumisen jälkeen PCR-levy siirrettiin toiseen laboratorioon, jossa Fragment Analyzer -laite oli valmiina näyteajoja varten. FA-laitteen näytelevy valmistettiin, ja siihen pipetointiin pipetointikartan mukaisesti PCR-tuotteet. Näyteajon pituus oli määritetty aikaisemmin selvittämällä, milloin UM (upper marker) oli detektoitu testiajoissa. Piikkiprofileista ensisijaisesti tarkistettiin, olivatko ne onnistuneita, ja jos jotkin näytteet olivat epäonnistuneet, niin ne merkittiin uudelleenanalysoitaviksi. Tulokset analysoitiin ja muokattiin parempaan muotoon myöhemmin. Seuraavat analyysit, uudelleenajot tai kapillaarien puhdistukset tarvittaessa tehtiin tarvittaessa.

7 Tulokset ja niiden tulkinta

Tuloksissa käydään läpi testiajojen tulokset ja tehdaskohtaisesti varsinaisten ajojen tulokset. Kuvissa näkyy näytteiden piikkiprofiilien suurennos n. 300–400 bp, eikä kokonaisuudessa, jossa näkyisi markerit ja alukedimeerit. Ilman zoomausta piikkiprofiilien erot eivät näkyisi selvästi. Kuvassa 5 on piikkiprofiili ilman suurennosta.



Kuva 5. Esimerkki ei-zoomatusta piikkiprofiilista. Kuvassa näkyy edelleen alukedimeerien piikit LM:n vieressä, mutta niiden emäsparipituudet ja pinta-alat on poistettu. Kyseinen näyte kuvassa on testinäyte 1.

Lähes kaikki käsiteltävät piikit ovat yli toteamisrajan, jonka päätettiin olevan markerien korkeus. Pieni osuus, joka ei ole ylittänyt rajaa, on piikkejä, jotka voidaan luotettavasti toisin todeta oikeiksi piikeiksi. Esimerkiksi jos piikki on läsnä suurimmassa osassa saman tehtaan näytteissä ja eri ajoista, niin piikki todettiin luotettavaksi.

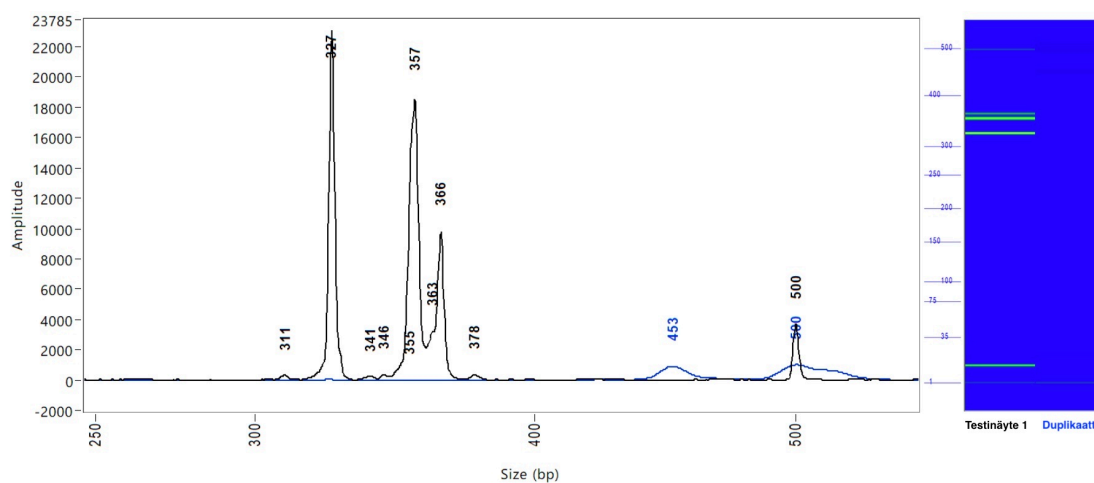
Piikkien pinta-aloista puhuttaessa ja vertailtaessa ei ole huomioitu alukedimeerien osuuksia, joka on pääosin aina yli 90 % koko piikkiprofiilin pinta-alan osuudesta ja vaikeuttaisi piikkien osuuksien vertailua toisiinsa. Esimerkiksi kuvassa 5 korkeimman piikin osuus 327 bp koko piikkiprofiilin pinta-alasta on

todellisuudessa vain 2,8 % ja alukedimeerien osuus on 92,1 % (piikkien yhteenlaskettu pinta-ala 7–32 bp). Piikki profiili on automaattisesti kuvattu laajalle alueelle, minkä vuoksi on vaikea hahmottaa piikkien leveyksiä. Tämä saattaa selittää, miksi alukedimeerin piikki on pienempi kuvassa 5 kuin piikki 327 bp, mutta on pinta-alaltaan osuudeltaan suurempi. Sen vuoksi markerit ja alukedimeerit ovat lopullisessa tulosten tulkinnassa merkityksettömiä.

Piikki profiilit ovat päällekkäin aseteltuja, ja jokaista näytettä kohti on oma väri (kuvat 6–29). Teoreettisen geelielektroferogrammin alla näkyy niitä vastaavat näytteiden nimet, ne on värikoodattu piikki profiilien mukaisesti. Geelielektroferogrammi ei ole reaaliaikainen kuva kapillaareista, vaan teoreettinen perustuen markkereihin ja kalibroituun kokostandardiin. Niistä näkyy piikkien raidat. Kuvien (verratut piikki profiilit) jälkeen vertaillaan myös välillä taulukoissa kyseisten näytteiden tunnistettuja bakteereita, niiden emäsparipituuksia, jotka on ennakoitu *in silico* Sequence Extractorin avulla (taulukko 1), FA:lta saatuja emäsparipituuksia, sekä niitä vastaavia bakteerien osuuksia piikki profiileista. Mahdolliset tiedossa olevia sekvensointitietoja (todellinen osuus mikrobiyhteisöstä) vertaillaan myös näytekohtaisesti FA:lta saatuihin osuuksiin.

7.1 Testiajojen tulokset

Kuvissa 6–9 on testinäytteiden ja niiden rinnakkaisnäytteet päällekkäin asetetut piikkiprofiileista. Duplikaatit, eli rinnakkaisnäytteet, ovat siis samasta DNA-näytteestä tehtyjä PCR-reaktioita, jolloin verrataan siis menetelmän vaihtelua. Samasta näytteestä tehtiin siis kaksi reaktiota ja niiden tuloksia vertaillaan toisiinsa. Näytteiden nimet on värikoodattu piikkiprofiilissa olevien värien mukaisesti.



Kuva 6. Kuvassa testinäyte 1 ja sen duplikaatin piikit. Piikit 300–400 bp tulevat näytteistä ja piikki 500 bp on ylämarkkeri.

Kuvassa 6 testinäyte 1 oli onnistunut: selvät piikit ja selvä ylämarkkeri. Rinnakkaisnäyte oli kuitenkin epäonnistunut eikä selviä piikkejä näkynyt. Ylämarkkeri 500 bp oli myös selvästi liian leveä. Syy saattoi olla kuplien esiintyminen kapillaarissa. Taulukossa 2 näkyy testinäytteen 1 relevantit tulokset ja *in silico* ennakoitua emäsparipituudet.

Taulukko 2. Taulukossa näkyy testinäytteen 1 bakteerien ennakoidut ja FA-tulosten emäsparipituudet sekä niiden osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofilissa.

| Taxon | <i>In silico</i> (bp) | Testinäyte 1 | |
|--------------------------------|-----------------------|---------------|-----------|
| | | FA-tulos (bp) | Osuus (%) |
| Sphingomonadaceae_unclassified | 317 | 327 | 34,8 |
| Tuntematon | | 355 | |
| Cloacibacterium | 343 | 357 | 41,6 |
| Tuntematon | | 363 | 4,1 |
| Dysgonomonas | 356 | 366 | 15,1 |

Mikrobiyhteisöstä tunnistettiin 3 eri bakteeria. Taulukossa 2 nähdään, että suurin osuus olisi *Cloacibacteriumilla* (357 bp), vaikka kuvassa 6 näkyy selvästi, että piikki 327 bp on suurempi kuin piikki 357 bp. Tämä johtuu siitä, että FA-laitte ei pysty erottamaan hyvin piikkejä, jotka eroavat alle 3 bp toisistaan. Tämän vuoksi *Cloacibacteriumin* (357 bp) osuus on suurempi, koska piikki 355 bp on liian lähellä. Piikin pinta-alan osuus kasvaa luonnollisesti piikin ollessa leveämpi. *In silico* bp-pituudet ovat lyhyemmät kuin FA:lta saaduista tuloksista, kuten oli odotettavissa perustuen edellisiin Kemiralla tehtyihin tutkimuksiin.

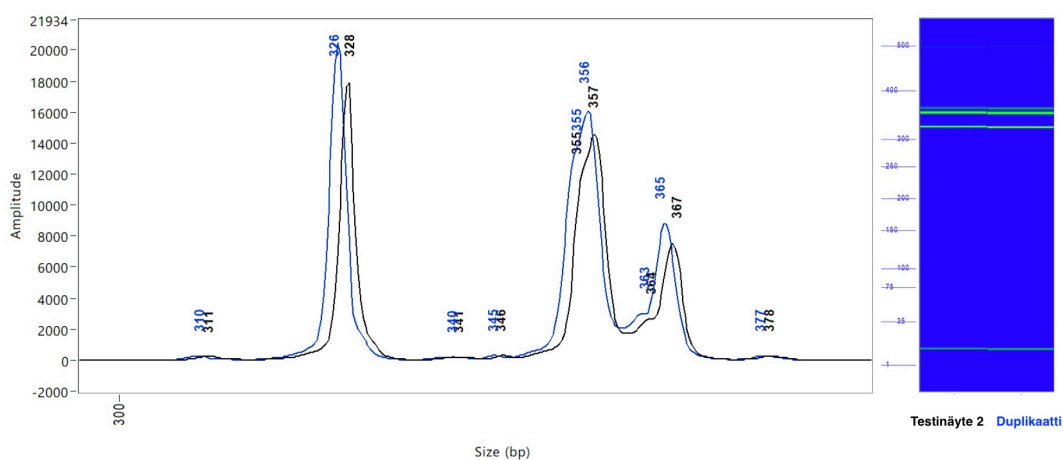
Näytteestä oli tehty sekvensointi, joten tiedossa oli todelliset bakteerien osuudet mikrobiyhteisöstä. Taulukossa 3 näkyy testinäytteen 1 relevantit tulokset ja sekvensointitieto.

Taulukko 3. Testinäytteen 1 tunnistetut bakteerit, niiden osuudet FA-laitteelta ja sekvensointitiedoista saadut todelliset osuudet mikrobiyhteisöstä. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofilissa.

| Taxon | Testinäyte 1 | | |
|--------------------------------|---------------|-----------|---------|
| | FA-tulos (bp) | Osuus (%) | Seq (%) |
| Sphingomonadaceae_unclassified | 327 | 34,8 | 41,7 |
| Cloacibacterium | 357 | 41,6 | 17,3 |
| Dysgonomonas | 366 | 15,1 | 8,7 |

Taulukosta 3 näkyy, että bakteerien osuuksien erot ovat pienemmät FA:lta saaduista tuloksista verrattuna sekvensoinnista saaduista tuloksista. On mahdollista, että monistuksessa bakteerit, joilla on pienempi osuus mikrobiyhteisöstä,

monistuvat hitaammin ja vähemmän. Tämä nostaa siten vallitsevien bakteerien osuuksia vielä enemmän näytteessä.



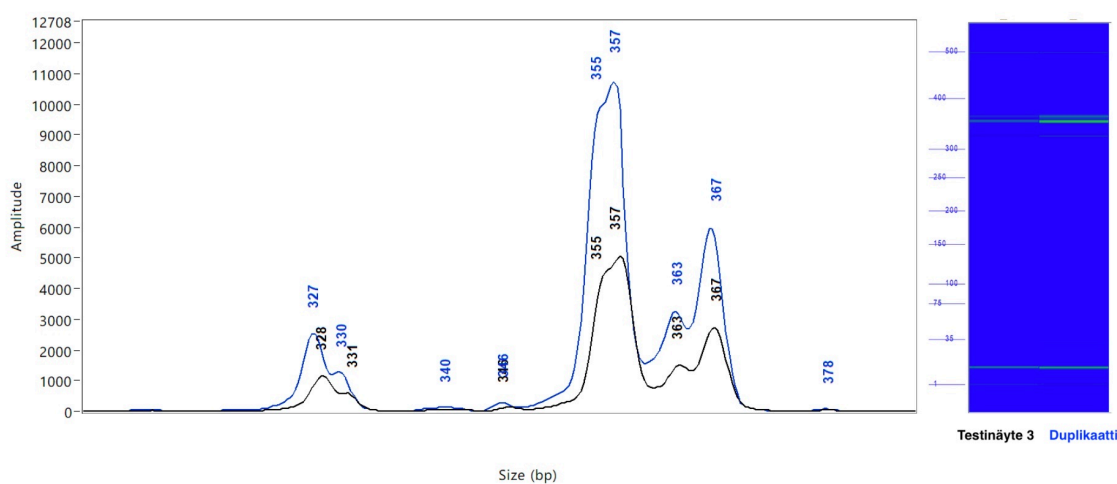
Kuva 7. Testinäytteen 2 ja sen duplikaatin piikkiprofiilit.

Kuvassa 7 testinäyte 2 oli onnistunut rinnakkaisnäytteensä kanssa. Ensimmäinen huomio on piikkiprofiilien epätasaisuus emäsparipituudessa. Rinnakkaisnäyte on systemaattisesti pääosin 1 bp lyhyempi, mutta muuten piikit ovat samankokoisia. Taulukossa 4 näkyy testinäytteen 2 ja sen rinnakkaisnäytteen relevantit tulokset ja *in silico* ennakoitujen emäsparipituudet.

Taulukko 4. Taulukossa näkyy testinäytteen 2 ja duplikaatin bakteerien ennakoitujen ja FA-tulosten emäsparipituudet sekä niiden osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | <i>In silico</i> (bp) | Testinäyte 2 | | Duplikaatti | |
|--------------------------------|--------------------------|------------------|--------------|------------------|--------------|
| | | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| Sphingomonadaceae_unclassified | 317 | 328 | 32,5 | 326 | 33,8 |
| Tuntematon | | 355 | | 355 | |
| Cloacibacterium | 343 | 357 | 38,8 | 356 | 40,4 |
| Tuntematon | | 364 | 4,3 | 363 | 4,6 |
| Dysgonomonas | 356 | 367 | 18,1 | 365 | 15,1 |

Taulukossa 4 nähdään, että piikkien osuudet ovat lähes samat testinäytteellä ja rinnakkaisnäytteellä. Tulokset ovat samat kuin testinäytteellä 1. Testinäytteen 2 kuvassa 7 nähdään kuitenkin selvemmin, kuinka piikki 355 bp leventää piikkiä 356/357 bp. Piikille 355 bp ei pystytä määrittämään omaa pinta-alaa, joten piikin 357 bp pinta-ala sisältää käytännössä piikin 355 bp.



Kuva 8. Testinäytteen 3 ja sen duplikaatin piikkiprofiilit. Duplikaatin piikit ovat huomattavasti korkeammat.

Kuvassa 8 testinäytteen 3 piikit ovat pienemmät kuin sen rinnakkaisnäytteellä, ja FA-laite ei tunnistanut testinäytteestä 3 piikkejä 340 bp ja 378 bp. Mahdollisesti FA-laite ei tunnista pienempiä piikkejä, jos amplitudi on liian alhainen. Kyseessä voi olla ero pipetointitarkkuudessa, kun pipetoitava tilavuus FA-laitteelle on 3 µl PCR-tuotteista ja amplitudi selvästi korkeampi rinnakkaisnäytteellä. Vaihtelua on myös emäsparipituuksissa, mutta vähemmän kuin testinäytteessä 2. Taulukossa 5 näkyy testinäytteen 3 ja sen rinnakkaisnäytteen relevantit tulokset ja *in silico* ennakoitujen emäsparipituudet.

Taulukko 5. Taulukossa näkyy testinäytteen 3 ja duplikaatin bakteerien ennakoitujen ja FA-tulosten emäsparipituudet sekä niiden osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | <i>In silico</i> (bp) | Testinäyte 3 | | Duplikaatti | |
|--------------------------------|--------------------------|------------------|--------------|------------------|--------------|
| | | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| Sphingomonadaceae_unclassified | 317 | 328 | 8,7 | 327 | 10,7 |
| Tuntematon | | 355 | | 355 | |
| Cloacibacterium | 343 | 357 | 51,6 | 357 | 47,1 |
| Tuntematon | | 363 | 11,3 | 363 | 9,9 |
| Dysgonomonas | 356 | 367 | 19 | 367 | 17,6 |

Taulukon 5 tulokset korreloivat edellisten testinäytteiden tulosten kanssa. Taulukoista voidaan nähdä vain piikkien osuudet piikkiprofiilin kokonaispinta-alasta, minkä vuoksi ei pystytä hahmottamaan esim. kuvassa 8 piikkiprofiilien koeroa.

