

Bruk av massiv parallell sekvensering for påvisning og identifikasjon av mikrober i galle hos pasienter med akutt kolecystitt og akutt kolangitt

Masteroppgave

Randi Monsen Nygaard

Masterprogram i helsevitenskap
Studieretning for radiografer og bioingeniører

Institutt for global helse og samfunnsmedisin
Det medisinske fakultet
Universitetet i Bergen

Høsten 2017

Forord

Masteroppgaven i helsevitenskap ved Institutt for global helse og samfunnsmedisin, Universitetet i Bergen ble utført i samarbeid med Mikrobiologisk avdeling, Haukeland universitetssjukehus. Avdelingsledelsen ved Mikrobiologisk avdeling takkes for tilrettelegging av arbeidstid og oppgaver i forbindelse med studiet.

Hovedveileder Øyvind Kommedal (Ph.d, overlege ved Mikrobiologisk avdeling) hadde ideen til prosjektet, og var den som fikk ballen til å rulle. Hans kunnskap og fantastiske entusiasme har vært til stor inspirasjon under hele arbeidet. Seksjonsoverlege Christoffer Lindemann har også bidratt på en fantastisk måte med god veiledning under hele prosessen. Begge fortjener en stor takk for sin tålmodighet og optimisme. Jeg har satt stor pris på at jeg har fått diskutere så mye, med så flinke og inspirerende veiledere.

Prøvematerialet som er benyttet i masteroppgaven er samlet inn av lege Ruben Dyrhovden ved Mikrobiologisk avdeling i forbindelse med et større forskningsprosjekt. Jeg vil takke for et fint samarbeid under både prøvepreparering og tolking av svar; det har vært herlig å kunne diskutere med andre som også blir lykkelige av å identifisere mikrober i galleprøver. Som student i et nytt og spennende fagfelt har jeg vært avhengig av flere gode hjelpere. Christine Drengenes, Tharmini Kalanathan og Laura Minsaas takkes for god hjelp med innføring i laboratorieteknikker og lån av utstyr.

Teknikkene som har blitt benyttet er kostbare; Helse Vest takkes for forskningsmidler som gjorde studien mulig. Bioingeniørfaglig institutt (NITO) og Mikrobiologisk avdeling, Haukeland universitetssjukehus takkes for økonomisk støtte til nødvendige kurs i løpet av studiet.

Når et tidkrevende masterstudium skal gjennomføres som deltidsstudium i tillegg til jobb, vil det nødvendigvis gå på bekostning av noe eller noen. Jeg vil rette en ekstra stor takk til min flotte familie, som har vært støttende gjennom hele prosessen. Min kjære Ernst har motivert meg til å holde ut i periodene når ting har gått litt trått. Han har godtatt at stuen har blitt tatt i bruk som lesesal, med veggene dekorert av nukleotidtabeller og fylogenetiske trær. De flotte ungdommene våre Jens, Ingeborg og Herman har blitt enda mer selvstendige mens jeg har fordypet meg i studier, og vært forståelsesfulle når det innimellom har vært litt kaotisk i hjemmet. Tusen takk!

Bergen, 12.09.2017

Randi Monsen Nygaard

Innhold

SAMMENDRAG	5
FORKORTELSER	6
1 INTRODUKSJON	1
1.1 Bakgrunn	1
1.1.1 Akutt kolecystitt og akutt kolangitt	2
1.1.2 Genteknologiske metoder for påvisning av mikrober	4
1.1.3 Massiv parallell sekvensering	5
1.1.4 Valg av gener for sekvenseringbasert identifikasjon av mikrober	8
1.2 Hensikt og problemstilling	10
1.2.1 Hensikt	10
1.2.2 Problemstillinger	10
2 METODE	11
2.1 Valg av forskningsdesign	11
2.2 Utvalg	12
2.3 Etablering av PCR-metoder for bruk i prøveprepareringen	12
2.3.1 Valg av primere og PCR-betingelser for 16S-PCR.....	13
2.3.2 Valg av primere og PCR-betingelser for ITS-2 PCR.....	13
2.3.3 Valg av primere og PCR-betingelser til <i>rpoB</i> -PCR for <i>Enterobacteriaceae</i>	14
2.3.4 Valg av primere og PCR-betingelser til <i>rpoB</i> -PCR for <i>Enterococcus</i> , <i>Streptococcus</i> og <i>Staphylococcus</i>	17
2.3.5 Etablering av PCR-primere med adaptersekvenser	20
2.4 Datainnsamling	21
2.4.1 Prøvetaking.....	21
2.4.2 Dyrkning.....	21
2.4.3 Nukleinsyreekstraksjon	22
2.4.4 Kontroll av tilstedeværelse av bakterielt DNA i prøvene	22
2.4.5 Preparering av genbibliotek.....	23
2.4.6 Analysering av DNA-sekvenser	26
2.4.7 Tolking av resultatene	27
2.4.8 Statistiske beregninger	28
2.5 Ethiske betraktninger	28
3 RESULTATER	29

3.1	Funn av mikrober i galleprøver fra pasienter med akutt kolecystitt og akutt kolangitt	29
3.2	Mikrobefunn ved PTHCS versus ERCP.....	32
3.3	MPS sammenlignet med dyrkning	36
3.3.1	RpoB-sekvensering	39
3.3.2	ITS-sekvensering.....	40
3.4	Utbytte og kvalitet	41
3.5	Innhold av mikrobielt DNA i negative kontroller	43
3.5.1	Ekskluderte funn	44
4	DISKUSJON.....	46
4.1	Etablering av PCR-metoder.....	46
4.2	Sekvensering og analysering av dataene	50
4.2.1	Identifikasjon av arter innen <i>Enterobacteriaceae</i> -familien.....	51
4.2.2	Identifikasjon av bakterier innen <i>Enterococcus</i> -, <i>Streptococcus</i> - og <i>Staphylococcus</i> -slektene.....	56
4.2.3	Kimærer.....	57
4.2.4	Vurdering av hvilke funn som er reelle.....	58
4.2.5	Mikrobefunn og tilknytning til prøveidentitet.....	60
4.2.6	Sekvenskvalitet.....	61
4.3	Dyrkning versus massiv parallell sekvensering.....	62
4.4	Gevinst av rpoB- og ITS-sekvensering	63
5	KONKLUSJON	67
6	REFERANSER	68

SAMMENDRAG

De fleste infeksjonssykdommer som forårsakes av bakterier og sopp kan behandles med antimikrobielle midler. Kjennskap til hvilke mikrober som forårsaker sykdommene er viktig for å kunne gi pasienten riktig behandling. Den vanligste metoden for påvisning og karakterisering av bakterier og sopp er dyrkning. Dyrkningsmetoden har imidlertid vesentlige begrensninger, da den er avhengig av levende mikrober som er i stand til å vokse under de betingelsene som kan tilbys i laboratoriet.

Vi har derfor etablert en dyrkningsuavhengig metode for identifisering av bakterier og sopp, basert på massiv parallell sekvensering av utvalgte gener. Sekvensering av *16S rRNA*-genet er en godt etablert metode for identifisering av bakterier i diagnostiske laboratorier, og kan i mange tilfeller identifisere bakterier til artsnivå. Det finnes imidlertid unntak, blant annet for arter innen *Enterobacteriaceae*-familien og innen *Enterococcus*-, *Streptococcus*- og *Staphylococcus*-slektene. For å løse dette har vi derfor innført sekvensering av *rpoB*-genet, som et supplement til *16S rRNA*. Sekvensering av ITS-2 segmentet er inkludert for å kunne påvise og identifisere sopp.

Tjue galleprøver fra nitten pasienter med kolecystitt og kolangitt ble undersøkt, og det ble totalt funnet 35 ulike arter. Det ble påvist signifikant flere mikrober ved massiv parallell sekvensering enn ved dyrkning, og det var en klar sammenheng mellom mengde påvist DNA og om mikroben lot seg dyrke. Sekvensering av *rpoB* og ITS-2 ga verdifull tilleggsm informasjon sammenholdt med sekvensering av *16S rRNA* alene, og de aller fleste mikrobene lot seg identifisere til artsnivå. Metoden som er etablert i denne studien kan også benyttes til andre typer prøvemateriale, og vil være spesielt verdifull for karakterisering av mikrobiologien ved ulike polymikrobielle infeksjoner.

FORKORTELSER

Oversikt over forkortelser som er brukt i teksten

bp	Basepar (nukleotider)
BLAST	Basic local alignment search tool
ct-verdi	Threshold Cycle, terskelverdi. Antall sykluser i en PCR-reaksjon som må til for at avlest signal når en gitt terskelverdi
DNA	Deoksyribonukleinsyre
DPO	Dual priming oligonucleotide principle
ERCP	Endoskopisk retrograd kolangiografi
ITS	Internal transcribed spacer
LNA	Locked nucleic acid
MIA	Mikrobiologisk avdeling
MPS	Massiv parallell sekvensering
NGS	Neste generasjons sekvensering
PCR	Polymerase kjede reaksjon (polymerase chain reaction)
PTHCS	Perkutan transhepatisk drenasje av galleblære/galleganger
rpoB	RNA polymerase, β -subenhet
rpoB_ent	RpoB-amplikon for <i>Enterobacteriaceae</i>
rpoB_ess	RpoB-amplikon for <i>Enterococcus</i> , <i>Streptococcus</i> og <i>Staphylococcus</i>
16S	16S rRNA

1 INTRODUKSJON

1.1 Bakgrunn

Dødelighet forårsaket av infeksjonssykdommer i Norge har hatt en nedgang siden begynnelsen av 1900-tallet. Dette skyldes en generell bedring av levekårene og helsetjenestetilbudet, og etter hvert også utvikling av vaksiner og antibiotika (Folkehelseinstituttet, 2013). Infeksjonssykdommer er likevel en viktig dødsårsak på verdensbasis. Nærmere en fjerdedel av alle dødsfall har mikrobiell årsak (World Health Organization, 2008).

De fleste infeksjoner som forårsakes av bakterier eller sopp kan behandles med antimikrobielle midler. Ulike mikrober trenger ulik behandling. Informasjon om hvilket medikament som skal benyttes i hvert tilfelle oppnås best ved at mikrobene fra pasienten identifiseres i et laboratorium. Ved alvorlige infeksjoner må noen ganger behandling startes før prøven tas, og i de fleste tilfeller før testresultatene foreligger. Det er utarbeidet nasjonale retningslinjer for valg av antibiotika ved de ulike infeksjonssykdommene, slik at behandleren kan velge det antibiotikum som vanligvis er best egnet for den enkelte infeksjonssykdom, såkalt empirisk behandling (Helsedirektoratet, 2013). De fleste forskningsdata som er lagt til grunn for retningslinjer for antibiotikabehandling baserer seg på dyrkning av mikrobene fra pasientprøver. Dyrkningsmetoden har vesentlige begrensninger da den er avhengig av levende mikrober som er i stand til å vokse under de betingelsene man tilbyr i laboratoriet. Noen mikrober er døde eller ute av stand til å formere seg allerede idet prøven tas, fordi pasienten har blitt behandlet med antibiotika. Anaerobe mikrober vil kunne dø dersom de utsettes for oksygen under prøvetaking eller transport til laboratoriet, og andre mikrober kan være særlig sårbare for kulde, uttørking eller mangel på bestemte næringsstoffer. I prøver med flere ulike mikrober (polymikrobielle infeksjoner) er det også en fare for at noen arter blir utkonkurrert av mer hurtigvoksende og hardføre arter. Det kan altså være mange arter tilstede i en prøve uten at disse påvises ved dyrkning. Dette gjelder spesielt prøver som tas fra dype infeksjoner som abscesser eller infeksjoner i tilknytning til tarm, der miljøet er anaerobt og vanskelig å gjenskape under transport og i laboratoriet. Intra-abdominale infeksjoner omfatter infeksjoner av ulik alvorlighetsgrad, og riktig behandling vil kunne være avgjørende for pasientens prognose (Lopez, Kobayashi, & Coimbra, 2011).

I motsetning til dyrkning er genteknologiske metoder ikke avhengig av levende mikrober, kun av intakt deoksyribonukleinsyre (DNA), og kan dermed være et supplement til dyrkning. Det siste 10-året har det skjedd en rivende utvikling innenfor DNA-sekvenseringsteknologien, slik at denne nå kan anvendes på komplekse polymikrobielle infeksjoner. Det foreligger imidlertid svært få studier som har benyttet dette til å beskrive større materialer fra ulike infeksjoner.

1.1.1 Akutt kolecystitt og akutt kolangitt

Galle dannes i leveren og skilles ut via gallegangene til galleblæren. Galleblæren, som er lokalisert under leverens fremre kant, fungerer som et reservoar for galle. Når matinnhold kommer til tolvfingertarmen utskilles et hormon, kolecystokinin, som stimulerer galleblæren til å tømme seg. Den tømmer seg via gallegangen som munner ut i tolvfingertarmen. Gallen blandes med matinnholdet der den, ved å spalte fett, bidrar til tarmens opptak av fettsyrer (*Keshav, 2013*)

Pasienter med akutt kolecystitt (inflamasjon og/eller infeksjon i galleblæren) og akutt kolangitt (infeksjon i gallegangene) utgjør en stor gruppe pasienter ved kirurgiske avdelinger. Antibiotikabehandling av tilstandene er basert på internasjonale retningslinjer (Yoshida et al., 2007). Retningslinjene ble første gang vedtatt på et internasjonalt konsensusmøte i Tokyo i 2007 ("Tokyo Guidelines for the management of acute cholangitis and cholecystitis. Proceedings of a consensus meeting, April 2006, Tokyo, Japan," 2007). Retningslinjene gir også forslag til definisjon av tilstandene akutt kolecystitt og kolangitt. I 2013 ble retningslinjene oppdatert (TG13) (Takada et al., 2013). Behandlingsforslagene i retningslinjene har lagt til grunn eldre, dyrkningsbaserte studier for hvilke mikrober som kan være tilstede. Både akutt kolecystitt og akutt kolangitt er dype intra-abdominale infeksjoner, og ofte polymikrobielle. Som beskrevet i innledningen har dyrkningsbasert diagnostikk av slike infeksjoner klare begrensninger når det for eksempel gjelder å påvise bakterier som vokser dårlig på dyrkningsmedier eller er anaerobe. Det er heller ingen norske studier med i grunnlaget for retningslinjene, og det kan diskuteres om internasjonale mikrobe- og resistensforhold er sammenfallende med norske forhold. I helsedirektoratets nasjonale retningslinjer for bruk av antibiotika i sykehus er det presisert at det ikke foreligger gode studier når det gjelder behandling av hverken akutt kolecystitt eller akutt kolangitt (Helsedirektoratet, 2013).

Akutt kolecystitt, betennelse i galleblæren, er en vanlig komplikasjon hos pasienter med gallestein. TG13 har definert følgende diagnostiske kriterier for tilstanden:

- A. Lokale inflammasjonstegn: (1) Smerter og ømhet under høyre kostalbue og (2) positiv Murphys test (pasienten puster dypt inn mens man kjenner under høyre kostalbue).
- B. Systemiske inflammasjonstegn: (1) feber, (2) forhøyet CRP og (3) forhøyet antall hvite blodlegemer.
- C. Bildefunn som er karakteristiske for akutt kolecystitt.

For at kolecystitt skal mistenkes må minst ett av kriteriene under punkt A og ett av kriteriene under B være oppfylt. For sikker diagnose må punkt C være oppfylt, i tillegg til minst ett av kriteriene under punkt A og ett av kriteriene under punkt B (Yokoe et al., 2012)

Akutt kolangitt er en alvorlig tilstand med høy mortalitet. TG13 har definert følgende diagnostiske kriterier for tilstanden:

- A. Systemisk inflammasjon: (1) Feber og/eller frostbyger og (2) biokjemiske tegn på systemisk inflammasjon.
- B. Gallestase: (1) Gulsott og (2) patologiske leverfunksjonsprøver.
- C. Billeddiagnostikk: (1) Utvidede galleganger og (2) påvist etiologi (stein, striktur, stent) på bilder.

For at akutt kolangitt skal mistenkes må ett av kriteriene under punkt A og ett av kriteriene under B eller C være oppfylt. For sikker diagnose må ett av kriteriene under hvert av punktene A, B og C være oppfylt.

I en omfattende studie, CIAOW-studien, (Complicated intra-abdominal infections worldwide observation study) ble det samlet data fra 68 medisinske institusjoner fra ulike deler av verden med den hensikt å beskrive klinikk, mikrobiologi og behandling ved kompliserte intra-abdominale infeksjoner (Sartelli et al., 2014). De mikrobiologiske dataene er basert på aerob og til dels anaerob dyrkning. Til tross for dyrkningsmetodens begrensninger bidrar studien til kunnskap om at det finnes et bredt spekter av bakterier og sopp som årsak til intra-abdominale infeksjoner, og at infeksjonene i mange tilfeller er polymikrobielle.

CIAOW-studien viste at bakterier i Enterobacteriaceae-familien som normalt finnes i tarmen er en viktig årsak til intra-abdominale infeksjoner (Sartelli et al., 2014), noe som støtter funnene fra en oversiktsartikkel over abdominale infeksjoner publisert i 2011 (Lopez et al., 2011).

Dyrkningsresultater viste at også enkelte sopparter er til stede ved intra-abdominale infeksjoner (Sartelli et al., 2014). En sammenligningsstudie på pasienter med akutt pankreatitt har vist at intra-abdominal soppinfeksjon er forbundet med høyere dødelighet enn intra-abdominal bakteriell infeksjon (Vege et al., 2009). Det finnes mange ulike sopparter som kan være årsak til infeksjoner i gallesystemet. En oversiktsartikkel fra 2013 viser at tidlig diagnostisering og behandling av disse kan være avgjørende for pasientens prognose (Rebolledo & Sarria, 2013).

En studie fra 1990 utført på galle fra 121 multiorgandonorer viste at galleveiene hos disse var sterile (Ikeda et al., 1990). Undersøkelsen ble gjort med dyrkning, og det kan ikke utelukkes at andre metoder kunne gitt andre resultater. Dersom det påvises mikrober i galle må prøvetakings- og analysemetode tas med i vurderingen av funnene.

Prøvetaking fra galleveiene kan utføres på flere måter; mest vanlig er *perkutan transhepatisk dreناسje* av galleblære/galleganger (PTHCS) og *endoskopisk retrograd kolangiografi* (ERCP). PTHCS utføres med et nålestikk gjennom hud og bukvegg, videre gjennom levervevet og inn i galleveier eller galleblære. ERCP er en endoskopisk undersøkelse med fremstilling av galleveier og pankreasgang. Et tynt kateter føres gjennom endoskopet slik at røntgenkontrast kan føres opp i de to gangene (Keshav, 2013). Prøve til mikrobiologisk undersøkelse kan tas via aspirasjon fra kateter før kontrastvæske sprøytes inn.

1.1.2 Genteknologiske metoder for påvisning av mikrober

I motsetning til dyrkning er genteknologiske metoder ikke avhengig av levende mikrober men kun av intakt DNA, og kan dermed være et attraktivt supplement til dyrkning (Kalland, Myrmel, & Nordbo, 2005; O. Kommedal, Kvello, Skjastad, Langeland, & Wiker, 2009). Spesifikk PCR (polymerase chain reaction) kan benyttes, men denne metoden er avhengig av at man vet hvilke mikrober man leter etter, og at det designes en PCR-reaksjon for hver enkelt art. Dersom det er mange potensielt patogene mikrober er metoden derfor lite egnet. Et alternativ er en universell PCR-reaksjon rettet mot et gen som er til stede hos alle mikrober, etterfulgt av DNA-sekvensering av PCR-produktet for identifikasjon.

Sanger-sekvensering er fremdeles den mest brukte sekvenseringsmetoden i diagnostisk mikrobiologi, men har begrenset bruksområde da den egner seg best til prøver med én eller få

mikrober (O. Kommedal, Karlsen, & Saebo, 2008; O. Kommedal et al., 2009). For polymikrobielle infeksjoner er massiv parallell sekvensering, også kalt «neste generasjons sekvensering» eller «NGS», en bedre egnet metode. Ved å ta i bruk massiv parallell sekvensering kan man karakterisere svært komplekse miljøer, da et stort antall amplikon kan sekvenseres samtidig

(Tucker, Marra, & Friedman, 2009). En studie på hjerneabscesser utført ved Mikrobiologisk avdeling (MIA) ved Haukeland universitetssjukehus viste at en fant tre ganger så mange bakterier med denne metoden som ved dyrkning (O. Kommedal et al., 2014). Studien viste også at noen bakterier alltid sameksisterer og kan virke avhengige av hverandre. Påvisning av denne typen sammenhenger er avhengig av at prøvene alltid blir fullstendig karakterisert.

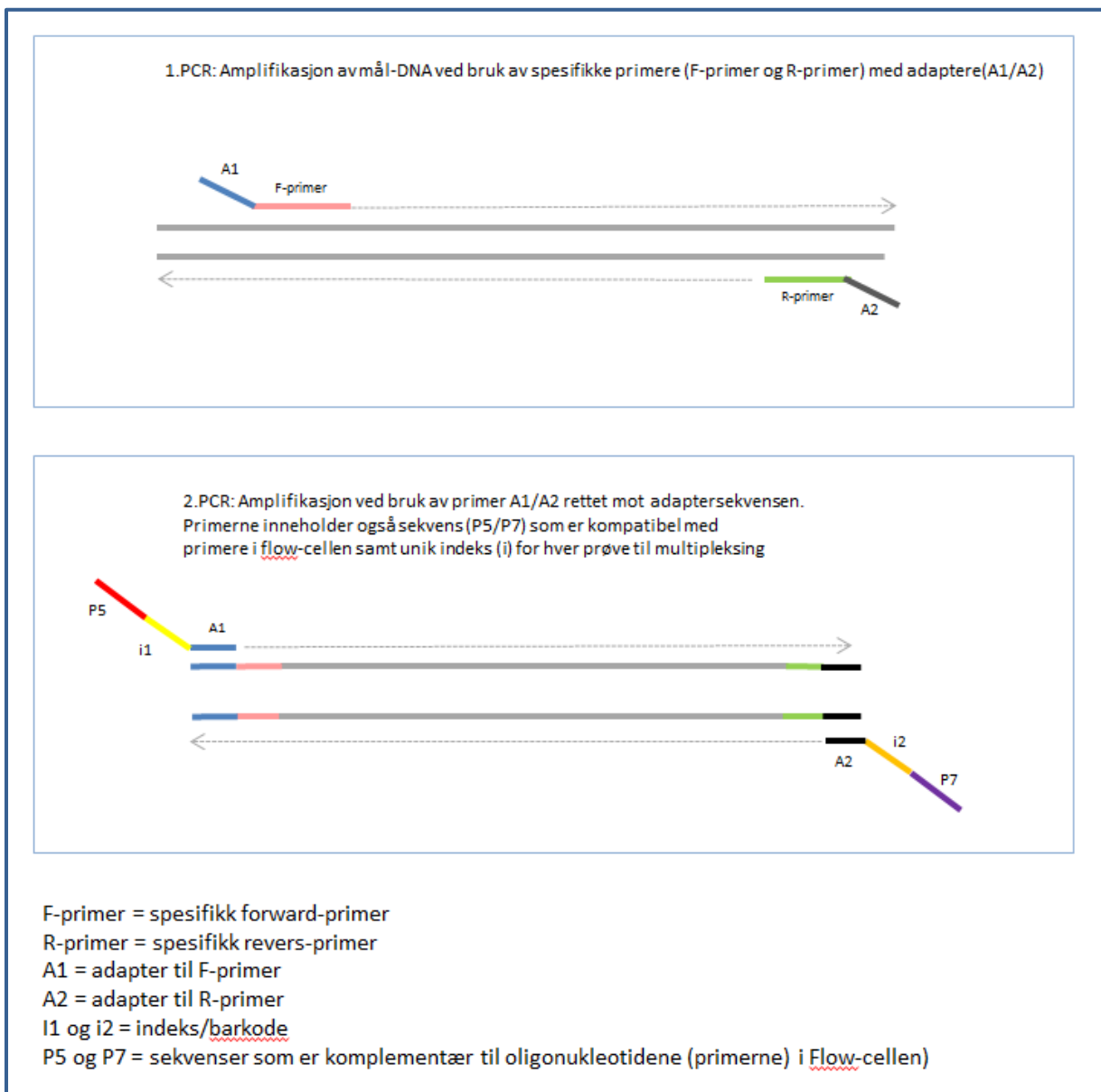
1.1.3 Massiv parallell sekvensering

Massiv parallell sekvensering gjør det mulig å sekvensere store mengder DNA på kort tid. Det finnes ulike analyseplattformer, med ulik kjemi. Felles for disse er at millioner av små DNA-fragmenter kan sekvenseres parallelt. Metoden kan benyttes til ulike formål som helgenomsekvensering eller dypsekvensering. Dypsekvensering benyttes for å sekvensere ulike varianter av et genområde i en og samme prøve, og er en godt egnet metode for identifisering av mikrober hos pasienter med polymikrobielle infeksjoner.

