

Typing av *Legionella pneumophila* med MALDI-TOF MS

Anders Halgunset

Bioteknologi (5 årig)

Hovedveileder: Per Bruheim, IBT

Norges teknisk-naturvitenskapelige universitet
Institutt for bioteknologi

Forord

Denne masteroppgaven ble utført som avsluttende oppgave på Bioteknologi – studiet ved institutt for Bioteknologi ved Norges teknisk-naturvitenskaplige universitet (NTNU). Studiet ble utført i samarbeid med Forsvarets forskningsinstitutt (FFI) og alt praktisk laboratoriearbeid er utført ved FFI på Kjeller, under veiledning av Jaran Strand Olsen.

Først ønsker jeg å takke Jaran for god veiledning og for deling av kunnskap gjennom hele perioden på FFI. Jeg vil også takk Ingjerd Thrane og Tone Aarskaug for all hjelp på laboratoriet, Marius Dybwad for gode råd og tips, Randi Skjermo Telstad for gjennomlesing og Professor Per Bruheim som var intern veileder ved NTNU. Også takk til alle de andre på Beskyttelse ved FFI, som har bidratt til et trivelig opphold og godt miljø.

Til slutt vil jeg takke familie og venner for god støtte gjennom hele studieperioden.

Forsvarets forskningsinstitutt

Kjeller, 15. mai 2012

Anders Halgunset

Sammendrag

Legionella pneumophila er fakultativ intracellulære bakterier som lever naturlig i akvatiske habitater. De kan replikere i makrofager og protozoer, bli overført til mennesker gjennom aerosoler og føre til smitte. Ved eventuelle utbrudd er det viktig å kunne identifisere kilden, noe som i dag bli gjort ved å typebestemme *L. pneumophila* med sekvensbasert typing (SBT). Samtidig har matrix-assisted laser desorption/ionization – time-of-flight (MALDI–TOF) mass spectrometry (MS) vist seg å kunne brukes til rask og effektiv identifisering av en rekke mikroorganismer på artsnivå, deriblant *Legionella*. Hos enkelte andre mikroorganismer har MALDI–TOF MS også vist å kunne skille mellom underarter. MALDI–TOF MS ble i denne oppgaven benyttet for å undersøke om det var mulig å skille *L. pneumophila* fra hverandre på stammenivå ved hjelp av MALDI Biotyper 3.0 programvare.

Standardprotokollen anbefalt av Bruker Daltonics for proteinekstraksjon fra mikroorganismer ble tilpasset/optimalisert for bruk på *L. pneumophila*, slik at reproducerbare massespektre ble oppnådd. Evalueringen viste at proteinekstraksjon fra ca 2 µl biologisk materiale med 10 µl acetonitrill og 10 µl maursyre, ga best scoreverdi opp mot Bruker Daltonics referansebibliotek. Den optimale temperaturen og varigheten for dyrkning ble vist å være henholdsvis 37 °C og 48 ± 2 timer. Det ble vist at lagring av skåler/kolonier med *L. pneumophila* ved 37 °C førte til forandringer i massespektrene, mens lagring ved 20 °C ikke medførte slike forandringer og at stabile massespektre ble oppnådd etter seks dagers lagring.

Det ble bygget opp et referansebibliotek i MALDI Biotyper 3.0 bestående av referansespekter/referansetopplister (MSP) fra til sammen 48 *Legionella* spp. hvorav 41 var *L. pneumophila* -isolater. Med referansebiblioteket ble det vist at MSP (dannet ved standard innstillinger) for mange *L. pneumophila* var for like hverandre til å gi identifisering mot egne MSP.

En klyngeanalyse ble brukt i sammenligning av *L. pneumophila* mellom SBT og MSP fra referansebiblioteket dannet med MALDI–TOF MS. Sammenligningen viste at diversiteten i proteinsammensetningen hos *L. pneumophila*, representert som MSP, ikke ga samme oppløselighet som gjennom sekvensvariasjonene fra genene i SBT - analysen.

Summary

Legionella pneumophila is facultative intracellular bacteria that live naturally in aquatic habitats. They can replicate in macrophages and protozoa, may be transmitted to humans through aerosols and cause infection in humans. In the event of an outbreak, it is important to identify the source, which is currently being done by determining the type of *L. pneumophila* with sequence based typing (SBT). Meanwhile, matrix-assisted laser desorption/ionization – time-of-flight (MALDI-TOF) mass spectrometry (MS) has been shown to be useful for rapid and efficient identification of a number of microorganisms at the species level, including *Legionella*. In some other micro-organisms MALDI-TOF MS also been shown to distinguish between subspecies. MALDI-TOF MS was used in this paper to investigate whether it was possible to distinguish *L. pneumophila* from each other on the strain level using MALDI Biotyper 3.0 software.

The standard protocol recommended by Bruker Daltonics for protein extraction from micro-organisms was adapted/optimized for use on *L. pneumophila*, allowing reproducible mass spectra to be created. The evaluation showed that protein extraction from about 2 µl biological material with 10 µl formic acid and 10 µl acetonitrile, gave the best score value against Bruker Daltonics reference library. The optimum temperature and duration of cultivation was shown to be respectively 37 °C and 48 ± 2 hours. It was shown that storage of colonies of *L. pneumophila* at 37 °C led to changes in the mass spectra, while storage at 20 °C did not lead to such changes and that stable mass spectra was reached after six days of storage.

It was built up a reference library in the MALDI Biotyper 3.0 consisting of the reference spectra/reference peak list (MSP) from a total of 48 *Legionella* spp. 41 of which were *L. pneumophila* isolates. With the reference library, it was shown that the MSP (formed by default) for many *L. pneumophila* was too similar to each other to provide identification with their own MSP.

A cluster analysis was used in the comparison of *L. pneumophila* between SBT and MSP from the reference library created with MALDI-TOF MS. The comparison showed that the diversity in protein composition in *L. pneumophila*, represented as MSP, did not give the same resolution as the sequence variations of genes in the SBT analysis.

Innholdsfortegnelse

Forord.....	I
Sammendrag	II
Summary	III
Forkortelser	VII
1. Introduksjon	1
1.1 <i>Legionella</i>	1
1.1.1 <i>Legionella pneumophila</i>	2
1.2 Motiv for oppgaven.....	2
1.3 MALDI-TOF MS	4
1.3.1 Grunnprinsippene ved analyse med MALDI-TOF MS til identifisering av mikroorganismer	4
1.3.2 Massespekter generert med et massespektrometer	5
1.3.3 Identifisering og klassifisering av mikroorganismer med MALDI-TOF MS	6
1.3.4 Preparering av analytt til MALDI-TOF MS identifisering	8
1.3.5 Analyser med MALDI Biotyper 3.0	9
1.4 Sekvensbasert typing av <i>L. pneumophila</i>	10
2. Materialer og Metoder	13
2.1 Dyrkingsforhold	13
2.2 Bakteriestammer	13
2.3 Prøvepreparering til MALDI-TOF -analyse	16
2.3.1 Protokoll for Proteinekstraksjon	16
2.3.2 Protokoll for Direkte – metode	17
2.4 Målinger med MALDI-TOF MS.....	18
2.5 Innledende forsøk.....	19
2.5.1 Evaluering av optimal mengde biologisk materiale.....	19

2.5.2	Evaluering av optimalt volum maursyre og acetonitril.....	20
2.5.3	Stabilitet av proteinekstrakter og kolonier ved lagring.....	20
2.6	Generering av massespekter med MALDI-TOF MS til MSP - dannelse	20
2.6.1	Analysering av massespekter og generering av MSP	21
2.6.2	Blindtest – identifisering av <i>L. pneumophila</i> med MSP i referansebibliotek	21
2.6.3	Fylogenetisk fremstilling av MSP og massespekter	22
2.7	SBT	23
2.7.1	DNA isolering	23
2.7.2	SBT gen og primere	24
2.7.3	PCR amplifisering.....	24
2.7.4	Verifisering av PCR produkt	25
2.7.5	Sekvensering og analysering av sekvenseringsdata.....	25
3.	Resultat	28
3.1	Innledende forsøk.....	28
3.1.1	Evaluering av optimal mengde biologisk materiale.....	28
3.1.2	Evaluering av optimalt volum maursyre og acetonitril.....	30
3.1.3	Stabilitet av massespekter ved lagring av proteinekstrakter og kolonier....	31
3.2	Typing av <i>L. pneumophila</i>	35
3.2.1	Blindtest -identifisering	35
3.2.2	Fylogenetisk sammenligning av SBT og MALDI–TOF MS.....	37
4.	Diskusjon	44
4.1	Innledende forsøk.....	44
4.2	Evaluering av referansebibliotek	46
4.3	Sammenligning og vurdering av typingsmetodene.....	48
4.4	Veien videre	50
5.	Konklusjon.....	52

Referanser	53
Vedlegg A: <i>L. pneumophila</i> oppføringer i databasen til Bruker Daltonics	A
Vedlegg B: Referansesekvenser til SBT – gen	B
Vedlegg C: Trestruktur av <i>Legionella</i> spp. MSP i referansedatabasen	C
Vedlegg D: MST -tre av SBT analyse	D

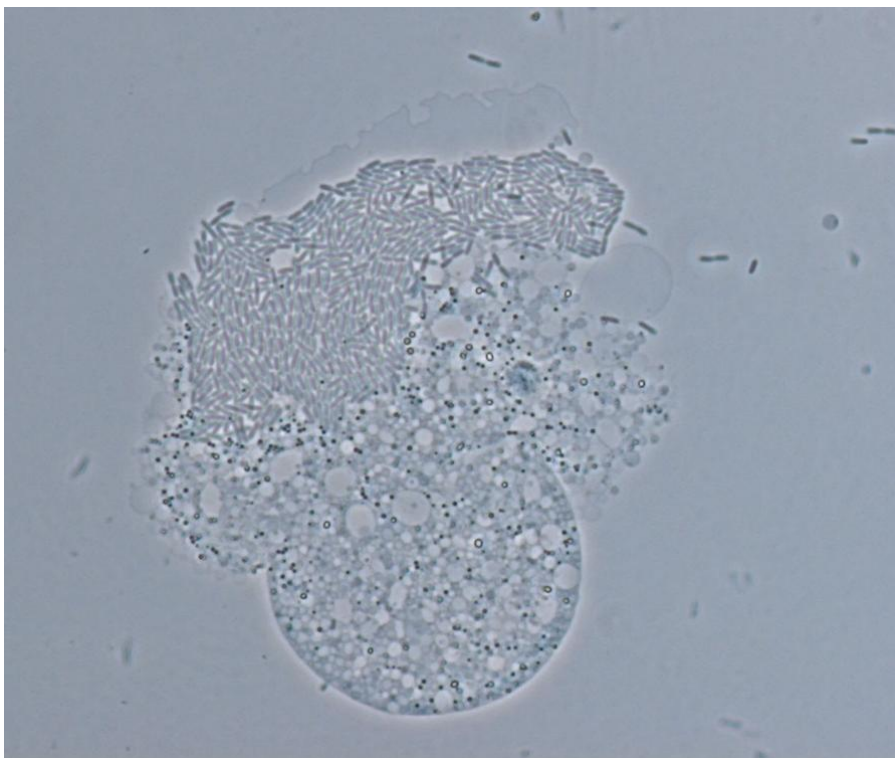
Forkortelser

Forkortelse	Forklaring
a. u.	Arbitrary units
B	Bakoverrettede (om primere)
bp	Basepar
BTS	Bakteriell teststandard
CCI	Composite Correlation Index
Da	Dalton
DNA	Deoksyribonukleinsyre
EWGLI	European Working Group for <i>Legionella</i> Infections
F	Foroverrettede (om primere)
FFI	Forsvaret forskningsinstitutt
HCCA	α -cyano-4-hydroxycinnamic acid
MALDI-TOF	Matrix assisted laser desorption/ionization – time of flight
MLST	Multilocus Sequencing Typing
MS	Mass Spectrometry
MSP	Main Spectrum
MST	Minimum Spanning Tree
m/z	Masse til ladning forhold
NJ	Neighbor Joining
NTNU	Norges teknisk-naturvitenskaplige universitet
PCA	Principal Component Analysis
PCR	Polymerase kjedereaksjon
PFGE	Pulsfelt gelelektroforese
ppm	Deler per million
SBT	Sekvensbasert typing
sg	Serogrupper
spp	Species
ST	Sekvenstype
TFA	Trifluoroacetic acid
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
z	Ladning

1. Introduksjon

1.1 *Legionella*

Legionella er aerobe, Gram-negative, stavlignende, gammaproteobakterier. De er kjemoorganotrofe og bruker aminosyrer som karbon- og energikilde og krever L-cystein og jernsalter for å kunne vokse optimalt. I kultur er den optimale veksttemperaturen $36\text{ °C} \pm 1\text{ °C}$. *Legionella* lever naturlig i akvatiske habitater og er allestedsnærværende i miljøer med temperaturer mellom 5 og 50 °C. Mer enn femti arter er blitt funnet. De er fakultative intracellulære parasitter som infiserer makrofager og protozoer og de assosieres naturlig med frittlevende amøber (figur 1.1), som kan fungere som overføringsvektorer til mennesker. (Garrity et al. 2005).



Figur 1.1: Lysering av amøben er *Acanthamoeba castellanii* (ATCC 30234) infisert med *Legionella pneumophila*. Foto: Tone Aarskaug ved FFI.

1.1.1 *Legionella pneumophila*

Den mest vanlige patogene arten for mennesker er *L. pneumophila* og derfor også den mest studerte (Garrity et al. 2005). Sykdom forårsaket av *Legionella* kalles gjerne legionellose, hvor de to mest utbredte sykdommene er legionærsyken og pontiacfeber. *L. pneumophila* har etter et utbrudd i 1976 vært kjent for å forårsake legionærsyken hos mennesker. Navnet *pneumophila*, som kan oversettes til lungeelskende, kommer av infeksjonen som ofte kan ha utspring i de alveolære makrofagene og monocyttene (Degré et al. 2008). *L. pneumophila* er vist å ha 17 serogrupper (sg), hvor majoriteten av sykdom hos mennesker er rapportert å stamme fra sg1 (Sobral et al. 2011). Bakterien smitter ikke mellom mennesker (Degré et al. 2008).

I de fleste tilfeller av *Legionella* – utbrudd skjer overføringen til mennesker gjennom inhalering av aerosoler som inneholder bakteriene. Aerosolene har blitt rapportert å stamme fra ulike aerosolproduserende innretninger som dusj, kjøletårn, boblebad, fontener (Gomez-Valero et al. 2009), hvor spredningen utendørs kan bli bestemt av meteorologiske påvirkninger (Blatny et al. 2008).

In vitro studier i kultur har vist at livssyklusen til *L. pneumophila* består av minst to adskilte faser. En overføringsfase; utenfor en vertscelle, hvor bakteriene er flaggelerte, bevegelige og virulente, og en reproducerende fase; hvor bakteriene er funnet i reproducerende vakuoler, hvor de er natriumresistente og uflaggelerte (Molofsky & Swanson 2004). Overgangen fra en reproducerende fase til overføringsfasen skjer som følge av aminosyre- og/eller fettsyremangel (Gomez-Valero et al. 2009).

1.2 Motiv for oppgaven

Forsvarets forskningsinstitutt (FFI) er et nasjonalt kompetansesenter innenfor trusselforståelse, sårbarhetsanalyser og beskyttelsestiltak. Ved instituttet arbeides det kontinuerlig med å forbedre eller se etter nye metoder for påvisning og identifisering av biologiske trusselstoffer.

I 2003 og 2008 ble det rapportert om tre utbrudd av legionærsyken i forbindelse med et biologisk behandlingsanlegg ved Borregaard Ind. Ltd., hvor FFI bidro til identifiseringen av kilden til utbruddet (Olsen et al. 2010).

For å kunne identifisere kilden til utbrudd av biologiske trusselstoffer er det viktig å kunne skille mellom stammer innenfor samme art av mikroorganismer (typingsmetoder). En typingsmetode bør ideelt sett være rask å utføre, gi et robust resultat og danne objektive data. En rekke ulike typingsmetoder blir benyttet og deles gjerne inn etter om de undersøker fenotypiske trekk eller genotypiske trekk. Fenotypiske typingsmetoder som for eksempel serotyping eller biotyping har fordelen av at de er relativt billige å utføre og resultatene kan sammenlignes og kategoriseres i et bibliotek. Bakdelene med dem er hovedsakelig hva som undersøkes, som kan påvirkes ved regulering, variabelt uttrykk og gir ikke nødvendigvis pålitelige identifiseringer av kilder til f. eks et utbrudd. Genotypiske typingsmetoder er til motsetning mer pålitelige og baserer seg på sammenligning av f. eks DNA -sekvenser eller DNA -mønstre (Hallin et al. 2012). Dagens internasjonale standard for typing av *L. pneumophila* er en sekvensbasert typingsmetode (SBT) utarbeidet av ”The European Working Group for *Legionella* Infections” (EWGLI) (The Health Protection Agency 2012). SBT er en metode basert på polymerase kjedereaksjon (PCR) som krever sekvensering av PCR – produktene, noe som gjør metoden tidkrevende og kostbar (Hallin et al. 2012).

Matrix-assisted laser desorption/ionization – time-of-flight (MALDI–TOF) mass spectrometry (MS) har de seneste årene vist seg å være en billig og effektiv metode for rask identifisering av mikroorganismer. MALDI–TOF MS har derfor blitt et viktig verktøy i mange kliniske mikrobiologiske laboratorier (Bizzini & Greub 2010). Det har tidligere blitt vist at det er mulig å gjøre pålitelige identifiseringer av *Legionella* på artsnivå med MALDI–TOF MS (Moliner et al. 2010). Oppløselighetene til MALDI–TOF MS har for *L. pneumophila* vist seg å være sammenlignbar med pulsfelt gelelektroforese (PFGE) (Fujinami et al. 2011).

Formålet med denne masteroppgaven var å danne et MALDI-TOF MS referansebibliotek på bakgrunn massespekter av *L. pneumophila* – isolater, ved hjelp av *MALDI Biotyper 3.0* (Bruker Daltonics, Bremen, Tyskland). Med referansebiblioteket skulle det bli undersøkt om det er mulig å skille mellom ulike stammer av *L. pneumophila*. Oppløselighetene av typingen ble sammenlignet med SBT gjennom en klyngeanalyse basert på en fylogenetisk fremstilling av dataene.

I tillegg ble det gjort forsøk for å tilpasse den generelle proteinekstraksjonsprotokollen som produsenten av Microflex MALDI-TOF MS instrumentet (Bruker Daltonics, Bremen, Tyskland) har foreslått, slik at protokollen ble optimalisert spesifikt for *L. pneumophila* -isolater.

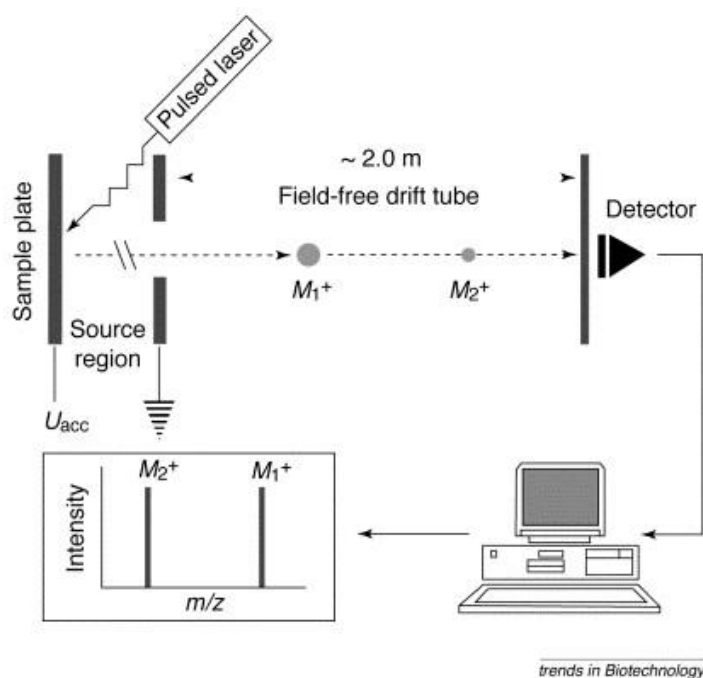
1.3 MALDI-TOF MS

Massespektrometeret har lenge vært et viktig instrument innen kjemien, men det var lenge et problem å måle massen til makromolekyler som proteiner og nukleinsyrer. Grunnen var at ”masse til ladning” -forholdet (m/z) må bli målt av molekyler i gassfase. For å få makromolekylene i gassfase krevde det tidligere gjerne varmebehandling eller andre typer behandling, noe som ofte førte til en rask dekomposisjon av de store molekylene. I 1988 kom det to forskjellige teknikker som løste problemet med makromolekylene (Nelson & Cox 2008). Den ene av dem var MALDI – teknikken, utviklet av Koichi Tanaka (Tanaka et al. 1988) og i 2002 mottok han Nobelprisen i kjemi. Prisen mottok han sammen med John B. Fenn for hans utvikling av elektrospray ionisering og Kurt Wüthrich for hans arbeid med kjernemagnetisk resonans i strukturbestemmelse av biologiske makromolekyler (Kungliga Vetenskapsakademien 2002).

1.3.1 Grunnprinsippene ved analyse med MALDI-TOF MS til identifisering av mikroorganismer

Ved MALDI-TOF MS blir analytten (ca 1 μ l med hele bakterieceller eller proteiner/peptider) som er applisert på en metallflate (targetplate) ionisert ved hjelp av en pulserende laser (frekvens = nanosekunder) (Hortin 2006). For at ioniseringen skal finne sted, må analytten være dekket av en lysabsorberende matriks (Griffin & Smith 2000). Matriksen absorberer energien fra laseren og overfører den eksplosivt videre til analytten, slik at molekylene blir ionisert i gassfasen (Watson & Sparkman 2007). Vanligvis blir hvert av molekylene ionisert med en positiv ladning ($z = 1$). De ioniserte molekylene blir overført som en gass/partikkelsky til et time-of-flight – rør (flyverør). I vakuum blir partikkelskyen akselerert i et elektrisk felt mot en detektor (figur 1.3.1). Fordi dannelsen av ioner skjer i vakuum, er analysen uavhengig av andre

faktorer enn massen, så lenge ladningen er $z = 1$. Ioner med lav m/z (M_2^+ , i figur 1.3.1) vil bli akselerert til en høyere hastighet og vil nå detektoren før ioner med høyere m/z forhold (M_1^+ i figur 1.3.1). Dette kan beskrives med formelen; $\text{tid} = k(m/z)^{0,5}$, hvor k er en konstant bestemt ved kalibrering opp mot en proteinstandard. Ionene når detektoren etter 0,01-1 ns for hvert enkelt laserskudd. Analyseringstiden vil derfor være kort for hver prøve, selv om flere laserskudd er nødvendig for å bygge opp et massespekter (Griffin & Smith 2000). MALDI-TOF – teknikken omtales derfor som en "high – throughput" - analyseringsmetode (Brown & Lennon 1995; Hortin 2006).