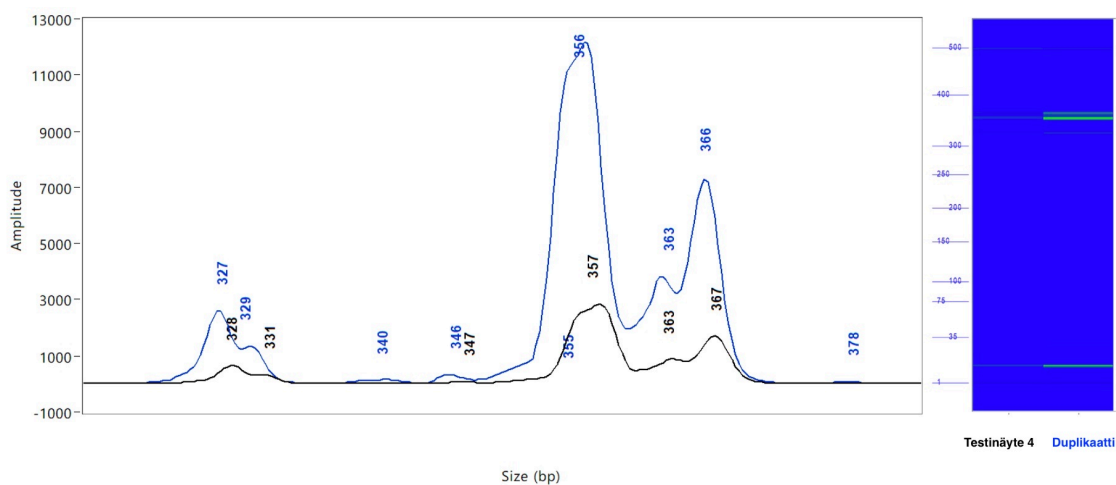
Näytteestä oli tehty sekvensointi, joten tiedossa oli todelliset bakteerien osuudet mikrobiyhteisöstä. Taulukossa 6 näkyy testinäytteen 3 relevantit tulokset ja sekvensointitieto.

Taulukko 6. Testinäytteen 3 tunnistetut bakteerit, niiden osuudet FA-laitteelta ja sekvensointitiedoista saadut todelliset osuudet mikrobiyhteisöstä. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | Testinäyte 3/Duplikaatti | | | Seq (%) |
|--------------------------------|--------------------------|--------------|-----------------|---------|
| | FA-tulos (bp) | Osuus T3 (%) | Osuus dupli (%) | |
| Sphingomonadaceae_unclassified | 328/327 | 8,7 | 10,7 | 16,9 |
| Cloacibacterium | 357 | 51,6 | 47,1 | 20,8 |
| Dysgonomonas | 367 | 19 | 17,6 | 15,2 |

Sekvensoinnissa osuudet näytteessä vallitsevien bakteerien osuudet olivat lähellä toisiaan, verrattuna FA:lta saaduista tuloksista, jossa on selvästi suuremmat erot. Tämän selittää kuitenkin sama päätelmä testinäytteen 1 kohdalla. Monistuksessa bakteerit, joilla on pienempi osuus mikrobiyhteisöstä, monistuvat hitaammin ja vähemmän. Tämän vuoksi tässä tilanteessa vallitsevin bakteeri

Cloacibacterium on monistunut eniten ja ottanut hallitsevan määrän PCR-tuotteessa.



Kuva 9. Testinäytteen 4 ja sen duplikaatin piikki profiilit.

Testinäytteen 4 tulokset kuvassa 9 ovat samat kuin testinäytteen 3. Rinnakkaisnäyte antaa suuremmat piikit ja emäsparipituudet heittävät hieman. Testinäytteessä 4 piti manuaalisesti identifioida piikit 331 bp ja 363 bp. Taulukossa 7 näkyy testinäytteen 3 ja sen rinnakkaisnäytteen relevantit tulokset ja *in silico* ennakoit emäsparipituudet.

Taulukko 7. Taulukossa näkyy testinäytteen 4 ja duplikaatin bakteerien ennakoit ja FA-tulosten emäsparipituudet sekä niiden osuudet piikki profiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikki profiilissa.

| Taxon | <i>In silico</i> (bp) | Testinäyte 4 | | Duplikaatti | |
|--------------------------------|--------------------------|------------------|--------------|------------------|--------------|
| | | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| Sphingomonadaceae_unclassified | 317 | 328 | 11,3 | 327 | 8 |
| Tuntematon | | 355 | | 355 | |
| <i>Cloacibacterium</i> | 343 | 357 | 51,6 | 356 | 52,1 |
| Tuntematon | | 363 | 11,7 | 363 | 11,7 |
| <i>Dysgonomonas</i> | 356 | 367 | 20,5 | 366 | 20,8 |

Tulokset testinäytteeltä 4 eivät eroa suuremmin muiden testinäytteiden tuloksiin.

Testiajojen perustella päätettiin, että ei ole tarpeellista tehdä rinnakkaisnäytteitä varsinaisissa ajoissa. Jos ajoon menevissä kaivoissa on ylimääräistä tilaa, valitaan tilanteen mukaan näyte, josta ajetaan myös rinnakkaisnäyte. Tulosten perusteella arvioitiin, että välillä voi olla tarvetta analysoida sama näyte uudelleen, mutta on tehokkaampaa analysoida enemmän näytteitä kuin vähemmän rinnakkaisnäytteiden kanssa per ajo.

Tulosten tulkinnessa pitää ottaa huomioon emäsparipituuksien vaihtelu ja epäselvät piikit, kuten piikki 355 bp, joka piti manuaalisesti tunnistaa kaikissa testinäytteissä. Näissä tilanteissa ei pystytä erottamaan kahden lähellä olevien piikkien pinta-aloja selvästi, joten riippuen tilanteesta, pinta-alat jaetaan kahtia, tai suuremmalle piikille jätetään koko piikin pinta-ala ja pelkästään pienempi piikki tunnistetaan.

Pinta-alojen osuudet ovat todellisuudessa amplikoneissa paljon pienempiä, muutamissa prosenteissa, koska alukedimeerien osuus oli näissä näytteissä yli 90 %. Tämä tulee pitämään paikkansa myös varsinaisissa ajoissa, joten katsottaessa bakteerien prosentuaalisia osuuksia piikkiprofilien tuloksissa siinä on poistettu alukedimeerien piikit ja niiden pinta-alojen osuudet, jotta piikkien suhteet toisiinsa olisi helpompi nähdä. Voidaan myös olettaa, että FA-laite ei tunnista kaikkia mahdollisia piikkejä, jos amplitudi piikkiprofililla on liian alhainen.

Testinäytteistä tunnistettiin 3 eri bakteeriryhmää, *Sphingomonadaceae_unclassified* (326–327 bp), *Cloacibacterium* (356–357 bp) ja *Dysgonomonas* (365–367 bp). Bakteerit tunnistettiin vertailemalla vallitsevat piikit piikkiprofileista sekvenssitietojen bakteerien osuuksiin ja vertailemalla näytteiden tuloksia toisiinsa. Myös *in silico* ennakoitujen emäsparipituuksia hyödynnettiin päättelyssä. *Sphingomonadaceae_unclassified* oli vallitsevin bakteeri testinäytteen 1 sekvenssitiedoissa, joten se todettiin olevan suurin piikki 327 bp testinäytteen 1 piikkiprofilissa kuvassa 6. Bakteeria ei todettu piikiksi 357 bp, koska testinäytteessä 3 vallitsevin piikki oli juuri 357 bp, mutta sen näytteen sekvenssitiedoissa vallitsevin bakteeri oli *Cloacibacterium*. *Cloacibacterium* todettiin siis piikiksi 356–357 bp, joka yhteni myös muihin tuloksiin. *In silico* bp

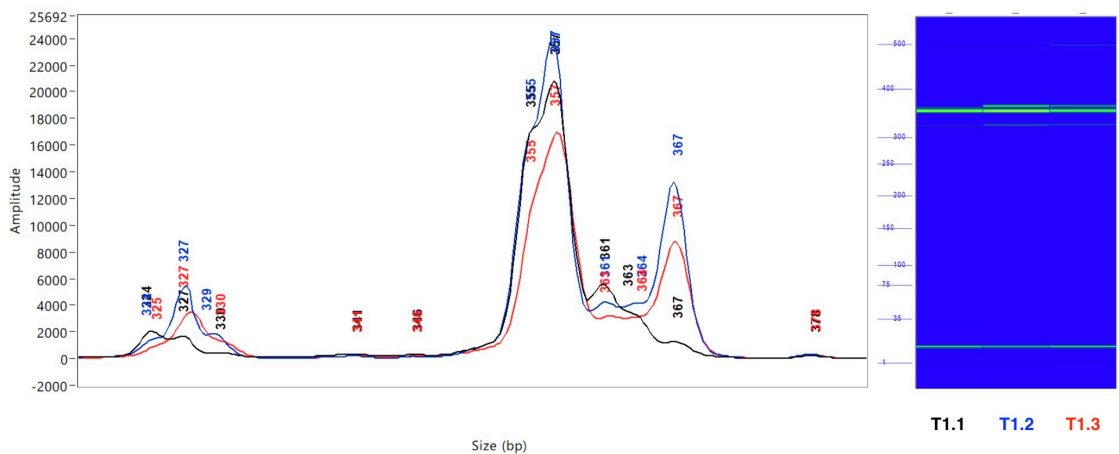
pituudet myös sopivat järkevämmiin, jos näin todettiin, etenkin kun *Dysgonomonas* todettiin piikiksi 365–367 bp, kun ei ollut muita merkitseviä bakteereita tiedossa (osuudeltaan).

7.2 Varsinaisten ajojen tulokset

Varsinaiset ajot ovat jaettu tehdaskohtaisesti alaotsikkoihinsa analysoitaviksi, paitsi tehtaan 6 ja 7 tulokset, joiden tulokset on analysoitu saman alaotsikon alla.

7.2.1 Tehdas 1

Alla olevassa kuvassa 10 on tulokset näytteistä T1.1, T1.2 ja T1.3. Sekvenssitietojen mukaan näytteissä T1.2 ja T1.3 oli kohonnut haitallisen *Dysgonomonas*-bakteerinsuvun osuus mikrobiyhteisöstä, ja sitä vertaillaan näytteeseen T1.1, jossa sen bakteerin osuus oli matalampi. Näyte T1.1 on noin 1 kk vanhempi kuin näytteet T1.2 ja T1.3, joilla on sama näytteenottopäivä, mutta eri näytteenottoaika.



Kuva 10. Näytteiden T1.1–T1.3 piikkiprofiilit. Vanhemman näytteen T1.1 piikkiprofiili poikkeaa tuorempien näytteiden piikkiprofiileista, jossa näkyy piikin 367 bp osuuden nousseen.

Kuvassa 10 nähdään, että näytteiden mikrobiyhteisöjen sormenjäljet ovat samanlaisia. Suurin ero näkyy piikin 367 bp kohdalla, jossa näytteen T1.1 osuus on erittäin matala, mutta näytteissä T1.2 ja T1.3 näkyy selvästi piikin osuuden nousevan. Tämä piikki tunnistettiin kuuluvan *Dysgonomonas*-bakteerisuvulle. Tulos korreloi testiajojen tunnistuksien kanssa. Tietyt piikit, kuten piikit 355 bp ja 363–364 bp piti manuaalisesti identifioida.

Taulukossa 8 on näytteiden T1.1–T1.3 relevantit tulokset ja *in silico* ennakoidut emäsparipituudet.

Taulukko 8. Taulukossa näkyy T1.1–T1.3 bakteerien ennakoidut ja FA-tulosten emäsparipituudet sekä niiden osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | <i>In silico</i> (bp) | T1.1 | | T1.2 | | T1.3 | |
|------------------------------------|--------------------------|------------------|--------------|------------------|--------------|------------------|--------------|
| | | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| Tuntematon | | 324 | 4,1 | 324 | 1,7 | 325 | 2 |
| Sphingomonadaceae_ unclassified | 317 | 327 | 3,1 | 327 | 8,3 | 327 | 8 |
| Tuntematon | | 330 | 0,9 | 329 | 2,3 | 330 | 2,2 |
| Tuntematon | | 355 | | 355 | | 355 | |
| Cloacibacterium | 343 | 357 | 63,1 | 357 | 52,3 | 357 | 53,6 |
| Tuntematon | | 361 | 9,8 | 361 | 5,5 | 361 | 5,2 |
| Tuntematon | | 363 | 5,6 | 364 | 5,6 | 364 | 4 |
| Dysgonomonas | 356 | 367 | 2,3 | 367 | 18,1 | 367 | 18,5 |

Näytteen T1.1 tulokset eroavat odotetusti hieman näytteiden T1.2 ja T1.3 tuloksista. Kaikissa on vallitsevasti *Cloacibacteriumia*, mutta näytteissä T1.2 ja T1.3 *Dysgonomonas*-piikki ilmestyy, kun näytteessä T1.1 sitä ei ole ollenkaan. Näytteiden T1.2 ja T1.3 bakteeriosuudet ovat myös erittäin lähellä toisiaan, minkä selittää näytteiden sama näytteenottopäivä. Taulukkoon 8 on kerätty myös piikkejä, joista ei ole tunnistettu bakteereita, koska ne esiintyvät säännöllisesti ja ovat nähtävästi osa normaalia mikrobiyhteisöä.

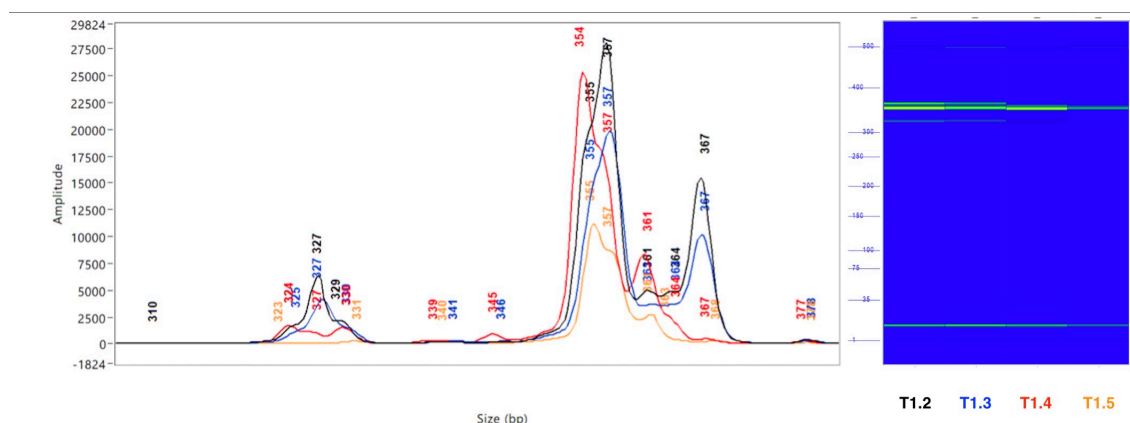
Taulukossa 9 näkyy näytteen T1.2 relevantit tulokset ja vertailu näytteen sekvenssitietoihin.

Taulukko 9. T1.2 tunnistetut bakteerit, niiden osuudet FA-laitteelta ja sekvensoititiedoista saadut todelliset osuudet mikrobiyhteisöstä. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | T1.2 | | Seq (%) |
|--------------------------------|---------------|-----------|---------|
| | FA-tulos (bp) | Osuus (%) | |
| Sphingomonadaceae_unclassified | 327 | 8,3 | 9,3 |
| Cloacibacterium | 357 | 52,3 | 43,1 |
| Dysgonomonas | 367 | 18,1 | 25,2 |

Sekvenssitietojen bakteeriosuudet ovat hyvässä linjassa FA:lta saatuihin osuuksiin. Tulokset tukevat aikaisempia johtopäätöksiä liittyen bakteerien tunnistukseen.

Kuvassa 11 vertaillaan näytteitä T1.2 ja T1.3 näytteisiin T1.4 ja T1.5. Näytteet T1.4 ja T1.5 ovat otettu kuusi kuukautta myöhemmin, kun mikrobiyhteisössä haitallisen *Dysgonomonas*-bakteerisuvun osuus on laskenut.



Kuva 11. Näytteiden T1.2–T1.5 piikkiprofiilit. Piikkiprofiileista näytteet T1.4 ja T1.5 ovat tuoreempia, ja odotetusti piikin 367 bp (*Dysgonomonas*) osuus on laskenut.

Kuvassa 11 näkyy, että piikki 367 bp (*Dysgonomonas*) on laskenut huomattavasti, kun verrataan näytteiden T1.4 ja T1.5 piikkiprofiileja näytteisiin T1.2 ja T1.3. Piikki 354/355 bp osuus on korkeampi uudemmista näytteistä verrattuna vanhempiin, jossa piikki 357 bp on vallitseva.