På Illuminas Miseq-plattform blir sekvensavlesningen utført mens nye DNA-tråder syntetiseres, «sequencing by synthesis». Før selve sekvenseringen må det produseres genbibliotek som i tillegg til mål-DNA inneholder «barkodesekvenser» som gjør det mulig å skille de ulike prøvene fra hverandre senere under behandling av dataene. Først utføres en PCR-reaksjon for amplifisering av den aktuelle sekvensen (Figur 1). Det dannes store mengder amplikon tilsvarende DNA-fragmentet. Oligonukleotidene som benyttes til amplifisering har flere funksjoner. Sekvensen i 3'enden er rettet mot DNA-tråden som skal amplifiseres, og fungerer som en ordinær PCR-primer. Sekvensen i 5'enden fungerer som en adapter. Den har en sekvens som er komplementære til primere som benyttes senere i prosessen, først til indeks-PCR og så til sekvenslesing. Alle PCR-produktene har altså like sekvenser i endene etter at PCR-reaksjonen er utført.

PCR-produktene benyttes som templat i en ny PCR-reaksjon, indeks-PCR. Oligonukleotidene som benyttes her har også flere funksjoner. Sekvensen i 3'enden er rettet mot

adaptersekvensen fra første PCR-reaksjon, og muliggjør dermed amplifikasjon av DNA-trådene. Den midterste delen av primerne inneholder indekssekvenser («barkoder») som gjør det mulig å samle flere prøver til et felles genbibliotek, og likevel skille resultatene fra hverandre under den senere dataanalysen. Hver prøve har en unik kombinasjon av to ulike indekser (dual indexing) med 8 nukleotider i hver, noe som gjør det mulig å koble hver enkelt sekvens til en bestemt prøve. Sekvensen i 5'enden (P5 og P7) er rettet mot oligonukleotidene som er bundet til flow-cellen der selve sekvenseringen foregår.



Figur 1: 1.PCR-reaksjon gir PCR-produkt med adapter i endene. Dette benyttes som templat i 2.PCR der DNA-trådene tilføres indeks (barkode) og sekvens som er komplementær til oligonukleotider i FLOW-cellen.

I Miseq overføres genbiblioteket til en flow-celle med kovalent bundne oligonukleotider. Oligonukleotidene har sekvenser som er komplementære til P5 og P7 i oligonukleotidene fra indeks-PCR. DNA-fragmentene hybridiseres til oligonukleotidene i flow-cellen før en isothermal polymerase amplifiserer fragmentene, og det dannes klynger for hver DNA-tråd. Etter dette starter selve sekvenseringen. Sekvenseringsprimerne er rettet mot adapterens 3'ende. Sekvensene amplifiseres på nytt, men nå med fire typer ulikt merkede nukleotider som alle har kjemisk inaktivert 3'OH. Kun en enkelt base inkorporeres dermed per syklus. For hver syklus gjøres en avlesning av hvilket nukleotid som er inkorporert før det fluorescerende molekylet fjernes og 3'enden reaktiveres (Voelkerding, Dames, & Durtschi, 2009) (Mardis, 2008; Metzker, 2010).

1.1.4 Valg av gener for sekvenseringbasert identifikasjon av mikrober

16S rRNA-genet finnes hos alle bakterier. Genet som består av ca. 1500 basepar koder for en komponent av prokaryot 30S subenhet. Genet består av vekselvis korte, høyt konserverte områder og lengre variable områder. Ved å designe oligonukleotider som passer med de konserverte områdene på hver side av ett eller flere variable områder får man en PCR som kan amplifisere opp den ønskede delen av *16S rRNA*-genet fra en hvilken som helst bakterie (en universell PCR). Amplikonet kan deretter sekvenseres slik at man kan bruke sekvensen i de inkluderte variable områdene for identifikasjon (Ø. Kommedal, 2012). Dette skjer ved at sekvensen sammenlignes med allerede kjente sekvenser som ligger i en database. Den største og mest brukte databasen er GenBank (*National Center of Biotechnology Information* NCBI). GenBank er en åpen database der alle kan legge inn sekvenser, men den blir i liten grad kurert. Kvaliteten på referansene som ligger der blir dermed høyst varierende.

	V1 69-99		V2 137-242		V3 433-497		V4 576-682		V5 822 - 829		V6 986- 1043		V7 1117 - 1173		V8 1243- 1294		V9
--	-------------	--	---------------	--	---------------	--	---------------	--	-----------------------	--	--------------------	--	-------------------------	--	---------------------	--	----

Figur 2: *16S rRNA* består av høyt konserverte områder (merket med blått) og variable områder merket V1-V9

Chakravorty med flere har beskrevet de variable områdene (V1-V9) på *16S rRNA*-genet og hvordan de kan benyttes til identifisering av ulike mikrober. Det vanligste området for sekvensering inkluderer de variable områdene V1-V3, men V3-V4 er også mye anvendt. Forskjellen har liten praktisk betydning, men oppløsningen (muligheten til å diskriminere mellom arter) på V3-V4 er noe lavere blant annet for streptokokker (Chakravorty, Helb, Burday, Connell, & Alland, 2007).

16S rRNA har en jevn og forutsigbar mutasjonsrate som sammenfaller godt med vårt behov for å dele bakterier inn på artsnivå. Genet fungerer derfor som et «molekylært kronometer» som også kan benyttes til å studere slektskapsforhold mellom organismer. Det er stort sett 99,5 % homologi innen en art og mindre enn 99,0 % homologi mellom ulike arter. Det finnes imidlertid unntak, og det kan være vanskelig å skille *Enterobacteriaceae*, *Staphylococcus*, *Streptococcus* og *Enterococcus* på artsnivå basert på sekvensering av *16S rRNA* (16S-sekvensering) alene. Mykobakterier er også vanskelig å identifisere på annet enn genus-nivå. (Janda & Abbott, 2007). Gener med høyere mutasjonsrate kan benyttes for å skille disse mikrobenes på artsnivå. Når mutasjonsraten er høy er det imidlertid ikke mulig å finne

konserverte områder egnet for universell PCR. Sekvensering av gener med høy mutasjonsrate vil derfor i praksis bare være et supplement til *16S rRNA* for utvalgte mikrober.

RpoB-genet som koder for β -subenheten i RNA-polymerase har høyere mutasjonsrate enn *16S rRNA*. Det er vist at genet er bedre egnet enn *16S rRNA* til å identifisere *Enterobacteriaceae* til artsnivå (Mollet, Drancourt, & Raoult, 1997). *RpoB* har også vist seg egnet til å skille mellom ulike *Staphylococcus*-arter (Drancourt & Raoult, 2002) og ulike *Streptococcus*-arter (Drancourt, Roux, Fournier, & Raoult, 2004). Genet har imidlertid få konserverte områder og er det er nødvendig å benytte ulike primere dersom det skal gjøres artstyping innen flere genus.

Bakterier er prokaryote mikrober mens sopp er eukaryote, og det er store forskjeller mellom organismenes arvestoff. Sopp har ikke *16S rRNA*-genet, og det må derfor benyttes et annet segment for påvisning og identifisering. I en metodestudie i 2009 ble ulike gener og DNA-sekvenser testet for påvisning og identifisering av ulike sopparter (Khot, Ko, & Fredricks, 2009). Studien viste at DNA-sekvenser fra ITS-2 (internal transcribed spacer 2) er det som egner seg best til påvisning av de vanligste soppartene og i tillegg skiller godt mellom disse. ITS-2 er ikke et gen da det ikke koder for et protein, men et segment som ligger mellom 5.8 S ribosomalt RNA og 28S ribosomalt RNA.

Massiv parallell sekvensering benyttes i økende grad til påvisning og identifisering av mikrober, og er spesielt nyttig for prøvemateriale der dyrkningsmetoden har begrensninger. Som tidligere nevnt er det gjort dyrkningsbaserte studier for påvisning av bakterier og sopp i galle hos pasienter med kolecystitt og kolangitt. For å få bedre kunnskap om mikrobiologien i galle hos disse pasientene er det også behov for dyrkningsuavhengige studier. I 2015 undersøkte Shen med flere galleprøver fra 15 pasienter med helmetagenomsekvensering, i tillegg massiv parallell sekvensering av genet som koder for *16S rRNA* (Shen et al., 2015). Pasientene hadde gallestein, men det var ikke tatt stilling til hvorvidt noen av dem hadde kolecystitt eller kolangitt. Det ble nylig også publisert en liten studie der galle fra 6 pasienter med kolecystitt ble undersøkt med MPS (Kujiraoka et al., 2017). Dataene fra denne studien ga ikke grunnlag for identifisering på artsnivå. Selv om det er stor fare for kontaminering er dette ikke adressert i noen av studiene. Begge studiene benyttet *16S rRNA*-genet. De gir dermed bare informasjon om bakterier, og ikke sopp, i prøvematerialet. Flere studier er derfor

nødvendig for å øke kunnskapen om mikrobiologien i galle hos pasienter med kolecystitt og kolangitt.

1.2 Hensikt og problemstilling

1.2.1 Hensikt

- Etablere en metode for påvisning og identifikasjon av bakterier og sopp ved bruk av massiv parallell sekvensering. Primært ønskes identifikasjon på artsnivå.
- Undersøke hvilke bakterier og sopp som finnes i pasienter med kolecystitt og kolangitt. Resultatene vil senere inngå i en større studie, der galle fra alle kolecystitt- og kolangittpasienter i en toårsperiode ved Haukeland universitetssjukehus skal undersøkes. Resultatene vil være et viktig bidrag til den kunnskapen som behøves når retningslinjer for antibiotikabehandling skal revideres.

1.2.2 Problemstillinger

- Hvilke mikrober er til stede i galleprøver hos et utvalg pasienter med kolecystitt og kolangitt?
- Er det forskjellige funn i prøver tatt ved perkutan transhepatisk drenasje av galleblære/galleganger (PTHCS) og prøver tatt ved endoskopisk retrograd kolangiografi (ERCP)?
- Påvises flere mikrober ved massiv parallell sekvensering enn ved konvensjonell dyrkning?
- Hvilken gevinst gir sekvensering av *rpoB*-genet som supplement til 16S-sekvensering?
- Hvilken gevinst gir dyrkningsuavhengig påvisning av sopp?

2 METODE

2.1 Valg av forskningsdesign

I studien er det undersøkt hvilke mikrober som er til stede hos pasienter med kolecystitt og kolangitt. Det er et deskriptivt tverrsnittstudium, der resultatene sier noe om tilstedeværelsen av ulike mikrober hos pasienter med disse diagnosene. Det er sjelden aktuelt med eksperimentelle studier når sykdom knyttet til infeksjose agens undersøkes. Når man undersøker potensielt skadelige påvirkninger av en eller flere faktorer vil man alltid måtte velge observerende studiedesign (Juul, 2012, p. 13).

Når tilstedeværelsen av ulike faktorer i prøver fra syke pasienter skal undersøkes, vil det i mange tilfeller være aktuelt å sammenligne resultatene med prøver fra samme pasient når han/hun er frisk. Dette forutsetter at det kan tas ny prøve etter pasienten har blitt frisk /eller at det er tatt prøve før sykdomsdebut. Slike data gir et godt utgangspunkt for å studere forskjellen på analyseparameteren hos friske og syke mennesker, men tilnærmingen kan kun benyttes dersom prøvetaking ikke får store konsekvenser for pasienten. Det kan for eksempel være aktuelt dersom analytten undersøkes i blod eller i overflatiske prøver fra hud og slimhinner. I kolecystitt- og kolangittstudien vil en slik studie derimot ikke være mulig. Prøvetakingen innebærer invasive prosedyrer som vanskelig kan forsvares utført på friske mennesker. Målingene utføres dermed kun når pasienten er syk, under inngrep i galleveiene som uansett skal utføres i rutinemessig diagnostisk eller terapeutisk hensikt. Det er ikke aktuelt med oppfølgingsprøver av pasientens galleveier. Dette gir et deskriptivt tverrsnittstudium der vi finner ut hva som er til stede i galleveiene på ett måletidspunkt, når pasienten er syk.

Det viktigste med denne masteroppgaven er etablering av en diagnostisk test som kan benyttes for påvisning og identifisering av bakterier og sopp i prøver fra pasienter med polymikrobielle infeksjoner. Den deskriptive studien er avhengig av en pålitelig test som gir informasjon om det er bakterier og/eller sopp i prøvene. Testen må også ha diskriminerende egenskaper og skille de ulike artene fra hverandre, også når der er mange ulike mikrober tilstede i ulike konsentrasjoner i en og samme prøve.

2.2 Utvalg

Målpopulasjonen er pasienter med kolecystitt og kolangitt. Det samles prøver fra alle pasienter som behandles eller utredes for tilstandene med perkutan galleblæredrenasje og/eller endoskopisk kolangiografi i en toårsperiode i forbindelse med en større studie ved MIA. Tjue av prøvene som ble samlet i perioden juni 2015- mars 2016 inngår denne masteroppgaven. Utvalget består av prøver som alle har fått påvist en eller flere mikrober ved dyrkning. Metodeutvikling er tidkrevende og kostbart, og det har derfor vært nødvendig å begrense antall prøver som skulle analyseres. Utvalget er for lite til å kunne gi en robust beskrivelse av mikrobefloraen ved de to infeksjonstilstandene, men ved å inkludere 20 dyrkningspositive pasientprøver kunne det forventes å finne mange ulike mikrober og få et inntrykk av testens egenskaper og eventuelt forbedringspotensial. 10 av prøvene ble tatt med PTCHS av pasienter med akutt kolecystitt, mens 8 prøver ble tatt med ERCP av pasienter med akutt kolangitt. I tillegg er det to prøver fra en og samme pasient. Den første av disse prøvene ble tatt med ERCP mens pasienten hadde gallegangsstein uten infeksjon, mens den andre prøve ble tatt med PTHCS etter at pasienten senere utviklet kolangitt.

2.3 Etablering av PCR-metoder for bruk i prøveprepareringen

Før pasientprøvene kunne analyseres måtte det etableres PCR-metoder for amplifisering av DNA til sekvensering. Når PCR-produktet skal sekvenseres er det avgjørende at fragmentet inneholder adaptersekvenser i endene. Til sekvenseringen var det derfor nødvendig med en PCR-metode der primere inneholdt adaptersekvenser i 5'enden. PCR-analysene ble likevel først etablert med primere uten adaptere. Valideringen ble gjort med sanntids-PCR (SYBRgreen) for å få god visualisering av resultatene. Retningslinjer for validering av nye metoder ved MIA ble fulgt, slik at metodene også kan benyttes i pasientdiagnostikken som selvstendige tester med Sanger-sekvensering av PCR-produktet.

Etter valideringen av PCR-metodene ble det bestilt nye primere med adaptersekvenser i 5'endene. Ny sanntids-PCR ble utført og resultatene sammenlignet med resultatene fra PCR uten adaptere. Justeringer i PCR-protokollene ble gjort ved behov, slik at resultatene fra primere med og uten adaptere ble tilnærmet like.

2.3.1 Valg av primere og PCR-betingelser for 16S-PCR

Til 16S-PCR ble det benyttet en modifisert variant av det Illumina foreslår i sin protokoll for Miseq (Illumina, 2013). Målområdet består av de hypervariable områdene V3 og V4.

Primerne (tabell 3) er valgt på bakgrunn av en evaluering gjort av Klindworth med flere i 2013 (Klindworth et al., 2013). Det er imidlertid gjort noen endringer i revers-primer. Tredje base fra 3'enden er endret fra T til K (T/G) for å unngå mismatch mot *Propionibacterium acnes* (O. Kommedal et al., 2009). I syvende og åttende base fra 5'enden er H (A/C/T) og V (A/C/G) redusert til henholdsvis C og A, da dette skal være tilstrekkelig for å detektere klinisk relevante mikrober. Et redusert antall basekombinasjoner vil kunne minke risikoen for kryssreagering med humant DNA. Primerparet har vist seg velegnet for deteksjon av humanpatogene mikrober direkte fra prøvemateriale i diagnostikken ved MIA gjennom mange år.

Sanntids-PCR utføres på SmartCycler (Cepheid) i 25 µl reaksjonsrør med 12,5 µl ExTaq SYBR master mix (TaKaRa), 0,4 µM av hver primer, 8,5 µl PCR-gradert vann og 2 µl ekstrahert DNA. Følgende PCR-protokoll ble benyttet: 30 sekunder ved 95 °C for initiell enzymaktivering etterfulgt av 40 sykluser med 10 sekunder ved 95°C, 15 sekunder ved 60 °C og 20 sekunder ved 72 °C. Smeltepunktanalyse utføres på PCR-produktet.

2.3.2 Valg av primere og PCR-betingelser for ITS-2 PCR

ITS-2 PCR ble etablert for påvisning av sopp. Primersekvensene (tabell 3) er tatt fra en publikasjon der ulike primersett ble testet spesifikt for identifisering av en lang rekke humanpatogene sopp direkte fra prøvemateriale (Khot et al., 2009). PCR-metoden ble optimalisert og verifisert med pasientprøver og kvalitetskontroller. I tillegg til at metoden skulle brukes til massiv parallell sekvensering i denne masteroppgaven var det et mål å etablere en universell PCR for sopp som kunne benyttes til Sanger-sekvensering i rutinediagnostikken ved MIA. All validering ble derfor gjort i henhold til avdelingens retningslinjer for validering av tester som skal brukes i pasientdiagnostikken (vedlegg E).

Ulike PCR-protokoller ble først testet med *Candida albicans* (ATCC 10231). Det ble valgt primerbindingstemperatur 2 °C lavere enn den høyeste temperaturen som fungerte. Dette ble gjort for å ta høyde for eventuell mismatch mellom primer og mål-DNA. Det er vesentlige forskjeller mellom ulike sopp i lengden på ITS-2, noe som kan ha betydning for hva som er

ideell tid på primerbindings-/ekstensjonstrinnet. Primerbindingstiden ble satt til 15 sekunder lenger enn den korteste tiden som fungerte optimalt for *Candida albicans*. To ulike enzymløsninger ble testet, ExTaq SYBR master mix (TaKaRa) og LightCycler® FastStart DNA MasterPLUS SYBR Green I (Roche). Reagensene fra Roche ga best resultater. I tillegg til å gi lavest ct-verdi for positiv kontroll, ga disse reagensene også størst avstand mellom positiv og negativ kontroll. Dette er en klar fordel når det skal vurderes om en prøve er positiv eller negativ, og LightCycler® FastStart DNA MasterPLUS SYBR Green I er derfor tatt i bruk for dyrkningsuavhengig påvisning av sopp ved MIA.

Følgende PCR-protokoll ble valgt: 10 minutter ved 95 °C for initiell enzymaktivering etterfulgt av 45 sykluser med 10 sekunder ved 95°C, 25 sekunder ved 58 °C og 15 sekunder ved 72 °C. Smeltepunktsanalyse utføres på PCR-produktet. Analysen utføres på SmartCycler i 25 µl reaksjonsrør med 18,5 µl LightCycler® FastStart DNA MasterPLUS SYBR Green I, 0.4µM av hver primer og 6.5 µl ekstrahert DNA.

Etter at PCR-protokoll var valgt ble det laget 10-folds fortynningsrekke i tripletter for å vurdere PCR-metodens effektivitet. 11 vevsprøver ble testet for å se etter kryssreagering med humant DNA. Kvalitetskontroller som inngår i eksternt kvalitetssikringsprogram ved MIA (Neqas 2016) ble undersøkt og PCR-produktene sekvensert med Sanger-sekvensering. De eksterne kontrollene inneholdt *Paecilomyces variotii*, *Geothricum candidum* og *Neoscytalidium* spp., som alle ble korrekt identifisert. Egen valideringsrapport for ITS-2 PCR ble skrevet og godkjent, slik at metoden kunne tas i bruk til analysering av pasientprøver ved laboratoriet.

2.3.3 Valg av primere og PCR-betingelser til rpoB-PCR for *Enterobacteriaceae*

RpoB-sekvenser for et utvalg *Enterobacteriaceae* ble sammenlignet for å kunne velge egnet område for PCR-primere. Utvalget er beskrevet i tabell 3 og representerer arter som forekommer i rutinediagnostikken ved MIA. DNA-sekvensene ble hentet i den amerikanske gendatabanken «Genbank», *National Center of Biotechnology Information* (NCBI). Det bioinformatiske verktøyet *Geneious* 9.1.2 (Biomatters, New Zealand) ble benyttet for lagring og behandling av sekvenser. Primersekvensene ble valgt manuelt ved hjelp av retningslinjer fra en forelesning ved Norges teknisk-naturvitenskaplige universitet (Bergh, 2015). I tillegg til å finne områder som var egnet for primere måtte det tas hensyn til at området mellom

primerne skulle ha tilstrekkelig variabilitet til å kunne diskriminere mellom de ulike arter. Primersekvensene ble lagt til de mest konserverte områdene, men det var likevel behov for degenererte primere. Degenererte primere inneholder flere nukleotidvarianter på ett eller flere steder i primeren, og betegnes etter et eget kodesystem laget av *International Union of Pure and Applied Chemistry* (IUPAC). Det ble designet tre ulike primersett for å øke sjansen for å finne en metode som fungerte godt. Primersekvensene er oppgitt i tabell 3.

De tre alternative amplikonsekvensene fra samtlige arter i tabell 3 ble undersøkt mot GenBank for å bekrefte at identifisering til artsnivå var mulig. En kontrollstamme med *Escherichia coli* (ATCC 25922) ble deretter brukt til uttesting og optimalisering av de ulike primerne i laboratoriet. Denne testingen viste at primersett 1 og primersett 3 ga lavere ct-verdi og brattere kurve enn primersett 2, og primersett 2 ble derfor forkastet. Søk i GenBank viste at primersett 3 var mindre egnet til å skille mellom ulike *Klebsiella*-arter, og primersett 1 ble derfor valgt. Primersett 1 har dessuten få degenererte baser. Revers-primer er lagt til et område som er så konservert at degenererte baser unngås helt.

Følgende PCR-protokoll ble valgt: 30 sekunder ved 95 °C for initiell enzymaktivering før 40 sykluser med 10 sekunder ved 95°C, 15 sekunder ved 60 °C og 20 sekunder ved 72 °C. Sanntids-PCR ble utført på SmartCycler (Cepheid) i 25 µl reaksjonsrør med 12,5 µl ExTaq SYBR master mix (TaKaRa), 0.4 µM av hver primer, 8.5 µl PCR-gradert vann og 2 µl ekstrahert DNA.

Tabell 1 Oversikt over Genbank-sekvenser som ble benyttet til design av rpoB-PCR for identifisering av *Enterobacteriace*-arter og sekvensen til hver av disse i primerområdene

Mikrobe	GB-nummer	F-primer 5'3	R-primer 5'3
Primer		rpoB_ent -F GAAGGTCCRAAYATCGGTCT	rpoB_ent -R TGCATGTTTCGCACCCAT
<i>Yersinia enterocolitica</i>	CP016944.1	GAAGGTCCAAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Citrobacter freundii</i>	JMTA01000076.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Citrobacter koseri</i>	MBTW01000008.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Klebsiella aerogenes</i>	CP014748.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Enterobacter cloacae</i>	CP022148.1	GAAGGTCCAAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Enterobacter hormaechei</i>	CP011662.1	GAAGGTCCAAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Esherichia albertii</i>	AP014857.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Escherichia coli</i>	CP002185.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Escherichia fergusonii</i>	CU928158.2	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Hafnia alvei</i>	CP014031.1	GAAGGTCCAAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Klebsiella oxytoca</i>	CP020358.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Klebsiella pneumoniae</i>	CANS01000041.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Pantoea ananatis</i>	CP022427.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Proteus mirabilis</i>	NJFA01000001.1	GAAGGTCCAAATATCGGTCT	TGCATGTTTCGCACCCAT
<i>Proteus vulgaris</i>	JPIX01000006.1	GAAGGTCCAAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Providencia stuartii</i>	NJEY01000002.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Raoultella ornithinolytica</i>	NJBC01000036.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Raoultella planticola</i>	AY367361.2	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Salmonella enterica typhi</i>	CP002099.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Salmonella marcescens</i>	NGUE01000001.1	GAAGGTCCAAACATCGGTCT	TGCATGTTTCGCACCCAT
<i>Morganella morganii</i>	NEYR01000005.1	GAAGGTCCGAACATCGGTCT	TGCATGTTTCGCACCCAT

Etter valg av primere ble metodeverifisering gjort i henhold til retningslinjer ved MIA.

Fortynningsrekke ble laget for å se på PCR-effektiviteten. Kryssreaktivitet med humant DNA ble testet ved å kjøre PCR på 11 vevsprøver med negativ 16S-PCR. Isolater fra pasientprøver med *Klebsiella oxytoca*, *Klebsiella pneumoniae*, *Klebsiella variicola*, *E.coli*, *Raoultella ornithinolytica*, *Raoultella planticola*, *Raoultella terrigena*, *Klebsiella aerogenes*, *Yersinia enterocolitica*, *Proteus mirabilis* og *Proteus vulgaris* ble undersøkt. Amplikon ble sekvensert for å validere at sekvensene ga ønsket informasjon om bakterienes artsnivå.