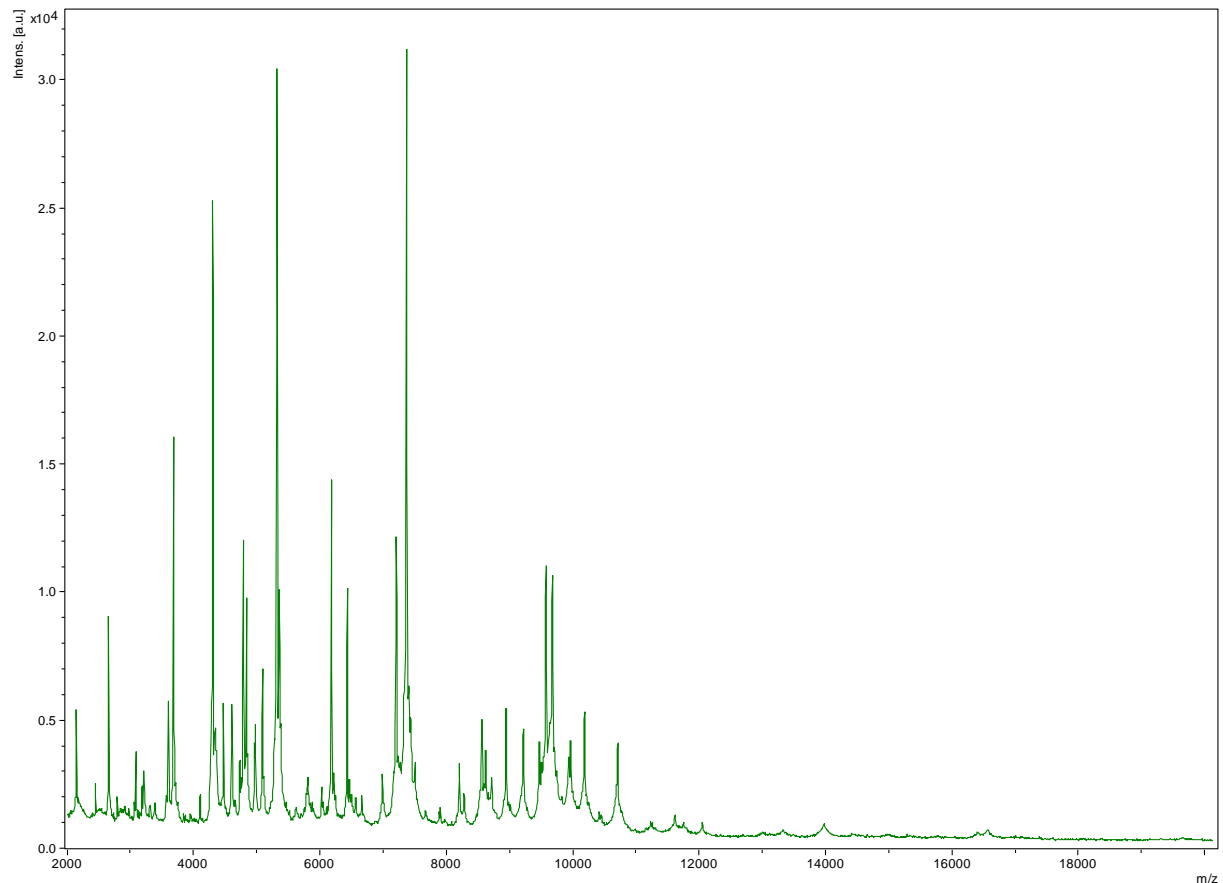


Figur 1.3.1: Skjematisk fremstilling av MALDI-TOF MS (Griffin & Smith 2000). M_1^+ og M_2^+ representerer analytter med henholdsvis høy og lav masse.

1.3.2 Massespekter generert med et massespektrometer

Et massespekter fra MS kan dannes når ionepartikler blir separert med hensyn på massen og antall ladninger, m/z forholdet (IUPAC 2011). Partikler som ikke er ionisert blir fjernet fra massespektrometer -systemet av pumpene som opprettholder vakuomet. X-aksen av et massespekter har benevnningen m/z (figur 1.3.2), som er basert på en atomisk skala. I den atomiske skalaen tilsvarer én masseenheter én

tolvtedel av isotopen ^{12}C , som tilsvarer én dalton (Da). Dette vil si at dersom $z = 1$, vil én m/z tilsvare én Da. Toppene i massespektrene er representert som intensiteten av mengden ioner som blir detektert ved en viss m/z i massespektrometret, y-aksen. Intensiteten (figur 1.3.2) har benevnningen ”arbitrary units” (a. u.), som indikerer at intensiteten ikke er et helt nøyaktig mål på mengden ioner, men samsvarer tilnærmet med dem (Watson & Sparkman 2007).



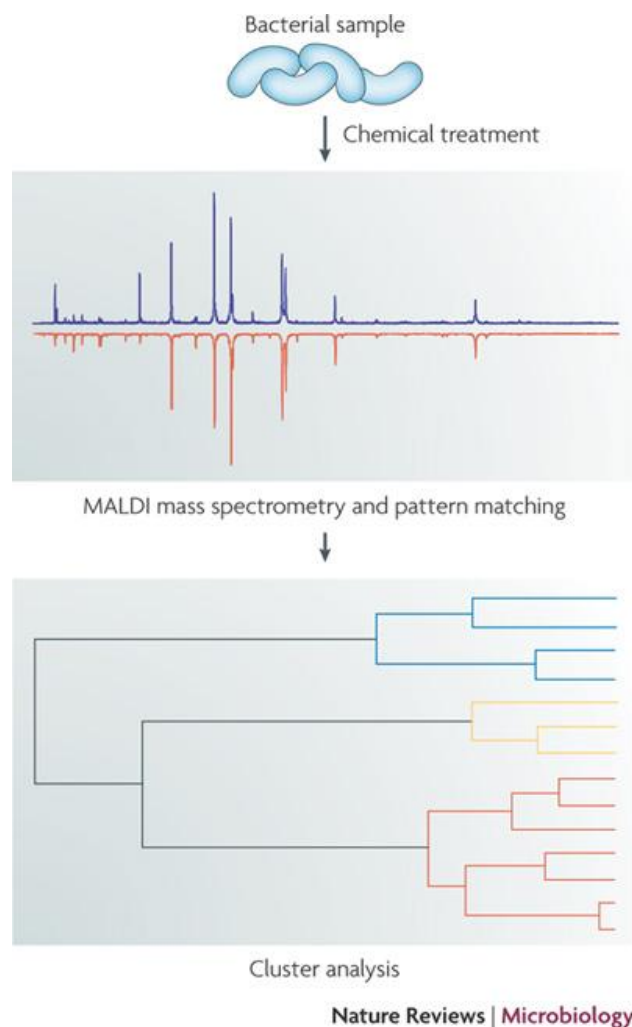
Figur 1.3.2: Råspekter av *L. pneumophila* (ATCC 33152) dannet i den oppgaven med MALDI-TOF MS. Hentet i flexAnalasys 3.3.

1.3.3 Identifisering og klassifisering av mikroorganismer med MALDI-TOF MS

MALDI-TOF MS -"fingerprinting" har vist seg å være en effektiv metode for identifisering av mikroorganismer på slekts-, arts- og i noen tilfeller underartsnivå. Metoden baserer seg på molekyler som er tallrike i en organisme og som derfor vil fremstå som klare topper i et massespekter. Massespektrene fra en organisme blir benyttet til å danne referansespekter og i et bibliotek vil de fungerer som et

fingeravtrykk i en sammenligning med et massespekter (Maier et al. 2006; Ryzhov & Fenselau 2001).

Referansespekterene vil i denne oppgaven vil bli omtalt som “Main Spectrum” (MSP). Et MSP er mer nøyaktig en referanse topp-liste som blir et molekylært fingeravtrykk av proteinsammensetningen for en art eller stamme. I programvaren MALDI Biotyper 3.0 danner MSP grunnlaget for klassifisering av organismer. Ved identifisering og klassifisering av ukjente prøver vil de massespektrene som blir dannet bli sammenlignet med referansebiblioteket (figur 1.3.3). Sammenligningen skjer ved hjelp av algoritmer for mønstergjenkjenning (Bruker Daltonics 2011).



Figur 1.3.3: Identifisering og klassifisering av mikroorganismer med MALDI–TOF MS (Sauer & Kliem 2010).

Ved en identifisering i MALDI Biotyper 3.0, vil massespektrene som er dannet fra analytten bli omregnet til en massetoppliste. I massetopplisten vil informasjonen om hvor massetoppene befinner seg og intensiteten av dem. Massetoppene under en definert terskelverdi vil ikke bli tatt med i listen. Massetoppene som overlapper mellom MSP og det eksperimentelle massespekteret vil telles. Deretter telles massetoppene som overlapper mellom det eksperimentelle massespekteret og MSP. Videre vil intensiteten av massetoppene som overlapper hverandre bli forbundet med hverandre gjennom en korrelasjonsmatrise. Totalt dannes tre verdier mellom 0 og 1 som multipliseres med hverandre og normaliseres til tusen. Scoreverdiene som blir dannet blir vist som logaritmen av resultatet. Den logaritmiske verdien som blir dannet ved sammenligningen blir behandlet av MALDI Biotyper 3.0 og satt opp mot definerte terskelverdier. Analytten blir gitt en scoreverdi, mellom 0 og 3 hvor 3 er nøyaktig likt, basert på sannsynligheten for korrekte klassifisering. En scoreverdi på >2 til 2.299 blir av MALDI Biotyper 3.0 bestemt til å være en sikker identifisering av slekt og antagelig identifisering av art. En scoreverdi >2.3 er bestemt til å være en høy sannsynlighet for korrekt artsidentifisering. Ved identifiseringen genereres en rapport over de 10 beste sammenligningene mot MSP i referansebiblioteket med tilhørende klassifiseringer. Ved sammenligning av nært beslektede underarter kan det i tillegg bli gitt en vektning av særegne topper i referansespektrene (Sauer et al. 2008) Dette blir i MALDI Biotyper referert til som ”subtyping” (Bruker Daltonics 2011). Referansebiblioteket fra Bruker Daltonics inneholder 4110 MSP fra ulike mikroorganismer i tillegg til 109 MSP i sikkerhetsdatabasen (12 Mars 2012).

1.3.4 Preparering av analytt til MALDI–TOF MS identifisering

Molekylene som blir detektert ved MALDI–TOF MS identifisering kommer hovedsakelig fra cytosol og deler generelle trekk som sterk basisitet, medium hydrofilitet og at de er tallrike. Studier har vist at i analyser av mikroorganismer ved bruk av MALDI-TOF MS, genereres store deler av massespektrene fra signaler fra ribosomale proteiner. Disse er ofte godt konserverte, men små sekvensvariasjoner forekommer også på underarts og stammenivå som kan bidra til at masseprofilene blir forskjellige (Ryzhov & Fenselau 2001; Teramoto et al. 2007).

Ulike prosedyrer for prøveopparbeidelse har blitt utviklet for ulike bakterieslekter (Claydon et al. 1996; Holland et al. 1996). Vanligvis er en enkelt bakteriekoloni tilstrekkelig analysemateriale ved MALDI-TOF MS -analyse. Studier har vist at ved bruk av forskjellige typer dyrkningsvilkår eller forskjellige kjemiske ekstraksjonsmetoder, kan samme bakterie gi forskjellige massespekter. Chen et al. 2008 mener ”forskjellene som ble sett ved ulik lengde på dyrkningen, kan skyldes et økt behov for ribosomale proteiner med lav masse i de tidlige stadiene av bakterieveksten”. Standardiserte prepareringsmetoder og godt kontrollerte vekstvilkår er derfor essensielt for å danne reproduerbare massespekter (Chen et al. 2008; Williams et al. 2003). Proteinekstraksjonsprotokollen som blir presentert i denne oppgaven er godt utprøvd og har tidligere blitt brukt til identifisering av blant annet; *Listeria* (Barbuddhe et al. 2008); gjær – isolater (Marklein et al. 2009); *Prototheca* (Murugaiyan et al. 2011). Protokollen har også vist, i sammenligninger med andre prepareringsmetoder til MALDI-TOF MS -analyse, å gi et høyt utbytte i konsentrasjon av ekstraherte proteiner (Drevinek et al. 2012).

1.3.5 Analyser med MALDI Biotyper 3.0

I tillegg til identifisering med referansebiblioteket, er det i MALDI Biotyper 3.0 mulig å gjøre tre forskjellige analyser med massespektrene; klassifisering med trestruktur, ”Composite Correlation Index” (CCI) -matrise og ”Principal Component Analysis” (PCA) (Bruker Daltonics 2011).

Med MSP er det i MALDI Biotyper 3.0 mulig å danne fylogenetiske trestrukturer. Trestrukturene gjør det mulig å vise hvor nært beslektet organismene er til hverandre med bakgrunn i de masseprofilene som er dannet i referansebiblioteket. Slektskapet i trestrukturen blir reflektert med et vilkårlig avstandsnivå (Bruker Daltonics 2011).

I likhet med trestrukturen, illustrerer CCI -matrisen hvor nært beslektet organismene som sammenlignes er hverandre. CCI -matrisen baserer seg ikke på MSP, men på massespekter og kan brukes til statistisk analyse av forholdet mellom massespektrene. ”CCI -matrisen blir i MALDI Biotyper 3.0 beregnet ved å dele spektrene inn i diskree intervaller og sammenligne disse intervallene over et datasett” BrukerDaltonics 2011. Avstanden mellom spektrene blir bestemt av hvordan intervallene overensstemmer med hverandre. CCI -analyse i MALDI Biotyper 3.0 kan

hovedsakelig bli brukt til å undersøke forholdet mellom nært beslektede organismer (Bruker Daltonics 2011).

PCA i MALDI Biotyper 3.0 gir informasjon om hetero/homogeniteten av et datasett. ”Hvor hensikten med PCA er å redusere dimensjonaliteten av datasettet hvor det er mange variabler knyttet til hverandre, samtidig som variasjonen opprettholdes” BrukerDaltonics 2011. Hensikten er å identifisere trender og sammenhenger i datasettet, som ellers kan være skjult, ved et koordinatsystem. PCA analysen kan i MALDI Biotyper 3.0 visualiseres i en trestruktur (Bruker Daltonics 2011).

1.4 Sekvensbasert typing av *L. pneumophila*

SBT er en universell betegnelse for en genotypingsmetode, hvor EWGLI har utviklet en spesifikk protokoll for *L. pneumophila* (Gaia et al. 2005). Metoden er basert på ”Multilocus sequencing typing” (MLST) først beskrevet for *Streptococcus pneumoniae* (Enright & Spratt 1998). SBT av *L. pneumophila* skiller seg fra MLST ved hvilke type gener som blir brukt. En MLST analyse er basert på husholdningsgener, mens SBT – protokollen som er brukt i denne oppgaven ser på gener med varierende selektivt press. Hensikten med SBT - metoden er å ha en standardisert genotypingsmetode og danne en database slik at kliniske og miljøisolater av *L. pneumophila* kan bli sammenlignet. DNA – sekvensen til spesifikke genfragmenter gir informasjon om bakteriens identitet (sekvenstype (ST)), og kan benyttes til smitteoppsporing ved utbrudd. ST bidrar også til at ulike laboratorier kan utveksle data uten å måtte utveksle isolater (Gaia et al. 2005).

SBT av *L. pneumophila* baserer seg på sammenligning av PCR -produkter fra genene *flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA* og *neuA*. Hvor de amplifiserte PCR produktene blir sekvensert med både foroverrettede (F) og bakoverrettede (R) primere. Nukleotidsekvensene som amplifiseres ved PCR har en lengde på mellom 245 og 648 basepar (bp) avhengig av hvilke gen det blir sett på (tabell 1.4.1). I EWGLI – databasen kan overensstemmende sekvenser bli angitt og sekvensene bli gitt individuelle allelnummer (sekvensene som blir angitt i databasen er forhåndsbestemte deler av PCR-produktene). Dette danner en profil basert på allelene (figur 1.4.1), sammensatt av 7 nummer separert av komma (eks. 3,4,1,1,14,9,1 for *L. pneumophila* ATCC 33152) i en forhåndsbestemt rekkefølge (*flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA* og

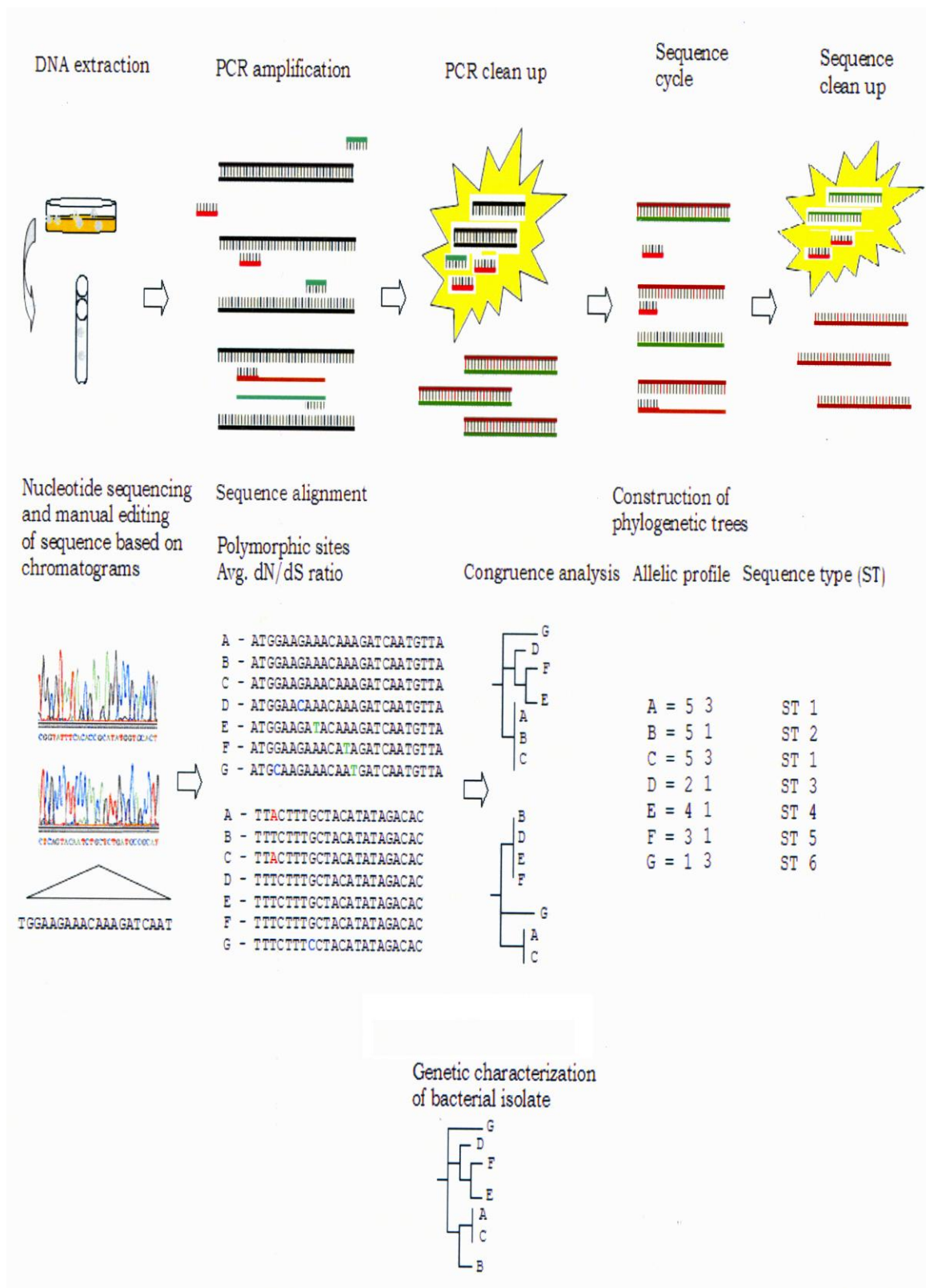
neuA). ST kan bli bestemt når alle de syv allelene er tilstede. Dersom et allelnummer ikke kan bli bestemt, blir det angitt med en 0 og det kan da heller ikke bli gitt en ST (eks. 3,4,1,1,14,9,0 = ST0) (The Health Protection Agency 2012). Allikevel kan allelprofilen med de allelnummerene som er bestemt være svært informativt (Ratzow et al. 2007).

Tabell 1.4.1: Informasjon om SBT – logi (The Health Protection Agency 2012)

Gen	Sørrelse på PCR-produkt (bp)	GeneBank-nummert til referansesekvens	Region av allel som blir angitt til EWGLI – databasen*	Størrelse på målregion (nukleotider)
<i>flaA</i>	245	X83232	653-834	182
<i>pilE</i>	459	AF048690	103-435	333
<i>asd</i>	575	AF034213	538-1010	473
<i>mip</i>	558	AJ496269	117-518	402
<i>mompS</i>	648	AF078136	523-874	352
<i>proA</i>	480	M32884	1134-1538	405
<i>neuA</i>	459	AJ007311	229-582	354

*med bakgrunn i referansesekvensen

Genene brukt i karakteriseringen av *L. pneumophila* med SBT er valgt ut på bakgrunn av den samlede genetiske diversiteten SBT- logi gir. Så langt har EWGLI – databasen 6002 innføringer, med 1167 ulike ST (april 2012) (The Health Protection Agency 2012).



Figur 1.4.1: Skjematisk fremstilling av stegene i en SBT – analyse (Tveten 2008).

2. Materialer og Metoder

2.1 Dyrkingsforhold

Legionella spp. som ble benyttet under forsøkene, ble dyrket opp på BCYE- agar (Oxoid 2010) ved 37 °C i 48 ± 2 timer. Dette ble gjort for å sikre at nok biologisk materiell var tilgjengelig og for å sørge for at bakteriene var i den reproduserende fasen (1.1.1).