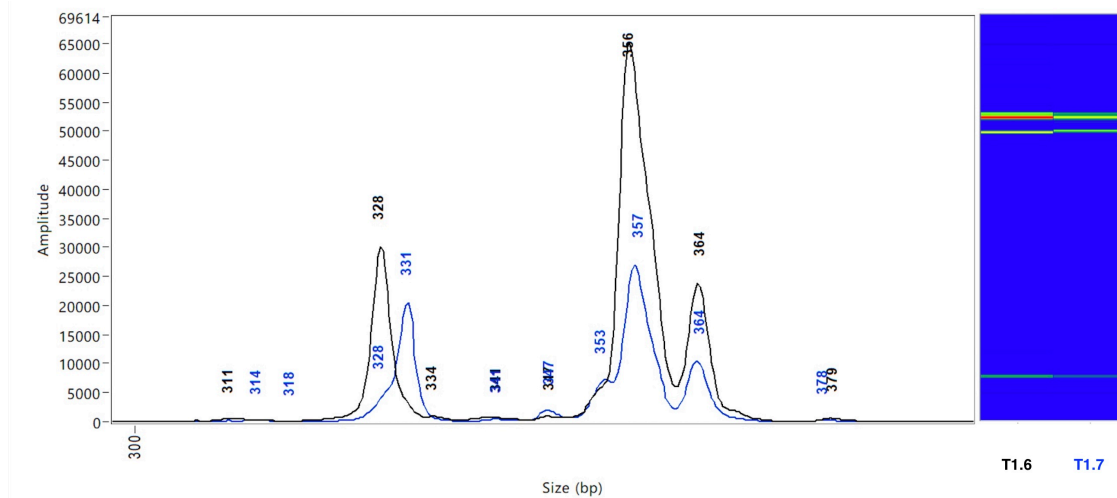
Taulukossa 10 on näytteiden T1.2–T1.5 relevantit tulokset ja *in silico* ennakoitujen emäsparipituudet.

Taulukko 10. Taulukossa näkyy näytteiden T1.2–T1.5 bakteerien ennakoitujen ja FA-tulosten emäsparipituudet sekä niiden osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofilissa.

| Taxon | <i>In silico</i> (bp) | T1.2 | | T1.3 | | T1.4 | | T1.5 | |
|--------------------------------|--------------------------|------------------|--------------|------------------|--------------|------------------|--------------|------------------|--------------|
| | | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| Tuntematon | | 324 | 1,7 | 325 | 2 | 324 | 4,2 | 323 | 0,1 |
| Sphingomonadaceae_unclassified | 317 | 327 | 8,3 | 327 | 8 | 327 | 1 | | |
| Tuntematon | | 329 | 2,3 | 330 | 2,2 | 330 | 3 | 331 | 1,5 |
| Tuntematon | | 355 | | 355 | | 354 | 52,9 | 355 | 50,3 |
| Cloacibacterium | 343 | 357 | 52,3 | 357 | 53,6 | 357 | 17,2 | 357 | 31,8 |
| Tuntematon | | 361 | 5,5 | 361 | 5,2 | 361 | 16,1 | 361 | 10,6 |
| Tuntematon | | 364 | 5,6 | 364 | 4 | 364 | 2,2 | 363 | 2,9 |
| Dysgonomonas | 356 | 367 | 18,1 | 367 | 18,5 | 367 | 0,6 | 368 | 1,1 |

Mikrobiyhteisön muutoksia voidaan näissä näytteissä seurata selvästi. Näytteiden T1.2 ja T1.3 osuudet ovat samanlaiset, kuten ovat myös näytteiden T1.4 ja T1.5. Poikkeuksena näytteen T1.5 *Cloacibacterium* piikki 357 bp, jonka osuus on suurempi kuin näytteen T1.4. Tämän selittää piikin 361 bp pienempi osuus näytteessä T1.5. Näytteissä T1.4 ja T1.5 on poikkeuksellisesti eroteltu piikille 355 bp oma pinta-ala halkaisemalla koko piikki kahtia keskeltä, koska siitä tulee hallitsevampi verrattuna normaalisti piikkiin 357 bp (*Cloacibacterium*).

Kuvassa 12 vertaillaan näytteiden T1.6 ja T1.7 piikkiprofiileja. Näytteiden mikrobiyhteisöt eivät eroa sen suuremmin muista näytteistä, mutta kuvassa halutaan korostaa näytteen T1.6 tuloksia.



Kuva 12. Näytteiden T1.6 ja T1.7 piikkiprofiilit. Näytteen T1.6 amplitudi on huomattavasti korkeampi kuin näytteen T1.7.

Kuvassa 12 voidaan nähdä näytteen T1.6 piikit, jotka ovat korkeammat kuin näytteen T1.7. Bakteeripitoisuus näytteessä ei ollut kuitenkaan suurempi, vaan näyte itse oli monistunut paljon paremmin verrattuna näytteeseen T1.7, vaikka näytteet oli ajettu samaan aikaan. Eroa mikrobiyhteisöissä ei ole muuta kuin että piikki 331 bp on vallitsevampi tuoreemmassa näytteessä T1.7 ja piikin 328 bp osuus on laskenut vanhemmassa näytteessä T1.6. Huomioitavaa että nämä ovat ainoat näytteet, jossa piikki 364 bp on huomattava, mutta mukana ei ole *Dysgonomonas* 366/367 bp-piikkiä.

Taulukossa 11 on näytteiden T1.6–T1.7 relevantit tulokset ja *in silico* ennakoitujen emäsparipituudet.

Taulukko 11. Taulukossa näkyy näytteiden T1.6–T1.7 bakteerien ennakoitujen ja FA-tulosten emäsparipituudet sekä niiden osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofilissa.

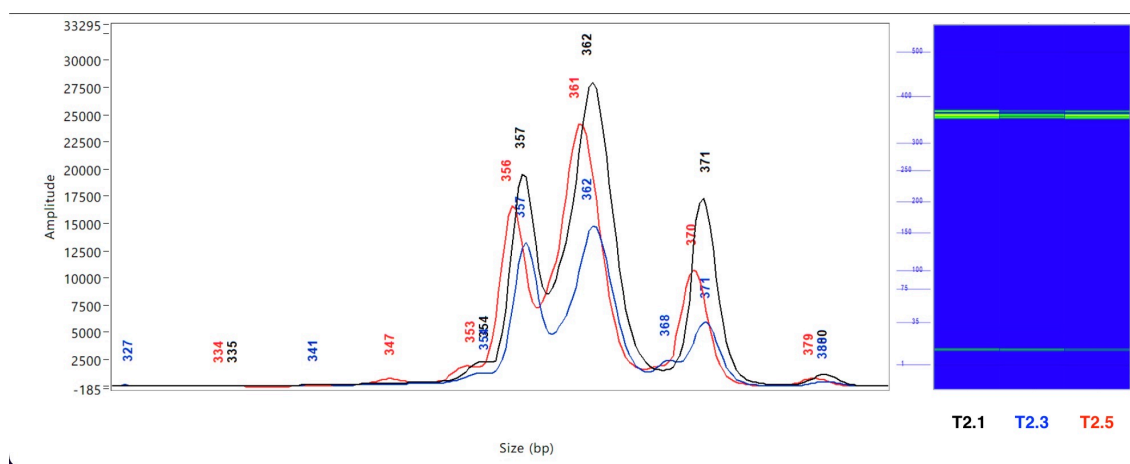
| Taxon | <i>In silico</i> (bp) | T1.6 | | T1.7 | |
|--------------------------------|--------------------------|------------------|--------------|------------------|--------------|
| | | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| Sphingomonadaceae_unclassified | 317 | 328 | 20,7 | 328 | 1,2 |
| Tuntematon | | | | 331 | 29,9 |
| Tuntematon | | 347 | 0,6 | 347 | 2,5 |
| Tuntematon | | | | 353 | 7,9 |
| Cloacibacterium | 343 | 356 | 59,1 | 357 | 43,1 |
| Tuntematon | | 364 | 16,5 | 364 | 13,7 |

Näytteissä T1.6 ja T1.7 ei ollut ollenkaan *Dysgonomonasta*. Kyseisistä näytteistä ei ollut tiedossa sekvenssitietoja, eikä muiden näytteiden tulosten avulla saatu selville, mitkä bakteerit mahdollisesti olisivat piikit 331 bp ja 364 bp.

Tehtaan 1 tulosten perusteella menetelmällä pystytään seuraamaan ainakin yksinkertaisia mikrobiyhteisöjä ja tunnistamaan vallitsevia bakteereja hyödyntämällä sekvensointitietoja ja *in silico* ennakoituja emäsparipituuksia. Näytteet olivat kahdesta eri näytepaikasta, mutta jos näytteet oli otettu samana päivänä, näytepaikka ei vaikuttanut melkein ollenkaan siihen, miltä mikrobiyhteisö näytti. Tuloksista tunnistettiin kolme eri bakteeriryhmää, *Sphingomonadaceae_unclassified* (326–327 bp), *Cloacibacterium* (356–357 bp) ja *Dysgonomonas* (365–367 bp). Varsinaisten ajojen tulokset tehtaalta 1 vahvistivat testiajojen tuloksia.

7.2.2 Tehdas 2

Tehtaan 2 näytteissä haluttiin nähdä menetelmän mahdollisuutta seurata samaa mikrobisyhteisöjä eri ajankohdilta. Ero tehtaan 1 tuloksiin oli, että kyseessä olevissa mikrobisyhteisöissä oli enemmän ja erilaisia vallitsevia bakteeriryhmiä perustuen tiedossa oleviin sekvenssitietoihin. Odotettavissa oli siis vaikeammin tulkittavia piikkiprofileja. Kuvassa 13 vertaillaan kiertokuidun varastotornin näytteitä kolmelta eri ajankohdalta.



Kuva 13. Kiertokuidun näytteiden piikkiprofilit. Mikrobisyhteisö ei ole merkityksellisesti muuttunut.

Sekvenssitiedoista nähtiin, että näytteissä on samat vallitsevat bakteeriryhmät. Piikkiprofilit yhtenevät sekvenssitietojen kanssa. Kuvassa 13 piikki 368 bp näytteessä T2.3 on mahdollisesti myös muissa näytteissä, mutta ei tunnistettavissa.

Taulukossa 12 on näytteiden T2.1, T2.3 ja T2.5 relevantit tulokset ja *in silico* ennakoidut emäsparipituudet.

Taulukko 12. Taulukossa näkyy näytteiden T2.1, T2.3 ja T2.5 bakteerien ennakoitujen ja FA-tulosten emäsparipituudet sekä niiden osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | <i>In silico</i> (bp) | T2.1 | | T2.3 | | T2.5 | |
|--|-----------------------|---------------|-----------|---------------|-----------|---------------|-----------|
| | | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| Tuntematon | | 354 | 3,3 | 354 | 2,9 | 353 | 2,9 |
| Ethanoligenens | 348 | 357 | 26,1 | 357 | 31,9 | 356 | 25,3 |
| Weeksellaceae_unclassified | 343/365 | 362 | 48,9 | 362 | 46 | 361 | 52,9 |
| Tuntematon | | | <0 | 368 | 4,8 | | <0 |
| Prevotellaceae_unclassified & Prevotella_7 | 354 | 371 | 19,8 | 371 | 12,7 | 370 | 16,5 |
| Enterococcus | 365 | 380 | 1,2 | 380 | 0,9 | 379 | 1 |

Näytteet ovat siis kolmelta eri ajankohdalta, ja mikrobiyhteisö on nähtävästi pysynyt samana taulukon 12 arvojen mukaan. Huomiona näkyy, että *Weeksellaceae_unclassified* bakteerin-*in silico* bp pituudet ei ns. sovi muiden näytteiden trendiin. Bakteerilla on kuitenkin nähtäväsi heimon sisällä paljon vaihtelua, koska valituista sekvensseistä Silvestä tuli 2 eri emäsparipituutta ennakoitua *in silico*.

Taulukossa 13 näkyy näytteen T2.1 relevantit tulokset ja vertailu näytteen sekvenssitietoihin. Tulokset tästä edustavat kuvan 13 kaikkien näytteiden tuloksia.

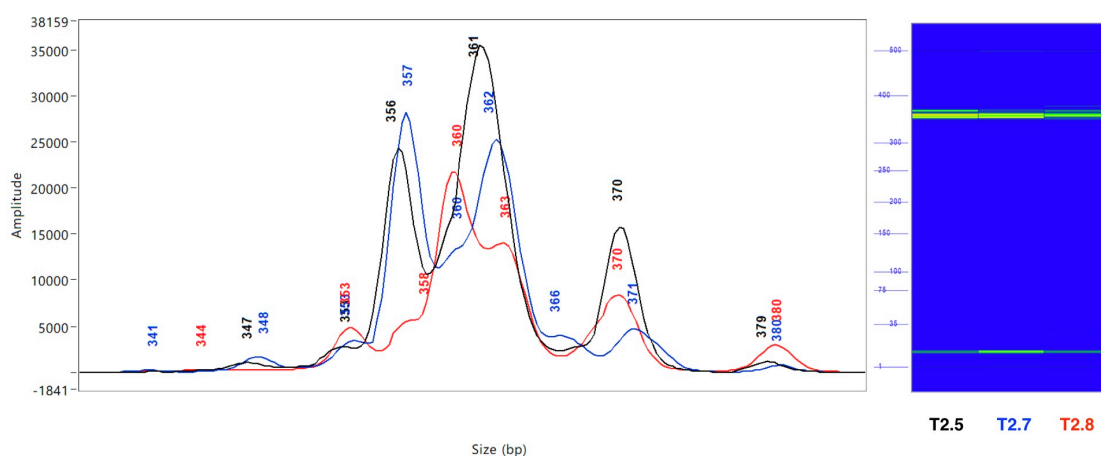
Taulukko 13. T2.1 tunnistetut bakteerit, niiden osuudet FA-laitteelta ja sekvenssitiedoista saadut todelliset osuudet mikrobiyhteisöstä. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | T2.1 | | |
|--|---------------|-----------|---------|
| | FA-tulos (bp) | Osuus (%) | Seq (%) |
| Ethanoligenens | 357 | 26,1 | 25,8 |
| Weeksellaceae_unclassified | 362 | 48,9 | 33,7 |
| Prevotellaceae_unclassified & Prevotella_7 | 371 | 19,8 | 15,0 |
| Enterococcus | 380 | 1,2 | 3,3 |

Sekvenssitiedot yhtenevät FA:lta saatujen tietojen kanssa. Sekvenssitiedoissa oli erillisesti *Prevotellaceae_unclassified* ja *Prevotella_7*, jossa ensimmäinen

bakteeri oli vallitsevampi. Ne ovat kuitenkin laskettu yhteen, koska mahdollisesti bakteerien emäsparipituudet ovat liian lähellä toisiaan, jotta FA-laite pystyisi niitä erottamaan toisistaan.

Kuvassa 14 vertaillaan myös kolmelta eri ajankohdalta kiertokuidun varastotorin näytteitä. Mukana edellisestä kuvasta 13 on näyte T2.5, joka edustaa ns. normaalia mikrobiyhteisöä. Näytteet T2.7 ja T2.8 ovat tuoreempia näytteitä.



Kuva 14. Tuorempien kiertokuidun näytteiden piikkiprofiilien vertailu. Piikkiprofiileissa selviä eroavaisuuksia.

Kuvasta 14 voidaan piikkiprofiilien ero huomata välittömästi. Näyte T2.5 muistuttaa vielä näytettä T2.7, mutta näytteen T2.7 piikkiprofiilissa vallitsevat piikit 357 bp (*Ethanolgenens*) ja 362 bp (*Weeksellaceae_unclassified*) ovat laskeneet pitoisuuksissa näytteeseen T2.8 verrattuna, jossa piikki 360 bp on noussut vallitsevaksi bakteeriryhmäksi. Samat bakteeriryhmät edelleen, mutta niiden osuuksissa voidaan nähdä selvä muutos. Huomioitavaa myös, että näytteissä T2.7 ja T2.8 piikki 360 bp oli tunnistettavissa, mutta muissa kiertokuidun näytteissä sitä ei tunnistettu, vaikka voitaisiin todeta, että kyseinen piikki on olemassa.

Taulukossa 14 on näytteiden T2.5, T2.7 ja T2.8 relevantit tulokset ja *in silico* ennakoidut emäsparipituudet.