2.3.4 Valg av primere og PCR-betingelser til *rpoB*-PCR for *Enterococcus*, *Streptococcus* og *Staphylococcus*

RpoB-PCR for *Enterococcus*-, *Streptococcus*- og *Staphylococcus* ble etablert etter samme mal som for *Enterobacteriaceae*. DNA-sekvenser fra bakteriene i tabell 4 ble hentet fra GenBank og sammenlignet med primersekvensene. Det var utfordrende å finne primerområder som var tilstrekkelig konserverte til å dekke alle tre genera. Vi endte derfor opp med å validere kun ett primerpar (tabell 3). I følge Illumina sin protokoll bør en tilstrebe primerbindingstemperatur på 60 °C, noe som var vanskelig med så begrensede muligheter for valg av primerområde. Både teori og empiri tilsa at optimal primerbindingstemperatur på den ene primeren var lavere enn ønskelig. Det ble derfor innført «Locked nucleic acids» (LNA™) for å heve smeltetemperaturen. LNA er et modifisert nukleotid der molekylets ribosering er «låst» i en formasjon som er stabiliserende for hybridiseringen mellom komplementære DNA-tråder. Det ble innført to slike nukleotider i F-primeren. Disse er angitt med «+» foran den modifiserte basen i tabell 3.

Det var nødvendig med degenererte primere for å dekke alle de viktigste mikrobene. For å redusere antall kombinasjoner ble det laget to ulike revers-primere.

Ulike temperaturprofiler ble testet. Følgende PCR-protokoll ble valgt: 30 sekunder ved 95 °C for initiell enzymaktivering etterfulgt av 40 sykluser med 10 sekunder ved 95°C, 15 sekunder ved 59 °C og 20 sekunder ved 72 °C. Sanntids-PCR ble utført på SmartCycler (Cepheid) i 25 µl reaksjonsrør med 12,5 µl ExTaq SYBR master mix (TaKaRa), 0.6 µM av F-primer, 0.4 µM av hver R-primer, 6.0 µl PCR-gradert vann og 2 µl ekstrahert DNA.

Tabell 2 Oversikt over Genbank-sekvenser som ble benyttet til design av rpoB-PCR for identifisering av *Streptococcus*-, *Staphylococcus*-, og *Enterococcus*-arter og sekvensen til hver av disse i primerområdene

Mikrobe	GB-nummer	F-primer 5'3	R-primer 5'3	
primere		GTTTRTAHCNTCCCADGTCAT	GCRACAGCRTGTATYCCRTTC GCDACAGCATGTATTCCWTTTC	R1 R2
<i>Enterococcus faecalis</i>	ACHK01000003	GTTGTAACCTTCCCA T GTCAT	GCA A CAGCATGTATTCC T TTC	R2
<i>Enterococcus hirae</i>	CP003504	GTTGTAACCTTCCCAAGTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Enterococcus durans</i>	CP012384	GTT A TACCCTTCCCAAGTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Enterococcus faecium</i>	LRHK01000001	GTTGTA C CTTCCCAAGTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Enterococcus gallinarum</i>	CP014067	GTTGTA C CTTCCCAAGTCAT	GCAACAGCTTGTATCC C A T TTC	R1
<i>Enterococcus casseliflavus</i>	CP004856	GTTGTAACCTTCCCAAGTCAT	GCAAC C GCATGTATTCC T TTC	R2
<i>Enterococcus dispar</i>	KE136354.1	GTTGTAACCTTCCCAAGTCAT	GCGA C TGCGTGTATTCC T TTC	R2
<i>Enterococcus villorum</i>	KE136488.1	GTTGTAACCTTCCCAAGTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus agalactiae</i>	CP006910	GTT A TACCCTTCCCAAGTCAT	GCAACAGCATGTATTCC T TTC	R2
<i>Streptococcus anginosus</i>	CP003860	GTTGTAACCTTCCCA T GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus constellatus</i>	CP003859	GTTGTAACCTTCCCA T GTCAT	GCGACAGCATGTATCC C G T TTC	R1
<i>Streptococcus gallolyticus</i>	FN597254	GTTGTAAC C A T CCCA T GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus intermedius</i>	CP003858	GTTGTAACCTTCCCA T GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus mutans</i>	AE014133	GTTGTAACCTTCCCAAGTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus suis</i>	FM252032	GTTGTAACCTTCCCA C GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus constellatus</i>	CP003859	GTTGTAACCTTCCCA T GTCAT	GCGACAGCATGTATCC C G T TTC	R1
<i>Streptococcus gallolyticus</i>	FN597254	GTTGTAAC C A T CCCA T GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus mitis</i>	FN568063	GTTGTAACCTTCCCAAGTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus oralis</i>	FR720602	GTTGTAACCTTCCCA T GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus parasanguinis</i>	CP002843	GTT A TACCCTTCCCAAGTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus pyogenes</i>	AP012491	GTT A TACCCTTCCCAAGTCAT	GCGAC G GCATGTATTCC T TTC	R2
<i>Streptococcus salivarius</i>	CP013216	GTTGTAACCTTCCCA T GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus thermophilus</i>	CP000023	GTTGTA C CTTCCCA T GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus vestibularis</i>	AEK001000002	GTTGTAACCTTCCCA T GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus pneumoniae</i>	FM211187	GTTGTAACCTTCCCAAGTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus sanguinis</i>	CDMW01000001	GTTGTAACCTTCCCA T GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus macedonicus</i>	HE613569	GTTGTAAC C A T CCCA T GTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Streptococcus dysgalactiae</i>	AP012976	GTTGTAACCTTCCCAAGTCAT	GCGACGGCATGTATTCC T TTC	R2
<i>Streptococcus uberis</i>	AM946015	GTTGTAACCTTCCCAAGTCAT	GCA A CAGCATGTATTCC T TTC	R2
<i>Streptococcus canis</i>	KP890284	GTT A TACCCTTCCCAAGTCAT	GCGAC G GCATGTATTCC T TTC	R2
<i>Streptococcus massiliensis</i>	CANO01000052	GTTGTA T CTTCC C AGGTCAT	GCGACAGCATGTATTCC T TTC	R2
<i>Staphylococcus caprae</i>	AF325868	GTT A TACC G TCCCAAGTCAT	GCGACAGCTTGTATTCC A TTC	R1
<i>Staphylococcus aureus</i>	KC707803	GTTGTAAC C G T CCCAAGTCAT	GCGACAGCATGTATTCC A TTC	R2
<i>Staphylococcus intermedius</i>	AF325869	GTT A TACC G TCCCAAGTCAT	GCGACAGCGTGTATCC C A T TTC	R1
<i>Staphylococcus epidermidis</i>	CP009046	GTTGTAAC C G T CCCA T GTCAT	GCA A CAGCATGTATTCC A TTC	R2
<i>Staphylococcus warneri</i>	CP003668	GTTGTAAC C A T CCCA T GTCAT	GC T ACAGCTTGTATTCC A TTC	R1
<i>Staphylococcus saprophyticus</i>	CP022093.1	GTTGTAAC C G T CCCAAGTCAT	GCGACAGCATGTATTCC T TTC	R2

Fortynningsrekke ble laget for å se på PCR-effektiviteten. Kryssreaktivitet med humant DNA ble testet ved å kjøre PCR på 11 vevsprøver som var negative på 16S-PCR. Isolater fra pasientprøver med *Enterococcus faecium*, *Enterococcus faecalis*, *Staphylococcus aureus*, og *Streptococcus agalactiae* ble undersøkt. Amplikon ble sekvensert for å bekrefte at sekvensene ga ønsket informasjon om bakterienes artsnivå.

Tabell 3 Primersekvenser for universell PCR av 16S rRNA, rpoB og ITS-2 samt spesifikk PCR for *Granulicatella adiacens* og *Raolutella ornithinolytica*

Mikrober som detekteres	primernavn	Primersekvens (5'→ 3')
Bakterier	16S-F	CCTACGGG N GGC W GCAG
	16S-R	GACTACCAGGGTATCTAA KCC
Sopp	ITS2-F	GTGAATCATCGA R TCTTTGAA
	ITS2-R	TATGCTTAAGTTCAGCGGGTA
<i>Enterobacteriaceae</i>	rpoB1_ent-F	GAAGGTCC R AAYATCGGTCT
	rpoB1_ent-R	TGCATGTT C GCACCCAT
<i>Enterobacteriaceae</i>	rpoB2_ent-F	CTGAYGTTAAAGAY T CTTCTC
	rpoB2_ent-R	G A RAT W ACAC CY TTGTTACC
<i>Enterobacteriaceae</i>	rpoB3_ent-F	CGTATGAACAT Y GGTCAGAT
	rpoB3_ent-R	TT S AGTTTCAGCATRTACAT
<i>Enterococcus</i>	rpoB_ess-F	GTT R TAAC+ CKT +CCCA W GT C AT
<i>Staphylococcus</i>	rpoB_ess-R1	GCGACAGCATGTAT Y CCGTT C
<i>Streptococcus</i>	rpoB_ess-R2	G C RACAGCATGTATT C WTT C
<i>Raolutella ornithinolytica</i>	Raoult1	TATTATTCCTTACCGCGGT
	Raoult2	GCTTTCTCGACGTAGACTT
<i>Granulicatella adiacens</i>	Granada1	AAGATAGAGTTGTTATTCAA
	Granada2	TTGGTCCTAATGTTACTT

2.3.5 Etablering av PCR-primere med adaptersekvenser

Etter at de fire PCR-metodene, 16S -PCR, ITS-2 PCR, rpoB_ent-PCR og rpoB_ess-PCR var etablert som sanntids-PCR metoder, ble de tilpasset bruk for massiv parallell sekvensering. Primersekvensene ble beholdt, men forlenget med en adaptersekvens i 5'enden.

Tabell 4: Oversikt over primere med adapteroverheng. Sekvensen til spesifikk primer for amplikon-PCR er uthevet

Primer	Primersekvens (5' → 3')
16S-F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG CCTACGGGNGGCWGCAG
16S-R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG GACTACCAGGGTATCTAAKCC
ITS2-F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG GTGAATCATCGARTCTTTGAA
ITS2-R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG TATGCTTAAGTTCAGCGGGTA
rpoB_ent-F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG GAAGGTCCRAAYATCGGTCT
rpoB_ent-R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG TGCATGTTCCGACCCAT
rpoB_ess-F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG GTTRTAAC+CKT+CCCAWGCAT
rpoB_ess-R1	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG GCGACAGCATGTATYCCGTTCC
rpoB_ess-R2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG GCRACAGCATGTATTCCWTTCC

PCR med adaptere ble sammenlignet med original PCR uten adaptere.

Dette ble gjort ved å analysere kontrollstammer som er i daglig bruk ved MIA. Kontrollene består av oppdyrket *E.coli*, *C.albicans*, *E.faecium*, *E.faecalis* og *S.aureus* som er allikvotert og frosset ned i små porsjoner. Nytt rør er hentet fra frysen for hver nye kjøring, slik at sammenlignen skal bli så riktig som mulig.

For 16S-PCR var forskjellen mellom primerne i utgangspunktet stor (5-6 sykluser). Ulike temperaturprofiler og primerkonsentrasjoner ble testet, og problemet ble løst ved å velge en protokoll med lang annealingtid, 45 sykluser med 20 sekunder ved 95°C, 60 sekunder ved 60 °C og 30 sekunder ved 72 °C.

Innføring av adaptersekvenser i primerne påvirket også ct-verdien for ITS-2 og rpoB-PCR. Tidene som ble valgt i protokollen for 16S rRNA ble testet, og fungerte fint for alle PCR-metodene. Primerbindingstemperaturene på 58 °C og 59 °C for henholdsvis ITS-2 og rpoB_ess ble imidlertid beholdt.

På grunn av to tilfeller med uventet diskrepans mellom dyrkning og MPS ble det etablert to artsspesifikke PCR-metoder for å avklare mulig tilstedeværelse av *Raoultella ornithinolytica*

og *Granulicatella adiacens*. Primersekvensene er oppgitt i tabell 3. Kryssreagering med andre arter ble sjekket både med søk i Genbank og med testing mot beslektede mikrober. Metodene ble også kontrollert på vevsprøver med negativ 16S-PCR for å se etter kryssreaktivitet mot humant DNA.

2.4 Datainnsamling

2.4.1 Prøvetaking

Prøvene ble tatt med enten perkutan transhepatisk drenasje av galleblære/galleganger (PTHCS) eller endoskopisk retrograd kolangiografi (ERCP) ved røntgenavdelingen, Haukeland universitetssjukehus. Prøveglass, samtykkeerklæringer og rekvisisjonsskjema var på forhånd plassert der prosedyren skulle foregå. Galle/puss ble aspirert før innsprøyting av kontrastmiddel og sprøytet i et sterilt prøverør som ble sendt MIA.

2.4.2 Dyrkning

Prøvene ble dyrket ved ankomst til Mikrobiologisk laboratorium. Samtlige prøver ble sådd ut manuelt etter laboratoriets prosedyre for dyrkning av galleprøver. Det ble benyttet blodskål og laktoseskål som ble inkubert i CO₂ atmosfære, mens Brain Heart buljong ble inkubert i vanlig atmosfære. FAA («Fastidious anaerobe agar») og KV-skål (blodskål med vancomycin og kanamycin) ble dyrket anaerobt. Dette er undersøkelser som i mange av tilfellene ville blitt gjort, uavhengig av studien. Det har ikke vært konsekvent bruk av «stafylokokk-stripe» ved utsæd på blodskål. «Stafylokokk-stripe» benyttes for å tilføre vekstfaktorer som *Haemophilus* trenger for å vokse.

Dyrkningen måtte skje fortløpende etter hvert som prøvene ble tatt, og ble derfor utført av bioingeniørene som hadde vakt da prøven ankom laboratoriet. Identifisering av kulturene er utført av bioingeniører som har erfaring med arbeidet, i de fleste tilfellene er identifisering gjort ved bruk av massespektrometri på MALDI-TOF (Bruker, Tyskland). Supplerende biokjemiske tester som *ornitin karboksylase test* har vært tatt i bruk ved behov. I tillegg er det utført resistensbestemmelse for alle mikrober som er tillagt betydning. Det ferdige dyrkningsresultatet ble rapportert til rekvirenten som dermed kunne benytte resultatene i vurderingen av behov for antimikrobielle behandling. Resultatene er hentet fra laboratoriets datasystem Unilab-700 (Alphasoft, Tyskland) og benyttet i denne studien.

2.4.3 Nukleinsyreekstraksjon

Nukleinsyreekstraksjon ble utført fortløpende etter hvert som prøvene ankom laboratoriet.. 400 µL prøvemateriale ble tilsatt 400 µL Bacterial Lysis buffer (Roche) og en blanding av keramikk- og glasskuler (SeptiFast Lysis kit, Roche). En negativ kontroll ble inkludert ved hvert renseoppsett. Prøvene og kontrollene ble behandlet med kraftig risting 2 x 45 sekunder i en homogenisator (MagNA Lyser, Roche). Etter dette ble prøvene sentrifugert (16000 x G, 5 minutter), før 400 µL av supernatanten ble overført til MagNa Pure Compact (Roche, Mannheim, Tyskland) for automatisert nukleinsyreekstraksjon og rensing. «MagNa Pure Compact Nucleic Acid Isolation Kit I» (Roche) ble benyttet med protokollen «Bacteria_DNA_V3_2».

2.4.4 Kontroll av tilstedeværelse av bakterielt DNA i prøvene

Det ble utført 16S -PCR for å bekrefte at det fantes bakterielt DNA i alle prøvene. Renset DNA ble så oppbevart ved -80 °C frem til videre analysering. Sanntids-PCR for 16S ble utført i 25 µL reaksjonsrør på instrumentet *SmartCycler* (Cepheid, USA) med primere som på det tidspunktet ble benyttet i rutinediagnostikken ved MIA. PCR-miksen besto av 12,5 µL ExTaq SYBR mastermix (TaKaRA, Japan), 0.4 µM av hver primer, 8.5 µL PCR-gradert vann og 2 µL ekstrahert DNA. PCR-profilen inkluderte et trinn for polymeraseaktivisering i 30 sekunder ved 95 °C før 50 sykluser med 95 °C i 10sekunder, 70 °C i 15 sekunder og 72 °C i 20 sekunder.

Tabell 5 Primerse som er benyttet for å kontrollere at det er bakterielt DNA i prøvene

Målgen	Primer	Sekvens (5'→3')
16S rRNA, V3-V4	16S-f_mod	CGGCCAGACTCCTACGGGAGGCWGCA
	16S-r_mod	GCGTGGACTACCAGGTATCTAAKCC

PCR-resultatene ble senere benyttet til å velge antall sykluser i prøveprepareringen til massiv parallell sekvensering. For prøver med Ct-verdi under 20 ble det kjørt 35 sykluser, mens det for prøver med Ct-verdi over 20 ble kjørt 45sykluser. 18 av 20 prøver hadde Ct-verdi under 20.

2.4.5 Preparering av genbibliotek

Det ble utført PCR for amplifisering av de fire ulike genområder som beskrevet over for de 20 pasientprøvene og tilhørende negative kontroller. Amplifiseringen ble gjort på LightCycler 480 (Roche, Mannheim, Tyskland) med primere som inneholder adaptersekvenser i tillegg til spesifikke sekvenser rettet mot målgenet.

F-primer: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-spesifikk forward-primer

R-primer: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-spesifikk revers-primer

Tabell 6 PCR-betingelser for ampikon-PCR.

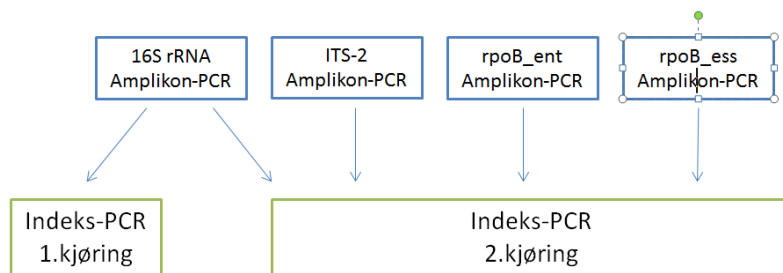
Målgen	Primer	Sekvens for primer som er spesifikk for målgenet (5'→3')	µM primer	Temperaturprofil
<i>16S rRNA</i> , V3-V4	16S-F	CCTACGGGNGGCWGCAG	0,4 µM	98 °C, 20s 60 °C, 60s 72 °C, 30s } x 35/45*
	16S-R	GACTACCAGGTATCTAAKCC	0,4 µM	
ITS-2	ITS2-F	GTGAATCATCGARTCTTTGAA	0,4 µM	98 °C, 20s 58 °C, 60s 72 °C, 30s } x 45
	ITS2-R	TATGCTTAAGTTCAGCGGGTA	0,4 µM	
<i>rpoB</i> , Enterobacteriaceae	rpoBent-F	GAAGGTCCRAAYATCGGTCT	0,4 µM	98 °C, 20s 60 °C, 60s 72 °C, 30s } x 40
	rpoBent-R	TGCATGTTCGCACCCAT	0,4 µM	
<i>rpoB</i> , <i>Staphylococcus</i> / <i>Streptococcus</i> / <i>Enterococcus</i>	rpoBess-F	GTTRTAAC+CKT+CCCAWGCAT	0,6 µM	98 °C, 20s 59 °C, 60s 72 °C, 30s } x 40
	rpoBess-R1	GCGACAGCATGTATYCCGTTC	0,4 µM	
	rpoBess-R2	GCRACAGCATGTATTCWTTC	0,4 µM	

«+» foran en base angir LNA (locked nucleic acid)

PCR ble utført i 96-brønnsplater i LightCycler 480 (Roche). PCR-miksen besto i tillegg til primere av 12,5 µL KAPA HiFi (KAPA biosystems), 5,5 µL PCR-gradert vann og 5 µL ekstrahert DNA.

Det ble først kjørt et oppsett der kun *16S rRNA*-genet ble sekvensert alene. Den totale analysekapasiteten ble fordelt på kun 24 prøver (20 pasientprøver og 4 kontroller) for å få informasjon om det som eventuelt finnes i meget små mengder i prøvene (dypere sekvensering). Det ble kjørt 45 PCR-sykluser for de negative prosesskontrollene og for de to prøvene med ct-verdi over 20, mens det ble kjørt 35 sykluser for resten av pasientprøvene.

I andre oppsett ble alle de fire PCR-produktene (ITS-2, rpoB_ent, rpoB_ess og 16S) sekvensert sammen. 16S-amplikon var med i begge oppsettene for å se konsekvensene av innføring av ekstra gener.



Figur 3: Flytskjema som viser forskjellen mellom de to analyseoppsettene.

Analysemetoden er meget følsom, og dermed utsatt for falskt positive resultater. Reagenser eller utsyr som er kontaminert med mikrobielt DNA vil feilaktig kunne bli tolket som funn i pasientprøvene. Kontrollene som har vært med fra nukleinsyrestraksjonstrinnet er med i hele analyseprosessen. Dersom hver enkelt blank skulle vært analysert separat ville det redusert antall prøver som kunne vært med i ett oppsett. Kontrollene ble derfor blandet etter utført amplikon-PCR. Etter at alle var blandet i ett rør ble innholdet splittet i fire. Det var da plass til å analysere 20 pasientprøver og 4 kontroller, med 4 gener hver $((20+4) \times 4 = 96)$.

PCR-produktene ble kontrollert med gelelektroforese. Det ble benyttet 1,5 % agarosegel, farget med etidiumbromid. Gelen ble undersøkt under UV-lys.

PCR-produktet ble rensert ved bruk av magnetkuler etter protokoll fra Illumina (Illumina, 2013). Det ble benyttet magnetiske kuler, Agencourt AMPure XP beads, fra Beckman Coulter Life Sciences sammen med 10mM Tris pH 8,5 –buffer (Qiagen) og 80% etanol. Rensingen ble utført i 96-brønnsplater. Amplikon som består av mer enn 100 nukleotider bindes til kulene, mens salter, enzymer og kortere nukleotider som primere, vaskes bort.

Indeks-PCR ble utført for individuell merking av hver prøve. Det ble benyttet 8 ulike forward-primere (indeks 2) og 12 ulike revers-primere (indeks 1) fra Illumina sitt *Nextera index kit*.

Tabell 7: Primere til indeks-PCR. Primerens 3' ende er komplementær til adaptersekvensen i produktet fra 1. PCR-reaksjon. 8 ulike indekssekvenser i5 og 12 ulike indekssekvenser i7 benyttes for å indeksere 96 ulike prøver samtidig.

	Sekvens (5'→3')
Indeks 1	CAAGCAGAAGACGGCATAACGAGAT[i7]GTCTCGTGGGCTCGG
Indeks 2	AATGATACGGCGACCACCGAGATCTACAC[i5]TCGTCGGCAGCGTC

PCR ble utført i 96-brønnsplater i LightCycler 480 (Roche). PCR-miksen besto av 5 µL indeks 1-primer, 5 µL indeks 2-primer, 25 µL 2 x KAPA HiFi (KAPA biosystems), 10 µL PCR-gradert vann og 5 µL PCR-produkt fra amplikon-PCR.

PCR-profilen inkluderte et trinn for polymeraseaktivering i 5 minutter ved 98 °C før 8 sykluser med 95 °C i 30 sekunder, 55°C i 30 sekunder og 72°C i 30 sekunder. Et ekstra elongeringstrinn ved 72 °C i 5 minutter var lagt til på slutten av protokollen.

Produktet etter indeks-PCR ble rensert med *Agencourt AMPure XP beads* på samme måte som produktet etter amplikon-PCR.

Størrelsen på PCR-produktene ble kontrollert ved å analysere hver prøve på Agilent 2100 Bioanalyzer. Det ble benyttet Bioanalyzer DNA 1000 chip som kan måle lengden til DNA-tråder mellom 25 og 1000 basepar.

Mengden DNA i hvert bibliotek (hver prøve) ble målt ved bruk av Qubit 3.0 Fluorometer (Fisher Scientific) der et fluorescerende fargestoff bindes til dobbeltrådig DNA (QubitR dsDNA HS Assay Kit *0.2–100 ng*).

Konsentrasjonen (nM) i hvert bibliotek ble kalkulert som beskrevet i Illuminas protokoll for 16S metagenomikk (Illumina, 2013). Det ble tatt utgangspunkt i en gjennomsnittlig produktlengde på 630 basepar (lengden på 16S-produktet inkludert primere og adaptere), etter følgende formel:

$$\frac{(\text{konsentrasjon i ng/}\mu\text{L}) \times 10^6}{660 \text{ g/mol} \times 630} = \text{konsentrasjon i nM}$$

Prøvene ble så fortynnet i buffer (10 mM Tris pH 8.5), slik at hver hadde en konsentrasjon på 4nM.

Bibliotekene ble så slått sammen til ett felles bibliotek med konsentrasjon 4nM. Denaturering ble gjort ved bruk av lut (NaOH). 5 μL bibliotek ble tilsatt 5 μL 0.2N NaOH. 10 μL denaturert bibliotek ble så blandet med 990 μL hybridiseringsbuffer (HT1, Illumina). Dette ga et 20 pM denaturert bibliotek i 1 mM NaOH. Biblioteket ble videre fortynnet til 4 pM, ved å blande 120 pM med 480 μL hybridiseringsbuffer.

PhiX-bibliotek fra Illumina ble fortynnet i buffer (10 mM Tris pH 8.5) til 4nM, denaturert med 0,2 N NaOH og så fortynnet med hybridiseringsbuffer til 4 pM.