Bacillus cereus og *Vibrio cholerae* som ble benyttet ble dyrket opp på Brain Heart Infusion (BHI) – agar (tabell 2.1.1) ved 37 °C i 24 timer.

Tabell 2.1.1: Innhold i BHI - agar

Forbindelse	g/L	Instruksjon
Brain Heart Infusion (CM1135)	37	BHI ble blandet med dH ₂ O og
Agar	15	Agar ved omrøring.
pH 7,4 ± 0.2		Autoklavert ved 120 °C i 20 min

2.2 Bakteriestammer

I denne studien ble totalt 46 *Legionella* spp. (tabell 2.2.1) benyttet, hvorav 41 var *L. pneumophila*. I tillegg ble isolater fra artene *L. bozemanii*, *L. brunensis*, *L. dumoffii*, *L. longbeachae*, *L. maceachernii*, *L. micdadei* benyttet i MALDI-TOF MS –arbeidet. I de innledende forsøkene ble et isolat fra *B. cereus* og *V. cholerae* benyttet til sammenligning mellom ulike bakteriarter.

Tabell 2.2.1:Stammeliste av *Legionella* spp.

Art	Kode/alias	Serogruppe	Fått/ kjøpt fra:	ST	Allelprofil	Allelprofil funnet i:	Antall massespekter i MSP
<i>L. pneumophila</i>	ATCC 33152	1	CRBIP	36	3,4,1,1,14,9,1	Denne oppgaven	34
<i>L. pneumophila</i>	ATCC 33154	2	CRBIP	39	3,5,1,7,14,9,8	Denne oppgaven	32
<i>L. pneumophila</i>	ATCC 33155	3	CRBIP	395	7,6,17,20,13,11,3	Denne oppgaven	33
<i>L. pneumophila</i>	ATCC 33156	4	CRBIP	0	11,14,16,25,7,13,0	(Visca et al. 2011)	32
<i>L. pneumophila</i>	ATCC 33216	5	CRIBP	0	11,14,16,18,15,13,0	Denne oppgaven	33
<i>L. pneumophila</i>	ATCC 33215	6	CRBIP	30	3,10,1,3,14,9,9	Denne oppgaven	36
<i>L. pneumophila</i>	ATCC 33823	7	CRBIP	0	2,6,17,14,12,8,0	Denne oppgaven	34
<i>L. pneumophila</i>	ATCC 35096	8	CRBIP	0	8,1,22,30,6,10,0	Denne oppgaven	33
<i>L. pneumophila</i>	ATCC 35289	9	CRBIP	390	1,4,3,28,1,1,6	(Visca et al. 2011)	35
<i>L. pneumophila</i>	ATCC 43283	10	CRBIP	17	2,10,3,28,9,4,6	Denne oppgaven	36
<i>L. pneumophila</i>	ATCC 43130	11	CRBIP	0	12,17,11,10,5,12,0	Denne oppgaven	34
<i>L. pneumophila</i>	ATCC 43290	12	CRBIP	187	3,10,1,28,14,9,3	(Visca et al. 2011)	29
<i>L. pneumophila</i>	ATCC 43736	13	CRBIP	0	6,10,5,10,9,1,0	(Visca et al. 2011)	34
<i>L. pneumophila</i>	ATCC 43703	14	CRBIP	0	2,10,3,28,9,4,0	Denne oppgaven	36
<i>L. pneumophila</i>	FFI-Leg 321	1	EWGLI	78	2,3,6,25,2,1,15	FFI	34
<i>L. pneumophila</i>	FFI-Leg 322	1	EWGLI	1	1,4,3,1,1,1,1	FFI	35
<i>L. pneumophila</i>	FFI-Leg 323	1	EWGLI	22	2,3,6,10,2,1,6	FFI	35
<i>L. pneumophila</i>	FFI-Leg 324	1	EWGLI	152	1,4,3,1,1,1,3	FFI	36
<i>L. pneumophila</i>	FFI-Leg 461	1	EWGLI	59	7,6,17,3,13,11,11	FFI	32
<i>L. pneumophila</i>	FFI-Leg 462	3	EWGLI	93	3,10,1,28,14,9,13	FFI	29
<i>L. pneumophila</i>	FFI-Leg 463	5	EWGLI	70	7,6,17,8,13,11,3	FFI	30
<i>L. pneumophila</i>	FFI-Leg 464	6	EWGLI	187	3,10,1,28,14,9,3	FFI	29
<i>L. pneumophila</i>	FFI-Leg 470	1	EWGLI	42	4,7,11,3,11,28,9	FFI	32
<i>L. pneumophila</i>	FFI-Leg 471	5	EWGLI	0			32
<i>L. pneumophila</i>	FFI-Leg 472	1	EWGLI	481	4,7,11,13,11,12,9	FFI	31
<i>L. pneumophila</i>	FFI-Leg 475		EWGLI	93	3,10,1,28,14,9,13	FFI	35
<i>L. pneumophila</i>	FFI-Leg 476		EWGLI	75	2,3,18,13,25,5,6	FFI	36
<i>L. pneumophila</i>	FFI-Leg 477		EWGLI	575	2,3,18,13,25,1,6	FFI	30
<i>L. pneumophila</i>	FFI-Leg 479		EWGLI	0			33

Tabell 2.2.1:Stammeliste av *Legionella* spp. (fortsettelse)

Art	Kode/alias	Serogruppe	Fått/ kjøpt fra:	ST	Allelprofil	Allelprofil funnet i:	Antall massespekter i MSP
<i>L. pneumophila</i>	B111		Rikshospitalet	40	3,6,1,14,14,9,11	Denne oppgaven	34
<i>L. pneumophila</i>	B607		Rikshospitalet	40	3,6,1,14,14,9,11	Denne oppgaven	36
<i>L. pneumophila</i>	B1094		Rikshospitalet	42	4,7,11,3,11,12,9	Denne oppgaven	33
<i>L. pneumophila</i>	B1481		Rikshospitalet	0	4,7,11,3,11,12,0	Denne oppgaven	28
<i>L. pneumophila</i>	B2639		Rikshospitalet	0	6,10,14,1,33,1,0	Denne oppgaven	36
<i>L. pneumophila</i>	B2760		Rikshospitalet	581	1,10,1,1,14,9,1	Denne oppgaven	33
<i>L. pneumophila</i>	B2802		Rikshospitalet	583	7,6,17,28,13,11,3	Denne oppgaven	35
<i>L. pneumophila</i>	B3760		Rikshospitalet	1045	4,10,11,10,11,12,28	Denne oppgaven	35
<i>L. pneumophila</i>	B5329		Rikshospitalet	1	1,4,3,1,1,1,1	Denne oppgaven	35
<i>L. pneumophila</i>	15		Borregaard (2005)	15	12,9,26,5,26,17,15	FFI	35
<i>L. pneumophila</i>	458		Borregaard (2007)	458	24,6,3,3,13,11,11	FFI	35
<i>L. pneumophila</i>	462		Borregaard (2008)	462	12,9,2,5,50,17,15	FFI	36
<i>L. bozemanii</i>	B1456		Rikshospitalet				31
<i>L. bozemanii</i>	ATCC 33217		CCUG				34
<i>L. dumoffii</i>	ATCC 33279		CCUG				32
<i>L. brunensis</i>	ATCC 43878		CCUG				36
<i>L. maceachernii</i>	ATCC 35300		CCUG				35
<i>L. micdadei</i>	ATCC 33218	1	CCUG				35
<i>L. longbeachae</i>	ATCC 33462	1	CCUG				34

2.3 Prøvepreparering til MALDI-TOF -analyse

I denne oppgaven benyttes to prepareringsmetoder til MALDI-TOF MS -analyse, anbefalt av leverandøren av Microflex -instrumentet (Bruker Daltonics, Tyskland); en proteinekstraksjonsbasert prosedyre og en direkteoverføring – metode. Ved begge prosedyrene ble prøvene overført til en targetplate med 96 spotter (figur 2.3.1) før analysen.



Figur 2.3.1: MALDI targetplate i polert stål med 96 spotter (Bruker Daltonics 2012).

2.3.1 Protokoll for Proteinekstraksjon

En standard ekstraksjonsprosedyre, foreslått av Bruker Daltonics GmbH, ble tilpasset og brukt til å preparere prøvene før de ble applisert på MALDI targetplatene. Tilpasningen i protokollen ble gjort gjennom de innledende forsøkene (2.5).

Ca 2 μ l med celler fra enkeltkolonier, ble suspendert i 300 μ l sterilt ione-fritt vann (LC-MS Chromasolv) i 1,5 μ l eppendorfrør. Løsningen ble homogenisert ved risting i 1 min med Reax top vortexmaskin (Heidolph, Tyskland). Suspensjonen ble tilsatt 900 μ l absolutt etanol, og deretter ristet i 1 min. Løsningen ble sentrifugert med Heraeus Pico 17 – sentrifuge (Thermo Scientific, Tyskland) ved 13000 RPM i 2 min og supernatanten fjernet ved forsiktig pippetering. Sentrifugeringen og pippeteringen ble gjentatt for å sikre at all etanol ble fjernet. Pelleten ble tørket i 5 min slik at mulig gjenværende etanol fordampet. 10 μ l med 70 % maursyre ble tilsatt og pelleten resuspendert til en homogen løsning ved risting i 1 min. Deretter ble et likt volum av

100 % acetonitril tilsatt forsiktig langs kanten av eppendorfrøret, for å unngå oksidering av prøven. Prøvene ble sentrifugert ved 13000 RPM i 2 min, før 1 µl av supernatanten ble applisert på MALDI targetplaten. Prøvene tørket så ved romtemperatur før 1 µl med α -cyano-4-hydroxycinnamic acid (HCCA)- matrix (tabell 2.3.1) ble applisert på spottene. Dette førte til en co-krystallisering mellom HCCA og prøven. De ferdig preparerte targetplatene tørket i avtrekkskap i ca 5 min, før de var klare for analyse med MALDI-TOF MS (Bruker Daltonics 2011).

Tabell 2.3.1: HCCA – matrix

Forbindelse	Volum (%)	Instruksjon
Sterilt ionefritt vann (LC-MS Chromasolv)	47,5	Tilsett TFA og acetonitril til vannet og bland ved risting.
Trifluoroacetic acid (TFA) 99,9 %	2,5	Tilsett 250 µl av løsningen til porsjonert HCCA.
Acetonitril 100 %	50	Rist til løsningen er homogen.
HCCA (porsjonert) (Bruker Care Online 2012b)	2,5 mg	
Ferdig HCCA - matrix lagres mørkt*		

*Ved evt. utfelling under lagring eller oppbevaring over 14 dager, ble løsningen kastet.

2.3.2 Protokoll for Direkte – metode

Bakterieceller fra agarplaten kan overføres direkte til targetplaten uten proteinekstraksjon.

En tannpirker ble dyppet forsiktig i en enkeltkoloni og smurt tynt utover på hver enkelt spott på targetplaten for å danne et tynt homogent lag med celler. Innen 10 minutter etter at prøven hadde tørket, ble 1 µl HCCA – matrix tilsatt hver spott. Når matrixen hadde tørket kunne analytten analyseres (Bruker Daltonics 2011).

MALDI-TOF analysen er meget følsom for mengden prøvemateriale applisert på targetplaten. Denne prepareringsmetoden kan derfor gi feilaktige resultater dersom for mye eller for lite biologisk materiale blir tilsatt (Bruker Daltonics 2011).

2.4 Målinger med MALDI–TOF MS

MALDI–TOF instrumentet brukt i denne oppgaven er en Microflex (Bruker Daltonics, Bremen, Tyskland). Instrumentet er designet for automatisert proteinidentifisering og -karakterisering, biomarkør-deteksjon, oligonukleotid kvalitetskontroll, og SNP (singel nucleotide polymorphism) genotyping (Bruker Daltonics 2004). Instrumentinnstillingene for Microflex -instrumentet ble ikke endret ved arbeidet med oppgaven (tabell 2.4.1).

Tabell 2.4.1: Instrumentinnstillinger, Microflex MALDI–TOF MS.

Instrumentparameter	
Laserfrekvens	60 Hz
Masserekkevidde	2000 til 20137 m/z
Ionekilde 1	20,04 kV
Ionekilde 2	16,73 kV
Linse	8,02 kV
Forsinkingstid på ekstraksjon	120 ns
Begynnende laserstyrke	30 %
Maksimal laserstyrke	40 %
Laserattenuator offset	23 % (rekkevidde 30 %)

Prøvene som ble analysert ble overført til en 96 spott targetplate. Hvert massespekter ble automatisk dannet ved hjelp av AutoXecute i programmet Flex control 3.3 (Bruker Daltonics, Leipzig, Tyskland). Programmet genererte et massespekter ved å summere massespektrene fra 240 laserskudd (et massespekter fra hvert skudd, 40 skudd i ulike posisjoner på hver spott, ved ”random walk movement”).

Ekstern kalibrering av maskinen ble gjort med en bakteriell teststandard (BTS) (Bruker Care Online 2012a) (tabell 2.4.2) som dekket en masse av molekyler på mellom 2 tusen til 20 tusen m/z (Bruker Daltonics 2011).

Tabell 2.4.2: BTS

Forbindelse	Volum (%)	Instruksjon
Sterilt ionefritt vann LC-MS Chromasolv	47,5	Tilsett TFA og acetonitril til vannet og bland ved risting.
Trifluoroacetic acid (TFA) 99,9 %	2,5	Tilsett 50 µl av løsningen til porsjonert BTS.
Acetonitril 100 %	50	Pipeter forsiktig opp og ned 20 ganger for å unngå oksidering.
BTS (porsjonert)		Gjenta pipeteringen etter 5 min dersom løsningen ikke er blakket.
BTS -løsningen ble lagret mørkt ved -18 °C *		

* Kastes etter tredje opptining.

2.5 Innledende forsøk

Før arbeidet med å danne biblioteket ble påbegynt ble tre separate forsøk utført for å evaluere henholdsvis; hvilke volum av biologisk materiale som ga høyest scoreverdi ved MALDI-TOF MS identifisering; hvilke volum av acetonitril og maursyre ved proteinekstraksjonen som ga høyest scoreverdi ved MALDI-TOF MS identifisering; og stabiliteten av lagrede proteinekstrakter og kolonier på agarplatene. Ved alle forsøkene ble MALDI Biotyper Realtime Classification Wizard i programvaren MALDI Biotyper Version 3.0 (Bruker Daltonics GmbH) brukt til å danne massespekter og klassifisere prøven opp mot det eksisterende referansebiblioteket i programvaren. Totalt ble tre massespekter dannet for hver prøve, hvor scoreverdien fra sammenligningene med best treff mot MSP ble brukt i vurderingen av resultatet.

2.5.1 Evaluering av optimal mengde biologisk materiale

Proteinekstraksjonsprotokollen (2.3.1) ble brukt i utførelsen av forsøket, hvor de eneste parametrene det ble justert på var volumet av biologisk materiale fra agarplatene. Til forskjell fra nevnte protokoll var også volumet av acetonitril og maursyre satt til 20 µl. Volumene av biologisk materiale som ble brukt tilsvarer ca ¼, 1, 2 og 6-8 kolonier, hvor koloniene hadde en diameter på ca 3-4 mm. Tre parallelle proteinekstraksjoner ble dannet for hvert volum hvorav tre pseudoparalleller fra hver ekstraksjon ble overført på targetplaten. Forsøket ble utført med *B. cereus*, *V. cholerae* og *L. pneumophila* (ATCC 33823 og ATCC 33152).

2.5.2 Evaluering av optimalt volum maursyre og acetonitril

I evalueringen av acetonitril- og maursyrevolumet ble proteinekstraksjonsprotokollen (2.3.1) brukt, de eneste parametrene det ble justert på var de nevnte volum. Volumene som ble brukt var 2, 5, 15, 30 og 50 µl. Tre parallelle proteinekstraksjoner ble dannet for hvert volum hvorav tre pseudoparalleller fra hver ekstraksjon ble overført til targetplaten. Evalueringen ble gjort med *L. pneumophila* (ATCC 33823).

2.5.3 Stabilitet av proteinekstrakter og kolonier ved lagring

For å teste stabiliteten av proteinekstrakter og kolonier ved lagring, ble det både gjort en proteinekstraksjon (2.3.1) og en direkteoverføring (2.3.2). Ved proteinekstraksjonen ble et volum på 30 µl av acetonitril og maursyre brukt. Forsøket ble utført med *B. cereus*, *V. cholerae* og *L. pneumophila* (ATCC 33823) over henholdsvis 9, 8 og 7 dager. Hver bakterie ble dyrket opp på to agarplater ved 37 °C og ved 20 °C, og lagret videre ved samme temperaturer til forsøket var over. For hver dag ble det gjort tre parallelle proteinekstraksjoner for hver agarplate, hvor totalt tre pseudoparalleller fra hver ekstraksjon ble overført på targetplaten. Samtidig ble tre parallelle fra hver agarplate overført på targetplaten ved hjelp av direkte metoden. Proteinekstraksjonen fra dag én ble lagret videre ved 4 °C og overført på targetplaten hver dag, for å teste stabiliteten ved lagring.

For å illustrere endringene i massespektrene ble ett massespekter fra hver av dagene plukket ut og sammenlignet. Massespektrene (totalt 5) ble hentet fra MALDI-TOF MS -identifiseringen fra *L. pneumophila* -koloniene dyrket ved 37 °C og preparert med proteinekstraksjon. Sammenligningen ble gjort i Flex Analysis 3.3 (Bruker Daltonics, Bremen, Tyskland).

2.6 Generering av massespekter med MALDI-TOF MS til MSP - dannels

For å preparere prøver til dannels av massespekter til MSP - referansebiblioteket, ble ekstraksjonsprotokollen (2.3.1) benyttet. Massespektrene ble direkte dannet i Flex control 3.3.

For hver bakterie ble det gjort 3 proteinekstraksjoner og for hver ekstraksjon ble 4 paralleller overført på targetplaten, til totalt 12 spotter. Av disse ble det dannet 3 massespekter per spott som totalt ga 36 massespekter per isolat.

2.6.1 Analysering av massespekter og generering av MSP

Kvaliteten til hvert massespekter ble analysert og manuelt bestemt gjennom programvaren FlexAnalysis 3.3. 36 ”rå-massespekter” fra hvert isolat ble først behandlet med en algoritme for grunnlinjesubstitusjon og en algoritme for å jevne ut massespektrene. Deretter ble avvikende massespekter i massespektersettene fjernet, basert på variasjon rundt toppene. I denne oppgaven var variasjonen bestemt til å være 500 deler per million (ppm), etter anbefaling fra Bruker Daltonics.

Et minimum av 20 massespekter var nødvendig for å danne ett MSP. Programvaren, MALDI Biotyper 3.0, konstruerte en peak/topp-liste fra hele datasettet og danner MSP ved å bruke informasjonen om massene til toppene, frekvensen av topper og intensitetsfordelingen til toppene (Bruker Daltonics 2011).

2.6.2 Blindtest – identifisering av *L. pneumophila* med MSP i referansebibliotek

For å undersøke om det var mulig en god identifisering med MSP i referansebiblioteket som ble konstruert, ble det satt opp en blindtest. 14 av *L. pneumophila* isolatene (tabell 2.2.1) ble tilfeldig valgt ut og blindet. Tre proteinekstraksjoner (2.3.1) ble utført for hvert isolat, hvor 3 paralleller fra hver ekstraksjon ble overført til targetplaten (totalt 9 spotter for hvert isolat).

Identifiseringen ble utført i MALDI Biotyper 3.0 opp mot det konstruerte referansebiblioteket (41 *L. pneumophila* oppføringer, tabell 2.2.1) og mot 11 *L. pneumophila* MSP som var oppført i referansebiblioteket til Bruker Daltonics (Vedlegg A, tabell A1) (totalt 52 MSP). Ett massespekter ble dannet for hver spott som totalt ga ni identifiseringer per isolat. I vurderingen av resultatene ble den høyeste scoreverdien for hvert massespekter i identifisering vurdert og treff mot eget MSP ble telt for hver identifisering.

Massespektrene fra blindtesten ble gjennom MALDI Biotyper 3.0 brukt til å danne nye MSP. MSP fra massespektrene ble identifisert mot referansebiblioteket for å undersøke om en bedre sammenligning ble oppnådd, enn det massespektrene hver for seg oppnådde. MSP fra massespektrene ble dannet slik beskrevet (2.6.1), med unntak av at totalt ni massespekter ble brukt i dannelsen av hver MSP.

2.6.3 Fylogenetisk fremstilling av MSP og massespekter

En fylogenetisk klassifisering basert på MSP som var blitt dannet ble konstruert i Biotyper 3.0 programvaren. Programvaren konstruerte en trestruktur etter standard innstillinger (tabell 2.6.1). Trestrukturen ble behandlet og tilpasset i Adobe Photoshop Elements 8 (Adobe, California, USA).

Tabell 2.6.1: Standardinnstillinger ved dannelse av trestruktur fra MSP i MALDI Biotyper 3.0

Score Oriented Dendrogram	
Distance Measure	Correlation
Linkage	Average
Score Threshold Value for a Single Organism	300
Score Threshold Value for a Related Organism	0

Samtidig ble det dannet et fylogenetisk tre basert på PCA (1.3.4) for å undersøke om det ble sett noen nye trender. PCA -analysen kan i MALDI Biotyper 3.0 bare utføres på massespekter. Det ble derfor valgt ut et massespekter fra hvert av *L. pneumophila* -isolatene som var benyttet til dannelsen av MSP (2.6). Massespektrene som ble valgt var analysert og behandlet slik det er forklart (2.6.1). Alle massespektrene var generert fra posisjon 6 på targetplaten (A6 – H6). Standard innstillinger (tabell 2.6.2), foreslått av Bruker Daltonics ble brukt til dannelsen av treet. Trestrukturen ble behandlet og tilpasset i Adobe Photoshop Elements 8.

Tabell 2.6.2: Standardinnstillinger ved dannelse av trestruktur med PCA i MALDI Biotyper 3.0

Method	Hierarchical
Distance Measure	Correlation
Linkage Algorithm	Average
Maximal Cluster	4
Maximal Number of Top Level Noedes	All

I den fylogenetiske fremstillingen fra både MSP og PCA ble det observert to tydelige forgreninger (figur 3.2.3 b og figur 3.2.4) hvor grenene er markert i rødt og blått. Den blå forgreningen besto i begge de fylogenetiske treene av de samme *L. pneumophila* -isolatene. Ett massespekter fra tre av isolatene i hver klynge ble valgt ut for undersøke om det kunne observeres en sammenheng i massetoppene. *L. pneumophila* -isolatene som ble valgt ut fra blå forgrening var; BRG-ST458; BRG-ST462; ATCC 33154, og fra rød forgrening; BRG-ST15; ATCC 33152; ATCC 33215 (tabell 2.2.1). Massespektrene som ble benyttet var på forhånd analysert i dannelsen av MSP (2.6.1) og brukt i dannelsen av PCA – treet som forklart ovenfor. Massespektrene ble illustrert gjennom FlexAnalasys 3.3 -programvaren.