Taulukko 14. Taulukossa näkyy näytteiden T2.5, T2.7 ja T2.8 bakteerien ennakoitua ja FA-tulosten emäsparipituudet sekä niiden osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | <i>In silico</i> (bp) | T2.5 | | T2.7 | | T2.8 | |
|--|--------------------------|------------------|-----------|------------------|-----------|------------------|-----------|
| | | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| Tuntematon | | 353 | 2,9 | 353 | 3,5 | 353 | 8,2 |
| Ethanoligenens | 348 | 356 | 25,3 | 357 | 31,5 | 357 | 6,7 |
| Tuntematon | | | <0 | 360 | 9,9 | 360 | 40,2 |
| Weeksellaceae_unclassified | 343/365 | 361 | 52,9 | 362 | 42,5 | 363 | 22,6 |
| Prevotellaceae_unclassified & Prevotella_7 | 354 | 370 | 16,5 | 371 | 5,8 | 370 | 17,5 |
| Enterococcus | 365 | 379 | 1 | 380 | 0,7 | 380 | 4,4 |

Taulukon 14 mukaan piikin 360 bp osuus on noussut 30 %-yksikköä näytteestä T2.7 näytteeseen T2.8 ja piikki 363 bp, *Weeksellaceae_unclassified*, osuus on laskenut. Lukuja käytetään suuntaa antavana, joten välttämättä niin suurta heittoa ei ole tapahtunut näytteenottoaikalla. Taulukosta voi myös nähdä, kuinka *Weeksellaceae_unclassified*-bakteerin osuus on laskenut ajan myötä. Mahdollista on kuitenkin, että näytteessä T2.5 piikissä 361 bp on sekä piikki 360 bp että bakteerin *Weeksellaceae_unclassified* piikki yhdistynyt.

Prevotellaceae_unclassified ja *Prevotella_7* osuus on myös ensimmäistä kertaa näytteessä T2.8 suurempi kuin *Ethanoligenensin*. Tämän yksityiskohdan voi nähdä kuvasta 14. Sen huomion ja *in silico* emäsparipituuksien takia ei uskota, että bakteerien piikit olisivat tunnistuksessa sekoitettu keskenään.

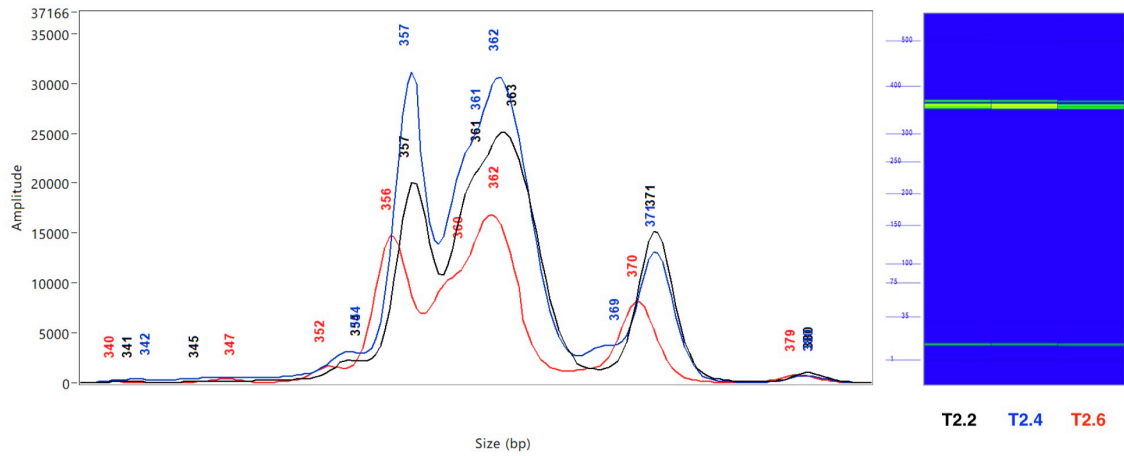
Taulukossa 15 näkyy näytteen T2.8 relevantit tulokset ja vertailu näytteen sekvenssitietoihin.

Taulukko 15. T2.8 tunnistetut bakteerit, niiden osuudet FA-laitteelta ja sekvenssitiedoista saadut todelliset osuudet mikrobiyhteisöstä. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | T2.8 | | Seq (%) |
|--|---------------|-----------|---------|
| | FA-tulos (bp) | Osuus (%) | |
| Ethanoligenens | 357 | 6,7 | 6 |
| Weeksellaceae_unclassified | 363 | 22,6 | 25,1 |
| Prevotellaceae_unclassified & Prevotella_7 | 370 | 17,5 | 18,1 |
| Enterococcaceae_unclassified | 380 | 4,4 | 1,2 |

Näytteessä T2.8 oli ongelmana, että vallitsevin piikki oli 360 bp. Vallitsevin bakteeri sekvenssien mukaan oli kuitenkin *Weeksellaceae_unclassified*, joka oli määritelty jo piikiksi 361–363 bp. Muuta syytä ei ole ristiriitaiselle tulokselle, kuin että mikrobiyhteisössä on monia bakteereja, joilla on sama emäsparipituus. Kyseiset bakteerit ovat sitten monistuneet ja PCR-tuotteessa sitten olleet vallitsevat monistujat, vaikka prosessissa näytteenottoaikalla ne eivät ole vallitsevassa asemassa mikrobiyhteisössä. Tehtaan näytteistä kaikista oli sekvenssitiedot tiedossa paitsi yhdestä, ja *Weeksellaceae_unclassified* on aina vallitseva sekvenssien mukaan ja FA:lta tuloksissa suurin piikki on aina 361–363 bp. *Ethanoligenens* piikki 357 bp on näytteessä T2.8 osuudeltaan laskenut ja bakteeri oli muissa näytteissä (T2.1–T2.7) vallitsevampi, joten piikin osuuden lasku sekä FA:n tuloksissa, että sekvenssitiedoissa vahvistaa kyseisen piikin kuuluvan bakteerille *Ethanoligenens*.

Kuvassa 15 on kartonkikoneen kirkassuodossäynteet.



Kuva 15. Kartonkikoneen kirkassuodossäynteet kolmelta eri ajankohdalta.

Piikkiprofiilit näytteiden välillä ovat käytännössä samanlaisia. Näyte T2.6 on emäsparipituudeltaan systemaattisesti hieman lyhyempi, mutta muuten ainoa huomioitava ero piikkiprofiilien välillä on näytteen T2.4 piikki 369 bp. Piikit 360–361 bp identifioitiin manuaalisesti.

Taulukossa 16 on näytteiden T2.2, T2.4 ja T2.6 relevantit tulokset ja *in silico* ennakoidut emäsparipituudet.

Taulukko 16. Taulukossa näkyy näytteiden T2.2, T2.4 ja T2.6 bakteerien ennakoitujen ja FA-tulosten emäsparipituudet sekä niiden osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | <i>In silico</i> (bp) | T2.2 | | T2.4 | | T2.6 | |
|--|--------------------------|------------------|-----------|------------------|-----------|------------------|-----------|
| | | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| Tuntematon | | 354 | 2,5 | 354 | 2,6 | 352 | 2,4 |
| Ethanoligenens | 348 | 357 | 21,8 | 357 | 27,1 | 356 | 22 |
| Tuntematon | | 361 | 14 | 361 | 13,3 | 360 | 14,3 |
| Weeksellaceae_unclassified | 343/365 | 363 | 46,3 | 362 | 43,5 | 362 | 46,2 |
| Tuntematon | | | <0 | 369 | 2,2 | | <0 |
| Prevotellaceae_unclassified & Prevotella_7 | 354 | 371 | 14,3 | 371 | 10,4 | 370 | 13,4 |
| Enterococcus | 365 | 380 | 0,9 | 380 | 0,5 | 379 | 1,1 |

Kartonkikoneen kirkassuodosten näytteiden mikrobiyhteisöä pystytään seuraamaan luotettavasti. Mikrobiyhteisö ei eroa suuresti kiertokuitu näytteistä T2.1, T2.3 ja T2.5 (kuva 13). Erona vain, että taulukossa 16 nähtävissä näytteissä on suurempi osuus piikillä 360–361 bp.

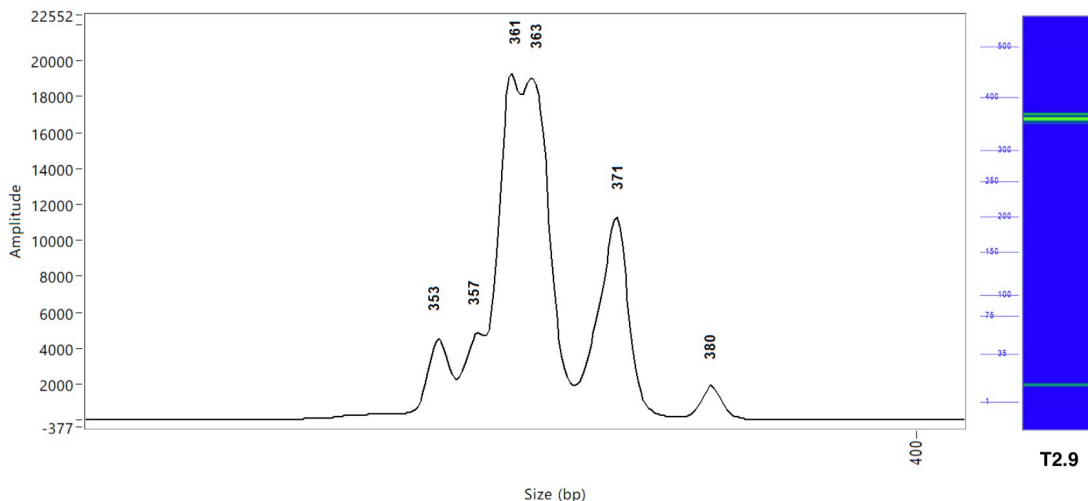
Taulukossa 17 näkyy näytteen T2.8 relevantit tulokset ja vertailu näytteen sekvenssitietoihin.

Taulukko 17. T2.4 tunnistetut bakteerit, niiden osuudet FA-laitteelta ja sekvenssitiedoista saadut todelliset osuudet mikrobiyhteisöstä. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

| Taxon | T2.4 | | |
|--|---------------|-----------|---------|
| | FA-tulos (bp) | Osuus (%) | Seq (%) |
| Ethanoligenens | 357 | 27,1 | 24,9 |
| Weeksellaceae_unclassified | 362 | 43,5 | 30,6 |
| Prevotellaceae_unclassified & Prevotella_7 | 371 | 10,4 | 14,3 |
| Enterococcus | 380 | 0,5 | 1,7 |

T2.4 FA-tulokset yhtenevät sekvenssitietojen kanssa, ja vahvistaa johtopäätöksiä, mitä tehty tähän mennessä. *Enterococcus* tunnistettiin piikiksi 380 bp hyödyntäen *in silico* ennakoituja emäsparipituuksia ja toisen tehtaan tuloksia.

Kuvassa 16 on pelkästään näyte T2.9. Näyte T2.9 on eri kohdasta prosessia kuin muut tehtaan 2 näytteet. Näytteen tulokset otettiin mukaan tulosten laadun vuoksi.



Kuva 16. Näytteen T2.9 piikkiprofilii. Piikkiprofilista voidaan nähdä selvästi 6 eri piikkiä.

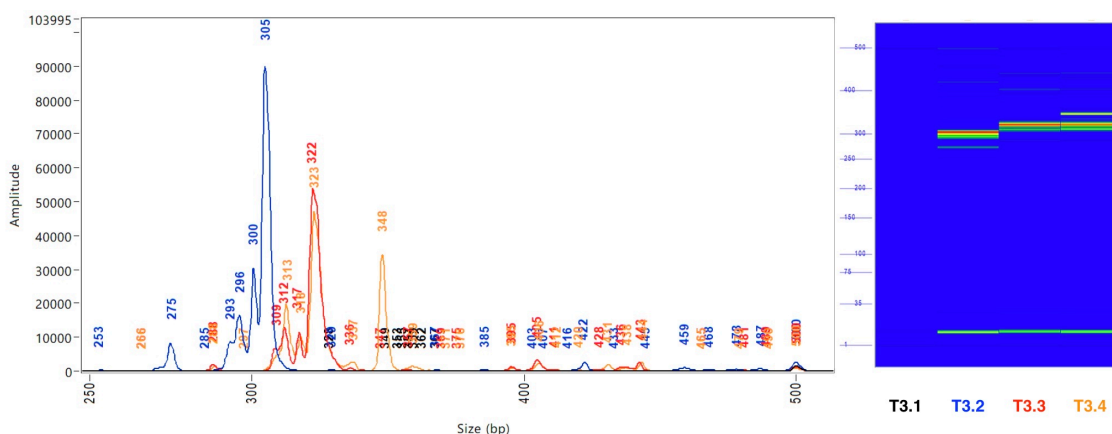
Kuvassa 16 näkyy, että näytteen T2.9 piikkiprofilissa on samat piikit verrattessa muihin tehtaan 2 näytteisiin. Piikit 361 bp sekä 363 bp ovat paljon selvemmät verrattuna muihin näytteisiin. FA-laite ei oletetusti pysty erottamaan piikkejä tai tunnistamaan niitä, jos ne eroavat toisistaan 1–3 emäsparilla. Kyseisessä näytteessä se on kuitenkin onnistunut, koska bakteerit ovat samoissa pitoisuuksissa näytteessä (piikit 361 bp ja 363 bp). Sekvenssitietoja ei ole olemassa näytteelle, minkä avulla voitaisiin saada selville piikin 361 bp identiteetti. Yksi spekulatio on, että kummatkin piikit olisivat *Weeksellaceae*-heimoisia bakteereja, mutta koska sekvensointia varten monistetaan 16S-rRNA-geenin-alue V4 tai alueet V4 – V5, niin bakteeria ei pystytä tunnistamaan sukutasolle saakka näistä geenin alueista. Tässä työssä monistetaan 16S-rRNA-geenin alue V1–V2, joten mahdollisesti kyseisessä bakteerissa taas olisi siellä geenissä enemmän vaihtelua lajitasolle saakka.

Tehtaan 2 näytteistä pystytään toteamaan, että mikrobiyhteisöjä pystytään seuraamaan luotettavasti. Vallitsevia bakteereja pystytään seuraamaan, vaikka

niitä on enemmän verrattuna tehtaan 1 näytteisiin, jotka olivat yksinkertaisempia mikrobiyhteisö mielessä. Fragment Analyzer -tulokset olivat myös paljon lähempänä sekvensointituloksia tehtaan 2 näytteissä verrattuna tehtaan 1 tuloksiin. On mahdollista, että jos mikrobiyhteisössä on enemmän vallitsevia bakteereja, niin niiden osuudet vääristyvät vähemmän monistuksessa. Poikkeuksena näyte T2.8, jossa epäselvänä on piikin 360 bp identiteetti ja että se on vallitseva piikki kyseisessä näytteessä. Tuloksista tunnistettiin 4 eri bakteerisukua, *Ethanoligenens* (356–357 bp), *Weeksellaceae_unclassified* (361–363 bp), *Prevotellaceae_unclassified* & *Prevotella_7* (370–371 bp) ja *Enterococcus* (380 bp).

7.2.3 Tehdas 3

Tehtaan 3 näytteet olivat kaikki matalan pitoisuuden näytteitä. Käytännössä kyseiset näytteet sisälsivät siis kokonaisuudessaan erittäin vähän DNA:ta. Näyte T3.1 on ensimmäinen ja näyte T3.4 viimeisin prosessin näytteenotto kohta. Tarkoituksena oli selvittää menetelmän tarkkuutta seuraamalla *Tepidimonas*- ja *Pseudoxanthomonas*-bakteerisukujen osuuden laskua prosessissa, minkä vuoksi kaikki piikkiprofiilit ovat yhdessä kuvassa. Kuvassa 17 on poikkeuksellisesti zoomattu laajemmalle alueelle piikkiprofiilissa. Piikkiprofiilista haluttiin korostaa pienien piikkien määrä, joka on lukumäärältään enemmän kuin normaalisti. Se johtuu todennäköisimmin pidemmästä qPCR-ajosta, koska matalan pitoisuuden DNA-näytteet monistuivat hitaammin, minkä vuoksi kaikkea muuta turhaa monistuu myös samalla. Nämä pienet piikit saattavat kuitenkin häiritä tulosten tulkintaa, ja voi aiheuttaa lisätyötä, kun piikit pitää yksi kerrallaan poistaa manuaalisesti.



Kuva 17. Näytteiden T3.1–T3.4 piikkiprofiilit UM (500 bp) saakka.

Kuvaa 17 tarkastellessa nähdään, että näytteestä T3.1 ei ole mitään piikkejä yli detektorirajan, eli mikään piikeistä ei ole siis luotettava. Näytteen piikit ovat mustalla ja ovat hädän tuskin nähtävissä. Yllättävää kuitenkin, että se on ainoa näyte, jonka piikit ovat odotettavissa emäsparipituuksissa. *Tepidimonas*-bakteerisuku tunnistettiin piikiksi 367 bp ja *Pseudoxanthomonas*-bakteerisuku tunnistettiin piikiksi 362 bp perustuen muiden tehtaiden tuloksiin. Näitä ei kuitenkaan sisällytetä lopullisiin tuloksiin. Muissa näytteissä piikkiprofiilit ovat systemaattisesti emäsparipituuksiltaan liian lyhyitä, eikä pystytä luotettavasti vertailemaan tuloksia muiden tehtaiden tuloksiin. Jos ei kuitenkaan kiinnitetä huomiota emäsparipituuksiin, vaan pelkästään piikkien kokoihin, huomataan, että näytteen T3.2 piikki 305 bp on suurin, verrattuna näytteen T3.3 sitä täsmäävään piikkiin 322 bp, ja näistä pienin piikki on näytteen T3.4 sitä vastaava piikki 323 bp. Nämä piikit tunnistettiin *Tepidimonas*-bakteerisuvuksi, ja jonkinlainen piikkien lasku koossa voidaan tässä nähdä. Näytteestä T3.2 bakteerin osuus laskee koko mikrobiyhteisöstä noin 20 % näytteeseen T3.3, ja siitä noin 10 % näytteeseen T3.4 sekvenssitietojen mukaisesti.