PhiX-biblioteket ble tilsatt biblioteket med prøvene, slik at det endelige biblioteket inneholdt 5 % PhiX. Dette ble tilsatt Miseq-kassetten (Miseq Reagent Kit V3 (600 cycles)) etter denaturering ved 96 °C i 2 minutter. Prosessen i Miseq tok ca. 2,5 døgn.

2.4.6 Analysering av DNA-sekvenser

Data fra Miseq ble levert som FASTQ-filer, et tekstbasert format som i tillegg til nukleotidsekvensen også inneholder informasjon om tilhørende kvalitetsskår.

FASTQ-filene ble behandlet med bruk av dataprogrammet AdapterRemoval 2.1.7 (Lindgreen, 2012), der adaptersekvenser ble fjernet, og forward og revers sekvens slått sammen til én sekvens. Sekvenser med lav kvalitetsskår ble fjernet.

Følgende kommandoer ble benyttet:

```
AdapterRemoval --file1 <reads_1.fq> --file2 <reads_2.fq> --basename  
<mymergedfile> --threads 7 --trimns --trimqualities --minquality 20 --collapse --  
adapter-list <adapters>.txt --gzip
```

Sekvensene for hver prøve ble så behandlet videre i programmet RipSeq (Pathogenomix Inc., California, USA). Først ble FASTQ-filene lastet opp i RipSeq Preprocessor der sekvensene dannet klynger basert på likhet.

Følgende parametere ble benyttet:

```
Primer check length: 12  
Primer check max errors: 4  
Trim ends with no 3'end primer: 10  
Sequence length threshold: 300 (200 for ITS-2)  
Copy number threshold: 20  
Max cluster variation: 1 %
```

Det er altså tillatt en maksimal variasjon på 1 % fra den dominerende sekvensen i hver klynge. For 16S tilsvarer dette 4 nukleotider, da amplicon har en lengde på 430-440 bp. RpoB_ent amplicon er 328 bp, og 1 % vil da tilsvare at alle med inntil 3 ulike baser samles i samme klynge. RpoB_ess amplicon varierer noe, og er i overkant av 400 bp for stafylokokker, men i underkant av 400 bp for streptokokker. Variasjonen i ampliconlengden til ITS-2 er enda større og de minste er under 300 basepar. Det tillates dermed kun 2 basepar forskjell i hver klynge.

Filene ble så overført til Ripseq NGS, der resultatene ble sammenlignet med sekvenser i Genbank. Det ble satt cutoff på 99 %, det vil si at det må være minst 99 % likhet med en mikrobe i Genbank for at resultatet skal betraktes som gyldig. Den eller de mikrobene med størst likhet (over 99 %) foreslås. Det legges i tillegg inn et krav om hvor stor avstand det skal være til neste mikrobe på trefflisten i Genbank. Denne ble satt til 0,8 % for 16S, og 2,0 % for rpoB_ent, rpoB_ess og ITS-2.

2.4.7 Tolking av resultatene

Når det utføres massiv parallell sekvensering med tilstrekkelig dybde, vil «bakgrunns-DNA» også sekvenseres. Dette kan være DNA fra kontaminerte reagenser eller miljø.

Krysskontaminering mellom prøver og feillesing av sekvenser kan også påvirke resultatene. Det må vurderes hvilke sekvenser som representerer DNA fra pasientprøven, og hvilke som

skyldes ulike former for bakgrunn. Kriterier for vurdering av gyldigheten til et funn må fastsettes. I denne oppgaven er funn i prøvene sammenlignet med funn i de negative kontrollene. Dersom det er påvist en mikrobe i en pasientprøve som også finnes i kontrollene, er funnet ikke ansett som gyldig. Den mest dominerende bakgrunnsmikroben i en prøve ble brukt som cutoff, slik at alle senere mikrobefunn i mindre mengder også ble ekskludert uavhengig av om de var til stede i de negative kontrollene.

2.4.8 Statistiske beregninger

Deskriptiv statistikk og beregninger av statistisk signifikans er gjort med «SPSS Statistics, versjon 24». Det er benyttet Mann-Whitney test for rangsum for sammenligning av ulike grupper, Wilcoxons tegnrangtest (Wilcoxons signed rank test) for sammenligning av parvise data og chi-kvadrattest for å studere sammenhengen mellom ulike kategorier.

2.5 Ethiske betraktninger

Studien er godkjent av Regional komite for medisinsk og helsefaglig forskningsetikk (referansenummer 2015/65). Pasientene som inngår i studien har fått muntlig og skriftlig informasjon om studien, og gitt skriftlig samtykke til deltakelse. De har når som helst kunnet trekke seg fra studien uten å måtte oppgi grunn for det, og uten at dette vil få konsekvenser for behandlingen. Utarbeidelse av prosjektplan og gjennomføring av studien er gjort i henhold til helseforskningsloven (Helseforskningsloven, 2008, sist endret 2015). Studiedeltakerne er sikret full anonymitet ved at datafilen med resultatene er aidentifisert.

Indikasjonen for ERCP eller PTHCS har vært helt uavhengig av studien, da pasienter kun er inkludert i studien dersom de allerede var henvist til prosedyrene. Innsamling av prøvemateriale skal ikke ha gitt noen økt skaderisiko for pasientene. Selve DNA-sekvenseringen og tolkningen av resultatene er gjort etter at prøveinnsamlingen er avsluttet, og resultatene har dermed ikke fått konsekvenser for den enkelte pasient og er ikke videreformidlet behandlende lege eller lagt inn i pasientjournal. Resultatene vil derimot komme andre pasienter til gode senere dersom de er med å bidra til at retningslinjene for behandling oppdateres.

3 RESULTATER

3.1 Funn av mikrober i galleprøver fra pasienter med akutt kolecystitt og akutt kolangitt

Det er påvist 35 ulike arter ved massiv parallell sekvensering av 20 galleprøver fra 19 pasienter. *16S rRNA*-genet og *ITS-2*-gensegmentet er benyttet til påvisning og identifisering av en rekke mikrober på slekts- og artsnivå. *RpoB*-sekvenser er benyttet til identifisering i mange av tilfellene der *16S*-sekvensering ikke har kunnet skille på artsnivå. Tabell 8 viser hvilke bakterier som er påvist i hver prøve, antall sekvenser hver mikrobe er representert ved og relativ fordeling av disse i hver prøve. Det er angitt i tabellen hvilke mikrober som også er påvist ved dyrkning, og om pasienten har fått antibiotika siste 14 dager før prøvetaking.

Tabell 8: Oversikt over bakterier som er påvist med MPS i 20 galleprøver fra 19 pasienter. Prøve nr. 19 og prøve nr. 20 er fra samme pasient

ID	Bakteriefunn	MPS sekvenser	%-andel	dyrkning	prøve- taking	Anti- biotika
1	<i>Escherichia coli/Shigella</i>	140156	100,00	+	PTHCS	+
2	<i>Escherichia coli/Shigella</i>	21654	53,12	+	PTHCS	+
	<i>Klebsiella oxytoca</i>	18822	46,18	+		
	<i>Streptococcus i mitisgruppen</i>	286	0,70	-		
3	<i>Enterococcus faecalis</i>	104606	88,28	+	PTHCS	+
	<i>Klebsiella variicola</i>	13792	11,64	+		
	<i>Intestinibacter barlettii</i>	90	0,08	-		
4	<i>Fusobacterium nucleatum/canifelium</i>	90283	42,12	-	PTHCS	+
	<i>Streptococcus massiliensis</i>	46387	21,64	+		
	<i>Escherichia coli/Shigella</i>	45177	21,08	+		
	<i>Bilophila wadsworthia</i>	31811	14,84	-		
	<i>Klebsiella oxytoca</i>	453	0,21	-		
	<i>Enterococcus faecalis</i>	220	0,10	-		
5	<i>Klebsiella pneumoniae</i>	114789	95,95	+	PTHCS	+
	<i>Streptococcus salivarius/</i> <i>vestibularis/thermophilus</i>	4296	3,59	-		
	<i>Escherichia coli/Shigella</i>	523	0,44	-		
	<i>Enterococcus faecalis</i>	32	0,03	-		
6	<i>Streptococcus salivarius</i>	256754	100,00	+	PTHCS	+
7	<i>Klebsiella pneumoniae/quasipneumoniae</i>	85567	99,97	+	PTHCS	+
	<i>Escherichia coli/Shigella</i>	24	0,03	-		
8	<i>Escherichia coli/Shigella</i>	177838	100,00	+	PTHCS	+
9	<i>Streptococcus salivarius</i>	77874	100,00	+	PTHCS	+
10	<i>Klebsiella pneumoniae/quasipneumoniae</i>	77012	99,97	+	PTHCS	+
	<i>Escherichia coli/Shigella</i>	26	0,03	-		

11	<i>Enterococcus faecium</i>	132865	99,93	+	ERCP	+
	<i>Escherichia coli/Shigella</i>	68	0,05	+		
	<i>Clostridium perfringens</i>	25	0,02	-		
12	<i>Escherichia coli</i>	48911	84,68	+	ERCP	+
	<i>Fusobacterium periodonticum</i>	2841	4,92	-		
	<i>Streptococcus sanguinis</i>	2779	4,81	+		
	<i>Veillonella dispar/parvula</i>	2205	3,82	-		
	<i>Actinomyces naeslundii</i>	593	1,03	-		
	<i>Bifidobacterium dentium</i>	176	0,30	-		
	<i>Campylobacter concius</i>	150	0,26	-		
	<i>Granulicatella adiacens</i>	102	0,18	-		
13	<i>Enterococcus faecium</i>	69508	58,79	+	ERCP	+
	<i>Haemophilus parainfluenzae</i>	27152	22,97	-		
	<i>Streptococcus parasanguinis</i>	20763	17,56	-		
	<i>Lactobacillus rhamnosus/casei</i>	809	0,68	-		
14	<i>Escherichia coli/Shigella</i>	117521	99,84	+	ERCP	+
	<i>Klebsiella pneumoniae/quasipneumoniae</i>	184	0,16	-		
15	<i>Streptococcus mutans</i>	185586	83,60	-	ERCP	+
	<i>Cellulosimicrobium cellulans/funkei</i>	12116	5,46	+		
	<i>Staphylococcus aureus/haemolyticus/pasteuri/warneri</i>	9351	4,21	-		
	<i>Lactobacillus paracasei</i>	8216	3,70	+		
	<i>Haemophilus parainfluenzae</i>	3699	1,67	-		
	<i>Corynebacterium nuruki</i>	3030	1,36	-		
16	<i>Enterococcus faecium</i>	31940	90,19	+	ERCP	-
	<i>Rothia mucilaginosa</i>	2328	6,57	-		
	<i>Streptococcus salivarius</i>	1148	3,24	-		
17	<i>Staphylococcus aureus</i>	208347	94,19	+	ERCP	+
	<i>Streptococcus anginosus/constellatus/intermedius</i>	12846	5,81	-		
18	<i>Escherichia coli/Shigella</i>	221713	100,00	+	ERCP	+
19	<i>Escherichia coli/Shigella</i>	50507	66,75	+	ERCP	-
	<i>Clostridium perfringens</i>	16769	22,16	+		
	<i>Clostridium baratii/sardiniense</i>	4307	5,69	-		
	<i>Bifidobacterium bifidum</i>	1952	2,58	-		
	<i>Veillonella parvula</i>	1710	2,26	-		
	<i>Serratia marcescens/ureilytica</i>	215	0,28	+		
	<i>Actinomyces naeslundii</i>	117	0,15	-		
	<i>Enterococcus faecium</i>	89	0,12	-		
20	<i>Clostridium perfringens</i>	68665	75,13	+	PTHCS	+
	<i>Bifidobacterium bifidum</i>	17992	19,69	-		
	<i>Enterococcus faecium</i>	3136	3,43	+		
	<i>Escherichia coli/Shigella</i>	1017	1,11	+		
	<i>Actinomyces naeslundii</i>	498	0,54	-		
	<i>Clostridium baratii/sardiniense</i>	86	0,09	-		

E.coli er den hyppigst forekommende mikroben, tilstede hos 12 av 19 pasientene. *E.coli* er også den mest dominerende mikroben i 7 av prøvene.

Klebsiella pneumoniae/quasipneumoniae, *Streptococcus salivarius* og *Enterococcus faecium* dominerer innholdet i mer enn én prøve, mens *Staphylococcus aureus*, *Streptococcus mutans* og *Enterococcus faecalis* dominerer innholdet i hver sin prøve.

Det er gjort fire funn av sopp med MPS. Kun én av disse er påvist med dyrkning. Ingen av pasientene har fått behandling mot sopp.

Tabell 9: Oversikt over sopp som er funnet ved MPS i galleprøvene

	Mikrober	sekvenser	dyrkning	pr.taking
11	<i>Candida albicans</i>	38020	+	ERCP
	<i>Saccharomyces spp.</i>	1540	-	
12	<i>Candida albicans</i>	9370	-	ERCP
16	<i>Saccharomyces spp.</i>	2610	-	ERCP

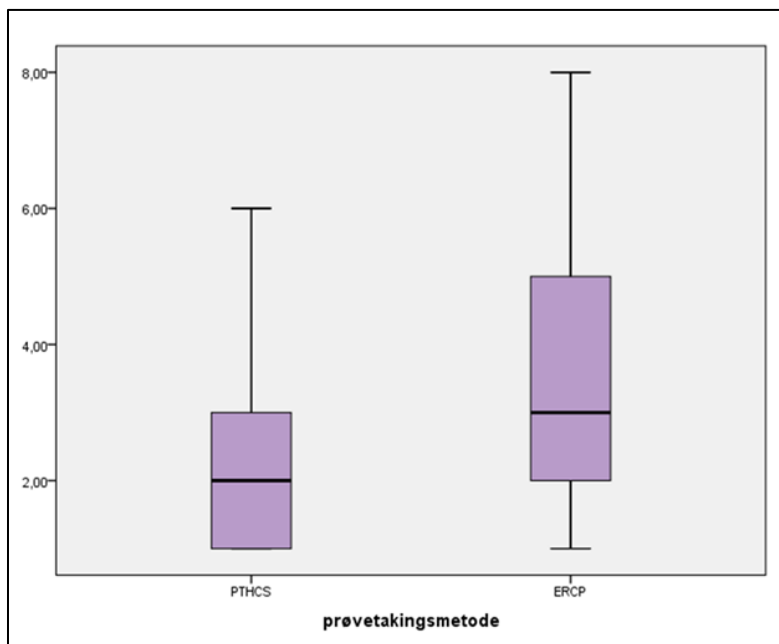
I tillegg til funnene som er presentert i tabellene er det identifisert to bakterieisolater ved dyrkning som ikke gjenfinnes ved MPS, *Raoultella ornithinolytica* i prøve 2 og *Granulicatella adiacens* i prøve 14. Det ble etablert spesifikke PCR-analyser for supplerende testing av de to prøvene. Begge analysene ga negativt resultat, og det ble konkludert med det ikke er DNA fra de aktuelle mikrobene i materialet som er undersøkt med MPS.

3.2 Mikrobefunn ved PTHCS versus ERCP

Prøvene som inngår i studien kan deles i to grupper, prøver tatt med PTHCS fra pasienter med akutt kolecystitt og prøver tatt med ERCP fra pasienter med akutt kolangitt. Materialet inkluderer også to prøver fra én pasient der første prøve var tatt med ERCP i forbindelse med kolecystitt, og den andre prøven med PTHCS i forbindelse med en post-ERCP kolangitt. Siden prøvene fra denne pasienten skiller seg fra de øvrige, både med hensyn til prøvetaking og etiologi, er de ekskludert fra den videre sammenligningen. Antall mikrober påvist i prøvene er sammenlignet for de to gruppene. Dataene presenteres i tabell 12 og i boksploTT (figur 4)

Tabell 10: Antall mikrober påvist med MPS, inndelt i grupper etter prøvetakingsmetode

<i>Antall mikrober per prøve</i>	<i>PTHCS</i>	<i>ERCP</i>
minimum	1	1
25-percentil	1	2
median	2	3
75-percentil	3	5
maksimum	6	8
Gjennomsnitt	2,4	3,6

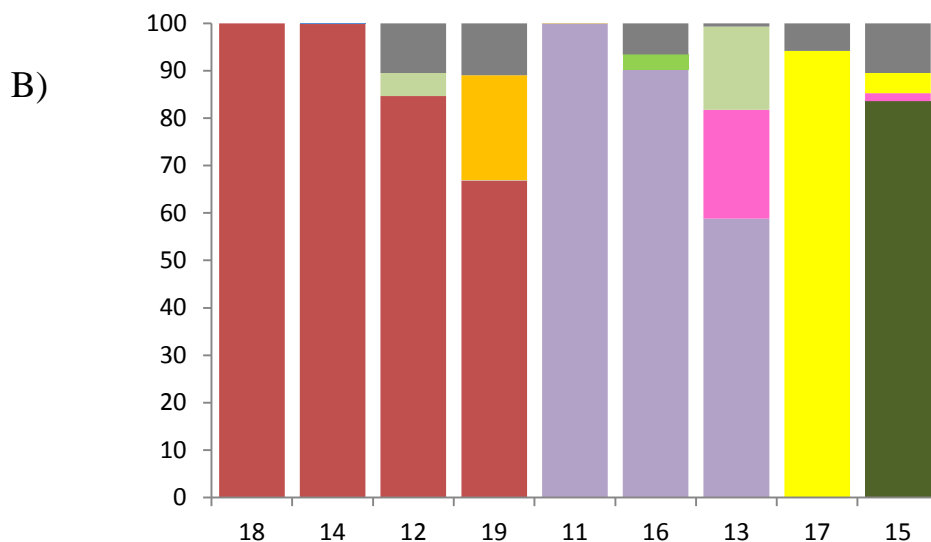
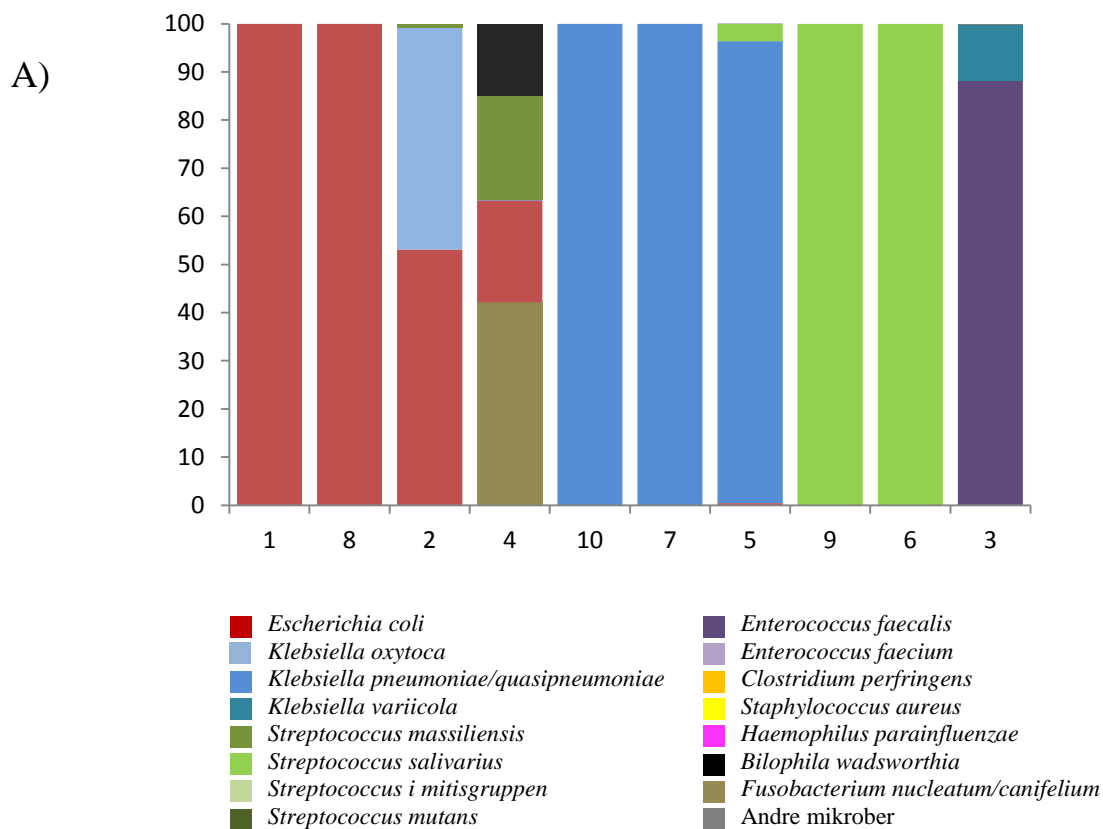


Figur 4: BoksploTTene gir en grafisk fremstilling av minimumsverdi, 25-percentil (Q1), median, 75-percentil (Q3) og maksimumsverdi. Median, Q1 og Q3 vises som henholdsvis midterste, nederste og øverste strek i boksen. Linjene som strekker seg ut av boksen går til største og minste observasjon.

Maksimumsverdi, median og gjennomsnittsverdi for antall påviste mikrober er høyest for gruppen av prøver tatt med ERCP. Ikke-parametrisk to-utvalgstest er benyttet for å undersøke om det er signifikant forskjell. Mann-Witney U-test av rangsum viser ikke at det er signifikant forskjell ($p = 0,237$).

Figur 5 viser hvilke bakterier som er til stede i galleprøvene fra de 19 pasientene, og prosentvis fordeling i den enkelte prøve. Som vist i figur 5a og 5b er *E.coli* dominerende i flere prøver, både blant dem som er tatt med ERCP og dem som er tatt med PTHCS. *Klebsiella pneumoniae/quasipneumoniae* og *Streptococcus salivarius* dominerer innholdet i henholdsvis tre og to prøver tatt med PTHCS, men gjenfinnes knapt i prøver tatt med ERCP. *Enterococcus faecium* dominerer i tre av prøvene som er tatt med ERCP, men er ikke til stede i noen av PTHCS-prøvene. Sopp er kun påvist i tre prøver; alle disse er tatt med ERCP.

Figur 5: Mikrober påvist med MPS i 19 prøver fra pasienter med akuttolecystitt og akutt kolangitt. Søylen viser fordelingen av ulike mikrober i hver prøve. Prøvene er gruppert etter prøvetakingsmetode.



Figur 5a: Bakterieinnhold i prøver tatt med PTHCS. Prøvene er sortert etter hvilke bakterier som dominerer. Mikrober som ikke har høyere andel enn 10 % i noen av prøvene er samlet under «andre mikrober».

Figur 5b: Bakterieinnhold i prøver tatt med ERCP. Prøvene er sortert etter hvilke bakterier som dominerer. Mikrober som ikke har høyere andel enn 10 % i noen av prøvene er samlet under «andre mikrober».

Det er også mange mindre populasjoner til stede i prøvene. Dersom en populasjon er til stede i mindre enn 10 % av en prøve, er den i denne studien definert som en liten populasjon.

Sammenhengen mellom prøvetaking og populasjonsstørrelse er studert ved å se på antall små populasjoner (tilstede i mindre enn 10 % av en prøve) og store populasjoner (tilstede i mer enn 10 % av en prøve) i prøver tatt med PTHCS hos pasienter med kolecystitt og ERCP hos pasienter med kolangitt. Dataene er presentert i krysstabell.

Tabell 11: Krysstabell for store/små populasjoner og ERCP/PTCHS

	<i>ERCP</i>	<i>PTHCS</i>	<i>Totalt</i>
Stor populasjon (>10 %)	10 (34,5 %)	15 (62,5 %)	25
Liten populasjon (<10 %)	19 (65,5 %)	9 (37,5 %)	28
Totalt	29 (100 %)	24 (100 %)	53

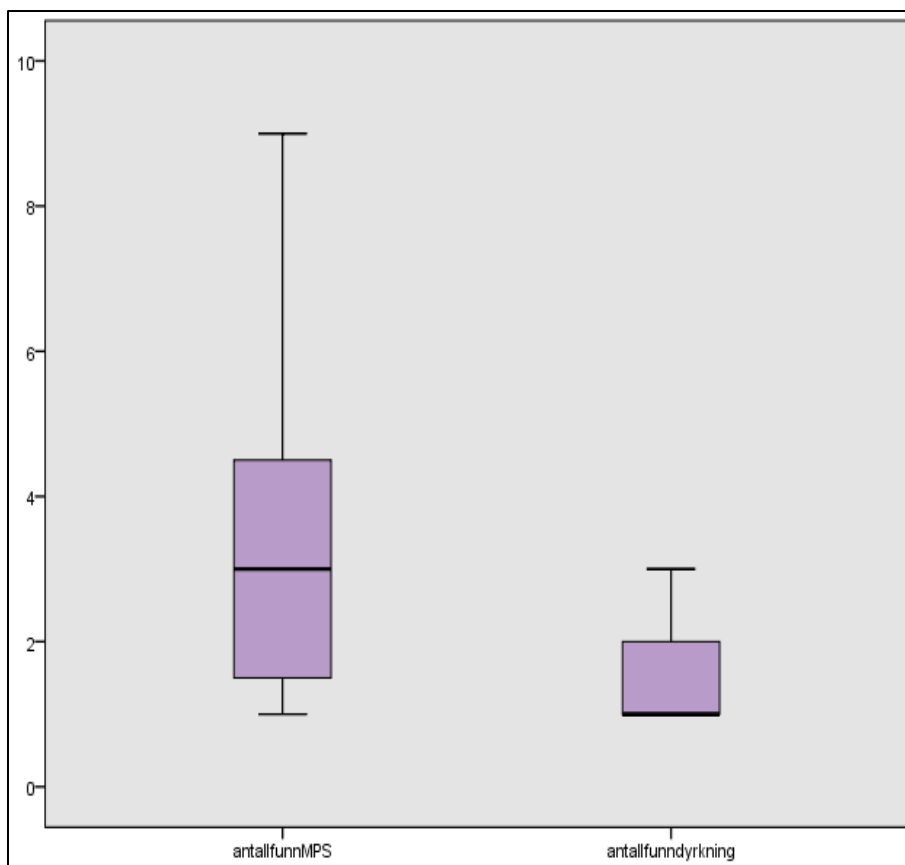
Kji-kvadrat-test (Pearson chi-square) er utført i SPSS, og viser at det er signifikant sammenheng mellom prøvetakingsmetode og de to populasjonsstørrelsene ($p=0,042$). Det er funnet flest små populasjoner i prøver tatt med ERCP. Prøvetaking via munnen og øvre delen av tynntarmen gjør disse prøvene mer utsatt for kontaminering, og funn av mange små populasjoner av munnhulebakterier kan tyde på at dette har vært tilfelle.