2.7 SBT

En SBT protokoll fra EWGLI til epidemiologisk typing av *L. pneumophila* ble benyttet til å bestemme ST av isolatene i en 7-nummer allel-profil. Allel - profilene ble videre brukt til å gjøre en fylogenetisk sammenligning av isolatene.

2.7.1 DNA isolering

DNA fra 19 *L. pneumophila* isolater (markert; Denne oppgaven, tabell 2.2.1) ble isolerte ved å danne kokepreparater. En suspensjon av 0,7 ml sterilt GIBCO-vann (DNase- og RNase – fritt) og ca 3 µl med celler fra enkelkolonier ble laget i eppendorfrør, og ristet. Suspensjonen ble varmet med Thermomixer confort – varmeblokk (Eppendorf, Tyskland) ved 99 °C i 15 min. Suspensjonen ble sentrifugert og lysatet brukt i PCR amplifisering. Kokepreparatene ble lagret ved -18 °C.

2.7.2 SBT gen og primere

Primerene brukt (tabell 2.7.1) rettet seg mot syv gener, *flaA*, *pilE*, *asd*, *mip*, *mompS*, *proA* og *neuA*. Primerene var alle M13 – primere, slik at de direkte kunne bli klonet inn i M13 – vektorer og gjøre sekvenseringen lettere.

Tabell 2.7.1: SBT primere brukt i denne studien.

Primer-navn	Mål-gen	Nukleotidsekvens (5'-3')
<i>flaA</i> -619F	<i>flaA</i>	TTT CTC TGG CGC AAG CTT CC
<i>flaA</i> -846R	<i>flaA</i>	GCT GCT TTG GCA TAG GCA G
<i>pilE</i> -12F	<i>pilE</i>	CAC AAT CGG ATG GAA CAC AAA CTA
<i>pilE</i> -453R	<i>pilE</i>	GCT GGC GCA CTC GGT ATC T
<i>asd</i> -487F	<i>asd</i>	CCC TAA TTG CTC TAC CAT TCA GAT G
<i>asd</i> -1039R	<i>asd</i>	CGA ATG TTA TCT GCG ACT ATC CAG
<i>mip</i> -58F	<i>mip</i>	GCT GCA ACC GAT GCC AC
<i>mip</i> -595R	<i>mip</i>	CAT ATG CAA GAC CTG AGG GAA C
<i>mompS</i> -430F	<i>mompS</i>	TTG ACC ATG AGT GGG ATT GG
<i>mompS</i> -1116R	<i>mompS</i>	TGG ATA AAT TAT CCA GCC GGA CTT C
<i>proA</i> -1090F	<i>proA</i>	GAT CGC CAA TGC AAT TAG
<i>proA</i> -1553R	<i>proA</i>	ACC ATA ACA TCA AAA GCC
<i>neuA</i> -196F	<i>neuA</i>	CCG TTC AAT ATG GGG CTT CAG
<i>neuA</i> -611R	<i>neuA</i>	CGA TGT CGA TGG ATT CAC TAA TAC

2.7.3 PCR amplifisering

En LightCycler[®] 480 Real-Time PCR System ble brukt sammen med LightCycler[®] 480 SYBR Green I Master Kit (Roche applied science) ved PCR amplifiseringen. PCR amplifiseringen ble gjort med 18 µl reaksjonsblanding (tabell 2.7.2) sammen med 2 µl templat DNA fra kokepreparatet (2.7.1) (utgjorde en 1:10 fortyning av det totale reaksjonsvolumet), totalt 20 µl reaksjonsvolum. Amplifiseringen ble utført under forholdene som er beskrevet (tabell 2.7.3).

Tabell 2.7.2: Innhold i reaksjonsblanding for PCR amplifisering (for hver reaksjonsbrønn)

Reagens	Volum (μ l)
dH ₂ O (GLYCO)*	4
Enzym miks**	10
Primer miks (5 μ M) *** (blanding av både F og R primer)	4

*DNase- og RNase- fritt.

**Består av FastStart Taq DNA Polymerase, reaksjonsbuffer, dNTP miks (med dUTP i stede for dTTP), SYBR Green I dye, og MgCl₂.

***Konsentrasjonen utgjør 1 μ mol i det totale reaksjonsvolumet på 20 μ l.

Tabell 2.7.3: Programmet brukt ved PCR amplifiseringen.

Trinn	Temperatur ($^{\circ}$ C)	Sykluser	Tid
Denaturering	95	1	300 sek
Amplifisering	Denaturering		30 sek
	Avspenning	35	30 sek
	Forlenging		30 sek
Forlengning	72	1	420 sek
Smeltekurve	58 \rightarrow 95	1	0,11 $^{\circ}$ C/sek
Avkjøling	37		

2.7.4 Verifisering av PCR produkt

Smeltepunktskurven som ble dannet under PCR amplifiseringen (2.7.3) ble brukt til å indikere spesifisiteten til PCR produktet, ved å se etter dannelser av dimere mellom primerene.

I tillegg ble Aligent 2100 BioAnalyzer[®] (Aligent Technologies) brukt til å utføre elektroforese, med DNA 1000 series II kit, ved å følge instruksjonene fra fabrikanten. Dette ble gjort for å bestemme om det var ett PCR produkt og om nukleotidelengden tilsvarte det som var forventet, slik at ikke ufullstendige PCR produkter ble sendt til sekvensering.

2.7.5 Sekvensering og analysering av sekvenseringsdata

Rensingen og sekvenseringen av PCR produktet ble utført av Eurofins MWG Operon, Tyskland.

Analyseringen av sekvenseringsdataene ble utført i gratisprogramvaren Staden – Package, versjon 1.7 (Staden 1996a). I Staden Package ble de rådataene fra de komplementære F og R sekvensene, som var sekvensert, lastet opp som ab1-filer til programmet Pregap4 versjon 1.5. Pregap4 gjorde en automatisk kvalitetskontroll av de komplementære sekvensene ved å klippe bort deler av sporingsdataene (kromatogrammene som ble dannet ved sekvenseringen) som lå under kvalitetsgrensen. Utputt – filene fra Pregap4 ble dannet som AUX - format og ble videre brukt i programmet Gap4. I Gap4 ble det gjort en manuell analyse og bearbeiding av de sammenstilte komplementære sekvensene. Det ble sjekket at konsensussekvensene stemte overens med kromatogrammene for de komplementære sekvensene og at det ikke var innført gap eller baseendringer (Staden 1996b).

Konsensussekvensene fra hvert av isolatene ble sortert inn i txt-filer og omgjort til FASTA – format, én fil for hvert av SBT – genene (1.4), og analysert i programmet MEGA 5 (Tamura et al. 2011). I FASTA-filene ble også referansesekvensen (Vedlegg B, tabell B1) inkludert for hvert av SBT – genene. MEGA 5 er et program som ved hjelp av det innebygde programmet ClustalW gjør det mulig å sammenstille sekvenser. MEGA 5 tillater også bearbeiding av sekvenser ved; innsetting og sletting av baser, og fjerning av uønskede ender. Ved sammenstillingen ble sekvensen orientert etter referansesekvensen (5'-3') og alle forlengninger på konsensussekvensene ble fjernet for å gjøre alle sekvensene like lange. Sekvensene ble lagret i FASTA – format for videre analyse.

Allel – numrene (1.4) ble bestemt fra sekvensene i FASTA – filene ved hjelp av EWGLI sitt ” Check a sequence for its allele number” -verktøy, tilgjengelig på EWGLI sin nettside (The Health Protection Agency 2012). BioNumerics versjon 6.6 (Applied Maths, Ghent, Belgium) ble brukt til å danne en database av *L. pneumophila* isolatene (tabell 2.2.1). BioNumerics er en databaseplattform med ulike verktøy for analyser av biodata. I denne oppgaven ble et MLST-plugin brukt til å gjøre en klyngeanalyse av allel – profilene (Applied Maths 2012). Allel – profilene funnet i denne oppgaven ble sammen med allel – profilene allerede tilgjengelig fra FFI og litteratur (tabell 2.2.1), lastet inn og koblet til isolatene i databasen. Deretter ble det dannet trestrukturer basert på diversiteten i alle allel – profilene. Ikke alle allelprofilene hadde informasjon om *neuA* – genet, derfor ble de fylogenetiske treene dannet med allelnummerene fra *flaA*, *pilE*, *asd*, *mip*, *mompS* og *proA* – genene.

Treene ble dannet med algoritmene; ”Unweighted Pair Group Method with Arithmetic Mean” (UPGMA), “minimum spanning tree” (MST); og “neighbor joining” (NJ). UPGMA er en hierarkisk klustringsmetode hvor det antas en konstant hastighet på evolusjonen. Et MST – tre representerer den korteste avstanden mellom dataen, den korteste evolusjonære veien. I NJ antar minst mulig evolusjon mellom dataen og vises som den totale lengden i treet.

3. Resultat

I denne oppgaven ble det utført tre innledende forsøk (3.1) for å standardisere/optimalisere protokollen for proteinekstraksjon fra *L. pneumophila* for bruk i MALDI-TOF MS -analyse(r). Det ble bygget opp et referansebibliotek i MALDI Biotyper 3.0 bestående av MSP fra til sammen 48 *Legionella* spp. hvorav 41 var *L. pneumophila* -isolater. Referansebiblioteket ble evaluert i en blindtest. En klyngeanalyse ble brukt i sammenligningen mellom typingen med MALDI-TOF MS og SBT -analysen (3.2).

3.1 Innledende forsøk

3.1.1 Evaluering av optimal mengde biologisk materiale

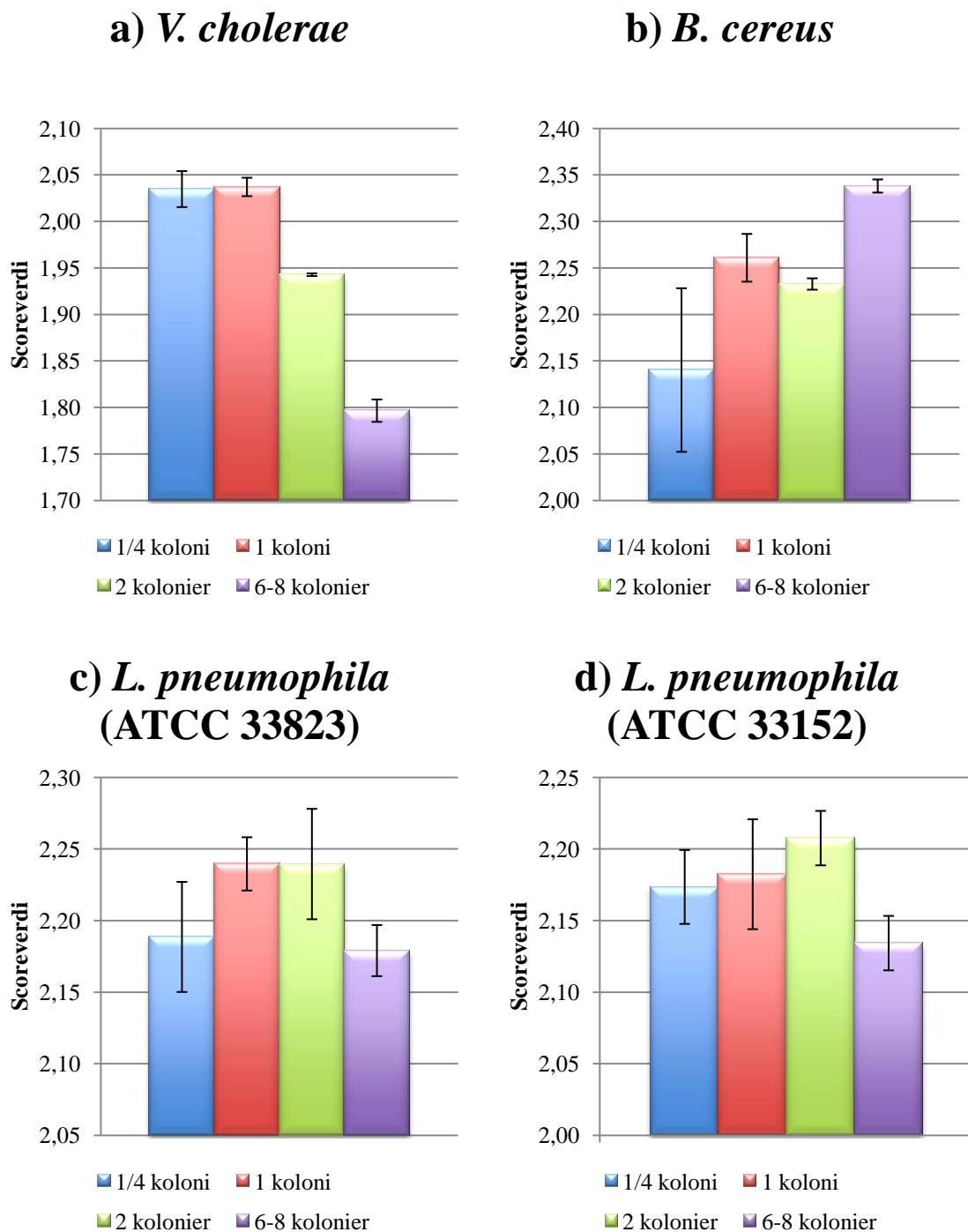
Evalueringen ble utført for å undersøke om mengden biologisk materiale i proteinekstraksjon kunne endre massespektrene som igjen kunne påvirke identifiseringen med MALDI Biotyper RTC (2.5.1). Scoreverdiene som er presentert er de gjennomsnittlige logaritmiske scoreverdiene ved MALDI-TOF MS -identifisering, for de tre proteinekstraksjonene som ble ekstrahert for hvert isolat. Standardavviket illustrerer variasjonen mellom de ulike proteinekstraksjonene utført ved hvert forsøk.

Proteinekstraksjonene fra $\frac{1}{4}$ - 1 koloni av *V. cholerae* (figur 3.1.1 a) resulterte i de høyeste scoreverdiene når isolatene ble sammenlignet mot *V. cholerae* i Bruker Daltonics -referansebibliotek. Variasjonen mellom hver ekstraksjon illustrerte liten forskjell i massespektrene ved de fire volumene som ble undersøkt.

Hos *B. cereus* (figur 3.1.1 b) ble det observert høyest scoreverdi ($>2,3$) for massespektrene dannet med biologisk materiale tilsvarende 6 – 8 kolonier.

Sammenligningen av massespektrene dannet fra *L. pneumophila* ATCC 33823 (figur 3.1.1-c) resulterte i større variasjon sammenlignet med *V. cholerae* og *B. cereus*. 1 – 2 kolonier med biologisk materiale resulterte i de høyeste scoreverdiene sammenlignet med *L. pneumophila* i Bruker Daltonics -referansebibliotek.

Da vi observerte at massespektrene mellom de ulike artene ble påvirket av mengden biologisk materiale ønsket vi å undersøke om dette også gjaldt mellom ulike stammer av *L. pneumophila*. *L. pneumophila* ATCC 33152 (figur 3.1.1 d) viste høyest scoreverdi for massespektrene dannet ved proteinekstraksjon med biologisk materiale tilsvarende 1 – 2 kolonier.

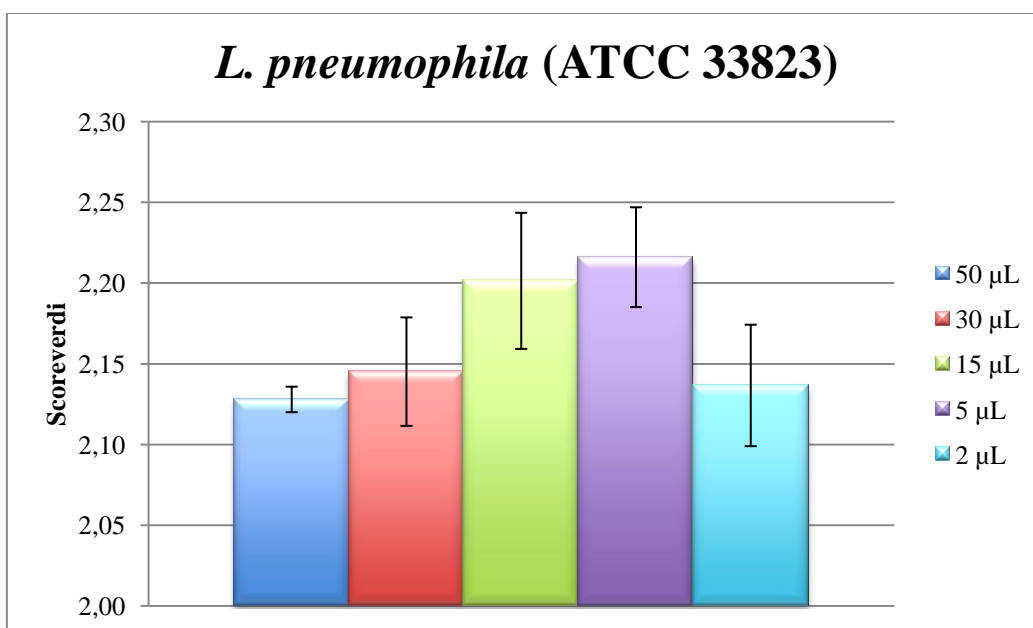


Figur 3.1.1: Scoreverdiene som funksjon av mengden biologisk materiale. a) *V. cholerae* b) *B. cereus* c) *L. pneumophila* (ATCC 33823) d) *L. pneumophila* (ATCC 33152). Usikkerhetene er angitt som standardavviket av tre proteinekstraksjoner.

3.1.2 Evaluering av optimalt volum maursyre og acetonitril

Evalueringen ble utført for å undersøke om ulike volum av maursyre og acetonitril ville endre massespektrene som igjen kunne påvirke identifiseringen med MALDI Biotyper RTC (2.5.2).

Forsøket ble utført med 2 – 50 µl av hver reagens på *L. pneumophila* ATCC 33823 (figur 3.1.2). Proteinekstraksjonen med 5 – 15 µl av maursyre og acetonitril ga massespektre med høyest scoreverdi mot *L. pneumophila* i Bruker Daltonics - referansebibliotek. Variasjonen mellom scoreverdiene (<0,1) viste en liten forskjell mellom massespektrene.



Figur 3.1.2: Scoreverdi som funksjon av tilsatt maursyre/acetonitril -volum i proteinekstraksjonsprosedyren. *L. pneumophila* (ATCC 33823). Forsøk utført med 2, 5, 15, 30 og 50 µl av hver reagens.

3.1.3 Stabilitet av massespekter ved lagring av proteinekstrakter og kolonier

Evalueringen ble utført for å undersøke om lagring av kolonier på skål og lagring av proteinekstrakter ville endre massespektrene som igjen kunne påvirke identifiseringen med MALDI Biotyper RTC (2.5.3). Forsøket ble utført med proteinekstraksjonsprotokollen (2.3.1) og direkteoverførings – metoden (2.3.2) for å undersøke om de to metodene påvirket massespektrene forskjellig.

For *V.cholerae* (figur 3.1.3 a) resulterte direkteoverførings – metoden i de høyeste scoreverdiene for massespektrene i sammenligningen med *V.cholerae* i Bruker Daltonics referansebibliotek. Scoreverdiene som ble oppnådd viste at massespektrene fra koloniene dyrket ved 20 °C (merket; 20 °C direkte, i figur 3.1.3 a) ga stabile massespekter med en scoreverdi >2,2 frem til dag tre ved dyrkingen. Massespektrene fra koloniene dyrket ved 37 °C (merket; 37 °C direkte, i figur 3.1.3 a) sank jevnt i scoreverdi fra >2,2 på dag én til <1.3 på dag åtte. Variasjonen mellom overføringene per dag illustrerte en usikkerhet i massespektrene, når direkte metoden ble brukt. Usikkerheten i massespektrene ble større etter hvert som koloniene ble eldre.

Massespektrene fra proteinekstraksjonen dannet fra koloniene dyrket ved 20 °C hos *V.cholerae* (merket; 20 °C ekstraksjon, i figur 3.1.3 a) resulterte i en scoreverdi >2,0 frem til dag tre. Stabiliteten til massespektrene fra proteinekstraksjon dannet fra koloniene dyrket ved 37 °C (merket; 37 °C ekstraksjon, i figur 3.1.3 a) resulterte i en scoreverdi >2,0 frem til dag to, og sank til en scoreverdi på 1,5 etter tre dager lagret på skål.

De lagrede proteinekstraktene (merket; 37 °C lagret og 20 °C lagret, i figur 3.1.3 a) resulterte begge i massespekter med scoreverdi >2,0 etter én dags lagring, og stabiliserte seg på en scoreverdi rundt 1,8 etter fem dagers lagring.

Hos *B. cereus* resulterte massespektrene fra proteinekstraksjonen fra kolonier dyrket ved 20 °C (merket; 20 °C ekstraksjon, i figur 3.1.3 b) i en scoreverdi på >2,2 frem til dag fire. Massespektrene fra proteinekstraksjon fra koloniene dyrket ved 30 °C (merket; 30 °C ekstraksjon, i figur 3.1.3 b) oppnådde høyest scoreverdi (>2,2) etter at koloniene var lagret i to dager og sank jevnt i scoreverdi til rundt 1,9 ved dag ni.