Taulukossa 18 on näytteiden T3.2–T3.4 FA-tulokset bakteereille *Tepidimonas* ja *Pseudoxanthomonas*.

Taulukko 18. Taulukossa näkyy näytteiden T3.2–T3.4 bakteerien *Tepidimonas* ja *Pseudoxanthomonas* osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

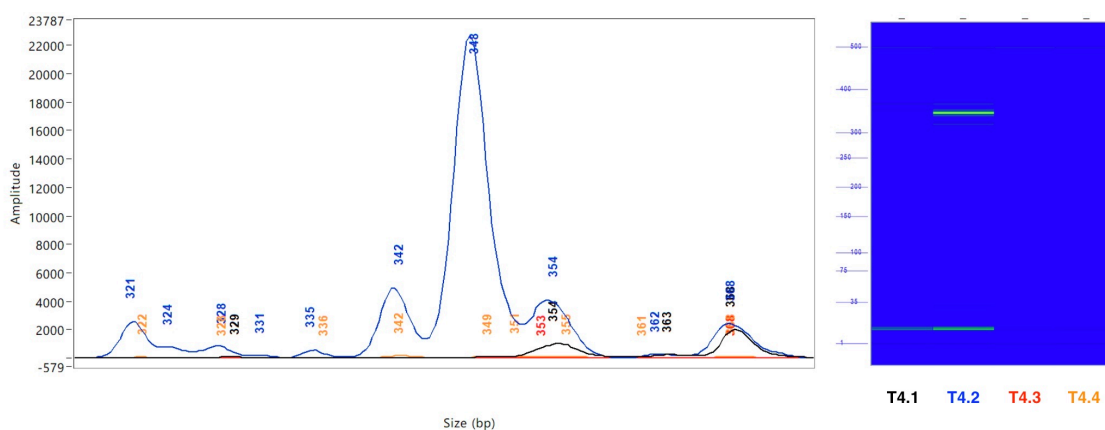
| Taxon | T3.2 | | T3.3 | | T3.4 | |
|--------------------------|---------------|-----------|---------------|-----------|---------------|-----------|
| | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| <i>Pseudoxanthomonas</i> | 296 | 11,5 | 312 | 11,5 | 313 | 17,2 |
| <i>Tepidimonas</i> | 305 | 59,3 | 322 | 67,8 | 323 | 47,2 |

Taulukon 18 osuusarvojen mukaan ei nähdä samanlaista laskua bakteereilla, mikä voidaan todeta kuvaa 17 katsomalla. Näytteissä T3.3 ja T3.4 on myös enemmän vallitsevia bakteereja mikrobiyhteisössä verrattuna näytteeseen T3.2 sekvenssitietojen mukaan, joten oli odotettavissa, että näytteessä T3.2 olisi eniten *Tepidimonasta* ja *Pseudoxanthomonasta*. Näin ei kuitenkaan ole, kun puhutaan puhtaasti bakteerien piikkien osuuksista piikkiprofiileista.

Tehtaan 3 näytteet ovat emäsparipituuksista puhuttaessa epäluotettavia, eivätkä ne ole onnistuneita. Sama päätelmä voidaan tehdä puhuttaessa bakteerien osuuksien seuraamisesta tarkemmin mikrobiyhteisössä. qPCR-ajo oli onnistunut eikä monistumisessa näkynyt olevan ongelmaa, joten ongelma on itse näytteissä tai näytteenkäsittelyssä, minkä vuoksi emäsparipituudet ovat liian lyhyitä. Matalan pitoisuuden näytteet monistuvat hitaammin, joten se selittää, miksi piikkiprofiileissa normaalia enemmän pieniä piikkejä. Luotettavia vertailtavia piikkejä ei ollut tehtaan piikkiprofiileissa, joten tehtaasta ei sisällytetä tuloksia yhteenvetoon. Tunnistettavia bakteereja kuitenkin oli, kuten *Tepidimonas*, *Pseudoxanthomonas* ja *Ruminococcaceae_unclassified* (näyte T3.4, piikki 348 bp).

7.2.4 Tehdas 4

Tehtaan 4 näytteissä haluttiin selvittää haitallisen *Meiothermus*-bakteerisuvun piikki. Tehtaalla esiintyi myös muista tehtaista paljon enemmän eroava mikrobiyhteisö, ja haluttiin selvittää, olisiko niitä mahdollista analysoida FA-laitteella. Näytteet T4.1 ja T4.3 ovat referenssinäytteitä, jotka olivat siis puhtaita kartonkinäytteitä. Näytteet T4.2 ja T4.4 ovat kartonkinäytteitä, jossa esiintyi punaisia läikkiä, viitaten kontaminaatioon. Kaikki näytteet olivat kartonkeja ja matalan pitoisuudennäytteitä. Referenssinäytteistä ei odotettu tulevan piikkiprofiiliin merkittäviä piikkejä, kun taas läikkänäytteissä odotettiin asian olevan toisin. Kuvassa 18 kartonkinäytteiden piikkiprofiilit.

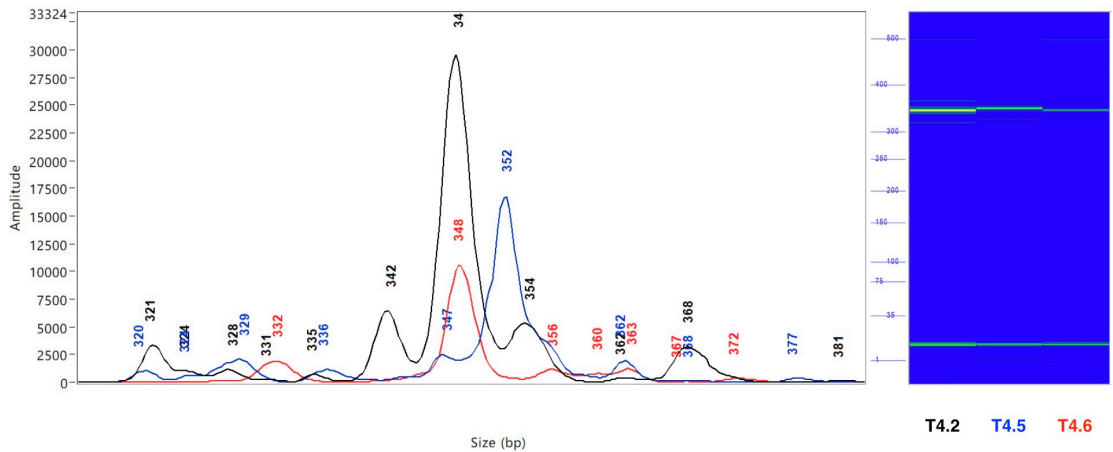


Kuva 18. Näytteiden T4.1–T4.4 piikkiprofiilit. Lopputuotenäytteiden piikkiprofiilien vertailua.

Referenssinäytteet T4.1 ja T4.3 odotettavasti eivät tuottaneet merkittäviä piikkejä, kuten kuvassa 18 voidaan nähdä. Näyte T4.4 ei tuottanut tuloksia. Ainoastaan näytteestä T4.2 tuli tuloksia, missä vallitseva piikki 348 bp tunnistettiin *Meiothermukseksi*. Sekvenssitiedoissa *Meiothermus* kattaa näytteestä T4.4 mikrobiyhteisöstä 51,4 %. FA-tuloksissa piikki 348 bp kattaa piikkiprofiilista 58,2 %, joten vahvasti oletettiin piikki 348 bp *Meiothermukseksi*. Muut bakteerit/piikit olivat 10 %:n paikkeilla sekä sekvenssitiedoissa että FA-tuloksissa.

Tehtaalta otettiin myös 2 prosessinäytettä, näytteet T4.5 ja T4.6. Näytettä T4.6 vertailtiin näytteeseen T4.2, jonka tulos oli ainoa järkevä kyseiseltä FA-ajolta.

Tämä varmistaisi vielä näytteen T4.2 tulosten luotettavuuden. Näytteellä T4.5 oli erilainen mikrobiyhteisö, joten odotettiin sen piikkiprofiilin poikkeavan muiden näytteiden piikkiprofileista. Kuvassa 19 on edellä mainittujen näytteiden piikkiprofiilien vertailu.



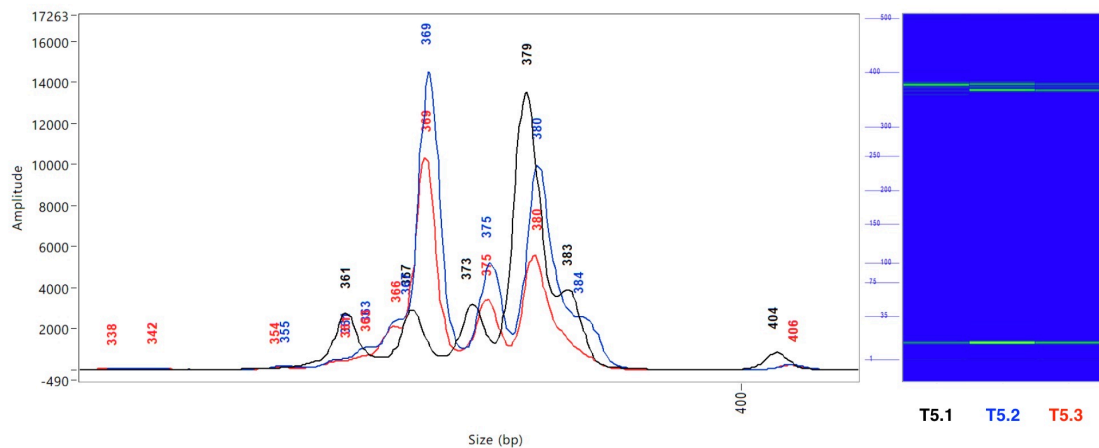
Kuva 19. Näyte T4.2 on läikkänäyte, jonka piikkiprofiilia vertaillaan kahden prosessinäytteen piikkiprofiilien kanssa.

Näytteiden T4.2 ja T4.6 vallitseva bakteerisuku sekvenssitietojen mukaan oli *Meiothermus*. Piikkiprofiileissa kuvasta 19 voidaan todeta siis, että *Meiothermus* tuottaa piikin 348 bp. Näyte T4.5 erosi suuresti muihin näytteisiin mikrobiyhteisöllisesti (myös kaikkien muiden tehtaiden), eikä pystytty määrittämään luotettavasti mitään bakteeriryhmiä piikkiprofiilista.

Tehtaasta 4 saatiin määritettyä haittabakteerin *Meiothermus* identiteetti piikkiprofiileista, joka on piikki 348 bp. Näytteissä esiintyi myös muita piikkejä, joten nähtävästi on mahdollista analysoida kyseistä mikrobiyhteisöä. Muita bakteereja ei pystytty tunnistamaan luotettavasti pelkästään näistä tuloksista.

7.2.5 Tehdas 5

Tehtaasta 5 valittiin kolme näytettä, joista näytteet T5.1 ja T5.3 olivat matalan pitoisuuden näytteitä. Analyysit olivat kuitenkin onnistuneita verrattuna tehtaan 3 tuloksiin. Kuvassa 20 piikkiprofiilien vertailu.



Kuva 20. Tehtaan 5 näytteiden piikkiprofiilit. Piikkiprofileista nähdään näytteen T5.1 mikrobiryhmän eroavan muista, mutta ei suuresti.

Tehtaan 5 näytteet valittiin, koska mikrobiryhmät niissä sisälsivät bakteerisukuja, joita oli muissa tehtaissa ja niiden tunnistaminen auttaisi muiden tulosten tulkintaa. Kuvassa 20 esimerkiksi piikki 369 bp tunnistettiin *Tepidimonas*-bakteerisuvuksi, jota esiintyi tehtaan 6 näytteissä ja sekvenssitietojen mukaan myös tehtaassa 3. Oli myös ongelmia, kuten *Thermicanus*-bakteerisuvun tunnistus. Sen todettiin olevan piikki 373 bp näytteessä T5.1 ja sitä vastaavan piikin olevan yhdistynyt piikkiin 375 näytteessä T5.2. *Thermicanusta* esiintyi myös tehtaan 6 ja tehtaan 7 mikrobiryhmäisissä.

Taulukossa 19 on näytteiden T5.1, T5.2 ja T5.3 relevantit sekvenssitulokset.

Taulukko 19. Taulukossa näkyy näytteiden T5.1, T5.2 ja T5.3 relevanttien bakteerien sekvenssitiedot. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on mikrobiyhteisöstään.

| | T5.1 | T5.2 | T5.3 |
|-------------------------------|--------------|------|------|
| Taxon | Sequence (%) | | |
| Clostridium_sensu_stricto_N | 14,4 | 4,4 | 4,9 |
| Tepidimonas | 5,5 | 21,1 | 29,6 |
| Thermicanus | 10,6 | 11,6 | 2,9 |
| Pseudomonadaceae_unclassified | 1,3 | 16,5 | 18,3 |
| Anoxybacillus | 20,2 | 17,6 | 11,7 |
| Bacillales_unclassified | 17,6 | 0 | 4,5 |
| Bacillus | 0,9 | 7,7 | 0,1 |

Taulukon 19 tietojen avulla on päätelty taulukon 20 tulokset. Taulukossa 20 on näytteiden T5.1, T5.2 ja T5.3 relevantit tulokset ja *in silico* ennakoitujen emäsparipituudet.

Taulukko 20. Taulukossa näkyy näytteiden T5.1, T5.2 ja T5.3 bakteerien ennakoitujen ja FA-tulosten emäsparipituudet sekä niiden osuudet piikkiprofiileista. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa.

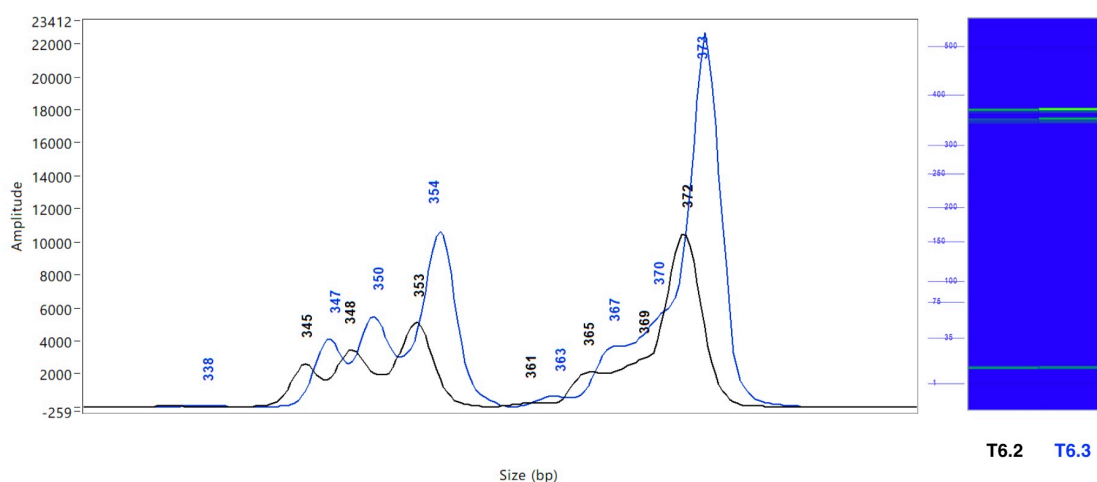
| | | T5.1 | | T5.2 | | T5.3 | |
|-------------------------------|--------------------------|------------------|--------------|------------------|--------------|------------------|--------------|
| Taxon | <i>In silico</i> (bp) | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| | | 361 | 12,1 | 361 | 1,6 | 361 | 1,9 |
| Clostridium_sensu_stricto_N | 333/336 /354 | 367 | 10,7 | 367 | 5,9 | 366 | 7,2 |
| Tepidimonas | 344/351 /355 | | <0 | 369 | 36,6 | 369 | 40,9 |
| Thermicanus | 358/368 | 373 | 11,6 | | <0 | | <0 |
| Pseudomonadaceae_unclassified | 345 | | <0 | 375 | 14 | 375 | 14,8 |
| Anoxybacillus | 357 | 379 | 49,6 | 380 | 31,8 | 380 | 30 |
| Bacillales_unclassified | 357 | 383 | 12,6 | | <0 | | <0 |
| Bacillus | 327/364 | | <0 | 384 | 5,8 | | <0 |

Kaikki päätelmät hyödyntävät taulukon 19 tietoja. Taulukosta 20 näkyy, että *Anoxybacillusta* oli kaikissa näytteissä, minkä avulla tunnistettiin kyseinen piikki 379–380 bp sille. Seuraavaksi tunnistettiin *Tepidimonas*, joka oli vallitseva näytteissä T5.2 ja T5.3. Sen piikki 369 bp on kuvissa suurempi kuin piikki 379–380 (*Anoxybacillus*), mutta osuuksien perusteella ja siten missä näytteissä piikki esiintyy, pääteltiin piikki 369 bp *Tepidimonakselle*. *Pseudomonadaceae_unclassified* pääteltiin piikiksi 375 bp samalla tavalla kuin *Tepidimonas*. Piikkien ei kuitenkaan uskota menneen sekaisin, johtuen niiden osuuksista ja siitä, miten ne täsmäytyvät sekvenssitietojen kanssa. *Thermicanus* esiintyi näytteissä T5.1 ja T5.2 piikkinä 373 bp. Piikkiä ei kuitenkaan pystytä erottamaan näytteessä T5.2, koska se on yhdistynyt piikkiin 375 bp (*Pseudomonadaceae_unclassified*). Tehtaista 6 ja 7 tunnistettiin bakteeri samoihin emäsparipituuksiin, minkä vuoksi ei uskota, että bakteeri olisi piikki 383 bp näytteessä T5.1 ja piikki 384 bp näytteessä T5.2. *Clostridium_sensu_stricto_N* uskotaan olevan piikki 367 bp, koska trendi, jota bakteeri tekee FA-tuloksissa, täsmää trendin kanssa, mitä sekvenssitiedoissa näkyy. Sen ei uskota olevan piikki 361 bp, koska mikrobiyhteisössä on bakteereja, joilla sama trendi, mutta pienemmät pitoisuudet. Mahdolliset bakteerit piikille 361 bp ovat siis *Ruminococcus* tai *Melioribacteraceae*. Näytteessä T5.1 jää jäljelle piikki 383 bp, jonka ainoa mahdollinen kandidaatti on *Bacillales_unclassified* ja näytteessä T5.2 jää piikki 384 bp, jonka ainoa kandidaatti on *Bacillus*.