3.3 MPS sammenlignet med dyrkning

Det er påvist mikrober i samtlige prøver med både dyrkning og MPS. Ved dyrkning ble det påvist 1-3 ulike mikrober i hver prøve, mens det ved MPS ble påvist 1-9 mikrober.

Tabell 12: Antall ulike mikrober påvist per prøve ved dyrkning og MPS

<i>antall ulike mikrober per prøve</i>	<i>MPS</i>	<i>Dyrkning</i>
minimum	1,0	1,0
25-percentil	1,5	1,0
median	3,0	1,0
75-percentil	4,5	2,0
maksimum	9,0	3,0
Gjennomsnittsverdi	3,4	1,6



Figur 6: Boksplottene gir en grafisk fremstilling av minimumsverdi, 25-percentil (Q1), median, 75-percentil (Q3) og maksimumsverdi for antall funn per prøve. Median, Q1 og Q3 vises som henholdsvis midterste, nederste og øverste strek i boksen. Linjene som strekker seg ut av boksen går til største og minste observasjon. Q1 og medianen sammenfaller for antall dyrkningsfunn

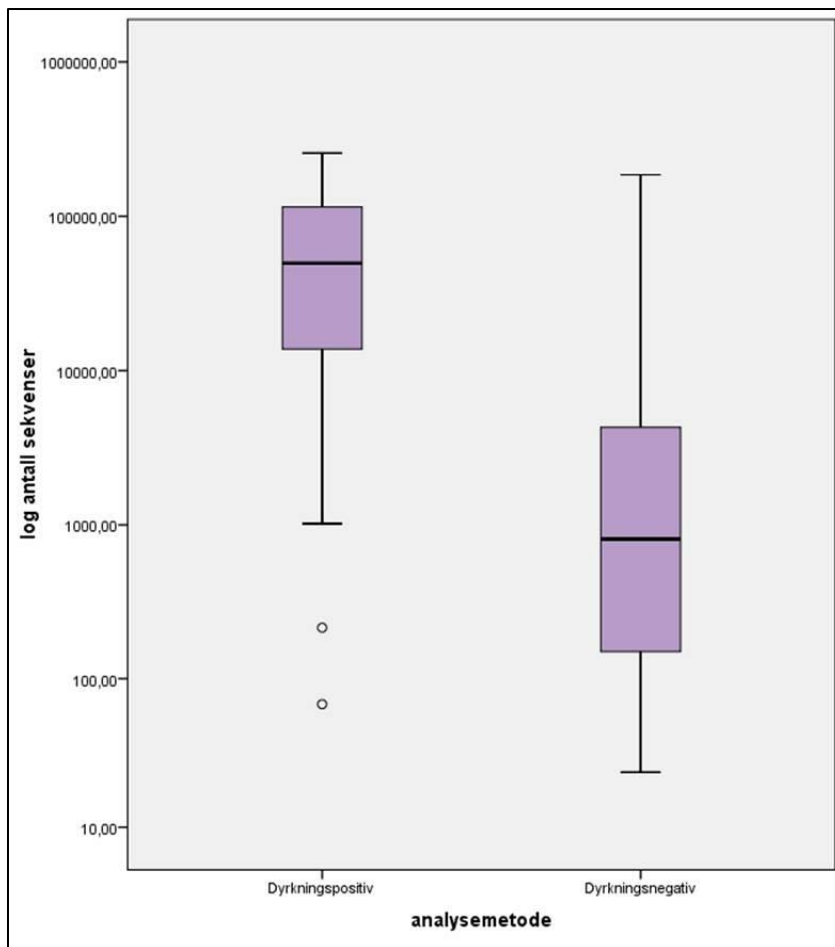
Ikke-parametrisk ettutvalgstest er benyttet for å undersøke forskjellen mellom antall mikrober i analyseparene (Altman, 1991) Det påvises flere mikrober i prøvene ved MPS enn ved dyrkning. Wilcoxons tegnrangetest (Signed rank test) for parvise data viser at forskjellen er signifikant. ($p = 0,002$).

I tillegg til forskjell mellom analysemetodene i antall påviste mikrober forekom også diskrepans mellom identifikasjonen som er gjort av isolatene i rutinediagnostikken ved MIA, og identifiseringen som er gjort med MPS som del av denne studien. Det ble påvist *Klebsiella variicola* i prøve 3 ved analysering av DNA-sekvensene. *Klebsiella*-isolatet i dyrkningsprøven var imidlertid identifisert som *Klebsiella pneumoniae*. Identifikasjon av isolatet var basert på massespektrometri (MALDI-TOF), og ved nærmere undersøkelse viste seg at databasen i MALDI-TOF ikke inneholdt data om *Klebsiella variicola* på det tidspunktet prøven opprinnelig ble analysert. Databasen ble oppdatert våren 2017, blant annet med informasjon om *Klebsiella variicola*. Massespekteret ble hentet frem og analysert på nytt, og resultatene viste nå treff med *Klebsiella variicola*.

Det er stor variasjon i antall sekvenser mikrobene er representert med. For å undersøke om det er sammenheng mellom antall DNA-sekvenser og dyrkningsresultat er to grupper sammenlignet, mikrober påvist ved dyrkning og mikrober som ikke er påvist ved dyrkning. Det er stor forskjell i både median og gjennomsnittsverdi for de to gruppene.

Tabell 13: Oversikt over antall sekvenser for dyrkningspositive og dyrkningsnegative mikrober

<i>Antall sekvenser</i>	<i>Dyrknings- positive</i>	<i>Dyrknings- negative</i>
minimum	68	24
25-percentil	13792	150
maksimum	256754	185586
75-percentil	114789	4296
median	49709	809
Gj.snitt	72491	11562



Figur 7: Boksplottene gir en grafisk fremstilling av minimumsverdi, 25-percentil (Q1), median, 75-percentil (Q3) og maksimumsverdi. Median, Q1 og Q3 vises som henholdsvis midterste, nederste og øverste strek i boksen. Linjene som strekker seg ut av boksen går til største og minste observasjon. Sirklene representerer observasjoner som ligger mer enn 1,5 interkvartilbredde under Q1.

Ikke-parametrisk to-utvalgstest er benyttet for å vurdere dataene. Mann-Whitney U-test for rangsum viser at det er signifikant forskjell i antall sekvenser mellom gruppen av mikrober som påvises ved dyrkning og gruppen som ikke påvises ved dyrkning ($p < 0,001$). Mikrobene som er påvist ved dyrkning er representert ved et større antall sekvenser enn mikrobene som ikke er påvist ved dyrkning.

Det er likevel tilfeller med høyt sekvensantall som ikke lar seg dyrke. MPS viser at 83,6 % av innholdet i prøve 15 består av DNA fra *Streptococcus mutans* og at 42,12 % av innholdet i prøve 4 består av DNA fra *Fusobacterium nucleatum*/*Fusobacterium canifelium*. Ingen av disse er påvist ved dyrkning.

Det ble påvist 9 ulike obligat anaerobe mikrober ved MPS. Kun en av disse, *Clostridium perfringens*, ble påvist ved dyrkning. Det ble påvist *Haemophilus parainfluenzae* i to prøver, uten at disse ble påvist ved dyrkning. Av 10 streptokokkfunn var kun 4 påvist ved dyrkning. Fullstendig oversikt over hvilke mikrober som vokste finnes i Tabell 8 og Tabell 9.

3.3.1 RpoB-sekvensering

Det er ikke mulig å identifisere alle bakteriene på artsnivå ved bruk av 16S-sekvensering. RpoB-sekvensering har derfor blitt utført som et supplement. Dette bidro til identifisering på artsnivå av ytterligere 14 bakteriefunn. RpoB-sekvensering bidro også til reduksjon av antall alternativer for noen av bakteriene som ikke lot seg identifisere på artsnivå

Tabell 14: Oversikt over bakterier der rpoB-sekvensering har hatt betydning for identifiseringen

16S rRNA	rpoB
<i>Streptococcus oralis</i> / <i>Streptococcus parasanguinis</i> / <i>Streptococcus australis</i> / <i>Streptococcus oligofermentans</i> / <i>Streptococcus rubneri</i> / <i>Streptococcus cristatus</i>	<i>Streptococcus parasanguinis</i>
<i>Streptococcus salivarius</i> / <i>Streptococcus vestibulari</i> / <i>Streptococcus thermophilus</i>	<i>Streptococcus salivarius</i>
<i>Staphylococcus aureus</i> / <i>Staphylococcus simae</i>	<i>Staphylococcus aureus</i>
<i>Enterococcus durans</i> / <i>Enterococcus faecalis</i> / <i>Enterococcus faecium</i> / <i>Enterococcus hirae</i> / <i>Enterococcus ratti</i> / <i>Enterococcus villorum</i>	<i>Enterococcus faecium</i>
<i>Lactobacillus casei</i> / <i>Lactobacillus paracasei</i> / <i>Lactobacillus rhamnosus</i> / <i>Lactobacillus zeae</i> / <i>Lactobacillus lactis</i>	<i>Lactobacillus paracasei</i>
<i>Veillonella parvula</i> / <i>Veillonella tobetsuensis</i>	<i>Veillonella tobetsuensis</i>
<i>Escherichia albertii</i> / <i>Escherichia coli</i> / <i>Escherichia fergusonii</i> / <i>Shigella boydii</i> / <i>Shigella flexneri</i> / <i>Shigella sonnei</i>	<i>Escherichia coli</i> / <i>Shigella</i>
<i>Klebsiella pneumoniae</i> / <i>Klebsiella quasipneumoniae</i> / <i>Klebsiella variicola</i>	<i>Klebsiella pneumoniae</i> / <i>Klebsiella quasipneumoniae</i>
<i>Klebsiella pneumoniae</i> / <i>Klebsiella variicola</i>	<i>Klebsiella variicola</i>
<i>Klebsiella oxytoca</i> / <i>Enterobacter gergoviae</i> / <i>Salmonella enterica</i>	<i>Klebsiella oxytoca</i>
<i>Klebsiella aerogenes</i> / <i>Enterobacter asburiae</i> / <i>E. cancerogenes</i> / <i>Enterobacter cloacae</i> / <i>Enterobacter hormaechei</i> / <i>Enterobacter ludwigii</i> / <i>Enterobacter xiangfangensis</i> / <i>Klebsiella pneumoniae</i> / <i>Klebsiella oxytoca</i> / <i>Klebsiella michiganensis</i> / <i>Klebsiella variicola</i>	<i>Klebsiella pneumoniae</i> / <i>Klebsiella quasipneumoniae</i>

RpoB-sekvensering gjorde det mulig med identifisering til artsnivå for alle *Enterococcus*, *Streptococcus* og *Staphylococcus* som hadde positiv PCR for rpoB_ess (resultater fra

Bioanalyser finnes i vedlegg A). RpoB-sekvenser manglet imidlertid for seks av funnene som var påvist med 16S-sekvensering. Dette gjaldt funn som var til stede i liten mengde i prøvene. Sensitiviteten til rpoB_ess er trolig ikke optimal og bør optimaliseres.

Det var i flere tilfeller også mulig å identifisere *Enterobacteriaceae* til artsnivå ved hjelp av rpoB. Det var ikke mulig å skille mellom *Escherichia coli* og *Shigella* eller mellom *Klebsiella pneumoniae* og *Klebsiella quasipneumoniae*, men antall ulike alternativer ble redusert i forhold til 16S-sekvensering. I tillegg til identifisering av bakteriene metoden var designet for, ble også *Lactobacillus paracasei* og *Veillonella tobetsuensis* identifisert med rpoB-sekvensering.

3.3.2 ITS-sekvensering

Det ble påvist sopp med MPS i tre av prøvene (Tabell 9). Én prøve inneholdt *Saccharomyces*, én inneholdt *Candida albicans* og én prøve inneholdt både *Saccharomyces* og *Candida albicans*. *Candida albicans* lot seg kun dyrke fra prøven med høyest antall ITS-sekvenser.

3.4 Utbytte og kvalitet

Dypsekvensering ble gjort i to oppsett. I første oppsett ble kun *16S rRNA*-genet i de 20 prøvene undersøkt, i tillegg til 4 negative kontroller. I andre oppsett ble 16S, rpoB_ent, rpoB_ess og ITS-2 i hver av de 20 prøvene undersøkt, med halvparten av kapasiteten benyttet til 16S-sekvensering og den andre halvdelens likt fordelt på sekvensering av de tre andre genområdene. Første oppsett hadde derfor flere sekvenser per prøve/gen, og ga dermed en dypere sekvensering.

Tabell 15: Antall leste sekvenser per prøve i de to oppsettene, og hvor mange av disse som regnes som gyldige sekvenser etter bearbeiding med AdapterRemoval og Ripseq.

	1.oppsett			2.oppsett		
	Totalt antall sekvenser	Antall sekvenser benyttet	Andel sekvenser benyttet	Totalt antall sekvenser	Antall sekvenser benyttet	Andel sekvenser benyttet
Prøve 1	516393	304627	59,0 %	172343	140156	81,3 %
Prøve 2	204109	99507	48,8 %	60103	40762	67,8 %
Prøve 3	2267758	1392192	61,4 %	153867	118488	77,0 %
Prøve 4	1776863	979443	55,1 %	333781	214331	64,3 %
Prøve 5	507685	271833	53,5 %	167799	119640	71,4 %
Prøve 6	504536	248849	49,3 %	338663	256754	75,8 %
Prøve 7	298070	163768	54,9 %	115433	85591	74,1 %
Prøve 8	377758	220231	58,3 %	225017	177838	79,0 %
Prøve 9	358940	199179	55,5 %	97669	77874	79,7 %
Prøve 10	440849	250067	56,7 %	106272	77038	72,5 %
Prøve 11	484267	297407	61,4 %	164075	132958	81,0 %
Prøve 12	149745	63714	42,5 %	84725	57757	68,2 %
Prøve 13	375532	183034	48,7 %	203010	120219	59,2 %
Prøve 14	414469	235939	56,9 %	152870	117725	77,0 %
Prøve 15	336807	162261	48,2 %	305241	222767	73,0 %
Prøve 16	163401	52068	31,9 %	81491	39497	48,5 %
Prøve 17	242034	129970	53,7 %	286171	221213	77,3 %
Prøve 18	464299	271225	58,4 %	272763	221713	81,3 %
Prøve 19	510095	251087	49,2 %	108516	75666	69,7 %
Prøve 20	615672	322827	52,4 %	615672	322827	52,4 %
gjennomsnitt	550464,1	304961,4	55,4 %	202274	142054	70,2 %

Første oppsett ga mer enn dobbelt så mange 16S-sekvenser som andre oppsett. Andelen sekvenser som benyttes etter bearbeiding med AdapterRemoval og Ripseq er derimot større i andre oppsett. Mens det i første oppsett var mulig å benytte 52,8 %, var det i andre oppsett

mulig å benytte 70,7 % av sekvensene. Det kan tyde på at dypere sekvensering gir flere sekvenser som ikke representerer mikrobene i prøven, noe som også støttes av analyseresultatene. De to analyseoppsettene gir like resultater med hensyn til hvilke mikrober som finnes i prøvene. Det ble riktignok funnet DNA-sekvenser fra noen flere mikrober i det første oppsettet, men alle disse var også påvist i de negative kontrollene og dermed ikke betraktet som gyldige funn.

Sannsynligheten for at et gitt nukleotid identifiseres riktig i Miseq oppgis av instrumentet som «Phred Quality score» (Q-score). Den gjennomsnittlige Q30-skåren sier noe om hvor stor andel av nukleotidene som har en forventet feilrate på mindre enn 1/1000. For de to Miseq-oppsettene er det oppgitt at henholdsvis 65, 1 og 65, 6 % av nukleotidene og har en verdi \geq Q30. Sekvensklyngetettheten er oppgitt til mellom 500 og 700 K/mm².

I Miseq er sannsynligheten for at et gitt nukleotid identifiseres riktig i oppgitt som «Phred Quality score» (Q-score). Den gjennomsnittlige Q30-skåren sier noe om hvor stor andel av nukleotidene som har en forventet feilrate på mindre enn 1/1000. 65,5 % av nukleotidene har skår <Q30. Selv om dette er noe lavere enn ønskelig betyr det ikke at kvaliteten på sekvensene som er benyttet i analyseringen er dårlig. Nukleotidene med dårlig skår er fjernet under behandling av sekvensene i AdapterRemoval og Ripseq. Om mange sekvenser er fjernet vil det imidlertid bety at utbyttet av sekvenser er redusert, noe som kan gå ut over sensitivitet.

3.5 Innhold av mikrobielt DNA i negative kontroller

Poolen av negative kontroller ble splittet i fire før indeks-PCR, ble etter dette behandlet og analysert hver for seg. Sekvensering viste at alle fire inneholdt DNA fra en rekke mikrober. De samme mikrobenes ble funnet i alle kontrollene, med tilnærmet lik relativ andel av hver. Selv om materialet er lite tyder dette på god repeterbarhet for delen av analysen som skjer etter PCR-trinnet, altså indeks-PCR, denaturering og sekvensering i Miseq.

Tabell 16: Oversikt over bakterier som ble påvist med MPS i de 4 negative kontrollene

Bakterier påvist i negative kontroller	kontroll 1		kontroll 2		kontroll 3		kontroll 4	
	reads	%-del	reads	%-del	reads	%-del	reads	%-del
<i>Sediminiabacterium salmoneum</i>	19827	32,39	44426	33,97	37573	33,54	52517	34,73
<i>Aquabacterium citratiphilum</i>	10783	17,61	24827	18,99	20321	18,14	28312	18,72
<i>Phenylobacterium</i> spp.	6652	10,87	11245	8,60	10907	9,74	13487	8,92
<i>Propionebacterium acnes</i>	4415	7,21	9424	7,21	8000	7,14	10591	7,00
<i>Ralstonia</i> spp.	3779	6,17	8005	6,12	6704	5,99	9276	6,13
<i>Bradyrhizium</i> spp.	3023	4,94	6274	4,80	5568	4,97	7027	4,65
<i>Pelomonas aquatica</i>	2822	4,61	6004	4,59	5171	4,62	6980	4,62
<i>Pseudomonas</i> spp.	2354	3,85	5010	3,83	4221	3,77	5620	3,72
<i>Staphylococcus</i> *	1539	2,51	3285	2,51	2844	2,54	3347	2,21
<i>Variovorax</i> spp.	1224	2,00	2599	1,99	2165	1,93	2863	1,89
<i>Microbacterium</i> spp.	860	1,40	1989	1,52	1661	1,48	2215	1,46
<i>Rudaea</i> spp	685	1,12	1426	1,09	1218	1,09	1582	1,05
<i>Agrobacterium tumefaciens</i>	652	1,06	1244	0,95	1104	0,99	1502	0,99
<i>Phyllobacterium</i>	453	0,74	946	0,72	833	0,74	1155	0,76
<i>Lactococcus lactis</i>	361	0,59	753	0,58	657	0,59	874	0,58
<i>Corynebacterium kroppenstedtii</i>	353	0,58	631	0,48	581	0,52	766	0,51
<i>Haemophilus haemolyticus/influenzae</i>	341	0,56	504	0,39	463	0,41	631	0,42
<i>Delftia acidovorans</i>	335	0,55	700	0,54	591	0,53	721	0,48
<i>Phenylobacterium koreense</i>	303	0,49	484	0,37	506	0,45	542	0,36
<i>Streptococcus</i> **	138	0,23	321	0,25	307	0,27	366	0,24
<i>Hydrotaea flava</i>	109	0,18	208	0,16	218	0,19	296	0,20
<i>Brevundimonas</i>	106	0,17	227	0,17	220	0,20	278	0,18
<i>Bacillus fusiformis/Lysinibacillus sphaericus</i>	65	0,11	122	0,09	107	0,10	154	0,10
<i>Afipia broomeae/felis</i>	42	0,07	79	0,06	70	0,06	103	0,07
<i>Acinetobacter</i>	0	0,00	29	0,02	0	0,00	30	0,02
Totalt	61221	100	130762	100	112010	100	151235	100

* 100% treff med *Staphylococcus capitis/caprae/epidermidis/hominis/warneri*

** 100% treff med *Streptococcus dentisani/infantis/mitis/oligofermentans/oralis/pneumoniae/pseudopneumoniae/tigurinus*

Tabell 17: Oversikt over sopp som ble påvist med MPS i de 4 negative kontrollene

Sopp påvist i negative kontroller	Kontroll1		kontroll 2		kontroll3		kontroll 4	
	reads	%-del	reads	%-del	reads	%-del	reads	%-del
<i>Malassezia restricta</i>	2419	43,20	2485	47,02	2767	45,66	3733	44,95
<i>Corallomycetella repens</i> / <i>Orbilbia luteorubella</i> / <i>Sarocladium kiliense</i> / <i>Sarocladium strictum</i> / <i>Ustilagoidea</i>	521	9,31	477	9,03	575	9,49	799	9,62
<i>Cryptococcus pseudolongus</i>	451	8,06	427	8,08	480	7,92	613	7,38
<i>Penicillium buchwaldii</i>	374	6,68	258	4,88	349	5,76	584	7,03
<i>Chrysosporium carmichaelii</i>	282	5,04	208	3,94	262	4,32	362	4,36
<i>Trichoderma hamatum</i> / <i>Trichoderma harzianum</i> / <i>Trichoderma longibrachiatum</i> / <i>Trichoderma orientale</i> / <i>Trichoderma</i>	282	5,04	226	4,28	299	4,93	406	4,89
<i>Phlebiopsis gigantea</i> / <i>Fibroporia gossypium</i>	278	4,97	314	5,94	325	5,36	417	5,02
<i>Capnodiales</i> sp. / <i>Chalaropsis</i> sp. / <i>Toxicocladosporium</i> sp.	206	3,68	204	3,86	248	4,09	304	3,66
<i>Cladophialophora bantiana</i> / <i>Cladosporium cladosporioides</i> / <i>Cladosporium langeronii</i>	187	3,34	136	2,57	84	1,39	264	3,18
<i>Aspergillus versicolor</i> / <i>Cladophialophora minourae</i> / <i>Cladosporium Ophiocordyceps sinensis</i> / <i>Sphaerulina polyspora</i> / <i>Passalora</i>	148	2,64	118	2,23	161	2,66	221	2,66
<i>Pseudogymnoascus pannorum</i> / <i>Pseudogymnoascus appendiculatus</i>	138	2,46	100	1,89	166	2,74	199	2,40
<i>Mucor plumbeus</i> / <i>Mucor racemosus</i>	137	2,45	102	1,93	113	1,86	134	1,61
<i>Candida famata</i> / <i>Debaryomyces fabryi</i> / <i>Debaryomyces hansenii</i> / <i>Debaryomyces prosopidis</i> / <i>Debaryomyces subglobosus</i> / <i>Debaryomyces vindobonensis</i> / <i>Candida psychrophila</i> / <i>Debaryomyces tyrocola</i>	104	1,86	94	1,78	147	2,43	152	1,83
<i>Cladophialophora carrionii</i> / <i>Cladosporium cladosporioides</i> / <i>Cladosporium halotolerans</i> / <i>Cladosporium herbarum</i> / <i>Cladosporium sphaerospermum</i> / <i>Lacazia loboi</i>	72	1,29	136	2,57	84	1,39	116	1,40
Totalt	5599	100,00	5285	100	6060	100	8304	100

3.5.1 Ekskluderte funn

Alle resultatene ble sammenlignet med funn i de negative kontrollene. Mikrobene som ble påvist i kontrollene ble definert som bakgrunnsflora (kontaminanter). Når det ble gjort funn av tilsvarende mikrober i pasientprøvene ble det tolket som bakgrunn. Alle mikrober som var til stede i pasientprøver i mindre mengde enn en «bakgrunnsmikrobe» ble også ekskludert.

I to av prøvene ble det gjort funn av DNA fra mikrober som også ble påvist i de negative kontrollene. I prøve 15 var det små mengder *Aquabacterium citratiphilum* som det finnes store mengder av i kontrollene. Prøve 16 inneholdt en rekke mikrober hvorav flere også er påvist i kontrollene. *Phenylobacterium koreense*, og alle mikrober som var til stede i mindre mengde enn denne, ble ekskludert.