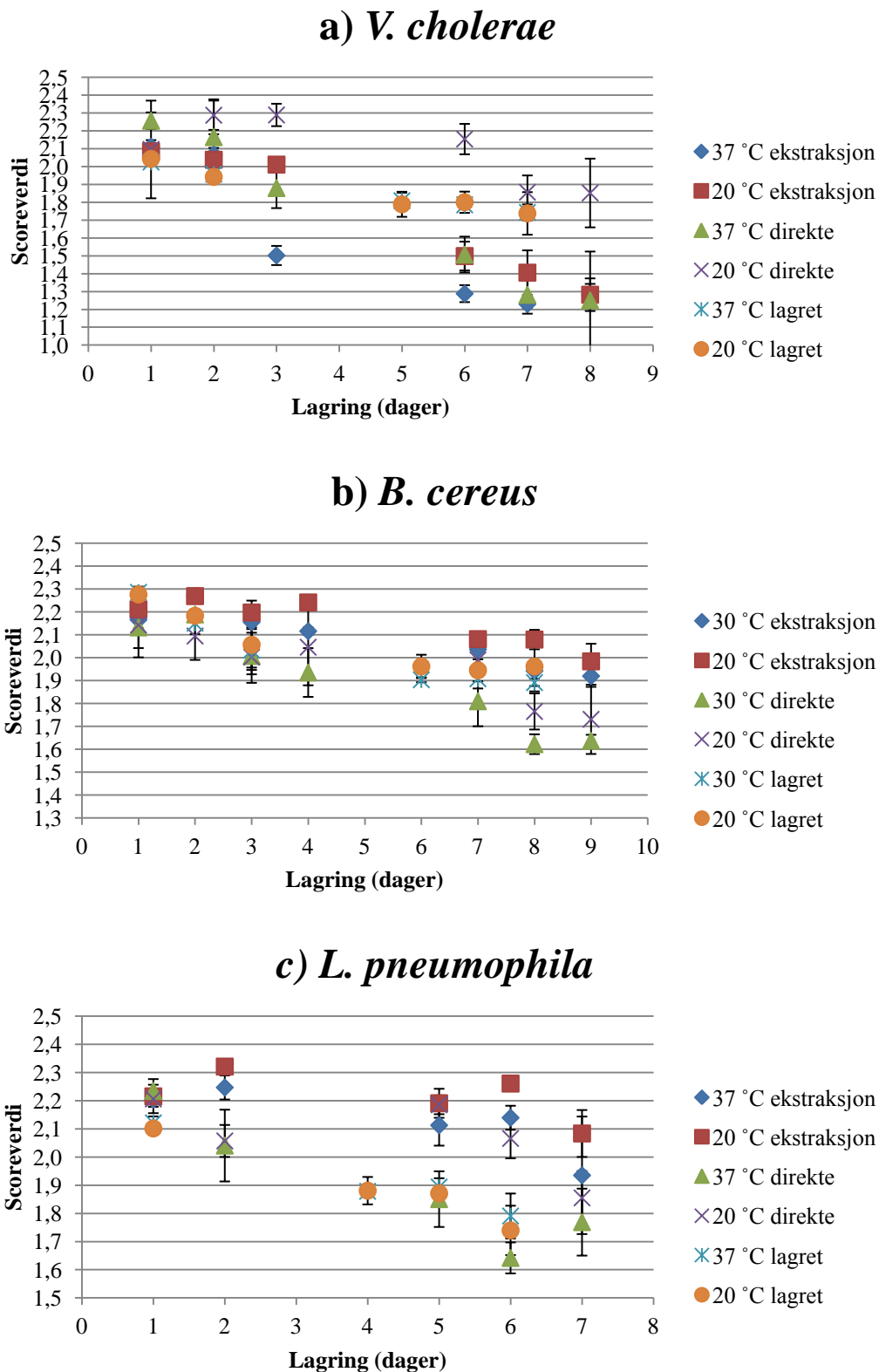
De direkteoverførte koloniene (merket; 20 °C direkte og 30 °C direkte, i figur 3.1.3 b) resulterte i massespektre med en scoreverdi rundt 2,1 frem til dag to og sank jevnt i scoreverdi (<1,8) frem til dag ni.

Massespektrene fra de lagrede proteinekstraktene (merket; 20 °C lagret og 30 °C lagret, i figur 3.1.3 b) oppnådde høyest scoreverdi (>2,2) ved dag én, scoreverdien var høyere enn den dagen ekstraksjonen fant sted (merket; 20 °C ekstraksjon og 30 °C ekstraksjon, dag én, i figur 3.1.3 b). Massespektrene sank deretter jevnt ned til en scoreverdi på rundt 1,9 og stabiliserte seg ved dag seks.

Fra *L. pneumophila* ATCC 33823 (figur 3.1.3 c), oppnådde massespektrene høyest scoreverdi (>2,2) fra proteinekstraksjon dannet fra kolonier lagret på skål i to dager (merket; 20 °C ekstraksjon og 37 °C ekstraksjon, i figur 3.1.3 c). Massespektrene fra koloniene dyrket ved 20 °C hadde en scoreverdi >2,3 ved to dagers lagring og var stabile med en scoreverdi >2,2 etter seks dagers lagring.

Massespektrene fra koloniene dyrket ved 20 °C som var direkteoverført (merket; 20 °C direkte, i figur 3.1.3 c) oppnådde den høyeste scoreverdien fra koloniene dyrket ved dag én og hadde en scoreverdi like over 2,0 frem til dag seks. Massespektrene fra koloniene dyrket ved 37 °C og var overført direkte (merket; 27 °C direkte, i figur 3.1.3 c) sank jevnt i scoreverdi fra >2,2 på dag én til <1,9 på dag fem. Variasjonen mellom parallellen per dag økte jevnt etter hvert som koloniene ble eldre.

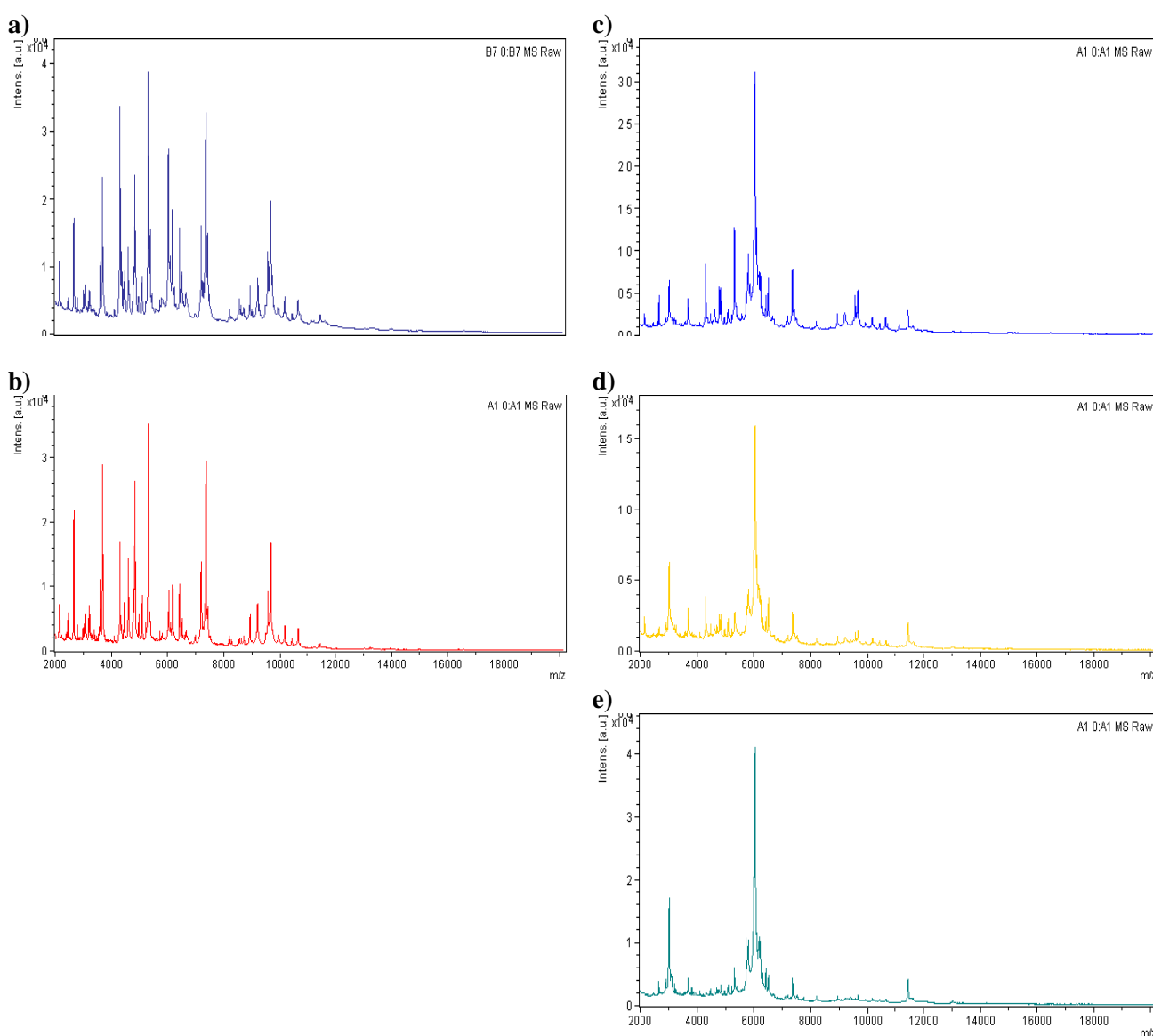
De lagrede proteinekstraktene (merket; 20 °C lagret og 37 °C lagret, i figur 3.1.3 c) fra *L. pneumophila* oppnådde massespektre med scoreverdi ved 2,1 etter én dags lagring og sank jevnt i scoreverdi til <1,8 etter seks dagers lagring.



Figur 3.1.3: Scoreverdi som funksjon av lagringstid av bakterieceller på skål/proteinekstrakter lagret i kjøleskap. a) *V. cholerae*; b) *B. cereus*; c) *L. pneumophila* (ATCC 33823). Variansen er oppgitt med standardavvik, n=3. De lagrede proteinekstraktene er fra proteinekstraksjonene dannet fra kolonier på dag én og lagret ved 4°C.

For å illustrere tapet av informasjon i massespektrene ved lagring, ble massespektrene som ble dannet fra proteinekstraksjon av kolonier fra *L. pneumophila*, sammenlignet. Massespektrene er hentet fra koloniene dyrket ved 37 °C (merket; 37 °C ekstraksjon, i figur 3.1.3 c) og er representert med ett massespekter per dag.

Massespekteret med den høyeste scoreverdien mot Bruker Daltonics referansebibliotek fra kolonier dyrket i to dager (figur 3.1.4 b) dannet tydelige definerte massetopper, slik det er illustrert. Fra de eldre koloniene (figur 3.1.4 c-e) blir det observert tydelig intensitetsforskjeller og forskyvninger i massetoppene i forhold til de massespektrene fra én og to dager gamle kolonier (figur 3.1.4 a-b).



Figur 3.1.4: Endringer i massespekter fra proteinekstraksjon av kolonier lagret på skål ved 37 °C. *L. pneumophila* (ATCC 33823): a) Dag 1; b) Dag 2; c) Dag 5; d) Dag 6; e) Dag 7.

3.2 Typing av *L. pneumophila*

Formålet med oppgaven var å danne et referansebibliotek bestående av MSP som representerte proteinprofilene til *L. pneumophila* -isolatene (tabell 2.2.1). Hvor referansebiblioteket kunne bli benyttet til å gjøre en klyngeanalyse og sammenlignet med en klyngeanalyse basert på SBT.

Hvert MSP ble i denne oppgaven dannet fra mellom 28 og 36 massespekter (tabell 2.2.1) avhengig av kvalitetskravet på 500 ppm i variasjon rundt hver massetopp i massespektrene (2.6). Et fylogenetisk tre (vedlegg C, figur C1) av alle *Legionella* – isolatene det ble dannet MSP for illustrer hvordan *L. pneumophila* – isolatene klynger seg sammen og er fylogenetisk forskjellig fra andre *Legionella* -arter. I det fylogenetiske treet ble også 11 *L. pneumophila* -MSP (Vedlegg A, tabell A1) fra Bruker Daltonics referansebibliotek inkludert og viser at de klynger seg sammen med de andre *L. pneumophila* -MSP dannet i denne oppgaven.

3.2.1 Blindtest -identifisering

For å undersøke om det var mulig å gjøre en korrekt identifisering med referansebiblioteket som ble utarbeidet gjennom denne oppgaven, ble det utført en blind – identifisering (2.6.2). 14 av isolatene som ble brukt til dannelsen av MSP til referansebiblioteket ble tilfeldig valgt ut og blindet, og identifisert mot referansebiblioteket.

Ved en identifisering av et massespekter i MALDI Biotyper 3.0, genereres det en rapport over de ti beste sammenligningene (1 til 10) mot de MSP som er tilstedet i referansebiblioteket (1.3.3). Massespektrene som ble identifisert med beste sammenligning (høyest scoreverdi) mot et MSP som representerte sitt eget isolat i referansebiblioteket, ble telt. Massespekter identifisert som andre til tiende beste sammenligning mot et MSP som representerte eget isolat i rapportene, ble på samme måte telt. Scoreverdiene som er presentert er de gjennomsnittlige logaritmiske scoreverdiene ved MALDI–TOF MS -identifisering, for de høyeste scoreverdiene gitt ved ni identifiseringer per isolat. Standardavviket viser variasjonen mellom høyeste scoreverdi fra ni identifiseringer per isolat.

Massespektrene fra isolatene i blindtesten (tabell 3.2.1) oppnådde alle, med unntak av massespektrene fra *L. pneumophila* ATCC 33216 (scoreverdi; 2,359) og ATCC 43130 (scoreverdi; 2,280), en gjennomsnittlig best scoreverdi >2,4. Variasjonen mellom massespektrene i de ni identifiseringene per isolat varierte i scoreverdi mellom 0,052 og 0,152 for de ulike isolatene i blindtesten.

Av de ni massespektrene som ble identifisert for hvert isolat ble *L. pneumophila* ATCC 33216 og ATCC 33823 identifisert med eget MSP henholdsvis syv og fem ganger med høyest scoreverdi mot eget MSP i referansebiblioteket.

Massespektrene fra ATCC 35096 – isolatet (scoreverdi; 2,423 i tabell 3.2.1) ble ikke identifisert med eget MSP i referansebiblioteket en eneste gang, heller ikke i andre til tiende beste sammenligning.

Massespektrene som ble dannet i blindtesten ble benyttet til å konstruere nye MSP for å undersøke om det ville føre til en bedre identifisering mot referansebiblioteket. Identifiseringen viser, med unntak for ATCC 43130, en økning i scoreverdi sett i forhold til den gjennomsnittlige scoreverdien ved identifisering mot massespektrene. Kun MSP dannet fra *L. pneumophila* ATCC 33216 og B2760 ble korrekt klassifisert med sitt MSP i referansebiblioteket.

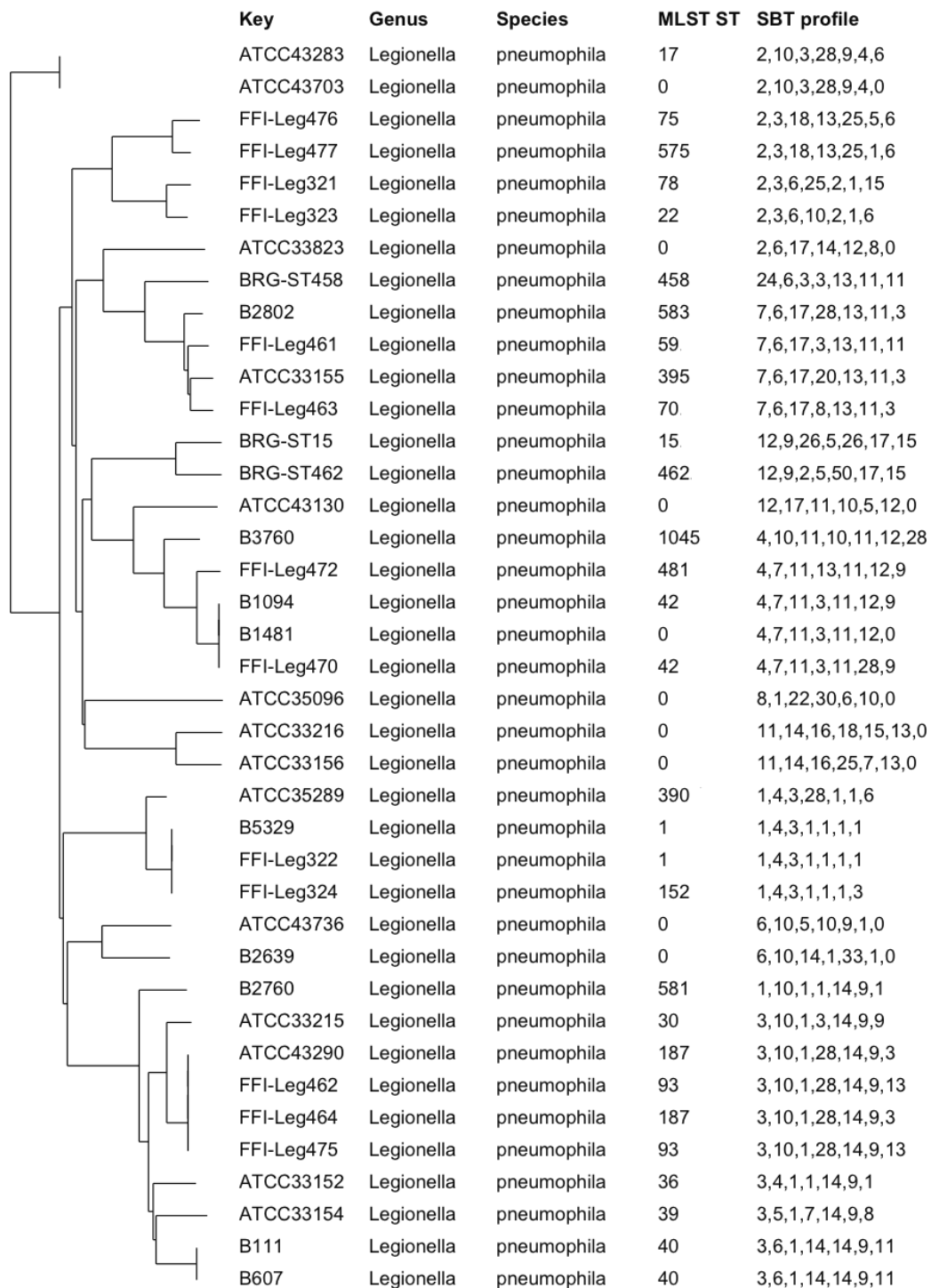
Tabell 3.2.1: MALDI–TOF MS -identifisering av 14 *L. pneumophila* -isolater ved blindtest.

<i>L. pneumophila</i> -isolat	Massespekter identifisert med MSP fra eget isolat som identifisering nr 1 av 10 (n=9)	Massespekter identifisert med MSP fra eget isolat som identifisering 2 til 10 (n=9)	Gjennomsnitt -scoreverdi av beste identifisering mot referanse - biblioteket	Standard – avvik av scoreverdi *	MSP fra massespektre identifisert med MSP fra eget isolat som identifisering nr 1 av 10 (n=1)	MSP fra massespektre identifisert med MSP fra eget isolat som identifisering 2 til 10 (n=1)	Scoreverdi for beste identifisering med MSP mot referanse - bibliotek
ATCC 35096	0	0	2,423	0,073	0	0	2.553
ATCC 33823	5	4	2,435	0,080	0	1	2.511
ATCC 33215	3	6	2,475	0,044	0	1	2.633
ATCC 33216	7	2	2,359	0,066	1	0	2.487
ATCC 43130	1	5	2,280	0,131	0	1	2.237
ATCC 33152	1	7	2,466	0,059	0	1	2.623
ATCC 33155	1	5	2,410	0,069	0	0	2.552
B111	1	3	2,404	0,152	0	1	2.574
B607	0	2	2,349	0,055	0	0	2.355
B1094	1	7	2,415	0,103	0	1	2.490
B2760	3	5	2,428	0,052	1	0	2.603
B2802	0	2	2,428	0,071	0	0	2.479
B3760	1	5	2,380	0,056	0	1	2.418
B5329	0	3	2,444	0,110	0	0	2.514

*Standardavviket er vist som variasjonen mellom høyeste scoreverdiene gitt ved ni identifiseringer

3.2.2 Fylogenetisk sammenligning av SBT og MALDI–TOF MS

Fra referansebiblioteket som ble utarbeidet i denne oppgaven ble det gjort en klyngeanalyse (2.6.3) av MSP - dataene ved å konstruere et fylogenetisk tre i MALDI Biotyper 3.0. Klyngeanalysen ble sammenlignet med klyngeanalysen dannet fra et fylogenetisk tre konstruert basert på allelprofilene, analysert med SBT ved å benytte BioNumerics (2.7.5). For *L. pneumophila* ATCC 33216, ATCC 33823, ATCC 35096, ATCC 43130, ATCC 43703, B1481 og B2639 resulterte PCR analysen (i SBT – analysen) ikke i amplifisering av *neuA* – gen. I henhold til EWGLI, blir disse allelene angitt med allelnummer 0 (1.4). For å oppnå en god genetisk sammenligning hvor alle isolatene har resultert i amplifisering av alle gener, ble den fylogenetiske analysen i BioNumerics derfor gjennomført for kun seks gener (*flaA*, *pilE*, *asd*, *mip*, *mompS* og *proA*). Allelprofilene som ble funnet (markert; Denne oppgaven, figur 2.2.1) resulterte i stabile fylogenetiske trær ved et NJ – tre (figur 3.2.1), UPGMA – tre (3.2.2 b) og MST – tre (Vedlegg D, figur D1).



Figur 3.2.1: NJ – tre av *L. pneumophila* -allelprofiler

Resultatene fra den fylogenetiske sammenligningen viser at *L. pneumophila* – isolatene klynger seg sammen ulikt ved bruk av MALDI–TOF MS og SBT. SBT – analysen (figur 3.2.2 a) resulterte i dannelsen av fire større klynger (A – D, figur 3.2.2 a). Isolater med identisk allelprofil er presentert med en lodderett strek foran stamnavnene i

trestrukturen. Fra trestrukturen konstruert fra MALDI-TOF MS MSP – data (figur 3.2.1 b), ble det dannet to hovedgrupper illustrert med blå og røde forgreninger. Disse kan videre deles i fem klynger, angitt som A – E. Klyngene i de to treene viser enkelte likhetstrekk mellom hverandre etter hvordan isolatene er fordelt, men samtidig noen tydelige hovedforskjeller.