Tehtaasta 5 tunnistettiin seitsemän eri bakteeria, joihin kuului myös tutkittavana olleet bakteerit *Tepidimonas* ja *Thermicanus*. Tulokset olivat onnistuneita, vaikka kaksi näytteistä oli matalan pitoisuuden näytteitä. Näytteistä tehtiin myös rinnakkaisnäytteet, mutta tuloksissa käsitelty näytteet, joissa ei ollut systemaattisia emäsparieroja. Ne on kuitenkin otettu huomioon johtopäätöksissä, kun on määritelty rajat emäspareilla bakteerien tunnistamiseen. *In silico*-emäsparipituuksissa on suurta vaihtelua, joten niistä ei paljon ollut apua bakteerien identifiomisessa.

7.2.6 Tehdas 6 ja 7

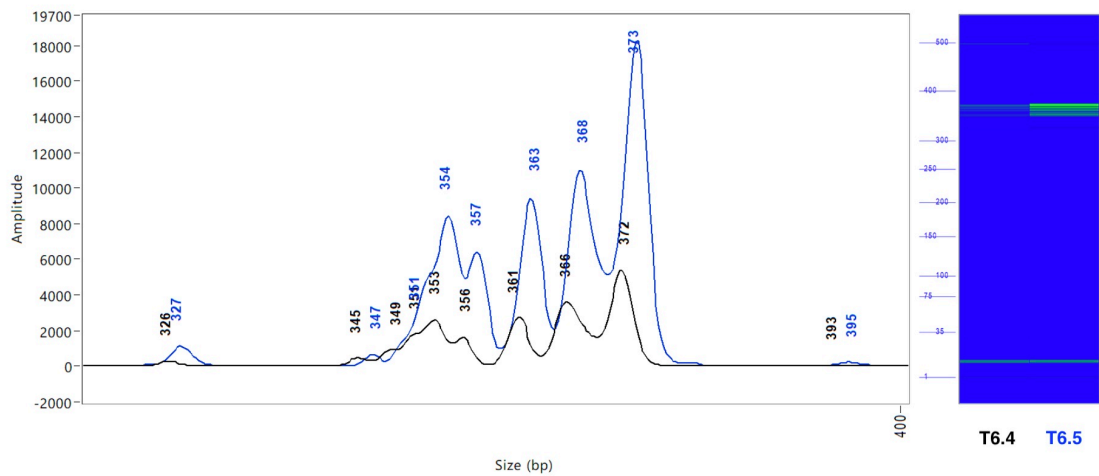
Kaikki näytteet ovat eri kohdista prosessia. Näytteet T6.2 ja T6.3 ovat kuitenkin samasta kartonkikoneesta. Näytteet T6.4 ja T6.5 ovat samalta tehdasalueelta, mutta toiselta kartonkikoneelta. Voidaan odottaa siis systeemin sisällä olevien näytteiden tuottavan samanlaisia piikkiprofiileja. Kiinnostuksena oli tunnistaa tiettyjä bakteerisukuja, kuten *Thermus*, *Thermicanus* ja *Prevotella_7*. Tehtaasta 7 valittiin vain yksi näyte, näyte T7, joka sisälsi samoja bakteereja kuin tehtaasta 6 näytteet, kuten *Thermus* ja *Thermicanus*. Nämä ovat vallitsevia bakteerisukuja näytteen mikrobiyhteisössä. Kuvassa 21 vertaillaan näytteiden T6.2 ja T6.3 piikkiprofiileja.



Kuva 21. Näytteiden T6.2 ja T6.3 piikkiprofiilit. Näytteet ovat samasta kartonkikoneesta.

Näytteet T6.2 ja T6.3 ovat erittäin samanlaisia piikkiprofiilien perusteella, kuten oli odotettu. Näytteen T6.2 emäsparipituudet ovat hieman lyhyempiä ja piikit ovat pienempiä, mutta muuten piikkiprofiilit ovat identtisiä. Selvästi oli havaittavissa viisi vallitsevaa bakteeria mikrobiyhteisössä.

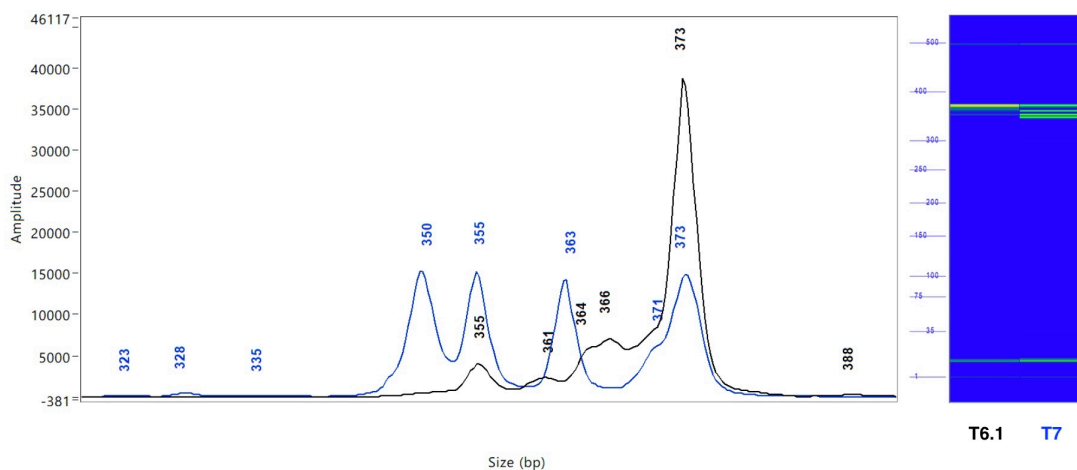
Kuvassa 22 on näytteiden T6.4 ja T6.5 piikkiprofiilit.



Kuva 22. Näytteiden T6.4 ja T6.5 piikkiprofiilit.

Kuvassa 22 voidaan nähdä, että piikkiprofiilit näytteissä ovat samanlaisia. Tulokset ovat jotenkin samanlaiset kuin kuvassa 21, eli toisen näytteen emäsparipituudet ovat systemaattisesti lyhyemmät kuin toisen ja piikit pienemmät (näyte T6.4). Näytteissä on myös selvästi viisi vallitsevaa bakteeriryhmää, mutta ei samat kuin kuvan 21 näytteissä T6.2 ja T6.3. Näytteen T6.4 piikkiä 372 bp vastaava on näytteessä T6.5 piikki 373 bp, ja piikki tunnistettiin *Thermicanukseksi*.

Kuvassa 23 vertaillaan näytettä T6.1 tehtaasta 6, joka sisältää paljon *Thermicanusta* sekvenssitietojen mukaan, näytteeseen T7 tehtaasta 7, jolla myös paljon *Thermicanusta*, mutta myös *Thermusta*.



Kuva 23. Näytteiden T6.1 ja T7 piikkiprofiilit. Piikkiprofiilista oli helppo tunnistaa etenkin näytteen T7 piikkejä vastaavat bakteeriryhmät.

Kummastakin näytteestä tuli selvät piikkiprofiilit (kuva 23). Suurin osa mikrobiyhteisöstä saatiin identifioitua kummastakin näytteestä, kuten oli odotettavissa. Taulukossa 21 on näytteiden T6.1 ja T7 relevantit FA-tulokset ja sekvenssitiedot.

Taulukko 21. Taulukossa näkyy näytteiden T6.1 ja T7 relevantit FA-tulosten emäsparipituudet, niiden osuudet piikkiprofiileista ja sekvenssitiedot. Mitä sinisempi solu, sitä korkeampi osuus bakteerilla on piikkiprofiilissa/mikrobiyhteisössä.

| Taxon | T6.1 | | | T7 | | |
|-------------------|---------------|-----------|---------|---------------|-----------|---------|
| | FA-tulos (bp) | Osuus (%) | Seq (%) | FA-tulos (bp) | Osuus (%) | Seq (%) |
| Thermus | | <0 | 0,1 | 350 | 25,5 | 19,8 |
| Tepidiphilus | 355 | 7,7 | 4,7 | 355 | 22,5 | 11,4 |
| Pseudoxanthomonas | | <0 | 0,7 | 363 | 19,6 | 12,3 |
| Prevotella_7 | 366 | 12,6 | 16,8 | | <0 | 0 |
| Anoxybacillus | | | 7,8 | 371 | 6,6 | 9,5 |
| Thermicanus | 373 | 69,3 | 48,4 | 373 | 24,2 | 29,0 |

Piikki 355 bp tunnistettiin *Tepidiphilus*-bakteerisuvuksi, jota esiintyi kummassakin mikrobiyhteisössä. Piikki 373 bp bakteeriksi *Thermicanukseksi*, koska kyseinen piikki on vallitsevampi verrattuna piikkiin 355 bp. Näytteessä T6.1 piikki 366

bp tunnistettiin *Prevotella_7*-bakteeriksi, koska näyte T6.1 on ainoa, joka sisälsi kyseistä bakteeria merkittävästi, eikä muilla näytteillä ollut kyseistä piikkiä.

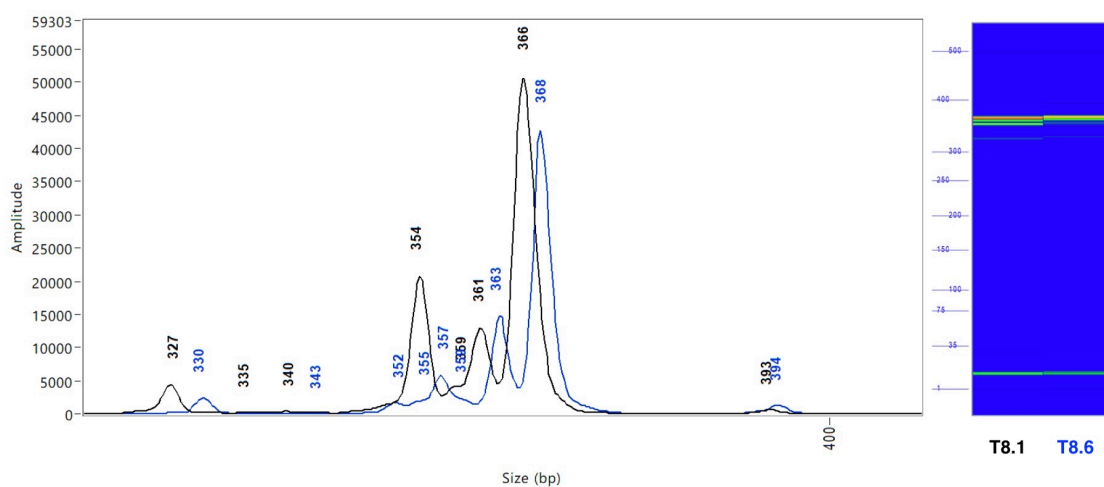
Pseudoxanthomonas on ainoa kandidaatti piikille 363 bp samoin perustein kuin *Prevotella_7* piikille 366 bp. *Thermus* tunnistettiin näytteessä T7 piikiksi 350 bp, joka täsmää tehtaassa 6 muissa tuloksissa näytteissä T6.2–T6.5 (348–350 bp), jolloin *Anoxybacillus* todettiin piikiksi 371 bp näytteessä T7. Mahdollisesti se on eri *Anoxybacillus* kuin tehtaassa 5, koska emäsparipituudet eivät täysin täsmää. *Anoxybacillus*-piikin näytteessä T6.1 uskotaan yhdistyneen piikin 373 bp (*Thermicanus*) kanssa.

Tehtaasta 6 ja 7 tunnistettiin monia bakteereja piikkiprofileista, joten näiden näytteiden suhteen oli tavoitteet saavutettu. *Thermicanus* tunnistettiin piikiksi 373 bp, joka oli myös tunnistettu tehtaasta 5, mutta ei yhtä suurella varmuudella, joten nämä tulokset validoivat niitä tuloksia. Tehtaasta 6 tunnistettiin myös bakteerit *Niabella* (361–364 bp), *Tepidimonas* (365–369 bp) ja *Cloacibacterium* (356–359 bp), jotka olivat läsnä vain toisessa kartonkikoneessa (kuva 22). Tehtaassa 1 oli *Cloacibacteriumilla* samanlainen emäsparipituus ja tehtaassa 5 oli *Tepidimonaksella* myös vastaavanlainen emäsparipituus. Kyseisessä kartonkikoneessa oli *Thermus*-bakteeri myös vähissä, mikä auttoi tunnistamaan bakteerin piikiksi 348–350 bp.

7.2.7 Tehdas 8

Tehtaasta 8 ei ole ennen analysoitu näytteitä säännöllisesti, eikä ollut tietoa mikrobiyhteisöstä. Tiedossa oli kuitenkin Qubit-kvantifiointin tulokset, jossa näkyi, että uudemmat näytteet sisälsivät enemmän mikrobeja. Näytteitä otettiin 5:stä eri näytteenottoa paikasta, olettaen että niissä on mahdollisesti erilaisia mikrobiyhteisöjä, kun näytteenottoa paikat eivät varsinaisesti ole yhteydessä toisiinsa prosessissa. Näytteitä on kahdelta eri ajankohdalta, helmikuulta ja toukokuulta 2022, joten piikkiprofiileja vertaillaan näytteenottoa paikkakohtaisesti. Vanhemmat näytteet ovat näytteet T8.1–T8.5 helmikuulta ja tuoremmat näytteet T8.6–T8.10 toukokuulta. Näyte T8.11 on lopputuotteesta eristetty näyte. Lopussa yhtä tuoretta näytettä verrataan lopputuotteeseen, kartonkiin, mikrobiyhteisön vertailua varten.

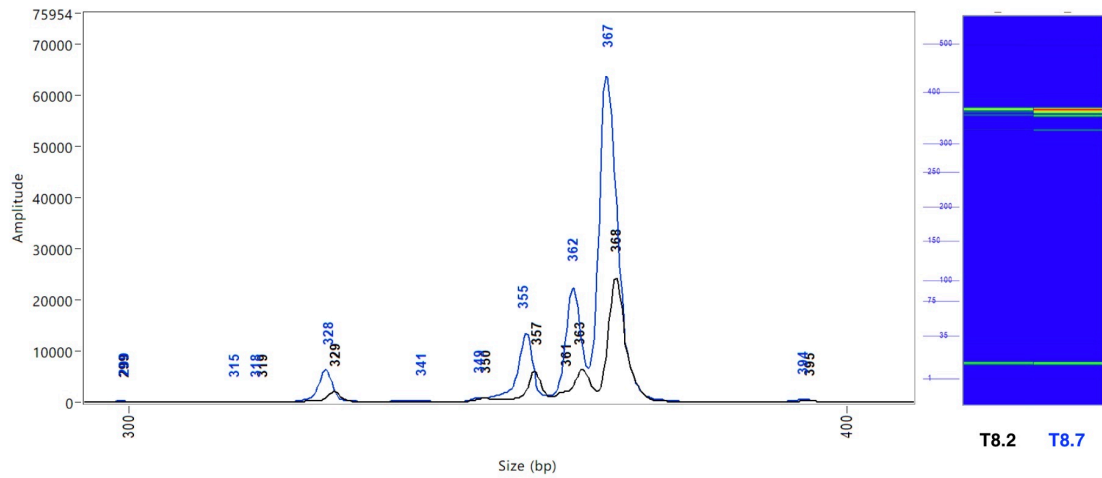
Kuvassa 24 on ensimmäisen näytteenottoa paikan piikkiprofiilit.