Tabell 18: Mikrober som er påvist i prøvene, men som ikke regnes som reelle funn

Prøve	Ekskluderte mikrobefunn	Antall sekvenser	Andel i prøven
15	<i>Aquabacterium citratiphilum</i>	20	0,01 %
16	<i>Phenylobacterium koreense</i>	1020	2,40 %
	<i>Bacillus halodurans</i> *	849	2,0 %
	<i>Actinomyces odontolyticus</i> *	762	1,79 %
	<i>Pseudomonas</i>	661	1,55 %
	<i>Lactobacillus</i>	216	0,51 %
	<i>Propionibacterium acnes</i>	31	0,07 %
	<i>Aquabacterium citratiphilum</i>	29	0,07 %

* Er ikke påvist i negativ kontroll, men er under cut-off pga. påvist «bakgrunnsmikrobe»

4 DISKUSJON

Det er mangelfull kunnskap om hvilke mikrober som er til stede i galle hos pasienter med akutt kolecystitt og akutt kolangitt. Pasienter med infeksjon i galleblære eller galleganger får empirisk behandling etter retningslinjer basert på dyrkningsresultater. Dyrkningsmetoden har begrenset verdi da den er avhengig av levende mikrober som kan dyrkes i laboratoriet. Dyrkningsuavhengig diagnostikk i form av genteknologiske metoder er viktig for økt kunnskap om mikrobiologien ved denne type sykdom.

I denne studien har massiv parallell sekvensering av flere universelle gener blitt benyttet. *16S rRNA* og *ITS-2* skiller henholdsvis bakterier og sopp godt på slektsnivå, og i mange tilfeller også på artsnivå. Arter innen *Enterobacteriaceae*, *Staphylococcus*, *Streptococcus* og *Enterococcus* kan ha svært like *16S rRNA*-gener, og *rpoB*-sekvensering ble derfor etablert som et supplement. Nukleotidsekvensene i *rpoB*-genet varierer såpass mye mellom ulike slekter at det er umulig å designe ett sett med primere som detekterer alle. Det ble derfor designet ett primerpar for amplifikasjon av DNA fra bakterier i *Enterobacteriaceae*-familien og ett for amplifikasjon av DNA fra bakterier i *Staphylococcus*-, *Streptococcus*- og *Enterococcus*-slektene.

4.1 Etablering av PCR-metoder

Det er avgjørende for vellykket amplikonsekvensering at PCR-metodene fungerer slik at DNA fra alle mikrober i en prøve amplifiseres. Mens *16S*-PCR og *ITS-2* PCR er basert på publiserte og velprøvde primere, ble *rpoB*-PCR etablert som en del av denne studien, med egendesignede primere.

Mutasjoner i primerområdet vil alltid være en potensiell feilkilde for PCR-metoder. Konsekvensen vil blant annet avhenge av hvor i primerbindingsstedet mutasjonen er lokalisert. Det er avgjørende for en vellykket PCR-reaksjon at primerens 3'ende bindes til templatet slik at polymerasen kan katalysere dannelsen av ny DNA-tråd. Primerens lengde og nukleotidsammensetning har betydning for hvor høy temperaturen kan være når primeren bindes til templatet. Under metodevalideringen av *rpoB*- og *ITS*-PCR ble det undersøkt hvilke primerbindingstemperaturer som fungerte. For å ta høyde for eventuelle mismatches mellom primer og templat ble primerbindingstemperaturen senket noe i forhold til høyest fungerende temperatur. For *rpoB_ess* ble det i tillegg innført to LNA-baser i den ene primeren. Alt dette

ble gjort for å øke testens sensitivitet med tanke på antall arter som kan detekteres. Når sensitiviteten økes for å detektere DNA med mismatch mot primerne, må det også forventes uspesifikke PCR-reaksjoner og redusert spesifisitet. I og med at PCR-produktet skal sekvenseres ble det vurdert som akseptabelt med en viss reduksjon av primerspesifisiteten. Sekvensene til PCR-produktet vil uansett bli sammenlignet med gendatabaser, og sekvenser som ikke stammer fra mikrober vil kunne fjernes under analysering av data.

Det var ulike utfordringer knyttet til etablering av *rpoB*_ent-PCR og *rpoB*_ess-PCR. For *rpoB*_ent var det krevende å finne et område med høy nok diversitet til å skille mellom alle aktuelle arter. Til tross for at *rpoB*-genet har høyere mutasjonsrate enn *16S rRNA*-genet er den genetiske likheten stor mellom enkelte arter i *Enterobacteriaceae*-familien. Det var mulig å finne høyt konserverte områder til primere, slik at behovet for degenererte baser i primerne var lite, og vi lyktes med å etablere en PCR-metode med god effektivitet. Metodevalideringen viste en robust metode som fungerte ved ulike primerbindingstemperaturer. Det var derfor mulig å benytte primerbindingstemperatur på 60 °C i PCR-reaksjonen, som anbefalt i Illumina sin protokoll for amplikonsekvensering. Det var ikke problemer med kryssreagering til humant DNA, hverken under metodevalideringen eller i analyseringen av galleprøvene. Det var derimot noe kryssreaktivitet med andre bakterier i galleprøvene, *Haemophilus parainfluenzae*, *Enterococcus faecium* og *Veillonella*. Dette hadde ikke negativ effekt på identifiseringen av *Enterobacteriaceae*, men muligheten for bedret identifisering av flere bakterier enn planlagt.

Det var ønskelig med et felles primerpar for enterokokker, stafylokokker og streptokokker for å begrense kompleksiteten i laboratoriearbeidet. Det er stor variasjon mellom disse i *rpoB*-genet, og dermed vanskelig å finne konserverte områder til primere. Det var nødvendig med degenererte primere, noe som øker risikoen for kryssreaktivitet mot humant DNA. For å redusere antall primerkombinasjoner ble det laget to ulike reversprimere. Dette var mulig fordi flere av nukleotidkombinasjonene gikk igjen hos flere arter. Da *rpoB*-genet var et supplement til *16S rRNA*-genet, og heller ikke skulle brukes til kvantifisering, ble enkelte mismatches tillatt for å redusere kompleksiteten i primerdesignet. Dette gjaldt nukleotid nr. 16 fra 3' enden for *Streptococcus pyogenes*, *Streptococcus canis*, *Enterococcus casseliflavus*, *Enterococcus dispar* og *Streptococcus dysgalactiae*, som trolig ikke har stor betydning for PCR-reaksjonen. Det er også mismatch med *Staphylococcus caprae* og *Streptococcus suis*. Denne er lokalisert nærmere 3' enden (nukleotid nr.4), og har trolig større betydning for PCR-

reaksjonen. Disse mikroben er imidlertid knyttet til henholdsvis geit og gris, og til tross for at de sporadisk er påvist også hos mennesker ble de nedprioritert ved design av rpoB-PCR.

Sensitiviteten til rpoB_{ent}-PCR og rpoB_{ess}-PCR har betydning for hvilke bakterier som kan identifiseres til artsnivå. Utilstrekkelig sensitivitet vil ha alvorlige konsekvenser for 16S- og ITS-PCR som er fundamentet for påvisning av mikroben. Dette er også begrunnelsen for at det i denne studien ble valgt å benytte tre ganger så stor sekvenseringskapasitet på 16S som for hvert av de to rpoB-områdene.

Kryssreagering med humant DNA vil kunne påvirke sensitiviteten. For *16S rRNA*-genet er V3-V4 er valgt framfor V1-V3 i denne studien, til tross for at V1-V3 er kjent for å skille bedre mellom enkelte arter, blant annet streptokokker (Chakravorty et al., 2007). Dette ble gjort nettopp fordi faren for kryssreagering med humant DNA er større for V1-V3. Chun med flere viste i 2007 at bruk av «dual priming oligonucleotide principle» (DPO-primere) kan redusere problemet (Chun et al., 2007). Det er vist at DPO-primere kan benyttes til påvisning og sekvensering av V1-V3, og at området dermed kan benyttes uten stor fare for kryssreagering med humant DNA (O. Kommedal, Simmon, Karaca, Langeland, & Wiker, 2012). DPO-primere ble vurdert under planlegging av denne studien, men forkastet på grunn av usikkerhet rundt eventuell påvirkning av poly-inosin på sekvenseringen. Innføring av rpoB-sekvensering har dessuten redusert behovet for at 16S-sekvensering skal skille blant annet streptokokker til artsnivå.

16S rRNA-genet er en såkalt evolusjonær klokke, med en regelmessig mutasjonsrate som kan benyttes til å studere fylogeni, i tillegg til at genet er egnet for identifisering av bakterier. Dette er ikke tilfellet for ITS, «nuclear ribosomal internal transcribed spacer», som er et ikke-kodende område. ITS-segmentet er langt mer utsatt for større genetiske hendelser som innsetting og tap av lengre nukleotid-elementer, og både amplikon-lengde og variabilitet kan variere betydelig mellom ulike soppgrupper. Den varierende og uforutsigbare mutasjonsraten gjør at ITS ikke er egnet til fylogenetiske studier, men likevel kan benyttes som en «barkode» for identifikasjon. ITS-fragmentet er så variabelt at det ikke er mulig å finne konserverte områder til primere. Primerne er lagt til konserverte områder av henholdsvis 5S rRNA- og 28S rRNA-genene på hver sin side av ITS-fragmentet. Det er vanskelig å finne et gen som skiller alle sopper på artsnivå, og selv om ITS-2 har vist seg å fungere godt i de fleste tilfeller finnes det slekter der innen-artsvariasjonen for ITS-2 er liten. Dette gjelder blant annet

Aspergillus, der det kan være nødvendig med supplerende tester for å skille på speciesnivå (Schoch et al., 2012).

Valideringen av PCR-metodene viste i utgangspunktet at det var lite kryssaktivitet med humant DNA. Sekvenseringsresultatene viste likevel at RpoB_ess og ITS-2 hadde en tendens til å amplifisere humant DNA, men bare i prøver uten sopp og gram-positive kokker, det vil si uten det egentlige målgenet. Det er således snakk om beskjeden kryssreaktivitet som trolig ikke har påvirket effektiviteten i prøver som inneholder målgenet. Uspesifikke produkter av ulik lengde kan ha påvirket beregningene og fortynningene av de ulike prøvene i forbindelse med sammenslåing til felles genbibliotek. Før sammenslåingen måles DNA-konsentrasjonen i hver av prøvene, og det gjøres en normalisering slik at hvert bibliotek har en konsentrasjon på 4 nM. DNA-innholdet måles på Qubit, og er oppgitt som $\mu\text{g}/\mu\text{L}$. Molekylvekten må være kjent for at konsentrasjonen skal regnes om til nM. I denne studien ble det antatt at det var molekylene i hver prøve besto av gjennomsnittlig 630 nukleotider. Dette er lengden for produktet for 16S-PCR inkludert primere og adaptore. ITS-2 og rpoB-produktene er kortere enn dette, mens uspesifikke produkter som fragmenter av humant DNA kan ha ulik lengde. Resultatene fra Bioanalyser viser at mange prøver inneholder PCR-produkter som er lengre enn 630 nukleotider. Noen er helt i øvre måleområde (rundt 1500 nukleotider) og det er også mulig at noen er utenfor måleområdet. Dette kan ha gjort at normaliseringen som ble gjort for å oppnå 4 nM i hver prøve er basert på feil grunnlag, og at det er blitt skjev fordeling av prøvene i det felles biblioteket. Noen prøver kan dermed ha blitt sterkt fortynnet pga. innhold av humant DNA. For å kompensere for dette burde gjennomsnittlig molekylstørrelse vært målt for hver prøve på Bioanalyser, og benyttet i utregningene.

Under valideringen av rpoB_ess-PCR ga sanntids-PCR med SYBR master mix (TaKaRa) gode resultater. Galleprøvene ble undersøkt som en del av valideringen, og alle prøvene med gram positive kokker ga positiv rpoB-PCR. Da PCR ble utført med KAPA HiFi (KAPA biosystems) under prøveprepareringen til MPS var det kun synlige bånd på Bioanalyser for prøvene som hadde ct-verdi lavere enn 23 (vedlegg A og B). Dette kan tyde på at det har betydning hvilket enzym som benyttes. I etterkant av denne studien er det gjort forsøk med å erstatte KaPa HiFi med SYBR master mix (TaKaRa) i prøveprepareringen. Dette har gitt betydelig bedret sensitivitet for rpoB_ess.

4.2 Sekvensering og analysering av dataene

For at de store mengdene DNA-sekvenser skal kunne sorteres og identifiseres må det gjøres en rekke analyseinnstillinger. Det må settes krav til testens evne til å diskriminere mellom ulike mikrober. Dette involverer tekniske krav knyttet til inndeling i ulike klynger på bakgrunn av sekvenslikhet, men også krav til genetisk likhet med mikrober i referansematerialet (sekvensdatabaser) og til hvilken genetisk avstand det skal være mellom arter for at testen skal benyttes til å skille dem.

Prinsippet bak den bioinformatiske analysen er at lignende sekvenser samles i klynger, før den dominerende sekvenstypen i hver klynge sammenlignes med kjente sekvenser i en database med et standard BLAST-søk (Basic local alignment search tool).

Ved inndeling av DNA-sekvensene i separate klynger er det satt som kriterium i denne studien at det skal være maksimal variasjon på 1 %. Dette gjelder alle de fire genområdene som er benyttet. Variasjonen måles mot den dominerende sekvensen i klyngen. I tillegg til at dette tar høyde for innen-artsvariasjon vil det også kunne bidra i noen grad til å kamuflere feil leste sekvenser. Dersom det tolereres for stor variasjon i denne prosessen vil lignende arter kunne bli slått sammen i én sekvensklynge, og bare den mest dominerende bli identifisert. Dersom kravet til likhet settes for strengt vil en art gi flere klynger. Ved videre analysering med identifisering av mikroben vil dette oppdages. Det betyr altså ikke noe for resultatet, men øker arbeidsmengden .

Når sekvensene så skal tolkes for å identifisere mikroben må det settes krav til genetisk likhet med mikrober i referansematerialet (sekvensdatabaser) samt til hvilken genetisk avstand det skal være mellom arter for at testen skal kunne benyttes til å skille dem.

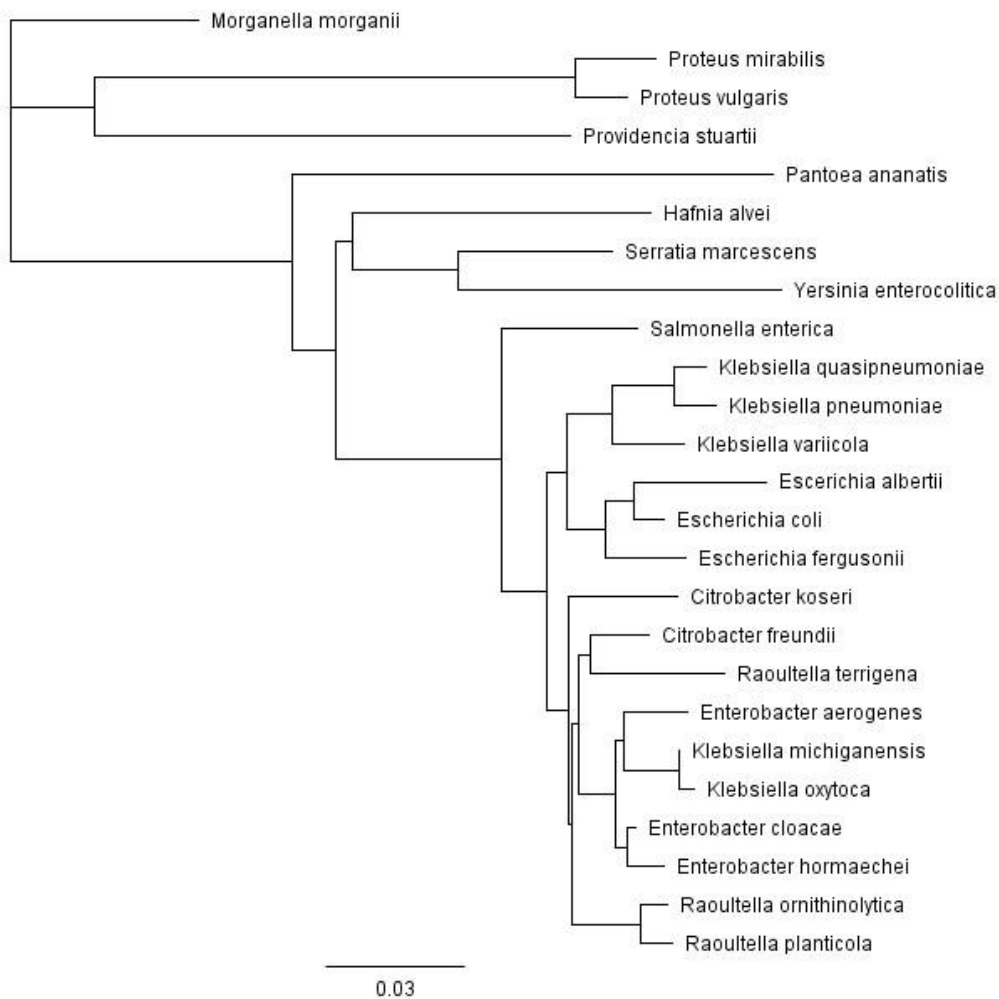
16S rRNA-genet er mye benyttet til sekvensering, og det finnes et stort antall sekvenser tilgjengelig i GenBank. Det er mye kunnskap om innen-artsvariasjon og om genetisk avstand mellom ulike arter. Ved MIA er det lang erfaring med analysering av *16S rRNA*, og analyseparameterne som benyttes for sekvenstolkning i rutinediagnostikken er også benyttet i denne studien. Cutoff for identifisering på artsnivå er satt til 99 %, mens krav til avstand mellom arter er satt til 0,8 %. Dette er også samme krav som anbefales av *Clinical and Laboratory Standards Institute* (CLSI, 2008).

For *rpoB*-genet finnes det langt færre sekvenser tilgjengelig, og kunnskapen om intra- og innen-artsvariasjon er mangelfull. Vi har derfor selv måttet definere hvilke grenser som skulle benyttes under sekvensanalysen. Da mutasjonsraten er høyere enn for *16S rRNA* kan en forvente at innen-artsvariabiliteten er større og at avstanden mellom ulike arter er høyere. Kravet til genetisk likhet ble likevel satt stringent til 99 %. Som tidligere nevnt risikerer en at ulike arter klynges sammen dersom kravet settes for lavt, og at en eller flere arter dermed kamufleres av en mer dominerende art. Vi anser det derfor som riktig å velge en trygg tilnærming når vi ikke har erfaring med metoden. Når det gjelder avstanden mellom mikrobene må denne antas å være høyere enn for 16S på grunn av høyere variabilitet i *rpoB*, og kravet til avstand bør derfor settes høyere enn 0,8 %. Vi har valgt å sette kravet til 2 % i denne studien. For *rpoB*_ent som har et PCR-produkt på 328 nukleotider tilsvarer dette en forskjell på 7 nukleotider. For *rpoB*_ess vil lengden av PCR-produktet variere noe. Det er store forskjeller mellom stafylokokker, streptokokker og enterokokker, ikke bare i baserekkefølge men også i lengden på genet. Det er bl.a. en innsats på 24 nukleotider i amplikonområdet hos stafylokokkene. Mens PCR-produktet for enterokokker og streptokokker er i underkant av 400 nukleotider er produktet i overkant av 400 nukleotider for stafylokokker, og avstanden på 2 % vil tilsvare henholdsvis 8 og 9 nukleotider. Når vi får mer erfaring med *rpoB*-testene kan det være aktuelt å gjøre endringer i kravene som er satt for sikker identifikasjon. Ervervet detaljkunnskap om reproduserbare forskjeller mellom nært beslektede arter kan gjøre det mulig å fravike disse generelle reglene.

4.2.1 Identifikasjon av arter innen *Enterobacteriaceae*-familien

Amplikonsekvenser for 27 arter innen *Enterobacteriaceae*-familien ble sammenlignet i et fylogenetisk tre generert i programvaren Geneious 9.1.2. Treet (Figur 8) viser at likheten mellom enkelte arter er stor. Det valgte genområdet er ikke egnet til å skille *Raoultella ornithinolytica* fra *Raoultella planticola*, *Serratia marcescens* fra *Serratia ureilytica* eller *Klebsiella pneumoniae* fra *Klebsiella quasipneumoniae*. Testen kan heller ikke benyttes til å skille fra *Enterobacter cloacae* fra *Enterobacter hormaechei*. For resten av mikrobene er forskjellen mellom 8 og 77 nukleotider og disse skal dermed kunne skilles med testen.

Figur 8: Fylogenetisk tre som viser forskjellen mellom de ulike arter basert på *rpoB*_ent-amplikon



For å undersøke om den utilstrekkelige forskjellen mellom enkelte arter skyldtes valget av genområde, ble det også laget et fylogenetisk tre basert på hele *rpoB*-genet (vedlegg C). Dette viste at det er mindre enn 2 % avstand mellom de aktuelle artene, også når hele genet sammenlignes.

Klebsiella-arter kan sjelden skilles fra hverandre ved bruk av 16S-sekvensering alene, noe som resultatene i denne studien også viser. Også *rpoB*-sekvensene var til dels vanskelig å klassifisere da det er store genetiske likheter mellom ulike arter innen *Klebsiella*, *Raoultella* og *Enterobacter*-slektene.

16S *rRNA*- og *rpoB*-genet ble sammenlignet for 9 *Klebsiella*-arter og 20 arter fra 11 andre slekter i *Enterobacteriaceae*-familien i en fylogenetisk studie i 2001. Basert på dette kunne *Klebsiella* deles inn i tre fylogenetiske linjer. Én linje (KI) besto av *Klebsiella pneumoniae*, én linje (KII) av *Klebsiella ornithinolytica*, *Klebsiella planticola* og *Klebsiella terrigena* mens den tredje linjen (KIII) besto av *Klebsiella oxytoca*. Det ble foreslått at KII burde skilles fra *Klebsiella* som egen slekt og reklassifiseres som *Raoultella ornithinolytica*, *Raoultella planticola* og *Raoultella terrigena*. I tillegg til at disse mikrobenes danner en egen fylogenetisk linje skiller de seg fra *Klebsiella* ved at de innehar egenskaper som muligheter for vekst ved 10 °C og at de kan benytte sorbose som karbonkilde (Drancourt, Bollet, Carta, & Rousselier, 2001).

Klebsiella pneumoniae har så blitt delt inn i fire grupper kalt KpI, KpII-A, KpII-B og KpIII (Brisse, van Himbergen, Kusters, & Verhoef, 2004; Brisse & Verhoef, 2001; Fevre, Passet, Weill, Grimont, & Brisse, 2005). I 2014 argumenterte Brisse med flere for å splitte *Klebsiella pneumoniae* opp i tre ulike arter og to nye underarter basert på disse gruppene. *Klebsiella pneumoniae* (KpI), *Klebsiella variicola* (KpIII), *Klebsiella quasipneumoniae* subsp. *quasipneumoniae* (KpII-A) og *Klebsiella quasipneumoniae* subsp. *similipneumoniae* (KpII-B) (Brisse, Passet, & Grimont, 2014).

Saha med flere beskrev *Klebsiella michiganensis* som egen art i 2013. De viste at nærmeste fylogenetiske slektning er *Klebsiella oxytoca* med 99 % likhet for 16S *rRNA* genot og at massespektrometriundersøkelse med MaldiToF viste størst likhet med *Klebsiella oxytoca*, *Klebsiella mobilis* og *Raoultella ornithinolytica*. (Saha, Farrance, Verghese, Hong, & Donofrio, 2013). Det finnes også noen arter som tidligere var regnet som *Klebsiella*, men som nå hører inn under andre slekter. I tillegg til reklassifiseringen av KII-arter som *Raoultella* gjelder dette *Klebsiella granulomatis* som har blitt reklassifisert som *Calymmatobacterium granulomatis*. I 2015 oppsummerte Janda *Klebsiella*-slekten til å bestå av følgende arter: *K.michiganensis*, *K.pneumoniae*, *K.quasipneumoniae*, *K.oxytoca*, og *K.variicola*. (Janda, 2015). Tindall med flere har imidlertid nylig hevdet at *Enterobacter aerogenes* egentlig skal klassifiseres som *Klebsiella aerogenes*. Det betyr også at *Klebsiella mobilis*, som er samme bakterie som *Enterobacter aerogenes*, også skal kalles *Klebsiella aerogenes*. (Tindall, Sutton, & Garrity, 2017) Det finnes flere underarter for *K.pneumoniae* og *K.quasipneumoniae*, men denne studien har ikke hatt som mål å skille disse.

RpoB-sekvensering gjorde det mulig å skille mellom *Klebsiella pneumoniae* (KI), *Raoultella*-arter (KII) og *Klebsiella oxytoca* (KIII). Det var også mulig å skille *Klebsiella variicola* (KpIII) fra *Klebsiella pneumoniae* (KpI) og *Klebsiella quasipneumoniae* (KpII), men basert på våre cutoff-verdier ikke mulig å skille mellom KpI og KpII. Imidlertid var det konsekvent en forskjell på 5 nukleotider mellom alle sikre *K. pneumoniae* og *K. quasipneumoniae* referanser i GenBank, så det er mulig man etter hvert, basert på erfaring, kan si at det er mulig å differensiere sikkert også mellom disse artene.