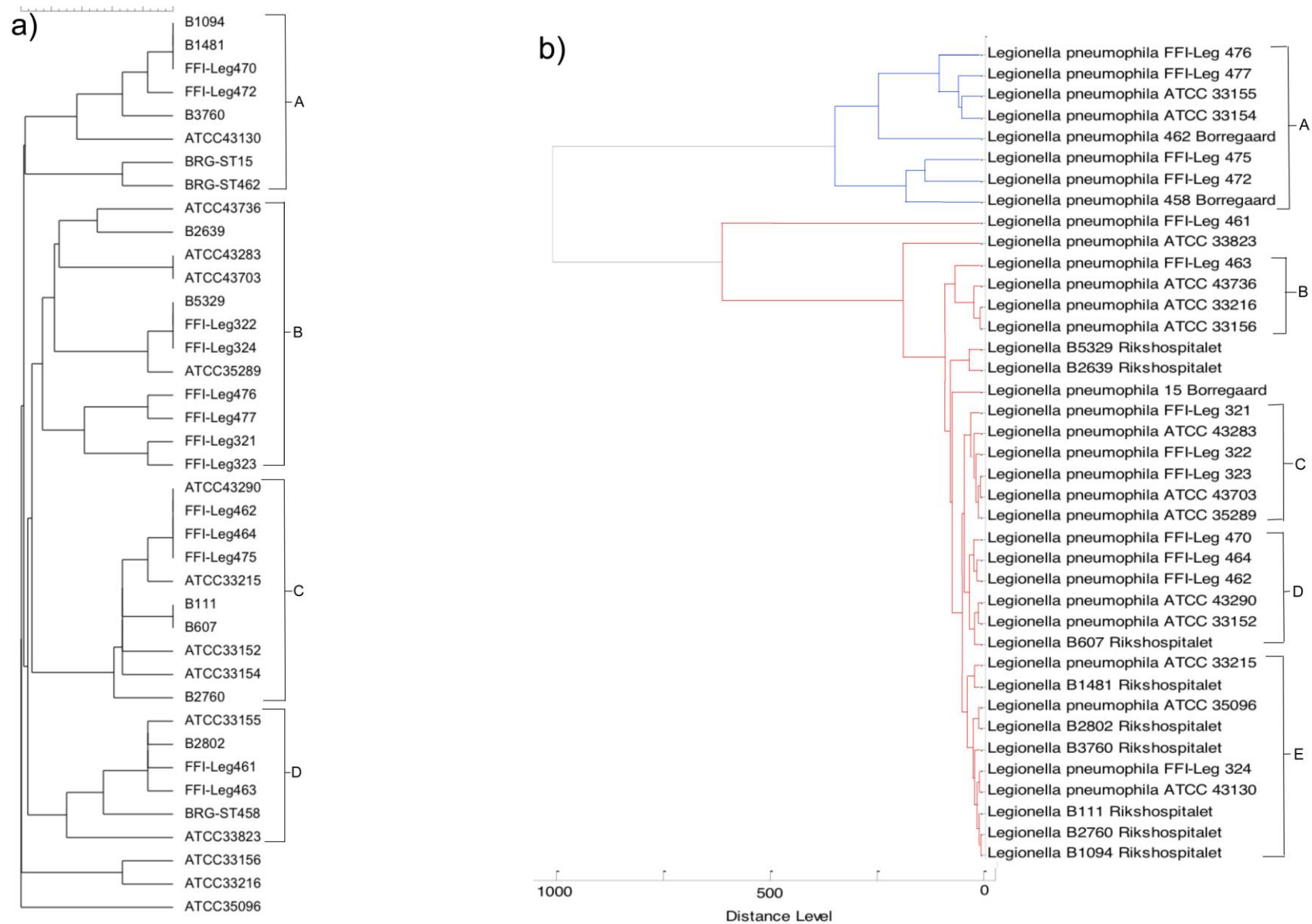
L. pneumophila stammene; FFI-Leg321, ATCC 43823, FFI-Leg322, FFI-Leg323, ATCC 43703 og ATCC 35289 fra klynge C i MSP - treet (figur 3.2.2 b), er alle representert i klynge B i UPGMA - treet (figur 3.2.2 a).

Alle stammene i klynge D i MSP - treet (figur 3.2.2 b); FFI-Leg464, FFI-Leg462, ATCC 43290, ATCC 33152 og B607, med unntak av FFI-Leg470, er representert i klynge C i UPGMA - treet (figur 3.2.2 a). I tillegg er ATCC 33215, B111 og B2760 i klynge E i MSP - treet, avstandsmessig ikke langt fra klynge D (figur 3.2.2 b), som også representert i klynge C i UPGMA - treet (figur 3.2.2 a).

Stammene ATCC 33156, ATCC 33216 og ATCC 35096 ligger i SBT –treet (figur 3.2.2 a) som utligger og roter treet. I MSP - treet (figur 3.2.2 b) har ATCC 33216 og ATCC 33156 gruppert seg sammen i klynge B, men roter ikke treet på samme måte. ATCC 35096 har i MSP - treet (figur 3.2.2 b) lagt seg i klynge E.

Den tydelige inndelingen i rød og blå forgrening i MSP - treet (figur 3.2.2 b) er ikke representert i SBT – treet (figur 3.2.2 a). Stammene i blå forgrening (klynge A, figur 3.2.2 b) er fordelt mellom klynge A – D i SBT – treet (figur 3.2.2 a)

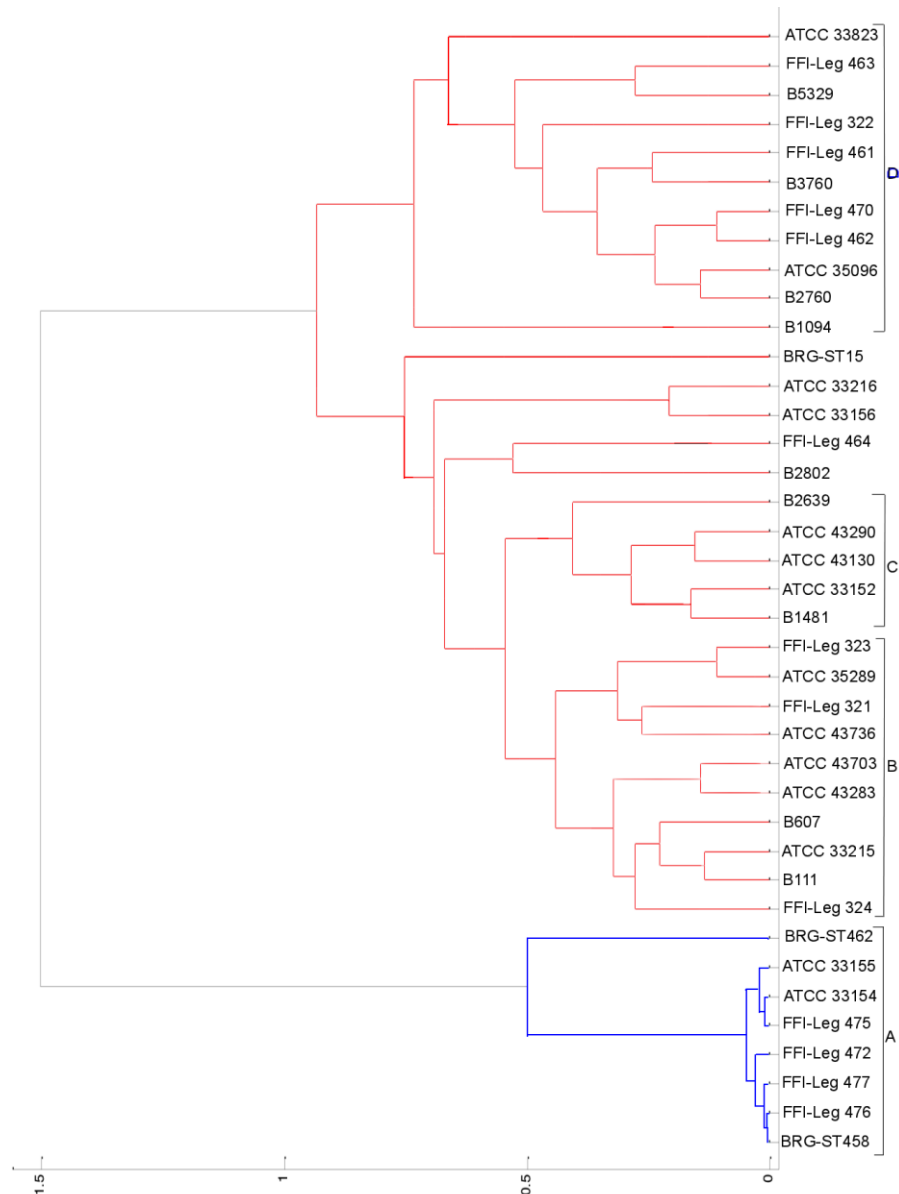
Av isolatene med samme allelprofil fra SBT – analysen (3.2.2 a) var det litt forskjellig hvor nært hverandre de hadde gruppert seg i MSP – treet (3.2.2 b). B111 og B607 lå avstandsmessig nært hverandre (figur 3.2.2 b), men i hver sin klynge, i henholdsvis E og D. ATCC 43290, FFI-Leg464 og FFI-Leg464 lå tett inntil hverandre i klynge D av MSP – treet (figur 3.2.2 b) mens FFI-Leg475 med den samme allelprofil fra SBT – analysen hadde lagt seg i klynge A av MSP - treet. B5329 og FFI-Leg322 ligger begge i rød forgrening, men ikke i samme klynge (figur 3.2.2 b). FFI-Leg470, B1481 og B1094 ligger avstandsmessig kort fra hverandre i klynge E og D (figur 3.2.2 b).



Figur 3.2.2: Klyngeanalyser av *L. pneumophila* – isolater. a) UPGMA – tre av SBT allelprofiler; b) Scoringsorientert tre fra MALDI-TOF MS MSP. De tydeligste klyngene markert (A – D (E)) illustrerer grupperingene som dannet seg.

Det ble utført PCA – analyse (2.6.3) av massespektrene fra *L. pneumophila* i tillegg til analysen basert på MSP, ved bruk av MALDI Biotyper 3.0 programvaren. Det var ønskelig å undersøke om de resulterende trestrukturen fra de to analysen var sammenfallende. PCA – analysen ble utført på ett massespekter fra hver av isolatene som er representert i referansebiblioteket.

Trestrukturene utledet fra PCA analysen (figur 3.2.3) og MSP – dataene (figur 3.2.2 b) resulterte i dannelsen av en identisk klynge, A med *L. pneumophila* isolatene; BRG-ST462, ATCC 33155, ATCC 33154, FFI-Leg475, FFI-Leg472, FFI-Leg477, FFI-Leg476 og BRG-ST458. *L. pneumophila* isolatene i klyngene B til D viste ikke noen tydelige likheter i forhold til mellom PCA trestrukturen (figur 3.2.3) og MSP trestrukturen (figur 3.2.2 b), eller mellom PCA trestrukturen og SBT – trestrukturen (figur 3.2.2 a).



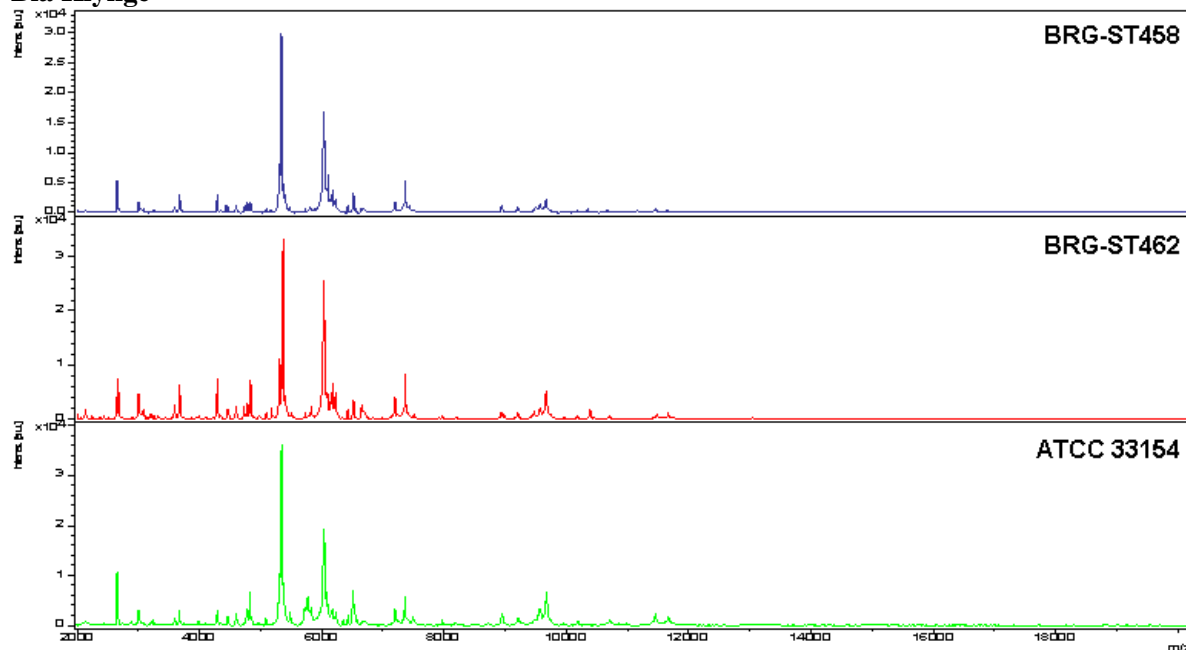
Figur 3.2.3: PCA tre fra MALDI-TOF MS massespekter. Hvert *L. pneumophila* –isolat er representert med ett massespekter. De tydeligste klyngene er vist som A – D.

Både MSP trestrukturen (figur 3.2.2 b) og PCA trestrukturen (figur 3.2.3) viste tydelig inndeling i to overordnede forgreninger, markert med blå og røde grener. For å undersøke om det var tydlige forskjeller mellom massespektrene i de to forgreningene, ble tre massespekter fra hver forgrening sammenlignet (2.6.3).

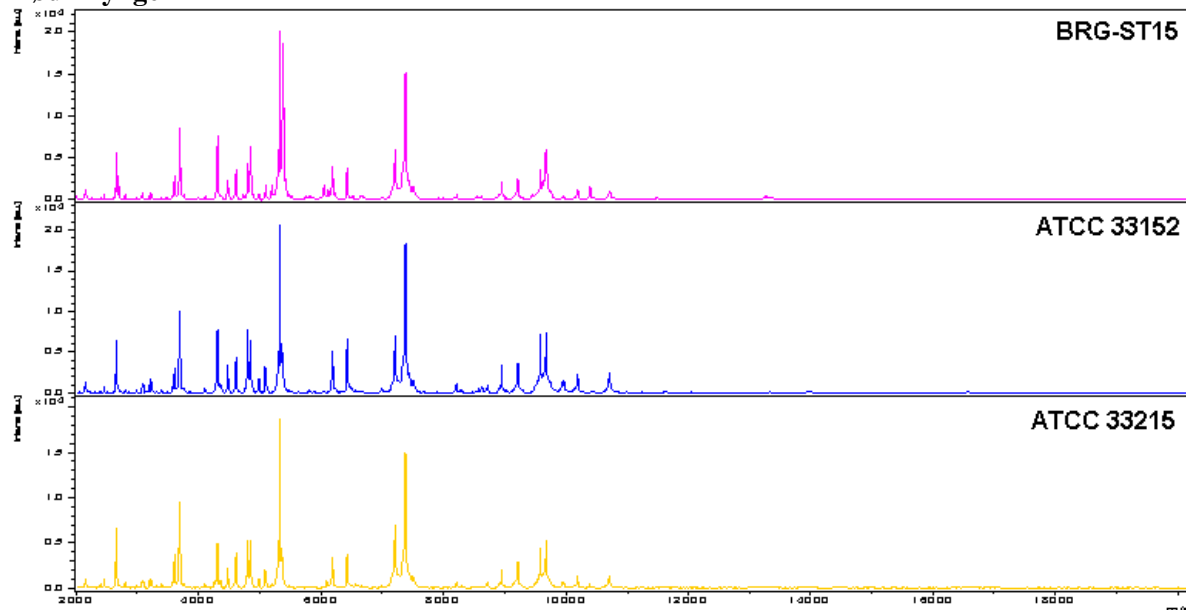
Sammenligningen (figur 3.2.4) viste at massespekterene fra isolatene i den blå forgreningen er like i forhold til hverandre, observert fra de mest tydelige massetoppene, med en liten variasjon i intensitet. Det samme kan observeres mellom

massespektrene fra isolatene i den røde forgreningen. Mellom massespektrene fra den blå og røde forgreningen er det en tydelig forskjell. Mellom henholdsvis 4000 og 6000 m/z og mellom 8500 og 11000 m/z, viste massespektrene fra rød forgreninger flere klare massetopper i massespektrene i forhold til massespektrene fra blå forgrening.

Blå Klynge



Rød Klynge



Figur 3.2.4: Sammenligning mellom massespekter fra rød og blå forgrening, observert i figur 3.2.2 b og 3.2.3. Massetoppene er vist mellom 2000 og 20000 m/z

4. Diskusjon

4.1 Innledende forsøk

I denne oppgaven ble det utført tre innledende forsøk for å tilpasse/optimalisere protokollen for proteinekstraksjon (2.3.1) fra *L. pneumophila* for bruk i MALDI-TOF MS -analyse(r). Hensikten med disse forsøkene var å finne parametrene som ga høyest mulig scoreverdien i en identifisering mot Bruker Daltonics referansebibliotek. Resultatene fra de inneledende forsøkene ble benyttet videre i arbeid med *L. pneumophila*, hvor vi ønsket å bygge opp et eget referansebibliotek av MSP fra ulike stammer av *L. pneumophila*.

Resultatene som er presentert i de innledende forsøkene (3.1) er alle basert på scoreverdien oppnådd ved MALDI-TOF MS -identifisering. En høyere scoreverdi vil vise til en bedre sammenligning mellom massespekter og Bruker Daltonics referansebibliotek (1.3.1). Vurderingene som har blitt gjort vil derfor reflektere en sammenligning med MSP fra Bruker Daltonics referansebibliotek og metoden de har benyttet i preparering av prøver, basert på scoreverdiene.

Variasjonene i scoreverdi funnet (figur 3.3.1, a-d) ved proteinekstraksjon fra ulike mengder biologisk materiale fra *V. cholerae*, *B. cereus* og *L. pneumophila* ATCC 33823 og ATCC 33152, indikerte at både for mye eller for lite materiale kan påvirke massespektrene og dermed klassifiseringen. Dette har også blitt observert tidligere (Drevinek et al. 2012). Det var derfor viktig å komme frem til en standardisert mengde biologisk materiale for å oppnå optimale scoreverdier når massespektrene fra kjente referansmaterialer (*V. cholerae*, *B. cereus* og *L. pneumophila* ATCC 33823 og ATCC 33152) ble sammenlignet mot MSP i Bruker Daltonics -referansebibliotek. For ekstraksjonsprotokollen som ble benyttet videre i denne oppgaven ble volumet bestemt til å være ca en koloni (ca 3 mm i diameter) tilsvarende ca 2 µl. Dette ble bestemt med bakgrunn i scoreverdiene som ble oppnådd for de to *L. pneumophila* (ATCC 33823 og ATCC 33152) isolatene. Resultatene for *B. cereus* (figur 3.3.1, b) indikerer imidlertid at mengden biologisk materiale som er optimalt for en organisme (for eksempel *L. pneumophila*) ikke nødvendigvis er optimalt for andre

mikroorganismer. For andre mikroorganismer bør det derfor vurderes å utlede et lignende forsøk for å finne den optimale mengden biologisk materiale, for å oppnå optimal scoreverdi i forhold til Bruker Daltonics -referansebibliotek. MSP i referansebiblioteket til Bruker Daltonics kan være utviklet av ulike personer og/eller med ulike metoder/prosedyrer som har variert over tid i forbindelse med oppbyggingen av biblioteket.

Volumet av maursyre og acetonitril ble også bestemt i forhold til det biologiske materialet. Massespektrene ble på samme måte som nevnt ovenfor noe påvirket ved bruk av ulikt volum av maursyre og acetonitril, noe som er naturlig i forhold til det biologiske materiale. Ut ifra resultatene ble det bestemt at volumet skulle tilsvare 10 µl fra hver reagens. Ved utviklingen av referansebiblioteket med massespektre for *L. pneumophila* skulle fire overføringer fra hver proteinekstraksjon (2.6) overføres til targetplaten. Et volum på totalt 10 µl av maursyre og acetonitril (som ga den høyeste scoreverdi) ble derfor vurdert til å være for lavt, om det skulle være behov for tilleggsanalyser.

Som det er nevnt i 1.3.4, vil lengde på dyrkningstiden påvirke massespektrene. Ut ifra resultatene (figur 3.1.3-c) og observasjoner gjort ved dyrking, ble det bestemt å gjøre en proteinekstraksjon etter to dagers (48 ± 2 timer) dyrking ved 37 °C for *Legionella* – isolatene. Dette til tross for at resultatene viser at en proteinekstraksjon etter to dagers dyrking ved 20 °C ga høyest scoreverdi for ATCC 33823. Grunnlaget for valget er at ikke alle isolatene som det var ønskelig å benytte i dannelsen av MSP – biblioteket, ville vokse ved 20 °C. Av de isolatene som vokste opp ved 20 °C etter to dager var det også en tydelig forskjell på hvor mye bakteriene hadde vokst, ut ifra morfologien på koloniene. Den mangelfulle dataen på to av dagene skyldes at det var helg, men det ble vurdert til at det var tilstrekkelig med informasjon i datasettet til å se en trend i forhold til stabiliteten som ble undersøkt. Den gradvise nedgangen i scoreverdi sett fra de lagrede proteinekstraktene skyldes endringer i proteinsammensetningen i bakteriene. Maursyre er en sterk organisk syre og bidrar til å løse opp celleveggen under ekstraksjonen. Acetonitrilen bidrar til ekstraksjonen av proteinene samtidig som den senker reaksjonshastigheten til maursyren, men stopper den ikke (Søren Lehmann, personlig meddelelse). Direktemetoden som ble brukt viste seg å gi best scoreverdi mot *V. cholerae*. Dette kan komme av at MSP for *V. cholerae* (totalt 3) som lå i sikkerhetsdatabasen (1.3.4) var dannet med en TFA ekstraksjonsmetode. Fra

direktemetoden ble det sett større variasjon mellom massespektrene hos *V. cholerae*, *B. cereus* og *L. pneumophila* ATCC 33823 -isolatene fra de ulike parallellene på targetplaten. Noe som bekrefter vanskeligheten med å danne et homogent tynt lag av biologisk materiale på targetplaten (2.3.2).