Kuva 24. Piikkiprofiilit hylkytornista. Emäsparipituudet vaihtelevat hieman ja pientä muutosta voidaan nähdä mikrobiyhteisössä.

Kuvassa 24 näytteet sisältävät samanlaisia ja samankokoisia piikkejä. Poikkeus on näytteen 36 piikki 354 bp, joka on sama piikki kuin 357 bp näytteestä T8.6, jossa näkyy selvästi olevan korkeampi piikki. Tämä piikki näkyy myös elektrofo-grammissa selvemmin.

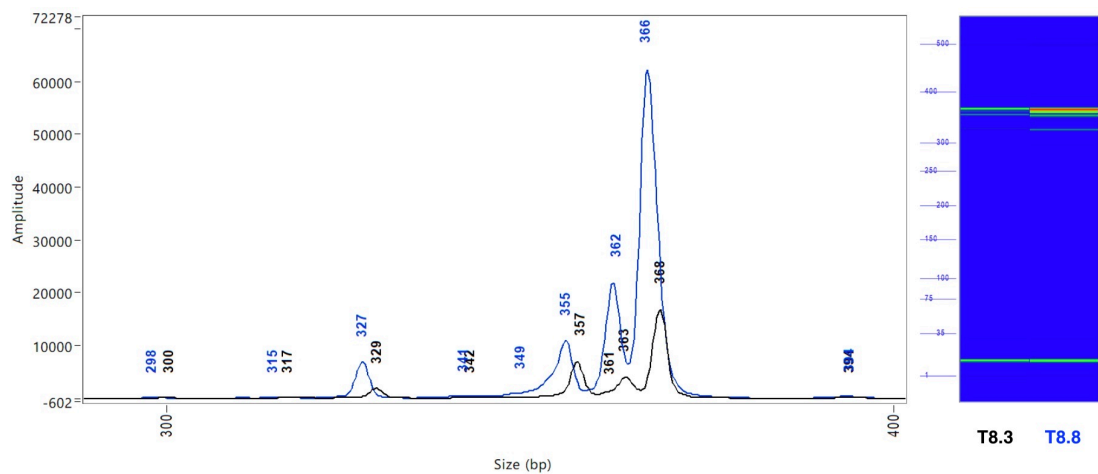
Kuvassa 25 on toisen näytteenottoapaikan piikkiprofiilit.



Kuva 25. Piikkiprofiilit mekaanisen massan annostelusäiliöstä.

Kuvassa 25 mikrobiyhteisö ei ole muuttunut näytteenottojen välillä. Samoja piikkejä kuin ensimmäisessä näytteenottoapaikassa, mutta ei selvää muutosta mikrobiyhteisössä.

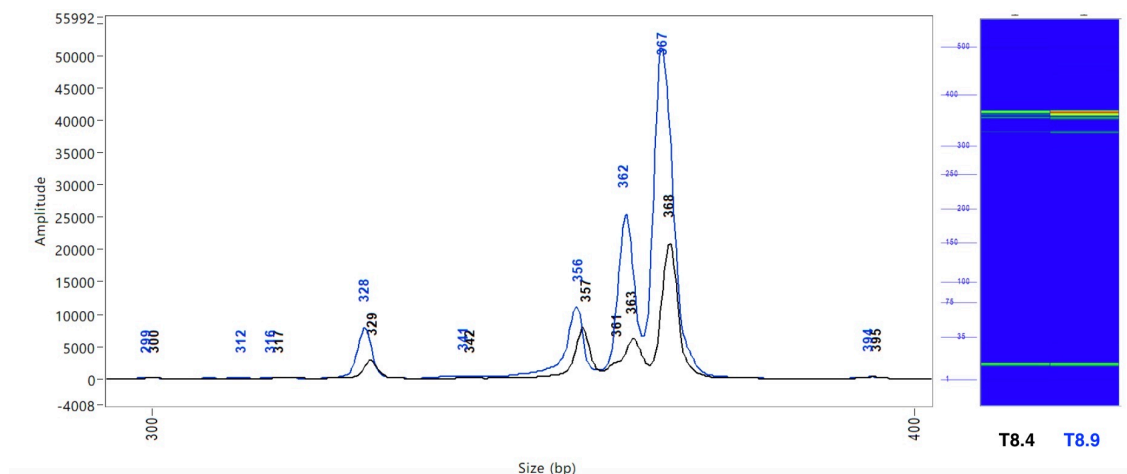
Kuvassa 26, jossa kolmannen näytteenottoapaikan piikkiprofiilit.



Kuva 26. Piikkiprofiilit mäntysellun tornista.

Kuvia 25 ja 26 vertailtaessa ei näy paljoa eroa. Voidaan todeta, että samanlainen mikrobiyhteisö on kyseessä ja samat erot näytteiden välillä näytepaikka-kohtaisesti.

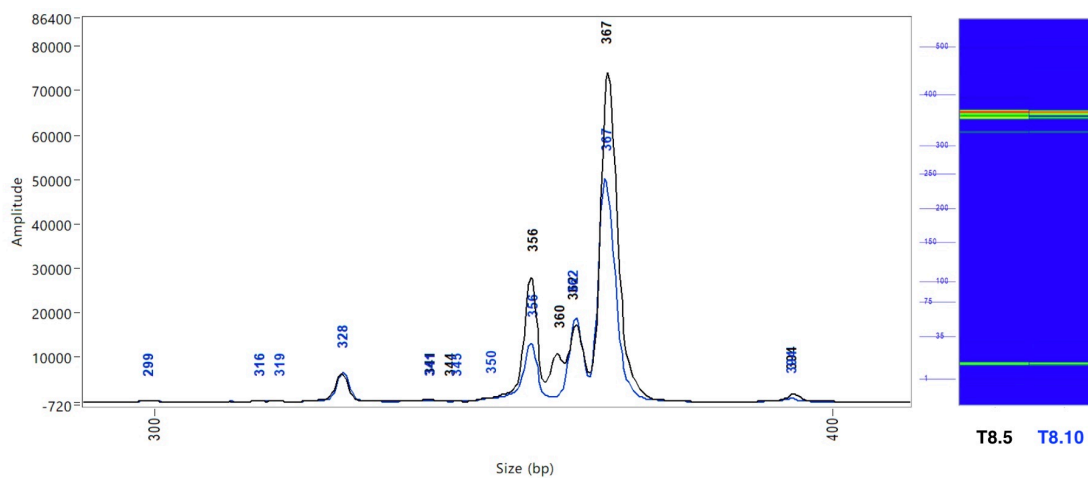
Kuvassa 27 on neljännen näytteenottoaikan piikkiä. Profiilit.



Kuva 27. Piikkiä koivusellun annostelusäiliöstä.

Neljännessä näytteenottoaikassa kuvassa 27 on samat piikit ja samat tulokset niin kuin kaikissa edellisissä näytteenottoaikoissa paitsi ensimmäisessä. Tuoreemmat näytteet tehtaalta eivät erota piikkejä 361 bp ja 363 bp, tai jompikumpi piikkejä edustavista bakteeriryhmistä on osuudeltaan laskenut.

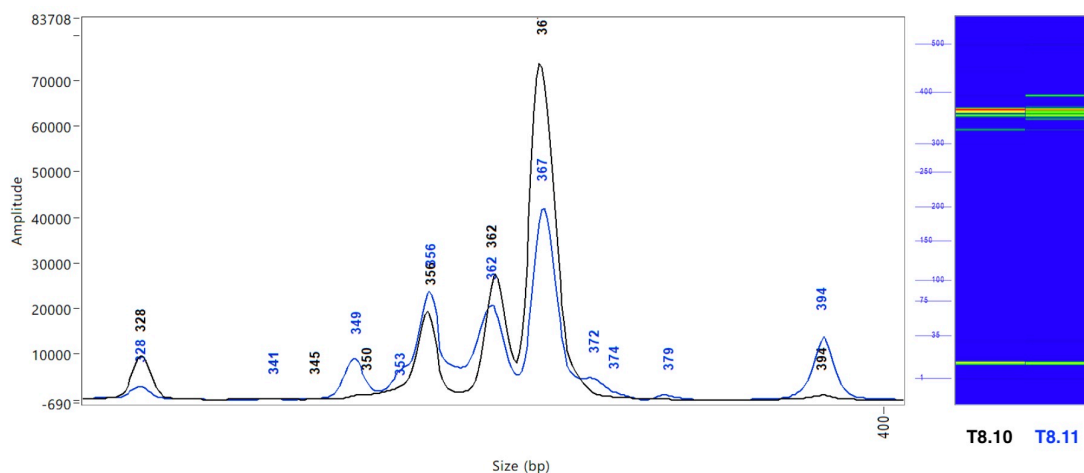
Kuvassa 28 on viidennen näytteenottoaikan piikkiä profiilit.



Kuva 28. Piikkiä profiilit 0-vesitorista.

Kuvassa 28 piikkiä profiilit ovat toistensa kanssa erittäin samanlaiset. Ero muihin näytteenottoaikoihin näkyy näytteen T8.5 piikeissä 360 bp ja 362 bp, jotka olivat yhden emäsparin pidempiä ja epäselvempiä muissa näytteenottoaikoissa. Näytteessä T8.10 piikki 360 bp ei ole tunnistettavissa, eli piikkiä edustava bakteeriryhmän osuus on laskenut. Emäsparipituudet ovat myös yhdenmukaiset.

Kuvassa 29 vertaillaan näytettä T8.10 näytteeseen T8.11. Näyte T8.10 edustaa mikrobiyhteisöjä prosessista ja näyte T8.11 on lopputuote (kartonki), joka on valmistettu samoihin aikoihin kuin prosessista otetut näytteet (toukokuussa). Koska kartonkinäyte valmistettiin samoihin aikoihin kuin prosessista on otettu näytteet, niin oletuksena oli, että piikkiä profiilit sisältäisivät samanlaisia piikkejä.



Kuva 29. Prosessinäytteen T8.10 vertailu kartonkinäytteen T8.11 piikki profiiliin. Prosessinäytteen ja lopputuotteen piikki profiilit sisältävät odotetusti samoja piikkejä.

Kuvassa 29 nähdään, että piikki profiileissa huomattavimmat erot ovat piikeissä 328 bp, 349 bp, 367 bp ja 394 bp. Kartonkinäytteellä T8.11 piikit 328 bp ja 367 bp ovat korkeammat kuin prosessinäytteellä T8.10, jolla on selvästi korkeammat piikit 349 bp ja 394 bp. Kartonkinäytteellä näkyy myös pienempinä pitoisuuksina esiintyviä piikkejä ovat 372 bp ja 379 bp, joita ei esiinny muissa näytteissä. Taulukossa 22 tämä voidaan nähdä tarkemmin.

Taulukko 22. Taulukossa nähdään näytteiden T8.10 ja T8.11 spekuloidut bakteerit, niiden *in silico* emäsparipituudet, FA-tulosten piikit ja niitä vastaavat osuudet piikki profiileista.

| Taxon | <i>In silico</i> (bp) | T8.10 | | T8.11 | |
|--------------------------------|--------------------------|------------------|--------------|------------------|--------------|
| | | FA-tulos (bp) | Osuus (%) | FA-tulos (bp) | Osuus (%) |
| Sphingomonadaceae_unclassified | 317 | 328 | 6,8 | 328 | 2,7 |
| Tuntematon | | 350 | 0,9 | 349 | 6,5 |
| Cloacibacterium | 343 | 356 | 15,4 | 356 | 24,3 |
| Weeksellaceae_unclassified | 343/365 | 362 | 18,2 | 362 | 19,2 |
| Tepidimonas | 344/351/355 | 367 | 57 | 367 | 32,2 |
| Tuntematon | | 394 | 0,8 | 394 | 9,6 |

Näytteistä ei ole sekvenssitietoja. Todetut bakteerit ovat vain spekulatioita perustuen mahdollisiin bakteereihin, jotka täsmäisivät kyseisiin emäsparipituuksiin. Tiedossa oli myös, että mikrobiyhteisö on terve, joten ei epäilty olevan haitta-bakteereita vallitsevassa asemassa. Taulukosta 22 näkyy, kuinka piikin 367 bp osuus laskee ja kuinka piikkien 349–350 bp, 367 bp ja 394 bp piikkien osuus nousee.

Tehtaan 8 piikkiprofililit ovat yksinkertaisia ja helppo seurata. Mikrobiyhteisön seurannassa ei ole ongelmaa ja yhteisöstä voidaan sanoa, että ei ole paljon muuttunut. Hylkytornin näytteet T8.1 ja T8.6 eroavat eniten muiden näytepaikkojen mikrobiyhteisöstä, mutta eivät paljon. Kartonkinäytteessä on hieman eroja prosessinäytteisiin, mutta mikrobiyhteisöt sisältävät samoja bakteereja, kuten oli odotettavissa. Kartonkinäytteen piikkiprofilili ei mitenkään korostetusti muistuta tietyn näytepaikan piikkiprofileja.

7.3 Projektin haasteet

Projektin toteutuksen aikana ilmeni muutamia ongelmia. Yksi niistä on Fragment Analyzer -laitteen herkkyys. Harvemmin ajossa olisi kaikki näytteet onnistuneet, että noin 1–2 näytettä olivat aina epäonnistuneet. Syy oli mahdollisesti kuplien pääsy kapillaareihin, mikä sitten estää näytteiden normaalin kulun kapillaareissa. Ongelman minimoimiseksi ei voi muuta kuin pipetoida ja sekoittaa reagenssit varovaisemmin. Tämä on osasy, miksi on todettu kaikkien tulosten jälkeen, että rinnakkaisnäytteet olisivat tarpeellisia.

Näytteet, joissa oli pitoisuuksiltaan vähän DNA:ta, aiheuttivat ongelmia tulosten tulokinnassa. Vaikka monistus voitiin todeta onnistuneeksi, esimerkiksi tehtaan 3 näytteiden tapauksessa, FA:lta tulleet tulokset eivät olleet luotettavia tai ne olivat muuten epäonnistuneita. Emäsparipituudet näytteissä olivat systemaattisesti liian lyhyitä tai piikit eivät ylittäneet markkereiden korkeutta, jolloin ne eivät olleet luotettavia. Ongelman saattoi aiheuttaa PCR-monistuksessa mahdollisesti alukedimeerien muodostus, joka virheellisesti viittasi näytteen monistumiseen, vaikka se olisi vasta alkamassa. Mahdollisesti myös DNA:n puute

monistuksessa saattoi edesauttaa ongelmaa. Kuitenkin jotkin matalan pitoisuuden DNA-näytteet olivat onnistuneita (esim. tehtaan 5 näytteet T5.2 ja T5.3), joten ongelma saattoi olla tehdaskohtainen.

Alukedimeerien osuus piikkiprofiileista oli miltei aina yli 90 %. Suositeltava osuus olisi korkeintaan 20 % kokonaispinta-alasta [21]. Ongelmana on silloin piikkien osuuksien hahmotus ja tulokset virheellisesti viittaavat, että eroja piikkien osuuksissa ei olisi paljon, vaikka erot pystyttäisiin hahmottamaan piikkiprofiileissa. Ongelman saa periaatteessa korjattua poistamalla alukedimeerien osuudet piikkiprofiileista, jolloin pelkästään relevanttien piikkien pinta-alat huomioidaan. Se ei kuitenkaan poista sitä ongelmaa, että mitä vähemmän haluttua tuotetta on monistettu, sitä huonommin FA-laite pystyy detektoimaan fragmentteja amplikoneista. Luotettavia piikkejä on myös silloin vähemmän (matalat RFU/Amplitude-arvot).

Tässä projektissa kaikki näytteet olivat monistettu samalla PCR-ajo-ohjelmalla. PCR-ohjelma keskeytetään, kun viimeinen näyte on monistunut tarpeeksi ohjeiden mukaisesti. Valitettavasti näytteet eivät monistu samaa tahtia, joten todellisuudessa jotkin näytteet ehtivät monistua enemmän kuin toiset ja siten jotkin näytteet sisältävät enemmän tuotetta kuin toiset. Tämä näkyy piikkiprofiileissa, jossa samalla näytteellä saattaa olla 2 erikokoista piikkiprofiilia. Piikkien koko ei siis peilaa todellisuutta siitä, kuinka suuri pitoisuus sillä on. Menetelmä on mahdollisesti erittäin herkkä kyseiselle ongelmalle, koska itse monistetun geenin osuus on usein vähemmän kuin 10 % PCR-tuotteesta. Pääosin tuote on vain alukedimeerejä, ja ne poistetaan tulosten tarkastelussa. Näiden ongelmien vuoksi eri ajojen tulosten vertailun luotettavuus kärsii.