Identifisering av *Klebsiella* til artsnivå var utfordrende, til tross for at *rpoB*-sekvensering altså skal kunne skille mellom de fleste artene. 16S-sekvensering av *Klebsiella* i prøve nr. 3 ga 100 % treff med både *Klebsiella pneumoniae* og *Klebsiella variicola*. Sekvensering av *rpoB*-genområdet ga i første omgang også treff mot begge disse. Nærmere undersøkelser viste at det i Genbank var 100 % treff med kun seks *Klebsiella pneumoniae* *rpoB*-referanser, mens det store flertallet av *Klebsiella pneumoniae* referansene skåret 97 %. Det var langt flere 100 % treff med *Klebsiella variicola*. De seks aktuelle *rpoB*-sekvensene fra *Klebsiella pneumoniae* i Genbank ble gjennomgått for å se etter eventuelle feilidentifiseringer.

Tabell 19: Referanser i Genbank oppgitt som *K.pneumoniae*

Identifikasjon i Genbank	Stamme	Accession, Genbank
<i>K.pneumoniae</i>	342	CP000946
<i>K.pneumoniae</i>	lsy1	KR269877
<i>K.pneumoniae</i>	NGB-FR73	LC049150
<i>K.pneumoniae</i>	NGB-FR75	LC049152
<i>K.pneumoniae</i>	SB1	AM051186
<i>K.pneumoniae</i>	KP5-1	CP008700

To av stammene var helgenomsekvensert og publisert som *K.pneumoniae* 342 og *K.pneumoniae* KP5-1 i henholdsvis 2005 og 2008. Begge disse stammetypene har senere blitt identifisert som *Klebsiella variicola* i en fylogenetisk studie av klebsiellaisolater i 2015 (Garza-Ramos et al., 2015)

Klebsiella variicola ble regnet som egen art fra 2003, og det ble funnet at bakteriene som Brisse og Verhoef hadde karakterisert som KpIII tilhørte *Klebsiella variicola*. Sekvensen fra SB1-stammen er publisert som *Klebsiella pneumoniae* KpIII i 2005. I en publikasjon fra 2014 er SB1 navngitt som *Klebsiella variicola* (Brisse et al., 2014)

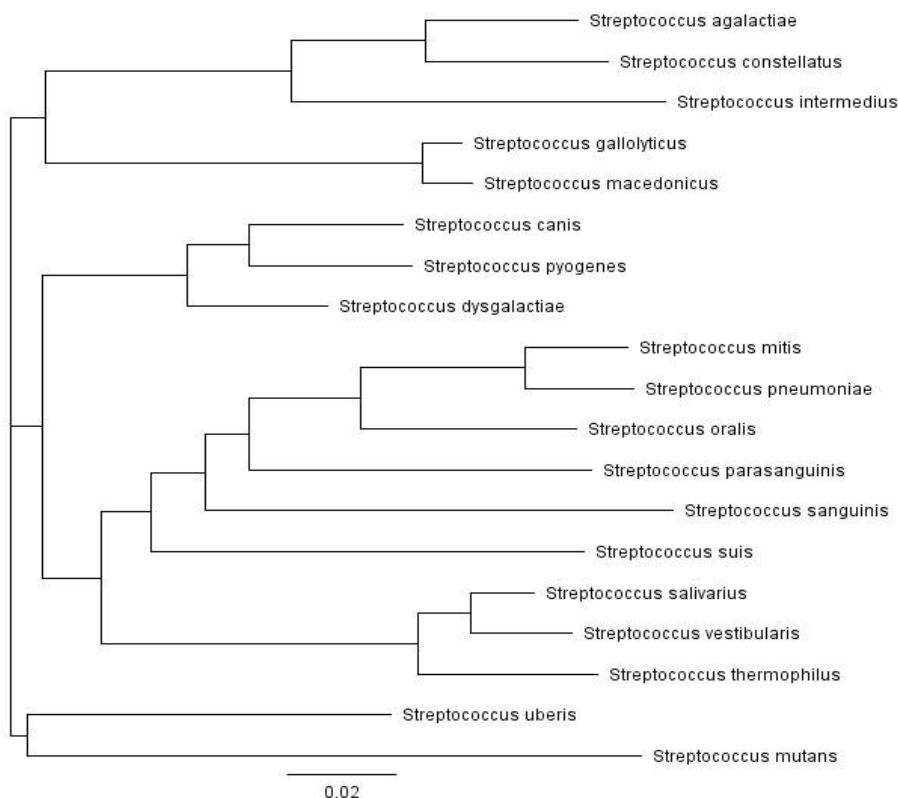
De tre siste sekvensene som er registrert i Genbank inngår ikke i publikasjoner. Alle er isolert fra planter, og har bare 97 % likhet med øvrige *Klebsiella pneumoniae*, inkludert typestammer. Det er derfor grunn til å konkludere med at disse også er registrert med feil navn i Genbank. Når de seks referansene fjernes fra søkelisten er det 100 % treff med *Klebsiella variicola*, og minimum 3 % avstand til neste mikrobe. Det konkluderes med at mikroben i prøve 3 er en *Klebsiella variicola*, og at metoden som er etablert for sekvensering av rpoB_ent er i stand til å skille mellom *Klebsiella variicola* og *Klebsiella pneumoniae*.

Eksempelet over er ett av flere som demonstrerer at de store taksonomiske endringene innenfor Enterobacteriaceae-familien har bidratt til at flere sekvenser i gendatabasene ikke har oppdaterte navn. Dette har gjort arbeidet med analysering av klebsiellasekvensene utfordrende. Sekvensene som har gitt treff i databasene har blitt studert grundig, blant annet med å lage fylogenetiske trær for å se på innbyrdes klynger og variasjon. Flere referanser har blitt forkastet da de trolig er gitt feil navn. I praksis viser det seg at det kun er mulig å stole på typestammer og stammer som er omtalt og vurdert i nyere publikasjoner.

4.2.2 Identifikasjon av bakterier innen *Enterococcus*-, *Streptococcus*- og *Staphylococcus*-slektene

Som tidligere beskrevet var sensitiviteten for rpoB_ess ikke optimal. Det manglet rpoB-sekvenser for seks av stafylokokkene og streptokokkene som ble identifisert til slektsnivå med 16S-sekvensering. Det er stor rpoB-variasjon mellom de ulike artene som ble oppgitt som alternativer ved 16S-sekvensering, og det ville vært mulig å identifisere samtlige til artsnivå dersom sekvenser hadde vært tilgjengelig. Det ble laget separate fylogenetiske trær for hver av de tre slektene for å vurdere avstanden mellom artene. Her representeres treet for streptokokkene. Fylogenetiske trær for enterokokker og stafylokokker finnes i vedlegg D.

Figur 9: Fylogenetisk tre som viser forskjellen mellom amplicon for ulike streptokokker



Det er store sekvensforskjeller mellom de aller fleste artene som er tatt med i sammenligningen, noe som viser at disse skal la seg identifisere med stor sikkerhet. Det kan imidlertid være vanskelig å skille *Streptococcus macedonicus* fra *Streptococcus gallolyticus*

og *Enterococcus gallinarum* fra *Enterococcus casseliflavus*. Ingen av disse problemstillingene var aktuelle i denne studien.

Det ble funnet små mengder av streptokokker i de negative kontrollene. Disse var med 16S-sekvensering identifisert til *Streptococcus dentisani /infantis /mitis /oligofermentans /oralis /pneumoniae /pseudopneumoniae /tigurinus*. Ved sammenligning med innholdet i pasientprøver kunne funnet ført til usikkerhet rundt gyldigheten av funn av streptokokker i mitisgruppen. DNA-sekvensene fra streptokokken i de negative kontrollene er imidlertid sammenlignet direkte med streptokokksekvensene i pasientprøvene; det er mer enn 4 nukleotider i forskjell, og dermed per definisjon ulike mikrober.

4.2.3 Kimærer

Kimærer kan komplisere analyseringen av sekvensene, og i verste fall føre til feilidentifisering. Kimærer er artefakter som består av sammenslåtte sekvenser fra to eller flere ulike genområder. Disse dannes som følge av ufullstendig ekstensjon i PCR-reaksjonen ved at uferdige amplicon bindes til et templat med lignende sekvens i neste PCR-syklus. Den fungerer da som primer og kimærer kan dannes. (Smyth et al., 2010). Lengre ekstensjonstid og høyere primerbindingstemperatur kan motvirke dette i noen grad. Det er benyttet 30 sekunder i denne studien, som anbefalt i produsentens protokoll. Primerbindingstemperaturen er 58 °C for ITS og 59 °C for rpoB_ess. Dette er lavere enn anbefalt arbeidstemperatur for enzymet, men det ble ikke observert større grad av kimæredannelse for disse enn for 16S og rpoB_ent der det har blitt benyttet primerbindingstemperatur på 60 °C. Det er trolig ikke mulig å bli helt kvitt problemet med kimærer, og en må derfor alltid ta dette med i betraktning når sekvensene analyseres. Særlig i prøver med flere nært beslektede bakterier kan andelen kimære bli høy.

Alle sekvenser som ikke har gitt 100 % treff med referansesekvenser har blitt studert grundig. Kimærer kan stort sett identifiseres ved at alle mismatch med referansesekvenser er lokalisert i samme ende av sekvensen.

Figur 10: Eksempel på prøve med to mikrober, mikrobe 1 og mikrobe 2. Dersom det dannes kimære av disse vil kimæren kunne identifisert som mikrobe 3, dersom denne gir treff med referansesekvenser.

Mikrobe 1	AGCTACCGTCA
Mikrobe 2	ACGTACCCACA
Mikrobe 3	ACGTACCGTCA

De fleste kimærene var til stede i små mengder, og forholdsvis enkle å identifisere da de ga dårlig treff mot referansesekvenser. I en av prøvene (prøve 2) var det imidlertid vanskelig å skille kimærene fra sekvenser som representerte reelle funn. Det ble påvist store klynger med *Escherichia coli* og *Klebsiella oxytoca*, men både 16S og rpoB-sekvensering genererte i tillegg mindre klynger med treff mot flere *Klebsiella*-, *Escherichia*- og *Kluyvera*-arter, med treffprosent mellom 97 % og 100 %. Alle sekvensene ble grundig studert, og det viste seg at alle mismatches mot referansesekvensene var lokalisert i samme område, og at reorganisering av sekvensene i dette området ga sekvenser med 100 % treff mot *E.coli* og *Klebsiella oxytoca*.

4.2.4 Vurdering av hvilke funn som er reelle

Massiv parallell sekvensering er en meget sensitiv analysemetode, noe som kan by på utfordringer når det skal avgjøres hvilke DNA-sekvenser som representerer reelle funn. Når universell PCR benyttes for å danne amplikon til sekvensering, vil det alltid være en fare for samtidig å amplifisere kontaminerende DNA i reagenser og miljø. Jo mer universell PCR-metoden er, jo større er sannsynligheten for dette. 16S-PCR og ITS-PCR er dermed spesielt utsatte metoder for falskt positive resultater.

Når funnene skal vurderes må det derfor benyttes kriterier for hvilke funn som eventuelt skal ekskluderes. I denne studien er negative kontroller benyttet som grunnlag for en slik «cutoff». Alle mikrober som er påvist i de negative kontrollene er regnet som kontaminanter, og funn av tilsvarende mikrober i pasientprøver er ekskludert. Det ble kun gjort funn av slike i to pasientprøver. I prøve 15 ble det påvist meget små mengder av *Aquabacterium citratiphilum*. Denne mikroben er en av de mest dominerende bakteriene i de negative kontrollene, og kan derfor regnes som en kontaminant i pasientprøven. Bakterien er heller ikke kjent som humanpatogen.

I prøve 16 ble det funnet flere bakterier som også er påvist i de negative kontrollene. 2,4 % av 16S-DNA besto av *Phenylobacterium koreense*, som er en av de mest dominerende bakteriene i de negative kontrollene. Alle bakteriefunn som var representert med mindre enn 2,4 % ble ekskludert på grunn av usikkerheten beheftet med disse, også de som ikke ble gjenfunnet i de negative kontrollene. *Bacillus halodurans* (2,0 %) og *Actinomyces odontolyticus* (1,8 %) ble dermed ikke regnet som gyldige funn. Det er altså valgt en streng tilnærming til godkjenning av funn. Det at disse to bakteriene ikke ble påvist i andre prøver kunne ha vært et argument for at kontaminering fra reagenser eller andre prøver er lite sannsynlig. Det er imidlertid brukt ulikt antall PCR-sykluser for prøvene i prøveprepareringen. Sanntids-PCR (16S) ble utført som kontroll på at det var DNA i prøven etter ekstraksjon, og ct-verdien ble benyttet til å bestemme hvor mange PCR-sykluser som skulle benyttes når det skulle dannes amplikon til sekvensering. Mens det ble benyttet 35 sykluser for prøver med ct-verdi under 20, ble det benyttet hele 45 sykluser for prøver med ct-verdi over 20. Prøve 16 var én av to prøver som ble preparert med 45 PCR-sykluser. Det er derfor mulig at sekvenseringen av denne prøven er «dypere» enn hos andre prøver, og at de svakt positive funnene dermed ikke kan sammenlignes med de andre pasientprøvene når det skal vurderes om de er reelle.

Det ble imidlertid benyttet 45 sykluser også på de negative kontrollene, og en kan tenke seg at alle kontaminanter derfor også vil finnes i disse. Som tidligere nevnt er det også svakheter forbundet med kontrollenes evne til å detektere kontaminanter. Kontrollene er tatt med fra ekstraksjonstrinnet og gir ingen informasjon om eventuell kontaminasjon under prøvetaking eller forsendelse. Kontrollene har også begrenset verdi når det gjelder kontaminering under forbehandling og ekstraksjon, da de ikke inneholder biologisk materiale.

Ekstraksjonsprotokollen er ikke beregnet for cellefattig materiale og det er usikkert om alt DNA som er til stede i kontrollen kan gjenfinnes i det rensede materialet. En bedre løsning ville vært å innføre «positiv negativ kontroll». En slik kontroll inneholder en kjent bakterie som ikke forventes å være i prøven. På den måten får man tilført biologisk materiale i kontrollen som kan «bære» eventuelt kontaminerende DNA gjennom renseprosessen. Ekstraksjonen vil skje på lignende premisser som for pasientprøvene, og sjansen for å avsløre eventuelle kontaminanter i ekstraksjonstrinnet vil trolig øke.

Våre kontroller vil i hovedsak avsløre kontaminerende DNA fra enzymer og reagenser som inngår i amplifikasjons og sekvenserings-reagenser. Vanligvis er dette de største kontaminasjons-kildene. Det ble påvist en rekke mikrober i kontrollene. Etter PCR-reaksjonen ble kontrollene blandet til en «pool» som så ble delt i fire. Resultatene viste at innholdet i alle fire var identisk, noe som tyder på at kontaminantene har vært i prøvene siden før de ble slått sammen. Når flere prøver er analysert vil det bli bedre grunnlag for vurdering av hva som vanligvis er kontaminanter og hva som kan regnes som reelle funn. Ved implementering av metoden i rutinediagnostikken vil det være hensiktsmessig å ta i bruk skjema der lot-nummer for alle reagenser noteres sammen med funn i kontrollene. Dette kan gi informasjon om hvilke kontaminanter som knyttes til ulike reagenser og hva som eventuelt må regnes som «bakgrunnsflora» i vårt laboratorium. Det vil også kunne gjøre annen feilsøking enklere.

Prøvene kan også kontamineres under prøvetaking. Prøvene i denne studien er tatt med to ulike prøvetakingsmetoder. Innholdet i prøver tatt med PTHCS fra pasienter med akutt kolecystitt er sammenlignet med innholdet i prøver tatt med ERCP fra prøver med akutt kolangitt. Bakteriefunnene ble delt inn i to populasjonsstørrelser, de som utgjør < 10 % av en prøve og de som utgjør > 10 % av en prøve. Statistisk testing viser signifikant forskjell mellom gruppene, med flest små populasjoner påvist i prøvene tatt med ERCP. En mulig forklaring på dette kan være økt kontaminasjonsfare som følge av at prøven tas via munnen og øvre del av tynntarmen. Det er imidlertid ikke mulig å knytte forskjellen med sikkerhet til prøvetakingen, da prøvene er tatt fra pasienter med ulike infeksjonssykdommer (kolecystitt og kolangitt). En stor andel små bakteriepopulasjoner som utgjør <10 % av innholdet i en prøve er heller ikke nødvendigvis noen god indikator på kontaminering. Det må analyseres langt flere prøver for å kunne konkludere om kontaminering under prøvetaking er et større problem for prøver tatt ved ERCP enn PTHCS.

4.2.5 Mikrobefunn og tilknytning til prøveidentitet

DNA-sekvensene er knyttet til prøveidentitet ved bruk av indekssekvenser tilført i andre PCR-reaksjon (indeks-PCR). Disse fungerer som barkoder bestående av 8 nukleotider. Dersom det skjer feil under lesing av en barkodesekvens vil DNA-sekvensen kunne knyttes til feil prøve. Feilprosenten øker vanligvis med lengden på sekvensen. Lesing av de korte indeks-sekvensene gjøres derfor i en separat reaksjon med egne primere i Miseq. For å redusere problemet ytterligere er det benyttet «dual-indexing», altså indeksmerking av begge

sekvensene i et par. Dette er vist å redusere feilidentifiseringen ytterligere (Kircher, Sawyer, & Meyer, 2012). Det vil likevel alltid være en viss sjanse for feillesing av indekssekvenser, eller at klynger overlapper hverandre (står for tett) i flow-cellen og at sekvensene «blør» slik at de knyttes til feil prøveidentitet. I denne studien har sekvensklyngetettheten vært lav ($<700 \text{ K/mm}^2$), og sannsynligheten for blødning mellom sekvensklyngene regnes som liten.

4.2.6 Sekvenskvalitet

Faren for feillesing av nukleotider og blødning mellom sekvenser øker når det er høy sekvenshomologi i et analyseoppsett. 16S-sekvensering for identifisering av ulike bakterier kan derfor være utfordrende. Det vil være hyppig frekvens av sykluser med likt nukleotid i alle prøver, noe som kan påvirke deteksjonen i instrumentet. For å redusere problemet er det vanlig å tilsette DNA med kjent sekvens og konsentrasjon. PhiX, som er en bakteriofag med et lite genom, er benyttet i denne studien. Selve phiX-sekvensen er ca. 300 bp, mens den totale lengden på fragmentet er ca 420 bp. Fragmentet er ikke indeksert, og kan dermed analyseres i tillegg til de 96 indekserte prøvene. I teorien burde innføring av *rpoB* og ITS i samme oppsett som 16S eliminere problemet med lav heterogenitet, slik at phiX ikke var nødvendig. Dette ville frigjort kapasitet i kjøringen slik at vi kunne kjørt flere prøver eller oppnådd flere sekvenser per prøve (bedre dybde) i denne studien. Det ble valgt å benytte 5 % phiX som anbefalt til 16S -sekvensering, til tross for at fire gener ble benyttet, for å begrense antall endringer fra Illumina sin protokoll for sekvensering av *16S rRNA* i denne omgangen.

I Miseq er sannsynligheten for at et gitt nukleotid identifiseres riktig i oppgitt som «Phred Quality score» (Q-score). Den gjennomsnittlige Q30-skåren sier noe om hvor stor andel av nukleotidene som har en forventet feilrate på mindre enn 1/1000. 65,5 % av nukleotidene hadde skår $<Q30$. Selv om dette er noe lavere enn ønskelig betyr det ikke at kvaliteten på sekvensene som er benyttet for identifisering er dårlig. Nukleotidene med dårlig skår er fjernet under behandling av sekvensene i AdapterRemoval og Ripseq. Om mange sekvenser er fjernet vil det imidlertid bety at utbyttet av sekvenser er redusert, noe som kan gå ut over sensitiviteten.

4.3 Dyrkning versus massiv parallell sekvensering

Langt flere mikrober ble påvist ved massiv parallell sekvensering enn ved dyrkning. Statistisk testing viste at forskjellen var signifikant ($P=0,002$ ved beregning med Wilcoxon's tegrangtest). Det ble til sammen gjort ni funn av obligat anaerobe bakterier. Kun én av disse, en *Clostridium difficile*, ble påvist ved dyrkning. Ved sekvensering ble i tillegg *Bifidobacterium bifidum*, *Bifidobacterium dentium*, *Bilophila wadsworthia*, *Clostridium baratii*, *Fusobacterium periodonticum*, *Fusobacterium nukleatum/canifelium*, *Intestinibacter barlettii* og to ulike *Veillonella*-arter påvist. I en av prøvene (prøve 4) var *Fusobacterium nucleatum/Fusobacterium canifelium* den mest dominerende mikroben (42,1 %). Det var også en betydelig andel *Bilophila wadsworthia* (14,8 %) i tillegg til *E.coli* og *Streptococcus massiliensis* i prøven. Dyrkning viste kun rikelig forekomst av *E.coli* og middels rikelig forekomst av *Streptococcus massiliensis*.

Haemophilus parainfluenzae ble påvist med DNA-sekvensering i to prøver til tross for at mikroben ikke har latt seg dyrke i noen av prøvene. Det var vekst på blodskål, FAA-skål og Brain-heart for begge prøvene, men det er uklart om *H.parainfluenzae* vokste og ble oversett på grunn av andre mikrober, eller om mikrobene ikke var levedyktige da de ble forsøkt dyrket. *Haemophilus* vokser ikke på vanlig blodskål, med mindre vekstfaktoren V-faktor gjøres tilgjengelig for eksempel gjennom å påføre en «stafylokokk-stripe». Dette er ikke benyttet for disse prøvene, noe som har redusert muligheten for å kunne dyrke *Haemophilus*. Det kan i teorien ha vært vekst i Brain-heart, sammen med andre mikrober fra prøven, uten at *Haemophilus* har blitt identifisert.

Selv om det i samtlige prøver ble påvist et like høyt eller høyere antall mikrober ved MPS enn ved dyrkning, ble det påvist og identifisert to mikrober ved dyrkning som ikke ble funnet ved MPS. Dette gjaldt *Raoultella ornithinolytica* i prøve 2 og *Granulicatella adiacens* i prøve 14. Dyrkning i anrikningsbuljong er svært sensitivt så lenge det finnes levende mikrober i prøvematerialet, mens en PCR i praksis krever en mikrobemengde på minimum 100 bakterier per milliliter prøvemateriale. Det er derfor absolutt en teoretisk mulighet til stede for at dyrkning kan påvise bakterier som er tilstede i så lave konsentrasjoner at de ikke gjenfinnes ved amplikonsekvensering. Dersom det finnes lite av en mikrobe i en prøve vil det alltid være usikkerhet forbundet med påvisning av denne.

Når det gjelder *Raoultella ornithinolytica* og *Granulicatella adiacens* var det imidlertid oppgitt at det var rikelig vekst ved dyrkning. Selv om dyrkningsmetoden ikke er egnet for kvantifisering gir det likevel uttrykk for at det har vært en viss mengde av mikroben til stede. Årsaken til at de ikke ble påvist ved MPS kan være at metoden ikke har fungert godt nok eller at mikroben er feilidentifisert. PCR-metodene skal i teorien fungere for disse mikroben. 16S rRNA skal kunne benyttes til begge, og *rpoB_{ent}* og *rpoB_{ess}* har primere som passer for henholdsvis *Raoultella* og *Granulicatella adiacens*. Det er lite trolig at det har vært mutasjoner i primerområdet til både 16S og *rpoB*. Det ble likevel utført spesifikk PCR for *Raoultella* og *Granulicatella adiacens*. Begge var negative, noe som bekrefter at det ikke var DNA fra disse mikroben til stede i eluatet som er benyttet til PCR. Vi konkluderer med at det er svært lite sannsynlig at det har vært rikelig med *Raoultella ornithinolytica* eller *Granulicatella adiacens* i disse prøvene.

I prøven der *Raoultella ornithinolytica* var påvist med dyrkning var det i tillegg *Klebsiella oxytoca* påvist med både dyrkning og MPS. Denne er nært beslektet med *Raoultella*, og det ble undersøkt om *Raoultella*-sekvenser og *Klebsiella*-sekvenser kunne ha havnet i samme sekvensklynge. Dersom ulike arter havner i samme klynge vil kun mikroben med den dominerende sekvensen bli identifisert. Tillatt klyngevariasjon ble endret fra 1 % til 0,5 % og 0,2 %. Antall sekvensklynger økte, men ingen av dem representerte *Raoultella*.

Raoultella ornithinolytica og *Klebsiella oxytoca* var i dyrkningsprøven identifisert ved bruk av massespektrometri og ornitin dekarboksylase test. Massespektrometri viste at begge isolatene kunne være enten *Raoultella ornithinolytica* eller *Klebsiella oxytoca*. Det ble konkludert med en av hver, på grunnlag av ornitin dekarboksylase testen. Vi konkluderer med at identifikasjonen av *Raoultella ornithinolytica* er beheftet med betydelig usikkerhet. Kombinert med resultatene fra sekvensering og PCR tror vi det heller dreier seg om en feilklassifisering av *Klebsiella oxytoca*.