For å bekrefte at massespektrene faktisk endret karakter i form av massetopper etter hvert som koloniene på agarplaten ble eldre, ble massespektrene fra hver dag av ekstraksjon ved 37 °C sammenlignet. Intensiteten i mange av massetoppene synker kraftig. Noe som antagelig kan forklares som en nedgang i blant annet ribosomale proteiner (1.3.4), tilgangen til næringsstoffer vil begrenses og endre proteinsammensetningen i mikroorganismene (Alispahic et al. 2011). Dette bekrefter hvor viktig standardisert dyrkningstid/lagring er for å danne reproducerbare massespekter. Basert på resultatene i oppgaven er det vanskelig å si akkurat når proteinsammensetningen til *L. pneumophila* begynner å forandre seg, på grunn av mangelfull data på dag tre og fire. *L. pneumophila* -koloniene som ble lagret på skål ved 37 °C viste seg å endre karakter i form av proteinsammensetning. Mens resultatene fra kolonien lagret ved 20 °C på skål (figur 3.1.3 c), viser at massespektrene til og med dag seks gir scoreverdier på >2,2. Lagring av *L. pneumophila* ved 20 °C i opp til seks dager ser derfor ikke ut til å påvirke massespektrene for å oppnå en god identifisering mot Bruker Daltonics referansebibliotek.

4.2 Evaluering av referansebibliotek

Blindtesten i denne oppgaven ble utført for å undersøke om det var mulig å gjøre en korrekt identifisering med referansebiblioteket som ble bygget opp. Fra resultatene ble det sett at massespektrene som ble dannet fra *L. pneumophila* -isolatene, oppnådde, med unntak av ATCC 43130 (tabell 3.2.1), en gjennomsnitt scoreverdi over 2,3, som av Bruker Daltonics blir betegnet som høy sannsynlighet for korrekt artsidentifisering (1.3.3). Massespektrene som ble dannet fra hvert *L. pneumophila* -isolat viser med dette en god sammenligning med referansebiblioteket. Tidligere studier har også vist at *Legionella* på artsnivå kan identifiseres ved bruk av MALDI-TOF MS (Moliner et al. 2010).

Ut ifra resultatene i blindtesten (3.2.2) ble det også sett at isolatene, med unntak av ATCC 33216 og ATCC 33823, i liten grad ble identifisert med sitt eget MSP fra referansebiblioteket, ved høyeste scoreverdi gitt ved hver identifisering. Derimot var MSP fra isolatene ofte representert blant de ti beste sammenligningene som ble angitt i rapporten (1.3.3) ved identifiseringen. Dette kan være en indikasjon på at proteinprofilene til mange av *L. pneumophila* – isolatene benyttet i oppgaven er veldig like hverandre i proteinsammensetning og derfor vanskelig å skille mellom ved en identifisering. Avstandsnivået i det fylogenetiske treet (3.2.2 b) dannet fra MSP – data, bidrar til å støtte denne indikasjonen. I det fylogenetiske treet er avstandsnivået normalisert i forhold til de MSP som bygger det opp og viser, spesielt i den røde forgreningen, kort avstand mellom hver MSP. *L. pneumophila* ATCC 33216 og ATCC 33283 oppnådde i blindtesten (3.2.1) best treff mot eget MSP (henholdsvis 7/9 og 5/9 ganger) viser et større avstandsnivå fra de andre MSP i treet (figur 3.2.2 b) og kan forklare hvorfor de skiller seg ut ved en identifisering.

Blindtesten viste også at ved å danne MSP fra massespektrene av isolatene som ble undersøkt, ville scoreverdien øke i identifiseringen mot MSP – biblioteket. Identifiseringen av MSP dannet fra massespektrene førte derimot ikke til en bedring i identifisering mot MSP fra egen stamme (tabell 3.2.1). Hvorfor dette skjer, kan skyldes endringen av informasjon mellom massespekter og MSP, men uten å vite noen om algoritmen som benyttes av MALDI Biotyper 3.0 -programvaren i dannelsen av et MSP er det vanskelig å vite hvilke endringer som oppstår. Det er også viktig å nevne at scoreverdien som blir gitt er en logaritmisk verdi (1.3.3). Endringene i scoreverdi mellom intervallet fra ca 1,7 og 2,3 vil derfor tilsvare et en større variasjon i datasettet enn en endring fra f. eks 2,3 til 2,4.

Selv om forskjellene mellom MSP fra de ulike isolatene ikke klarer å gi korrekt klassifisering i forhold til stammene med Biotyper 3.0 – programvaren, ble det ut fra det fylogenetiske treet fra MSP – dataene (figur 3.2.2 b) vist at det er enkelte forskjeller i proteinprofilene mellom stammene i *L. pneumophila*. En økning i antall MSP fra *L. pneumophila* i Bruker Daltonics biblioteket vil derfor antagelig øke sannsynligheten for korrekt klassifisering på artsnivå. Dette har også blitt sett og foreslått i Moliner et al. 2009 for *Legionella*. Diversiteten i forhold til andre *Legionella* MSP er illustrert i vedlegg D, figur D1, selv om antallet av andre *Legionella* -arter er lavt.

Under arbeidet med oppgaven ble ”subtypings” - rutinen i programvaren Biotyper 3.0 testet ut, for å undersøke om klassifiseringen mellom stammene ble bedre. ”Subtypingen” tillater økt vektning av særegne topper i MSP for å bedre kunne differensiere mellom ulike stammer i en art (Bruker Daltonics 2011). Det ble valgt å benytte standardinnstillingene som var foreslått av Bruker Daltonics, hvor MALDI Biotyper 3.0 automatisk velger ut toppene som skal vektes ulikt. ”Subtypingen” førte til en høyere scoreverdi ved identifiseringen, men viste ingen endring i forhold til bedring i klassifiseringen. På lignende måte som det ble sett ved dannelsen av MSP fra massespektrene i blindtesten. I MALDI Biotyper 3.0 er det ikke mulig å danne et fylogenetisk tre basert på ”subtypings” – scoreverdiene. Det ble derfor valgt ikke å gå videre med arbeidet med ”subtypingen”.

4.3 Sammenligning og vurdering av typingsmetodene

En klyngeanalyse ble utført for å undersøke om oppløselighetene som blir oppnådd fra MALDI-TOF MS var sammenlignbar med SBT -analysen.

SBT, som er en vel utprøvd epidemiologisk typingsmetode for *L. pneumophila* (1.4) ble i denne oppgaven brukt som referansemetode. Den fylogenetiske fremstillingen i resultatet ble bare utført med *flaA*, *pile*, *asd*, *mip*, *mompS* og *proA* – genene. *neuA* ble i utgangspunktet inkludert av EWGLI i sin SBT – protokoll for å bedre kunne differensiere mellom sg1 (Ratzow et al. 2007). Det har i etterkant vist seg at enkelte stammer i sg 2 – 14 ikke blir amplifisert med primerene som er publisert for *neuA*, noe som skyldes at *neuA* er vist å avvike kraftig på både DNA- og aminosyrenivå i enkelte av stammene i nevnte sg (Farhat et al. 2011). Dette blir også bekreftet i denne oppgaven for stammene i sg 4, 5, 7, 8, 11, 13 og 14 (tabell 2.2.1), hvor det ikke var mulig å amplifisere *neuA* med nevnte primere hos enkelte av isolatene. Nye primere er foreslått i Farhat et al. 2011. En god genetisk sammenligning ble allikevel oppnådd for alle *L. pneumophila* isolatene fra genene som resulterte i amplifisering.

NJ – treet (figur 3.2.1) og UPGMA – treet (figur 3.2.2 a) som ble dannet fra allelprofilene var sammenfallende etter hvordan *L. pneumophila* – isolatene klynget seg sammen i forhold til hverandre, men enkelte forskjeller ble observert. NJ antar minst mulig evolusjon mellom dataene mens UPGMA antar en konstant evolusjon i dataene (2.7.5). Dette kan forklare hvorfor det ble observert enkelte rotasjoner rundt

grenene i treene i forhold til hverandre, og at utliggerne som roter treene er forskjellig. NJ – treet blir rotet av *L. pneumophila* ATCC 43283 og ATCC 43703 som i UPGMA – treet ligger i klynge B, men som det kan observeres er den horisontale avstanden stor i forhold til isolatene rundt. UPGMA – treet blir rotet av uteliggeren ATCC 35096 som i NJ treet på samme måte viser en lang horisontal avstand fra de andre isolatene.

I det fylogenetiske treet basert på MSP - data (figur 3.2.2 b) av *L. pneumophila* er det valgt å presentere trestrukturen ut fra standard innstillinger i MALDI Biotyper 3.0 programvaren. MALDI Biotyper 3.0 tillater i utgangspunktet justeringer i valg av metode, men anbefaler standard innstillinger som utgangspunkt (Bruker Daltonics 2011). I arbeidet med oppgaven ble en rekke innstillinger testet ut, hvor de største forandringene i trestrukturen som ble sett var endringer i avstandsnivå og rotasjoner rundt grener. Stammene i rød og blå forgreningene var gjennomgående de samme for de ulike metodene testet, og viste også i PCA -treet å være konservert. Standardinnstillingene ble derfor beholdt og er også benyttet i andre undersøkelser som har gjort analyser med MALDI Biotyper -programvaren (Murugaiyan et al. 2011).

Avstandsnivået i MSP - treet mellom den blå og røde forgreningen av isolater er markant markert i forhold til forgreningene som skjer utover i treet. Noe som tydet på at MSP fra hver forgrening skilte seg tydelig fra hverandre. Derfor ble massespektre fra blå og rød forgrening sammenlignet (figur 3.2.4). Det ble i sammenligningen antatt at ett massespekter kunne representere et MSP, siden hvert MSP ble dannet med et kvalitetskrav på 500 ppm i variasjon rundt hver massetopp (2.6.1), selv om intensiteten i hver massestopp er sett å kunne variere litt mellom massespektrene. Massespektrene som ble presentert i resultatet (figur 3.2.4) viste at det var store likheter mellom massespektrene innad i rød gruppe og innad i blå gruppe, men mellom massespektrene i de to forgreningene er det visse forskjeller. Dersom intensitetsforskjellene blir sett bort ifra er den tydeligste forskjellen i massespektrene, massetoppene ved 3000 og 6000 m/z i blå forgrening som er borte fra massespektrene i rød forgrening. Ut ifra de to forgreningene, den korte avstanden mellom MSP i den røde forgreningen og massespektrene undersøkt, kan det virke som *L. pneumophila* – isolatene deler seg inn i to ”overordnede” proteinprofiler. Fra referansebiblioteket av MSP og av *L. pneumophila* MSP i Bruker Daltonics sitt referansebibliotek (Vedlegg

A, tabell A1) ble det konstruert en fylogenetisk tre (Vedlegg D, figur D1) som illustrerer klyngene og diversiteten i forhold til andre *Legionella* spp. Fra det fylogenetiske treet blir det observert at MSP fra *L. pneumophila* fra referansebibliotek til Bruker Daltonics fordeler seg mellom de to markerte forgreningene som er observert tidligere (rød og blå, i figur 3.2.2 b).

Sammenligningen mellom SBT og MSP – typingen (3.2.2) viste at diversiteten i proteinsammensetningen hos *L. pneumophila*, representert som MSP fra MALDI–TOF MS, ikke er sammenlignbar med oppløseligheten som oppnås gjennom SBT. MSP -treet viser gjennom det horisontale avstandsnivået at mange MSP er for like til å kunne representere hver stamme i forhold til slik det blir sett i sekvensvariasjonene fra genene i SBT - analysen. Selv om det ble observert likheter mellom klynge C i MSP – treet (figur 3.2.2 b) og klynge B i UPGMA – treet (figur 3.2.2 a) i forhold til stammene som er representert i hver klynge, er det fortsatt tydelig forskjeller mellom treene. Ut fra stammene med den samme allelprofil fra SBT – analysen (3.2.2), ville det være forventet at MSP fra de samme stammene la seg i samme klynge i MSP – treet. Resultatene viser derimot at de ved de fleste tilfeller ikke gjør dette, noe som kan bety at oppbygningen av trestrukturen kan være tilfeldig i de delene hvor det horisontale avstandsnivået er kort.

I Molnier et al .2011 har det tidligere blitt vist at MALDI–TOF MS ikke viser noen tydelig sammenheng med sg -fordeling hos *L. pneumophila*. Ut fra informasjonen som var tilgjengelig om sg til *L. pneumophila* -stammene benyttet i denne oppgaven(tabell 2.2.1), ble det heller ikke sett noen sammenheng med sg i forhold til de fylogenetiske treene som ble konstruert.

4.4 Veien videre

I et foreløpig upublisert arbeid ble det på ”22nd European Congress of Clinical Microbiology and Infectious Diseases” (ECCMID) presentert (02. april 2012) en artikkel; ”Subtyping of *Legionella pneumophila* for epidemiological investigation by matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry”. I sammendraget som foreløpig (12. mai 2012) er tilgjengelig, er det presentert en sammenligning mellom SBT og MALDI–TOF MS subtyping av *L. pneumophila*. Resultatene de vil presentere, viser til å kunne dele inn *L. pneumophila*

inn i ni undergrupper med MALDI-TOF MS -analyse. Undergruppene ble sammenlignet med den samme SBT -protokollen som ble benyttet i denne oppgaven og viste 88 % likhet mellom resultatene fra MALDI-TOF og SBT. Massespektrene fra MALDI-TOF MS ble, av de to laboratoriene involvert, analysert mellom 5000 m/z og 11000 m/z (Jung et al. 2012). Hvilke metoder de har benyttet for å komme frem til resultatet er uvisst.

MALDI Biotyper -systemet er av Bruker Daltonics i utgangspunktet designet for rask identifisering av mikroorganismer på slekt, art og i noen tilfeller underartsnivå (1.3.3). Det er mulig at sensitiviteten i systemet mot underarter og stammer blir ofret på bekostning av å kunne gjøre identifiseringene raskt og på et bredt utvalg av mikroorganismer. Standardinnstillingene som i denne oppgaven har blitt benyttet både ved identifisering og i dannelsen av MSP kan derfor ha senket sensitiviteten som ville vært nødvendig for å bedre kunne skille mellom de ulike stammene hos *L. pneumophila*. Muligheten for ”subtyping” i MALDI Biotyper 3.0 ble som nevnt tidligere prøvd ut med standard innstillinger (4.2). For å øke sensitiviteten i MALDI Biotyper -systemet vil en mulig løsning være å identifisere særegne massetopper i massespektrene manuelt og bruke disse som biomarkører for de ulike *L. pneumophila* stammene. Identifisering av særegne massetopper har tidligere vist seg å være vellykket for *Salmonellae* for typing på underartsnivå (Dieckmann et al. 2008).

5. Konklusjon

I denne oppgaven ble det utført tre innledende forsøk for å standardisere protokollen for proteinekstraksjon fra *L. pneumophila* for bruk i MALDI-TOF MS -analyse(r). Fra resultatene ble vist at både for mye og for lite biologisk materiale og ulike volum av maursyre og acetonitril ville påvirket massespektrene ved MALDI-TOF MS -identifisering. Det ble vist at lagring av *L. pneumophila* -kolonier på skål ved 20 °C i opptil seks dager, ga stabile massespektrene dannet med MALDI-TOF MS.

Det ble bygget opp et referansebibliotek i MALDI Biotyper 3.0 bestående av MSP fra til sammen 48 *Legionella* spp. hvorav 41 var *L. pneumophila* -isolater. Med referansebiblioteket ble det vist at MSP (dannet ved standard innstillinger) for mange *L. pneumophila* var for like hverandre til å gi klassifiseringer mot egne MSP. Det ble foreslått at sensitiviteten mot underarter i MALDI Biotyper -systemet blir ofret til fordel for rask identifisering mot et bredt utvalg av mikroorganismer.

En klyngeanalyse ble brukt i sammenligning av *L. pneumophila* mellom SBT og MSP i referansebiblioteket dannet med MALDI-TOF MS. Sammenligningen viste at diversiteten i proteinsammensetningen hos *L. pneumophila*, representert som MSP, ikke ga samme oppløselighet som gjennom sekvensvariasjonene fra genene i SBT -analysen.

Referanser

- Alispahic, M., Christensen, H., Hess, C., Razzazi-Fazeli, E., Bisgaard, M. & Hess, M. (2011). Identification of *Gallibacterium* species by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry evaluated by multilocus sequence analysis. *International Journal of Medical Microbiology*, **301** (6): 513-522 s.
- Applied Maths. (2012). *BioNumerics* [online]. 6.6 utg. Lastet ned fra: <http://www.applied-maths.com/bionumerics/bionumerics.htm>. (Lastet ned 17.04.2012).
- Barbuddhe, S. B., Maier, T., Schwarz, G., Kostrzewa, M., Hof, H., Domann, E., Chakraborty, T. & Hain, T. (2008). Rapid Identification and Typing of *Listeria* Species by Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. *Applied and Environmental Microbiology*, **74** (17): 5402-5407 s.
- Bizzini, A. & Greub, G. (2010). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification. *Clin Microbiol Infect*, **16**: 1614-1619 s.
- Blatny, J. M., Reif, B. A. P., Skogan, G., Andreassen, O., Høiby, E. A., Ask, E., Waagen, V., Aanonsen, D., Aaberge, I. S. & Caugant, D. A. (2008). Tracking Airborne *Legionella* and *Legionella pneumophila* at a Biological Treatment Plant. *Environmental Science & Technology*, **42** (19): 7360-7367 s.
- Brown, R. S. & Lennon, J. J. (1995). Mass Resolution Improvement by Incorporation of Pulsed Ion Extraction in a Matrix-Assisted Laser Desorption/Ionization Linear Time-of-Flight Mass Spectrometer. *Analytical Chemistry*, **67** (13): 1998-2003 s.
- Bruker Care Online. (2012a). *Bruker Bacterial Test Standard* [online]. I: Inc., B. D. Tilgjengelig fra: <http://www2.bdal.de/modux3/modux3.php?pid=002,004,000,01,02,050,002,255343&rid=002,004,000,01,02,050,001,0> (lest 13.03.2012).
- Bruker Care Online. (2012b). *HCCA, portioned* [online]. I: Inc., B. D. Tilgjengelig fra: <http://www2.bdal.de/modux3/modux3.php?pid=002,004,000,01,02,050,002,255344&rid=002,004,000,01,02,050,001,0> (lest 12.03.2012).

- Bruker Daltonics. (2004). *microflex - User Manual, Version 1.2*: Bruker Daltonic GmbH. 1-36 s.
- Bruker Daltonics. (2011). *MALDI Biotyper 3.0 User Manual* [online]. I: GmbH, B. D. Lastet ned fra: <ftp://ftp.bdal.com/pub/tof/outgoing.to.customer/JNC/Install/Documentation/MBT-3-0-UserManual.pdf>. (Lastet ned 15.08.2011).
- Bruker Daltonics. (2012). *AnchorChip target with 96 spots* [online]. Tilgjengelig fra: <http://www.bdal.de/products/maldi/microflex-series/images.html> (lest 13.03.2012).
- Chen, P., Lu, Y. & Harrington, P. B. (2008). Biomarker Profiling and Reproducibility Study of MALDI-MS Measurements of *Escherichia coli* by Analysis of Variance–Principal Component Analysis. *Analytical Chemistry*, **80** (5): 1474-1481 s.
- Claydon, M. A., Davey, S. N., Edwards-Jones, V. & Gordon, D. B. (1996). The rapid identification of intact microorganisms using mass spectrometry. *Nat Biotech*, **14** (11): 1584-1586 s.
- Degré, M., Hovig, B. & Rollag, H. (2008). *Medisinsk mikrobiologi*, b. 3: Gyldendal Norsk Forlag AS.
- Dieckmann, R., Helmuth, R., Erhard, M. & Malorny, B. (2008). Rapid Classification and Identification of *Salmonellae* at the Species and Subspecies Levels by Whole-Cell Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. *Applied and Environmental Microbiology*, **74** (24): 7767-7778 s.
- Drevinek, M., Dresler, J., Klimentova, J., Pisa, L. & Hubalek, M. (2012). Evaluation of sample preparation methods for MALDI-TOF MS identification of highly dangerous bacteria. *Letters in Applied Microbiology*: no-no s.
- Enright, M. C. & Spratt, B. G. (1998). A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology*, **144** (11): 3049-3060 s.
- Farhat, C., Mentasti, M., Jacobs, E., Fry, N. K. & Lück, C. (2011). The N-Acylneuraminate Cytidyltransferase Gene, *neuA*, Is Heterogenous in *Legionella pneumophila* Strains but Can Be Used as a Marker for Epidemiological Typing in the Consensus Sequence-Based Typing Scheme. *Journal of Clinical Microbiology*, **49** (12): 4052-4058 s.
- Fujinami, Y., Kikkawa, H. S., Kurosaki, Y., Sakurada, K., Yoshino, M. & Yasuda, J. (2011). Rapid discrimination of *Legionella* by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Microbiological Research*, **166** (2): 77-86 s.

- Gaia, V., Fry, N. K., Afshar, B., Lück, P. C., Meugnier, H., Etienne, J., Peduzzi, R. & Harrison, T. G. (2005). Consensus Sequence-Based Scheme for Epidemiological Typing of Clinical and Environmental Isolates of *Legionella pneumophila*. *Journal of Clinical Microbiology*, **43** (5): 2047-2052 s.
- Garrity, G. M., Bell, J. A. & Lilburn, T. (2005). Order VI. Legionellales ord. nov. I: Brenner, D. J., Krieg, N. R., Staley, J. T. & Garrity, G. M. (red.) *Bergey's Manual of Systematic Bacteriology*, s. 210-236. New York: Springer.
- Gomez-Valero, L., Rusniok, C. & Buchrieser, C. (2009). *Legionella pneumophila*: Population genetics, phylogeny and genomics. *Infection, Genetics and Evolution*, **9** (5): 727-739 s.
- Griffin, T. J. & Smith, L. M. (2000). Single-nucleotide polymorphism analysis by MALDI-TOF mass spectrometry. *Trends in Biotechnology*, **18** (2): 77-84 s.
- Hallin, M., Deplano, A. & Struelens, M. J. (2012). Molecular Typing of Bacterial Pathogens: A Tool for the Epidemiological Study and Control of Infectious Disease. I: Morland, S., Beaudreau, F. & Cabaret, J. (red.) *New Frontiers of Molecular Epidemiology of Infectious Diseases*, s. 9-25: Springer Science.
- Holland, R. D., Wilkes, J. G., Rafii, F., Sutherland, J. B., Persons, C. C., Voorhees, K. J. & Lay, J. O. (1996). Rapid Identification of Intact Whole Bacteria Based on Spectral Patterns using Matrix-assisted Laser Desorption/Ionization with Time-of-flight Mass Spectrometry. *Rapid Communications in Mass Spectrometry*, **10** (10): 1227-1232 s.
- Hortin, G. L. (2006). The MALDI-TOF Mass Spectrometric View of the Plasma Proteome and Peptidome. *Clinical Chemistry*, **52** (7): 1223-1237 s.
- IUPAC. (2011). *Compendium of the Chemical Terminology - Gold Book* [online]. Tilgjengelig fra: <http://goldbook.iupac.org/> (lest 20.03.2012).
- Jung, J., Groß, B., Weinert, K., Maier, T., Kostrzewa, M. & Schubert, S. (2012). Subtyping of *Legionella pneumophila* for epidemiological investigations by matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry. Upublisert manuskript.
- Kungliga Vetenskapsakademien. (2002). *The 2002 Nobel Prize in Chemistry - Popular Information* [online]: Nobelprize.org. Tilgjengelig fra: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2002/popular.html (lest 08 Des. 2011).
- Maier, T., Klepel, S., Renner, U. & Kostrzewa, M. (2006). Fast and reliable MALDI-TOF MS - based microorganism identification. *Nature Methods*, **3**.