8 Yhteenveto

LH-PCR-menetelmällä onnistuttiin seuraamaan mikrobiyhteisöjä ja niiden muutoksia. Pelkästään yhden bakteerin osuuden vähentyminen mikrobiyhteisössä pystytään havaitsemaan, ja menetelmällä voidaan tunnistaa luotettavasti vallitsevia bakteerisukuja piikkiprofilien ja sekvenssitietojen avulla. Jos tiedossa on jo sekvenssitiedot vanhemmista näytteistä ja on tunnistettu bakteereita vastaavia piikkejä, rutiininäytteille ei ole tarvetta sekvensoinnille. Menetelmä on tarpeeksi luotettava mikrobiyhteisön muutoksen seurantaan. Muutaman emäsparipituuden vaihtelu ei myöskään estä mikrobiyhteisön seurantaan tai tietyn tutun bakteeriryhmän tunnistamista, koska se on erittäin systemaattista ja se voidaan helposti päätellä. Ellei suurempaa muutosta tapahdu, voidaan FA-laitteelta saatujen piikkiprofilien perusteella päätellä, onko mikrobiyhteisö kunnossa vai vaatiiko asiakkaan tehdä toimenpiteitä, jos mikrobiyhteisö on suuntaamassa ei-haluttuun suuntaan.

Vaikka projekti oli pääosin onnistunut, oli myös ongelmia. Etenkin alukedimeerien suuri osuus kokonaispinta-alasta piikkiprofileissa on ongelma. Se vähentää menetelmän tarkkuutta ja herkkyyttä, etenkin määriteltäessä eri piikkien, eli bakteerien, osuuksia mikrobiyhteisöstä. Olisi myös mahdollista saada harvinaisempien bakteerien piikit esille piikkiprofileissa, jos alukedimeerejä olisi vähemmän PCR-tuotteissa. Menetelmä on kuitenkin onnistuneesti pystytty suorittamaan ongelmasta huolimatta, johtuen Fragment Analyzer -laitteen herkkyydestä ja kyvystä detektoida pienetkin erot ampikonien sisällä. Piikkiprofileista harvoin myöskään näkyy, kuinka suuri osuus alukedimeereillä on kokonaispinta-alasta. Se kuitenkin ei ole suuri ongelma, koska näytteiden fragmenteista syntyy selvät signaalit ja ne luovat selvät korkeat piikit piikkiprofiliiin (usein korkeammat kuin alukedimeerien), vaikka eivät kattaisikaan suurta osaa kokonaispinta-alasta. Tämän vuoksi menetelmän voidaan katsoa soveltuvan rutiininäytteiden mikrobiyhteisön seurantaan. Menetelmällä voidaan myös seurata monimutkaisempiakin mikrobiyhteisöjä, mutta jos vallitsevat bakteerit eivät eroa emäsparipituuksiltaan tarpeeksi, piikkiprofilien tulkitsemisessa saattaa tulla ongelmia. Piikkejä pitää manuaalisesti tunnistaa, joita FA-ohjelma ei itse tunnista.

FA-laitteen herkkyys saattaa aiheuttaa myös ongelmia. Näyteajossa saattaa epäonnistua 1–2 näytettä, johtuen kuplien pääsemisestä kapillaareihin tai niissä on muuta tukosta, jonka vuoksi näytteet/markkerit eivät pääse kulkemaan normaalisti eteenpäin kapillaareissa. Tämä saattaa lisätä kuluja ja pidentää aikaa tulosten saamisessa. Kyseisen riskin vähentämiseksi rinnakkaisnäytteiden ajaminen voi auttaa, jolloin toisen rinnakkaisnäytteen onnistuminen riittäisi. Rinnakkaisnäytteiden teko on myös suositeltavaa, koska riippuen seuraavien ajojen näytteiden DNA-pitoisuuksista edellisessä ajossa epäonnistunut näyte saataan laimentaa eri pitoisuuteen PCR-ajossa. Eri pitoisuudessa oleva näyte monistuu eri vauhtia, mikä vaikuttaa siten myös bakteerien välisiin suhteisiin PCR-tuotteissa. Tämä ei ole ideaalista, jos halutaan vertailla tiettyjen näytteiden piikkiprofiileja toisiinsa (esim. eri näytteenottoaika, sama näytteenottoaika). Tässäkin auttaisi, jos alukedimeerien osuus olisi pienempi. Silloin vaihtelevuus ei olisi yhtä suurta vertailtaessa eri ajojen tuloksia, joissa näytteet ovat todennäköisesti olleet eri määrän syklejä monistumassa PCR:ssä.

Piikkiprofiilien ja niistä saatujen arvojen avulla voidaan määrittellä vallitsevat bakteerit mikrobiyhteisössä, mutta menetelmä olisi vielä hyödyllisempi, jos pystyisi identifioimaan esimerkiksi haitalliset bakteerit niiden piikeistä sekä luotettavasti sanomaan, kuinka suuri osuus sillä on mikrobiyhteisössä. Menetelmällä on välillä onnistuttu määrittämään oikeat osuudet bakteereille (samat tulokset siis kuin sekvenssitiedoissa). Siihen vaikuttaa kuitenkin, minkälainen mikrobiyhteisö kyseessä ja kuinka monta vallitsevaa bakteeria on. Jos esimerkiksi vallitsevia bakteereita on vain 3, on mahdollista, että vallitsevin bakteeri monistuu paremmin kuin 2 muuta bakteeria, jolloin suhteet vääristyvät. Tämä nähtiin tehtaalla 1 *Cloacibacterium*, 356–356 bp kohdalla ja tehtaalla 2 *Weeksellaceae_unclassified*, 361–363 bp kohdalla. Voi olla, että tällä menetelmällä se ei ole mahdollista ja tulokset ovat vain suuntaa antavia, ellei ole keinoa saada suurinta osaa mikrobiyhteisöstä monistumaan tasavertaisesti, jotta bakteerien suhteet näytteissä pysyisivät samana.

In silico ennakoitujen emäsparipituudet olivat vain välillä avuksi. Bakteerien suvun sisällä on paljon vaihtelua, joten tulevaisuudessa muissa projekteissa *in silico*

ennakointi bakteereille ei välttämättä ole tarpeellista. Pääosin amplikonien pituudet olivat pitempiä todellisuudessa, kuin simuloituina *in silico* Sequence Extractorissa.

Taulukossa 23 on kerätty kaikki tunnistetut bakteerit, niitä vastaavat emäsparipituudet ja *in silico* ennakoidut emäsparipituudet vastaaviin bakteereihin.

Taulukko 23. Bakteerit ja niitä täsmäävät *in silico* ennakoidut bp pituudet sekä FA-tulosten bp pituudet. FA-tulos kohtaisesti ilmoitettu myös, että mistä tehtästä tulokset ovat saatu.

| taxon | <i>In silico</i> (bp) | FA-tulokset (bp) | | |
|--------------------------------|-----------------------|------------------|-------------|-------------|
| Cloacibacterium | 343 | T1: 356-357 | T6: 356-359 | |
| Dysgonomonas | 356 | T1: 365-367 | | |
| Prevotella_7 | 354 | T2: 370-371 | T6: 366 | |
| Prevotellaceae_unclassified | 354 | T2: 370-371 | | |
| Weeksellaceae_unclassified | 343 or 365 | T2: 361-363 | | |
| Niabella | 348 | T6: 361-364 | | |
| Anoxybacillus | 357 | T5: 376-380 | T7: 371 | T6: 369-371 |
| Ethanoligenens | 348 | T2: 356-357 | | |
| Thermicanus | 358 or 368 | T5: 371-375 | T6: 372-375 | T7: 373 |
| Enterococcus | 365 | T2: 380 | | |
| Clostridium_sensu_stricto_N | 333 or 336 or 354 | T5: 363-367 | | |
| Bacillus | 327 & 364 | T5: 383-384 | | |
| Bacillales_unclassified | 357 | T5: 380-383 | | |
| Pseudoxanthomonas | 349 or 351 | T7: 363 | | |
| Tepidimonas | 344 or 351 or 355 | T5: 366-369 | T6: 365-369 | |
| Tepidiphilus | 313 & 344 & 362 | T6: 353-356 | T7: 355 | |
| Sphingomonadaceae_unclassified | 317 | T1: 326-328 | | |
| Pseudomonadaceae_unclassified | 345 | T5: 373-375 | | |
| Meiothermus | 332 & 335 | T4: 348 | | |
| Thermus | 341 or 345 | T7: 350 | | |

Kuten taulukosta 23 voidaan nähdä, *in silico* ennakoidut emäsparipituudet ovat lähes aina lyhyempiä kuin FA-laitteelta saatujen tuloksien emäsparipituudet. Suoranaista syytä on vaikea sanoa. Ehkä asiaan voi vaikuttaa, minkälaisia

geelimatriiseja on käytössä Fragment Analyzerilla tai ehkä sähkövirta on liian suuri, jolloin amplikonit etenevät kapillaareissa liian nopeasti. Tällöin systemaattisesti kaikki tulokset olisivat suurempia kuin teoreettisesti pitäisi olla.

Suosittelvat askeleet ovat seuraavaksi tutkia, onko olemassa olisi paremmin sopivia alukkeita, qPCR-ohjelman mahdollista optimointia tai erilaisen PCR-reaktioseoksen tekoa, jossa pärjäisi vähemmällä alukkeiden määrällä. Päämääränä on nostaa kohdegeenin osuutta PCR-tuotteissa. Tiedossa ei myöskään ole, että pystyykö Fragment Analyzerin ajo-ohjelmaa optimoida, mutta ideaa voisi tutkia, jotta ehkä löytyisi konkreettisempi syy, miksi PCR-tuotteet ovat pitempiä kuin *in silico* ennakoidut amplikonit. Jos menetelmää halutaan laajempaan käyttöön, suositellaan myös tutkimaan, vaikuttavatko näytteenottomenetelmät tuloksiin ja mikä voisi mahdollisesti vaikuttaa siihen, miksi välillä näytteet ovat merkittävästi emäsparipituuksiltaan lyhyempiä kuin niiden pitäisi olla.

Lähteet

- 1 Operational Taxonomic Unit (OTU) and OTU Clustering. Verkkoaineisto. CD Genomics. <<https://www.cd-genomics.com/microbioseq/operational-taxonomic-unit-otu-and-otu-clustering.html>>. Luettu 28.3.2022.
- 2 Company. Verkkoaineisto. Kemira Oyj. <<https://www.kemira.com/company/>>. Luettu 8.7.2022.
- 3 Flemming, Hans-Curt; Meier, Michael & Schild Tobias. 2013. Mini-review: microbial problems in paper production, *Biofouling*, 29:6, s. 683–696. Verkkoaineisto. <<https://www.tandfonline.com/doi/abs/10.1080/08927014.2013.798865>>. Luettu 13.7.2022.
- 4 Öqvist, Charlotta. 2008. Microbial life and deposits in paper machine circuits, s. 9–10 & 22–24. Helsinki, Yliopistopaino. Verkkoaineisto. <<https://helda.helsinki.fi/bitstream/handle/10138/20885/microbia.pdf;sequence=1>>. Luettu 14.7.2022.
- 5 What is Real-Time PCR (qPCR)?. Verkkoaineisto. Bio-Rad Laboratories, Inc. <https://www.bio-rad.com/en-fi/applications-technologies/what-real-time-pcr-qpcr?ID=LUSO4W8UU#What_Is_Real-Time_PCR>. Luettu 22.7.2022.
- 6 Kadner, Robert & Rogers, Kara. 2022. Bacteria. Verkkoaineisto. Encyclopedia Britannica. <<https://www.britannica.com/science/bacteria/DNA-based-methods>> 14.7.2022. Luettu 28.8.2022.
- 7 cdadmin. 2018. 16S rRNA, One of the Most Important rRNAs. Verkkoaineisto. CD-Genomics. <<https://www.cd-genomics.com/blog/16s-rrna-one-of-the-most-important-rrnas/>>. 17.10.2018. Luettu 15.9.2022.
- 8 Chakravorty, Soumitesh; Helb, Danica; Burday, Michele; Connel, Nancy & Alland David. 2007. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of Microbiological Methods*, Vol. 69, issue 2, s. 330–339.
- 9 Shai, Shailesh; Freedman, Samantha & Mangalam, Ashutosh. 2017. Gut Microbiome in Multiple Sclerosis: The Players Involved and the Roles They Play. *Gut Microbiomes*, Vol. 8, issue 6, s. 607–615. Verkkoaineisto. Taylor & Francis Online. <<https://www.tandfonline.com/doi/full/10.1080/19490976.2017.1349041>>. Luettu 15.9.2022.

- 10 Klindworth, Anna; Priesse, Elmar; Schweer, Timmy; Peplies, Jörg; Quast, Christian; Horn, Matthias & Glöckner, Frank. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*, Vol. 41, issue 1, s. e1.
- 11 Abellan-Schneyder, Isabel; Matchado, Monica; Reitmeier, Sandra; Sommer, Alina; Sewald, Zeno; Baumbach, Jan; List, Markus & Neuhaus, Klaus. 2021. Primer, Pipelines, Parameters: Issues in 16S rRNA Gene Sequencing, Vol. 6, issue 1. *Verkkoaineisto. mSphere*. <<https://journals.asm.org/doi/10.1128/mSphere.01202-20>>. 24.2.2021. Luettu 29.8.2022.
- 12 Grahn, Niclas; Olofsson, Margaretha; Ellnebo-Svedlund, Katarina; Monstein, Hans-Jürg & Jonasson, Jon. 2003. Identification of mixed bacterial DNA contamination in broad-range PCR amplification of 16S rDNA V1 and V3 variable regions by pyrosequencing of cloned amplicons. *FEMS Microbiology Letters*, Vol. 219, issue 1, s. 87–91. *Verkkoaineisto. Oxford University Press*. <<https://academic.oup.com/femsle/article/219/1/87/505487?login=false>>. Luettu 29.8.2022.
- 13 Bacterial community fingerprinting. *Verkkoaineisto. BioMérieux SA*. <<https://www.applied-maths.com/applications/bacterial-community-fingerprinting>>. Luettu 15.7.2022.
- 14 Sardaro, Maria; Perin, Luana; Bancalari, Elena; Neviani, Erasmo & Gatti, Monica. 2018. Advancement in LH-PCR methodology for multiple microbial species detections in fermented foods. *Food Microbiology*, Vol. 74, s. 113–119. *Verkkoaineisto. ScienceDirect*. <<https://www.sciencedirect.com/science/article/pii/S074000201730984X?via%3Dihub>>. Luettu 19.7.2022.
- 15 Mills, DeEtta; Entry, James; Gillevet, Patrick & Mathee, Kalai. 2007. Assessing Microbial Community Diversity Using Amplicon Length Heterogeneity Polymerase Chain Reaction. *Soil Science Society of America Journal*, Vol. 71, issue 2, s. 572–578. *Verkkoaineisto. American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America*. <<https://access.onlinelibrary.wiley.com/doi/epdf/10.2136/sssaj2006.0147>>. 1.3.2007. Luettu 19.7.2022.
- 16 Fragment Analyzer Software. *Verkkoaineisto. Agilent*. <<https://www.agilent.com/en/product/automated-electrophoresis/fragment-analyzer-systems/fragment-analyzer-systems-software/fragment-analyzer-software-1149185#specifications>>. Luettu 30.3.2022.
- 17 5200 Fragment Analyzer System. *Verkkoaineisto. Agilent*. <<https://www.agilent.com/en/product/automated-electrophoresis/fragment->

analyzer-systems/fragment-analyzer-systems/5200-fragment-analyzer-system-365720>. Luettu 28.3.2022.

- 18 Nucleic Acid Analysis for Sample Quality Assessment Using the Agilent Fragment Analyzer Systems. Verkkoaineisto. Agilent. <<https://www.agilent.com/cs/library/applications/application-nucleic-acid-qc-fragment-analyzer-5994-2813en-agilent.pdf>>. Luettu 4.9.2022.
- 19 About: Relative fluorescence units. Verkkoaineisto. DBpedia. <https://dbpedia.org/page/Relative_fluorescence_units>. Luettu 13.9.2022.
- 20 Best Sizing Practices with The Agilent Fragment Analyzer Systems. Verkkoaineisto. Agilent. <<https://www.agilent.com/cs/library/applications/application-nucleic-acid-fragment-sizing-fragment-analyzer-5994-0585en-agilent.pdf>>. Luettu 13.9.2022.
- 21 Fragment Analyzer interpretation. Verkkoaineisto. Cornell Institute of Biotechnology. <<https://www.biotech.cornell.edu/core-facilities-brc/facilities/genomics-facility/qc-troubleshooting/fragment-analyzer>>. Luettu 30.3.2022.