- Marklein, G., Josten, M., Klanke, U., Müller, E., Horré, R., Maier, T., Wenzel, T., Kostrzewa, M., Bierbaum, G., Hoerauf, A., et al. (2009). Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for Fast and Reliable Identification of Clinical Yeast Isolates. *Journal of Clinical Microbiology*, **47** (9): 2912-2917 s.
- Moliner, C., Ginevra, C., Jarraud, S., Flaudrops, C., Bedotto, M., Couderc, C., Etienne, J. & Fournier, P.-E. (2010). Rapid identification of *Legionella* species by mass spectrometry. *Journal of Medical Microbiology*, **59** (3): 273-284 s.
- Molofsky, A. B. & Swanson, M. S. (2004). Differentiate to thrive: lessons from the *Legionella pneumophila* life cycle. *Molecular Microbiology*, **53** (1): 29-40 s.
- Murugaiyan, J., Ahrholdt, J., Kowbel, V. & Roesler, U. (2011). Establishment of a matrix-assisted laser desorption ionization time-of-flight mass spectrometry database for rapid identification of infectious achlorophyllous green microalgae of the genus *Prototheca*. *Clin Microbiol Infect.*
- Nelson, D. L. & Cox, M. M. (2008). Investigating Proteins with Mass Spectrometry. I: Lehninger (red.) b. Fifth edition *Principles of Biochemistry*, s. 98 - 100. New York: W. H. Freeman and Company.
- Olsen, J. S., Aarskaug, T., Thrane, I., Pourcel, C., Ask, E., Johansen, G., Waagen, V. & Blatny, J. M. (2010). Alternative Routes for Dissemination of *Legionella pneumophila* Causing Three Outbreaks in Norway. *Environmental Science & Technology*, **44** (22): 8712-8717 s.
- Oxoid. (2010). *BCYE Agar* [online]. I: Oxoid-Limitede. Tilgjengelig fra: http://www.oxoid.com/UK/blue/prod_detail/prod_detail.asp?pr=CM0655&c=UK&lang=EN (lest 03.03.2012).
- Ratzow, S., Gaia, V., Helbig, J. H., Fry, N. K. & Lück, P. C. (2007). Addition of neuA, the Gene Encoding N-Acylneuraminate Cytidylyl Transferase, Increases the Discriminatory Ability of the Consensus Sequence-Based Scheme for Typing *Legionella pneumophila* Serogroup 1 Strains. *Journal of Clinical Microbiology*, **45** (6): 1965-1968 s.
- Ryzhov, V. & Fenselau, C. (2001). Characterization of the Protein Subset Desorbed by MALDI from Whole Bacterial Cells. *Analytical Chemistry*, **73** (4): 746-750 s.
- Sauer, S., Freiwald, A., Maier, T., Kube, M., Reinhardt, R., Kostrzewa, M. & Geider, K. (2008). Classification and Identification of Bacteria by Mass Spectrometry and Computational Analysis. *PLoS ONE*, **3** (7): e2843 s.

- Sauer, S. & Kliem, M. (2010). Mass spectrometry tools for the classification and identification of bacteria. *Nat Rev Micro*, **8** (1): 74-82 s.
- Sobral, D., Le Cann, P., Gerard, A., Jarraud, S., Lebeau, B., Loisy-Hamon, F., Vergnaud, G. & Pourcel, C. (2011). High-Throughput Typing Method To Identify a Non-Outbreak-Involved *Legionella pneumophila* Strain Colonizing the Entire Water Supply System in the Town of Rennes, France. *Applied and Environmental Microbiology*, **77** (19): 6899-6907 s.
- Staden, R. (1996a). *The Staden sequencing analysis package* [online]. I: Laboratory of Molecular Biology. Cambridge. Lastet ned fra: <http://staden.sourceforge.net/>. (Lastet ned 28.03.2012).
- Staden, R. (1996b). The Staden sequencing analysis package. *Molecular Microbiology*, **5** (3): 233 - 41 s.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). *MEGA5: Molecular Evolutionar Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods* [online]. Molecular Biology and Evolution. Lastet ned fra: <http://www.megasoftware.net/>. (Lastet ned 28.03.2012).
- Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T. & Matsuo, T. (1988). Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, **2** (8): 151-153 s.
- Teramoto, K., Sato, H., Sun, L., Torimura, M. & Tao, H. (2007). A Simple Intact Protein Analysis by MALDI-MS for Characterization of Ribosomal Proteins of Two Genome-Sequenced Lactic Acid Bacteria and Verification of Their Amino Acid Sequences. *Journal of Proteome Research*, **6** (10): 3899-3907 s.
- The Health Protection Agency. (2012). *The EWGLI SBT database for the typing of Legionella pneumophila* [online]. Tilgjengelig fra: http://www.hpa-bioinformatics.org.uk/legionella/legionella_sbt/php/sbt_homepage.php (lest 28.03.2012).
- Tveten, A.-K. (2008). Genotyping of *Clostridium botulinum*. Masteroppgave: Norges teknisk-naturvitenskaplige universitet Institutt for Bioteknologi. 1-94 s.
- Visca, P., D'Arezzo, S., Ramisse, F., Gelfand, Y., Benson, G., Vergnaud, G., Fry, N. K. & Pourcel, C. (2011). Investigation of the population structure of *Legionella pneumophila* by analysis of tandem repeat copy number and internal sequence variation. *Microbiology*, **157** (9): 2582-2594 s.

Watson, J. T. & Sparkman, O. D. (2007). *Introduction to Mass Spectrometry: Instrumentation, Applications, and Strategies for Data Interpretation*, b. 4th edition. Chichester: John Wiley & Sons.

Williams, T. L., Andrzejewski, D., Lay Jr, J. O. & Musser, S. M. (2003). Experimental factors affecting the quality and reproducibility of MALDI TOF mass spectra obtained from whole bacteria cells. *Journal of the American Society for Mass Spectrometry*, **14** (4): 342-351 s.

Vedlegg A: *L. pneumophila* oppføringer i databasen til Bruker Daltonics

Det er totalt 11 MSP oppføringer (Tabell A1) av *L. pneumophila* i Bruker Daltonics referansebibliotek som følger med programvaren Biotyper 3.0 (Bruker Daltonics, Tyskland).

Tabell A1: *L.pneumophila* oppføringer i programvaren MALDI Biotyper 3.0

Slekt	Art	Typestammekode	Diverse
<i>Legionella</i>	<i>pneumophila</i>	ATCC 33155 TW VUN	ssp <i>pneumophila</i>
<i>Legionella</i>	<i>pneumophila</i>	ATCC 35096 TW VUN	ssp <i>pneumophila</i>
<i>Legionella</i>	<i>pneumophila</i>	ATCC 33154 TW VUN	ssp <i>pneumophila</i>
<i>Legionella</i>	<i>pneumophila</i>	ATCC 35251 TW VUN	ssp <i>fraseri</i>
<i>Legionella</i>	<i>pneumophila</i>	ATCC 33216 TW VUN	ssp <i>fraseri</i>
<i>Legionella</i>	<i>pneumophila</i>	ATCC 43283 TW VUN	ssp <i>pneumophila</i>
<i>Legionella</i>	<i>pneumophila</i>	LGN_PNHL_2004_146 TW VUN	
<i>Legionella</i>	<i>pneumophila</i>	LGN_PNHL_2003_058 TW VUN	
<i>Legionella</i>	<i>pneumophila</i>	LGN_PNHL_2003_057 TW VUN	
<i>Legionella</i>	<i>pneumophila</i>	ATCC 33152T THL	ssp <i>pneumophila</i>
<i>Legionella</i>	<i>pneumophila</i>	DSM 7515T DSM	ssp <i>pascullei</i>

Vedlegg B: Referansesekvenser til SBT – gen

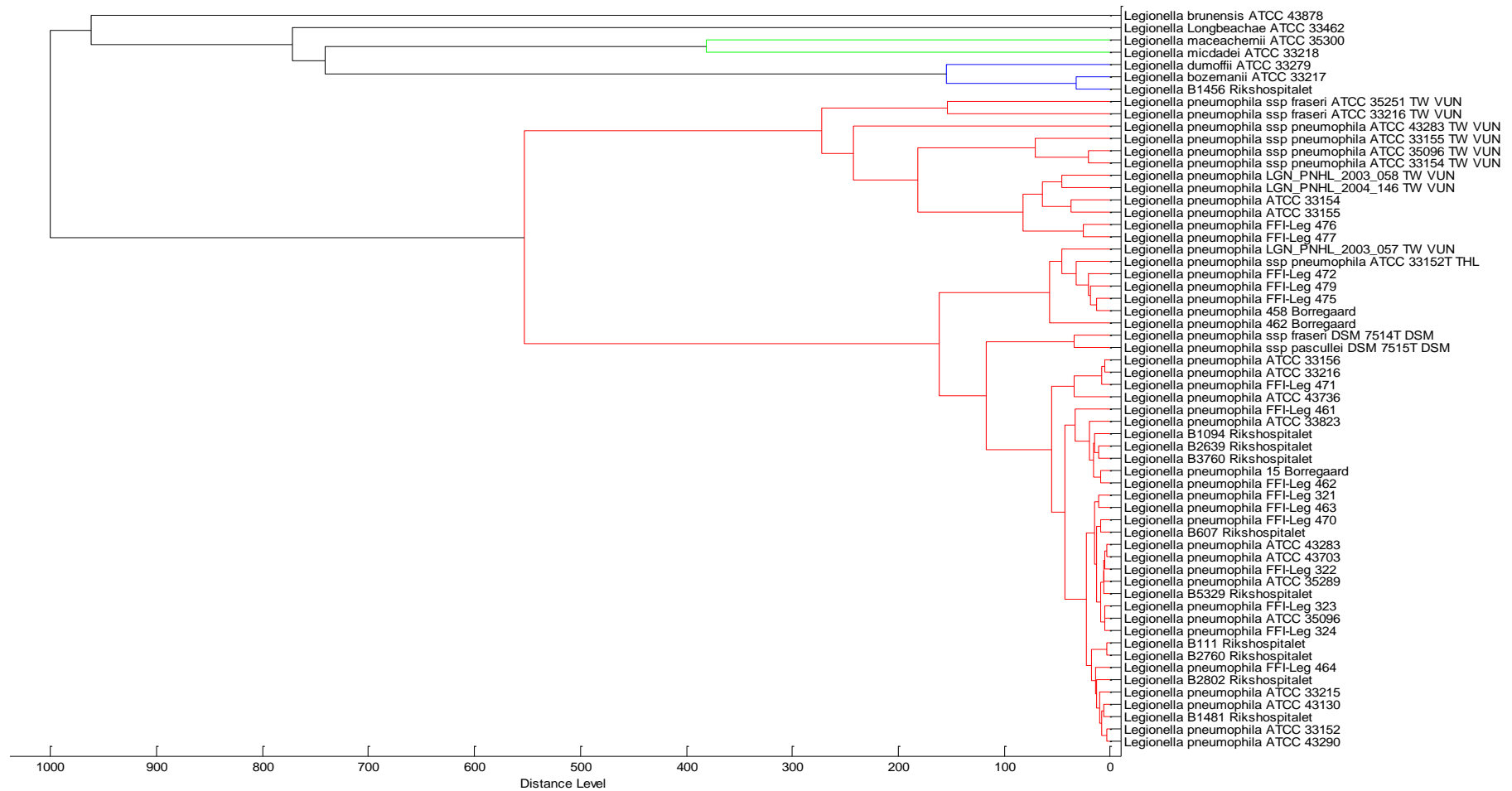
Ved analysering av konsensussekvensene beskrevet i 2.7.5, ble referansesekvensene (tabell B1) til SBT – genene brukt til å bestemme målregionen av genene som ble brukt i til å bestemme ST (1.4).

Tabell B: Referansesekvenser til SBT – gen (The Health Protection Agency 2012)

Gen (Lengde (nt))	Nukleotidsekvens
<i>flaA</i> (182)	TCAAACCAAA CCATCAATTT CAGTATCGGC AGCATAAAAAG CTTCTTCCAT TGGTGGTATT GCCACGGCAA CAGGAACAGA AGTAGCTGGT GCAGCAGCGA CAGATATCAC TATCGCAATT GGAGGAGGAG CAGCAACCAG TATTAACTCT TCTGCCAATT TTACCGGGGC ACTAAACGGG CA
<i>pilE</i> (333)	ATGCTCAGTC ATGTACATTT TATGAAGAAT CGCCGTATGA ACAAATCCCG GTTTACCCTG GTTGAAGTTC TGATCAGCAT GGTCATTATG GGCATTCTAG TTTCAATTGC CTATCCATCC TATTTACAAT ATATCCAAAA ATCCCGTCGT GCCGATGCTC ACGCCACATT GACACAAGAT CAAATTATTT TAGAACGCTG TTATTACAG AATTTTTCTT ATGTGCAGC GTGTGGCGCC CTACCAGCAT TTCCTCAAAC AACGCCGAAC GGGTACTATA CTATCAATAT TTCAAACCTG ACGGCCACAA CGTATACCTT AACCGCAACC CCT
<i>asd</i> (473)	CGCTGTTGGT ATTAGTCGTA TTAATGTGGC GACTTACCAG TCTGTTTTCCG GGACTGGCAA AAAAGCCATC AGCGAGCTGG TTGCTCAGGT AGGCGATCTT TAAATGGGA GACCAGCTAA TGTCAAGTT TATCCTCAGC AAATTGCTTT TAATGCGCTT CCTCATATTG ATCAGTTCGA AGACAATGGT TATACCCGAG AAGAGATGAA GATGGTCTGG GAAACCCGCA AGATTATGGA AGATGACAGC ATTATGGTTA ACCCTACAGC CGTCAGGGTT CCGTGTATTT ATGGGCATTC TGAAGCGGTT CATCTGGAGT TAAAAAGCC TTTGACGGCT GACGATGCTC GCGCGCTTTT GGCAAAGGCA CCCGGCGTTA CTGTAGTGGA TAATCTTCT AAAGCAAGTT ATCCACAGC AATTAAGAAT GCAGTCGGGC ATGATGATGT TTTGTAGGG CGCATAAGAC AGG
<i>mip</i> (402)	TGCCGATTG GGAAGAATT TAAAAATCA AGGCATAGAT GTTAATCCGG AAGCAATGGC TAAAGGCATG CAAGACGCTA TGAGTGGCGC TCAATTGGCT TTAACCGAAC AGCAAAATGAA AGACGTTCTT AACAAGTTTC AGAAAGATTT GATGGCTAAG CGTACTGCTG AATTCAATAA GAAAGCGGAT GAAAATAAAG TAAAAGGGGA AGCCTTTTTA ACTGAAAACA AAAACAAGCC AGGCGTTGTT GTATTGCCAA GTGGTTTGCA ATACAAAGTA ATCAATTCTG GAAATGGTGT TAAACCCGGA AAATCGGATA CAGTCACTGT CGAATATACT GGTGCTCTGA TTGATGGTAC CGTTTTGAC AGTACCGAAA AAAGTGGTAA GCCAGCAACG TT
<i>mompS</i> (352)	ATGACAGTGA TCACTGGGCT GATTTTGCTA ACTGGCACAA CTACAACAAC AAGTGGGATG CTGTTAATGC TGAATTAGGT CAATTCGTAG ATTTACGCGC TAACAAGAAA ATGCGTTTTCC ACGGCGGTGT TCAATACGCT CGCATTGAAG CTGATGTGAA CCGTTATTTT AATAACTTTG CCTTTAACGG GTTCAACTCT AAGTTCAATG GCTTTGGTCC TCGCACTGGT TTAGACATGA ACTATGTATT TGGCAATGGC TTTGGTGTAT ATGCTAAAGG CGCTGCTGCT ATTCGTGGTTG GTACCAGCGA TTTCTACGAT GGAATCAACT TCATTACTGG TTCTAAAAAT GC
<i>proA</i> (405)	TATGAAAACG CCTACTGGGA CGGCAAGCAA ATGACCTTTG GTGATGGTGA CACCATGATG TATCCTTTGG TTTCTCTAGG TGTAGGTGGT CATGAAGTTA GCCATGGATT CACTGAACAA CATTCTGGAC TTGAATATTT CGGTCAGTCT GGTGGCATGA ATGAGTCTTT TTCAGATATG GCTGCTCAAG CTGCTGAATA TTATTCAGTT GGAAAAAACA GCTGGCAGAT CGGCCGGAA ATAATGAAAAG AAGACAGTGG TTATGACGCA TTGCGTTACA TGGATAAACC AAGCCGTGAT GGTATGTCTA TTGATGTAGC TGATGATTAT TATGGCGGGT TAGATGTTCA TACTCCAGT GGAGTGATA ACCATTTATT CTACATATTA GCTAATCAAC CTAAT
<i>neuA</i> (354)	GATACTTCGG ATGTGATTCA TACTGTTATT GACCTCTGT TTAAGTTTCA GCAAATGGAT GTTTTTTTG ACAGTGTATT GCTGTTACAA CCAACTTCTC CATTTAGGAA GCCAGAAAACC ATAAGACATG CTGTTGAAAT ACATAAAGTA ACGGGGAAAA GTGTCGTTTC AGTTAGTCCC ATCTCTTTAA AGCCTTCTTG GTGTAGAAGC ATTGATAGCC AAGGCAATTT GGTTAAGCCA GAACTTTTC AGGATCTGGA AATTTATTGC AACGAGAATC CAATTTATAA ATTAATGGG TCAATTTATA TTGCAACAGC TAAACAAAT ATTGAAAATA AGAGTTTTTA TAGC

Vedlegg C: Trestruktur av *Legionella* spp. MSP i referansedatabasen

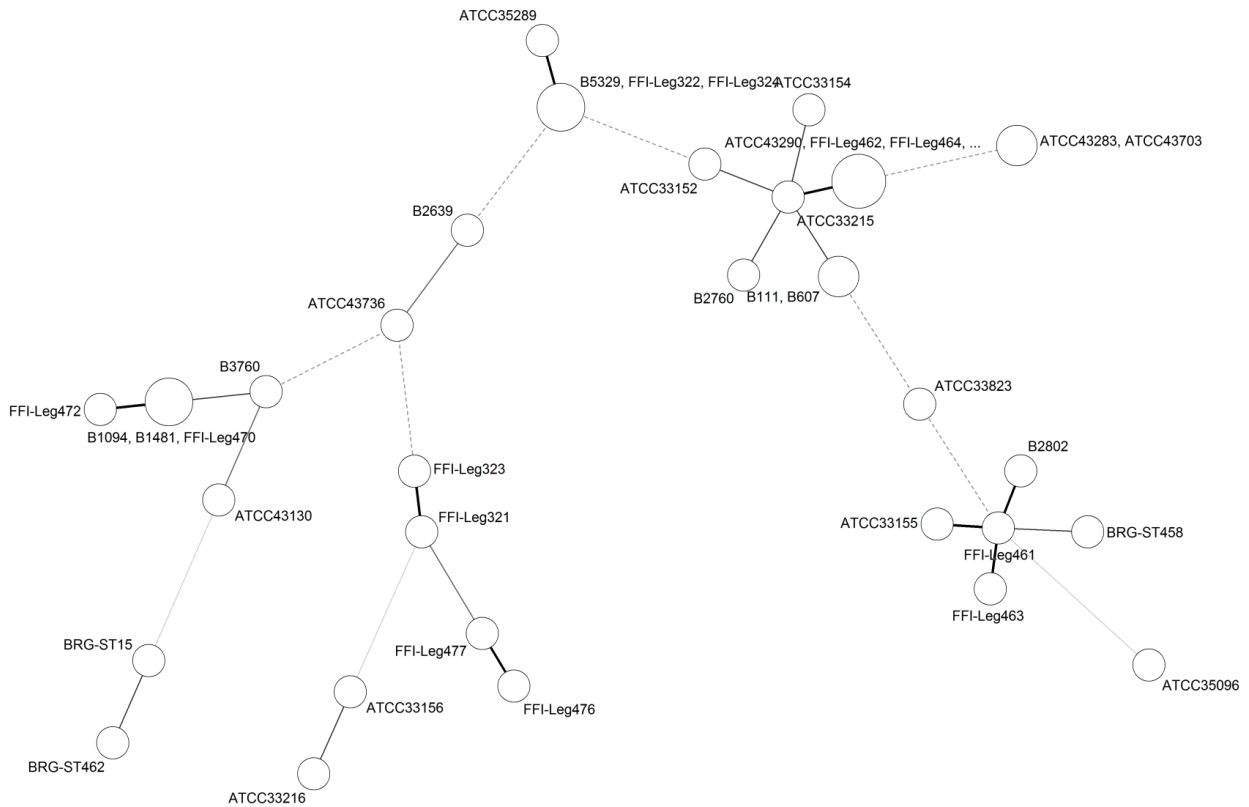
Fylegonetisk tre (figur C1) fra *Legionella* spp. av referansedatabasen som ble konstruert og 11 *L. pneumophila* MSP fra referansedatabasen til Bruker Daltonics



Figur C1: Scoringsorientert tre fra MALDI-TOF MS MSP av *Legionella* spp.

Vedlegg D: MST -tre av SBT analyse

Fra SBT - analysen av *L. pneumophila* ble det konstruert et MST - tre (figur D1) basert på allelprofilene, for å illustrere den korteste evolusjonære avstanden mellom dataen.



Figur D1: Fylogenetisk slektskap hos *L. pneumophila* isolatene basert på SBT – analysen, vist i et MST – tre.