4.4 Gevinst av *rpoB*- og ITS-sekvensering

De fleste mikroben kunne bestemmes til artsnivå i denne studien. *RpoB*-sekvensering som supplement til 16S-sekvensering har vært avgjørende for dette. Målet med *rpoB*-

sekvenseringen var å identifisere *Enterococcus*, *Streptococcus*, *Staphylococcus* og *Enterobacteriaceae* til artsnivå.

RpoB_ess fungerte fint til å skille mellom arter, når sekvenser var tilgjengelig. Sekvenser manglet imidlertid for seks av streptokokkene/stafylokokkene. Metoden bør optimaliseres for å øke sensitiviteten. Dette kan som tidligere nevnt oppnås ved å ta i bruk andre reagenser i PCR-reaksjonen.

Sensitiviteten for RpoB_ent-PCR var god. Det var sekvenser tilgjengelig for samtlige *Enterobacteriaceae*. Evnen til å skille mellom alle arter var imidlertid ikke optimal. RpoB-sekvensene kunne ikke benyttes til å skille *Klebsiella pneumoniae* fra *Klebsiella quasipneumoniae* eller *Serratia marcescens* fra *Serratia nematodiphila* og *Serratia ureilytica*.

Tolv av pasientene fikk påvist *Escherichia coli/Shigella* i galleprøvene. 16S-sekvensering skiller ikke mellom *Escherichia albertii*, *Escherichia fergusonii*, *Escherichia coli* og ulike *Shigella*-arter. RpoB-sekvensering utelukket *Escherichia albertii* og *Escherichia fergusonii*, men kan ikke benyttes til å skille *Escherichia coli* fra *Shigella*. Den genetiske likheten mellom *E.coli* og *Shigella* er stor og det var ikke forventet å kunne skille disse med sekvensering av amplikon fra en universell PCR. Selv om *Shigella* ikke kan utelukkes er sannsynligheten størst for at prøvene inneholder *E.coli*. Shigellose er meldepliktig i Norge, og i 2016 var det bare registrert 83 tilfeller (Folkehelseinstituttet, 2017).

Selv om rpoB-PCR ble designet for å skille mellom ulike *Enterobacteriaceae* og ulike *Staphylococcus/Streptococcus/Enterococcus* kunne metoden også brukes til å identifisere enkelte andre mikrober. *Veillonella parvula* ble amplifisert med primerne som er beregnet for *Enterobacteriaceae* mens *Lactobacillus paracasei* ble amplifisert ved hjelp av primerne som er beregnet for stafylokokker, streptokokker og enterokokker.

Det var ikke mulig å skille *Clostridium baratii* fra *Clostridium sardiniensis*, *Cellulosimicrobium cellulans* fra *Cellulosimicrobium funkei* eller *Fusobacterium nucleatum* fra *Fusobacterium canifelinum* med 16S-sekvensering. RpoB-primerne binder ikke til disse mikrobene, og nye primere må inkluderes dersom metoden skal skille mellom disse.

I denne studien ble alle prøver testet med *rpoB*-PCR parallelt med 16S-PCR. I forskningsammenheng vil det kunne være mer hensiktsmessig å først kjøre 16S og ITS, og deretter *rpoB* og eventuelt andre gener. På den måten vil sekvenseringskapasiteten på Miseq kunne utnyttes langt bedre. Dersom en skal analysere pasientprøver og ønsker svar på artsnivå uten unødig forsinkelse er det bedre å analysere alle genene samtidig. En måte å spare ytterligere tid og arbeid på kan være å kjøre den første PCR-reaksjonen som en sanntids-PCR med SYBRgreen, slik at en ikke trenger å kjøre gelelektroforese. Panelet som er benyttet i denne studien kan eventuelt utvides med universelle PCR-metoder for andre slekter. Det kan også på sikt bli aktuelt med innføring av PCR for ulike resistensgener for utvalgte mikrober.

RpoB-sekvenseringen har gitt nyttig informasjon om mangler ved diagnostikken som benyttes ved MIA. 16S-sekvensering skiller ikke mellom de tre *Klebsiella*-arter i gruppe KpI, *Klebsiella pneumoniae*, *Klebsiella quasipneumoniae* og *Klebsiella variicola*. MaldiTof som benyttes i rutinediagnostikken skilte inntil nylig heller ikke disse tre, men kan etter oppdatering av databasen våren 2017 skille *K. variicola* fra andre *Klebsiella*-arter. *K. quasipneumoniae* finnes fremdeles ikke i databasen til MaldiTof, og prøvene vil bli besvart som *K. pneumoniae* uavhengig av om det er *K. pneumoniae* eller *K. quasipneumoniae*. Dette gjelder uavhengig av om bakterien identifiseres med massespektrometri eller DNA-sekvensering. *K. quasipneumoniae* har aldri vært påvist ved MIA, noe som trolig skyldes manglende diagnostikk, og at problemstillingen ikke har vært belyst. Den var derfor ikke på listen over mikrober som ble benyttet under designing av *rpoB*-ent-PCR. Det er 5 baser i forskjell mellom typestammene når *rpoB*-sekvensen sammenlignes. Dette oppfyller ikke kravet om avstand på 2 %. Etter hvert som flere prøver analyseres og flere sekvenser er tilgjengelig kan grensen vurderes for å se om metoden kan benyttes for å skille disse to.

De vanligste metodene i mikrobiologiske laboratorier skiller altså ikke mellom alle *Klebsiella* på artsnivå, og den kliniske viktigheten av *K. quasipneumoniae* og *K. variicola* er trolig underestimert. En studie fra Stockholm i 2014 viste at dødeligheten hos pasienter med *K. variicola* i blodkultur var høyere enn hos pasienter med *K. pneumoniae* (Maatallah et al., 2014).

Sekvensering av ITS-2 gjorde det mulig å påvise sopp i tre prøver. Bare en av disse ble påvist ved dyrkning. Det var mulig å identifisere *Candida* til artsnivå, mens *Saccharomyces* kun lot seg identifisere til slektsnivå.

Metoden som er etablert i denne studien kan ha stor nytteverdi, både innen forskning og i pasientdiagnostikk. PCR-metodene som er etablert for sekvensering av rpoB-amplikon kan benyttes til identifisering av dyrkningsisolater, der konvensjonell diagnostikk ikke gir identifikasjon til artsnivå av bakterier innen *Enterobacteriaceae*-familien, *Enterococcus*-slekten, *Streptococcus*-slekten og *Staphylococcus*-slekten. Dette kan enklest gjøres ved Sanger-sekvensering av PCR-produktet.

5 KONKLUSJON

Vi har i denne studien etablert en metode for påvisning av mikrober, basert på massiv parallell sekvensering av amplikon. Sekvensering av *16S rRNA* og *ITS-2* er benyttet til påvisning og identifisering av henholdsvis bakterier og sopp. Supplerende testing med sekvensering av *rpoB*-genet gjorde det mulig å identifisere de fleste mikrobene til artsnivå.

Metoden fungerte meget godt til formålet, men *rpoB*-sekvensering av *Enterococcus*, *Streptococcus* og *Staphylococcus* bør optimaliseres for å oppnå bedre sensitivitet.

Identifiseringen er gjort ved at DNA-sekvensene sammenlignes med allerede kjente sekvenser i Genbank. Kvaliteten på referansene som finnes der er høyst varierende, blant annet på grunn av stadige reklassifiseringer, noe som har gjort identifiseringen krevende. Vi har sett at kompetanse på dette området er avgjørende for at resultatene skal kunne bli korrekte.

Det er benyttet to ulike prøvetakingsmetoder. Innholdet i prøver tatt med PTHCS fra pasienter med akutt kolecystitt er sammenlignet med innholdet i prøver tatt med ERCP fra prøver med akutt kolangitt. Resultatene indikerer at ERCP-prøvene kan være mer utsatt for kontaminering under prøvetaking. Utvalget er imidlertid lite, og de to prøvetakingsmetodene er dessuten benyttet i forbindelse med ulike infeksjonssykdommer.

Tjue galleprøver fra nitten pasienter med kolecystitt og kolangitt ble undersøkt, og det ble totalt funnet 35 ulike arter. Det ble påvist signifikant flere mikrober ved massiv parallell sekvensering enn ved dyrkning, og det var en klar sammenheng mellom mengde påvist DNA og om mikroben lot seg dyrke.

Metoden som er etablert i denne studien kan også benyttes til annet prøvemateriale, og vil være spesielt verdifull for karakterisering av mikrobiologien ved ulike invasive polymikrobielle infeksjoner.

6 REFERANSER

- Altman. (1991). *Non-parametric analysis. In Practical Statistics for medical Research*. London: Chapman & Hall.
- Bergh, K. (2015). *Primer og probe design : Kurs TH-28937 : Molekylærgenetiske metoder i medisinsk mikrobiologi 16.-20. mars 2015*. ppt-presentasjon. [Forelesning].
- Brisse, S., Passet, V., & Grimont, P. A. (2014). Description of *Klebsiella quasipneumoniae* sp. nov., isolated from human infections, with two subspecies, *Klebsiella quasipneumoniae* subsp. *quasipneumoniae* subsp. nov. and *Klebsiella quasipneumoniae* subsp. *similipneumoniae* subsp. nov., and demonstration that *Klebsiella singaporensis* is a junior heterotypic synonym of *Klebsiella variicola*. *Int J Syst Evol Microbiol*, 64(Pt 9), 3146-3152. doi: 10.1099/ijs.0.062737-0
- Brisse, S., van Himbergen, T., Kusters, K., & Verhoef, J. (2004). Development of a rapid identification method for *Klebsiella pneumoniae* phylogenetic groups and analysis of 420 clinical isolates. *Clin Microbiol Infect*, 10(10), 942-945. doi: 10.1111/j.1469-0691.2004.00973.x
- Brisse, S., & Verhoef, J. (2001). Phylogenetic diversity of *Klebsiella pneumoniae* and *Klebsiella oxytoca* clinical isolates revealed by randomly amplified polymorphic DNA, *gyrA* and *parC* genes sequencing and automated ribotyping. *Int J Syst Evol Microbiol*, 51(Pt 3), 915-924. doi: 10.1099/00207713-51-3-915
- Chakravorty, S., Helb, D., Burday, M., Connell, N., & Alland, D. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J Microbiol Methods*, 69(2), 330-339. doi: 10.1016/j.mimet.2007.02.005
- Chun, J. Y., Kim, K. J., Hwang, I. T., Kim, Y. J., Lee, D. H., Lee, I. K., & Kim, J. K. (2007). Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene. *Nucleic Acids Res*, 35(6), e40. doi: 10.1093/nar/gkm051
- CLSI. (2008). *Interpretive Criteria for Identification of Bacteria and Fungi by DNA Target Sequencing; Approved Guideline CLSI document*. Wayne, Pennsylvania: Clinical and laboratory Standards Institute.
- Drancourt, M., Bollet, C., Carta, A., & Rousselier, P. (2001). Phylogenetic analyses of *Klebsiella* species delineate *Klebsiella* and *Raoultella* gen. nov., with description of *Raoultella ornithinolytica* comb. nov., *Raoultella terrigena* comb. nov. and *Raoultella planticola* comb. nov. *Int J Syst Evol Microbiol*, 51(Pt 3), 925-932. doi: 10.1099/00207713-51-3-925
- Drancourt, M., & Raoult, D. (2002). *rpoB* gene sequence-based identification of *Staphylococcus* species. *J Clin Microbiol*, 40(4), 1333-1338.
- Drancourt, M., Roux, V., Fournier, P. E., & Raoult, D. (2004). *rpoB* gene sequence-based identification of aerobic Gram-positive cocci of the genera *Streptococcus*, *Enterococcus*, *Gemella*, *Abiotrophia*, and *Granulicatella*. *J Clin Microbiol*, 42(2), 497-504.
- Fevre, C., Passet, V., Weill, F. X., Grimont, P. A., & Brisse, S. (2005). Variants of the *Klebsiella pneumoniae* OKP chromosomal beta-lactamase are divided into two main groups, OKP-A and OKP-B. *Antimicrob Agents Chemother*, 49(12), 5149-5152. doi: 10.1128/AAC.49.12.5149-5152.2005

- Folkehelseinstituttet. (2013). Dødelighet og dødsårsaker i Norge gjennom 60 år 1951-2010 *Rapport / Nasjonalt folkehelseinstitutt* (pp. 210 s.). Oslo: Folkehelseinstituttet.
- Folkehelseinstituttet. (2017). *Smittevernveilederen : Shigellose - veileder for helsepersonell*. Oslo. Retrieved from <https://www.fhi.no/nettpub/smittevernveilederen/>
- Garza-Ramos, U., Silva-Sanchez, J., Martinez-Romero, E., Tinoco, P., Pina-Gonzales, M., Barrios, H., . . . Tellez-Sosa, J. (2015). Development of a multiplex-PCR probe system for the proper identification of *Klebsiella variicola*. *BMC Microbiol*, *15*, 64. doi: 10.1186/s12866-015-0396-6
- Helsedirektoratet. (2013). Nasjonal faglig retningslinje for bruk av antibiotika i sykehus (pp. digital nettsressurs). Oslo: Helsedirektoratet.
- Helseforskningsloven. (2008, sist endret 2015). *Lov om medisinsk og helsefaglig forskning*. Helse- og omsorgsdepartementet Retrieved from <https://lovdata.no/dokument/NL/lov/2008-06-20-44?q=helseforskningsloven>.
- Ikeda, T., Yanaga, K., Kusne, S., Fung, J., Higashi, H., & Starzl, T. E. (1990). Sterility of bile in multiple-organ donors. *Transplantation*, *49*(3), 653.
- Illumina. (2013). 16S Metagenomic Sequencing Library Preparation : Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System. from https://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html
- Janda, J. M. (2015). The Genus *Klebsiella*: An Ever-expanding Panorama of Infections, Disease-Associated Syndromes, and Problems for Clinical Microbiologists. *Clinical Microbiology & Case Reports*, *1*(4).
- Janda, J. M., & Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol*, *45*(9), 2761-2764. doi: 10.1128/JCM.01228-07
- Juul, S. (2012). *Epidemiologi og evidens* (2. udg. ed.). København: Munksgaard.
- Kalland, K. H., Myrmel, H., & Nordbo, S. A. (2005). [Nucleic-acid based diagnostics in clinical microbiology]. *Tidsskr Nor Laegeforen*, *125*(22), 3110-3114.
- Keshav, S. B., Adam. (2013). *The gastrointestinal system at a glance* (second edition ed.).
- Khot, P. D., Ko, D. L., & Fredricks, D. N. (2009). Sequencing and analysis of fungal rRNA operons for development of broad-range fungal PCR assays. *Appl Environ Microbiol*, *75*(6), 1559-1565. doi: 10.1128/AEM.02383-08
- Kircher, M., Sawyer, S., & Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res*, *40*(1), e3. doi: 10.1093/nar/gkr771
- Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glockner, F. O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res*, *41*(1), e1. doi: 10.1093/nar/gks808
- Kommedal, O., Karlsen, B., & Saebo, O. (2008). Analysis of mixed sequencing chromatograms and its application in direct 16S rRNA gene sequencing of polymicrobial samples. *J Clin Microbiol*, *46*(11), 3766-3771. doi: 10.1128/JCM.00213-08
- Kommedal, O., Kvello, K., Skjastad, R., Langeland, N., & Wiker, H. G. (2009). Direct 16S rRNA gene sequencing from clinical specimens, with special focus on polybacterial samples and interpretation of mixed DNA chromatograms. *J Clin Microbiol*, *47*(11), 3562-3568. doi: 10.1128/JCM.00973-09

- Kommedal, O., Simmon, K., Karaca, D., Langeland, N., & Wiker, H. G. (2012). Dual priming oligonucleotides for broad-range amplification of the bacterial 16S rRNA gene directly from human clinical specimens. *J Clin Microbiol*, *50*(4), 1289-1294. doi: 10.1128/JCM.06269-11
- Kommedal, O., Wilhelmsen, M. T., Skrede, S., Meisal, R., Jakovljević, A., Gaustad, P., . . . Helland, C. (2014). Massive parallel sequencing provides new perspectives on bacterial brain abscesses. *J Clin Microbiol*, *52*(6), 1990-1997. doi: 10.1128/JCM.00346-14
- Kommedal, Ø. (2012). *A culture-independent approach for diagnosis of bacterial infections. With special focus on the poly-bacterial specimens.* (Philosophiae doctor), University of Bergen.
- Kujiraoka, M., Kuroda, M., Asai, K., Sekizuka, T., Kato, K., Watanabe, M., . . . Kusachi, S. (2017). Comprehensive Diagnosis of Bacterial Infection Associated with Acute Cholecystitis Using Metagenomic Approach. *Front Microbiol*, *8*, 685. doi: 10.3389/fmicb.2017.00685
- Lindgreen, S. (2012). AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res Notes*, *5*, 337. doi: 10.1186/1756-0500-5-337
- Lopez, N., Kobayashi, L., & Coimbra, R. (2011). A Comprehensive review of abdominal infections. *World J Emerg Surg*, *6*, 7. doi: 10.1186/1749-7922-6-7
- Maatallah, M., Vading, M., Kabir, M. H., Bakhrouf, A., Kalin, M., Naucner, P., . . . Giske, C. G. (2014). *Klebsiella variicola* is a frequent cause of bloodstream infection in the stockholm area, and associated with higher mortality compared to *K. pneumoniae*. *PLoS One*, *9*(11), e113539. doi: 10.1371/journal.pone.0113539
- Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends Genet*, *24*(3), 133-141. doi: 10.1016/j.tig.2007.12.007
- Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nat Rev Genet*, *11*(1), 31-46. doi: 10.1038/nrg2626
- Mollet, C., Drancourt, M., & Raoult, D. (1997). *rpoB* sequence analysis as a novel basis for bacterial identification. *Mol Microbiol*, *26*(5), 1005-1011.
- Rebolledo, M., & Sarria, J. C. (2013). Intra-abdominal fungal infections. *Curr Opin Infect Dis*, *26*(5), 441-446. doi: 10.1097/01.qco.0000433309.21148.f7
- Saha, R., Farrance, C. E., Verghese, B., Hong, S., & Donofrio, R. S. (2013). *Klebsiella michiganensis* sp. nov., a new bacterium isolated from a tooth brush holder. *Curr Microbiol*, *66*(1), 72-78. doi: 10.1007/s00284-012-0245-x
- Sartelli, M., Catena, F., Ansaloni, L., Coccolini, F., Corbella, D., Moore, E. E., . . . Barnabe, R. (2014). Complicated intra-abdominal infections worldwide: the definitive data of the CIAOW Study. *World J Emerg Surg*, *9*, 37. doi: 10.1186/1749-7922-9-37
- Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., . . . Fungal Barcoding Consortium Author, L. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci U S A*, *109*(16), 6241-6246. doi: 10.1073/pnas.1117018109
- Shen, H., Ye, F., Xie, L., Yang, J., Li, Z., Xu, P., . . . Zhang, X. (2015). Metagenomic sequencing of bile from gallstone patients to identify different microbial community patterns and novel biliary bacteria. *Sci Rep*, *5*, 17450. doi: 10.1038/srep17450
- Smyth, R. P., Schlub, T. E., Grimm, A., Venturi, V., Chopra, A., Mallal, S., . . . Mak, J. (2010). Reducing chimera formation during PCR amplification to ensure accurate genotyping. *Gene*, *469*(1-2), 45-51. doi: 10.1016/j.gene.2010.08.009

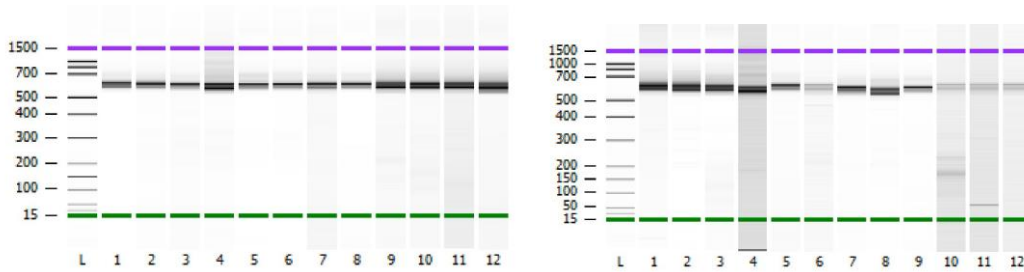
- Takada, T., Strasberg, S. M., Solomkin, J. S., Pitt, H. A., Gomi, H., Yoshida, M., . . . Tokyo Guidelines Revision, C. (2013). TG13: Updated Tokyo Guidelines for the management of acute cholangitis and cholecystitis. *J Hepatobiliary Pancreat Sci*, 20(1), 1-7. doi: 10.1007/s00534-012-0566-y
- Tindall, B. J., Sutton, G., & Garrity, G. M. (2017). *Enterobacter aerogenes* Hormaeche and Edwards 1960 (Approved Lists 1980) and *Klebsiella mobilis* Bascomb et al. 1971 (Approved Lists 1980) share the same nomenclatural type (ATCC 13048) on the Approved Lists and are homotypic synonyms, with consequences for the name *Klebsiella mobilis* Bascomb et al. 1971 (Approved Lists 1980). *Int J Syst Evol Microbiol*, 67(2), 502-504. doi: 10.1099/ijsem.0.001572
- Tokyo Guidelines for the management of acute cholangitis and cholecystitis. Proceedings of a consensus meeting, April 2006, Tokyo, Japan. (2007). *J Hepatobiliary Pancreat Surg*, 14(1), 1-121.
- Tucker, T., Marra, M., & Friedman, J. M. (2009). Massively parallel sequencing: the next big thing in genetic medicine. *Am J Hum Genet*, 85(2), 142-154. doi: 10.1016/j.ajhg.2009.06.022
- Vege, S. S., Gardner, T. B., Chari, S. T., Baron, T. H., Clain, J. E., Pearson, R. K., . . . Sarr, M. G. (2009). Outcomes of intra-abdominal fungal vs. bacterial infections in severe acute pancreatitis. *Am J Gastroenterol*, 104(8), 2065-2070. doi: 10.1038/ajg.2009.280
- Voelkerding, K. V., Dames, S. A., & Durtschi, J. D. (2009). Next-generation sequencing: from basic research to diagnostics. *Clin Chem*, 55(4), 641-658. doi: 10.1373/clinchem.2008.112789
- World Health Organization. (2008). The global burden of disease : 2004 update (pp. vii,146 s.). Geneva: WHO.
- Yokoe, M., Takada, T., Strasberg, S. M., Solomkin, J. S., Mayumi, T., Gomi, H., . . . Tokyo Guidelines Revision, C. (2012). New diagnostic criteria and severity assessment of acute cholecystitis in revised Tokyo Guidelines. *J Hepatobiliary Pancreat Sci*, 19(5), 578-585. doi: 10.1007/s00534-012-0548-0
- Yoshida, M., Takada, T., Kawarada, Y., Tanaka, A., Nimura, Y., Gomi, H., . . . Ker, C. G. (2007). Antimicrobial therapy for acute cholecystitis: Tokyo Guidelines. *J Hepatobiliary Pancreat Surg*, 14(1), 83-90. doi: 10.1007/s00534-006-1160-y

VEDLEGG

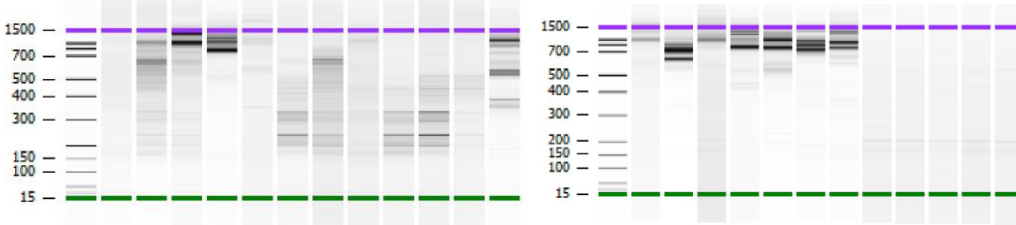
Vedlegg A

PCR-produktene fra indeks-PCR, visualisert ved bruk av Bioanalyser

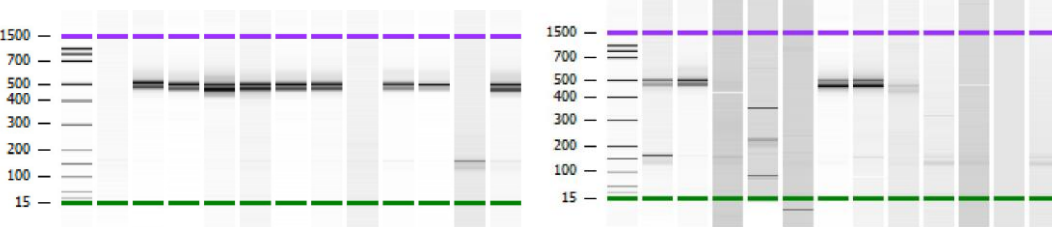
Indeks-PCR for 16SrRNA-PCR (negative kontrollen er i posisjon 6, 10, 11 og 12 på gel nr. 2)



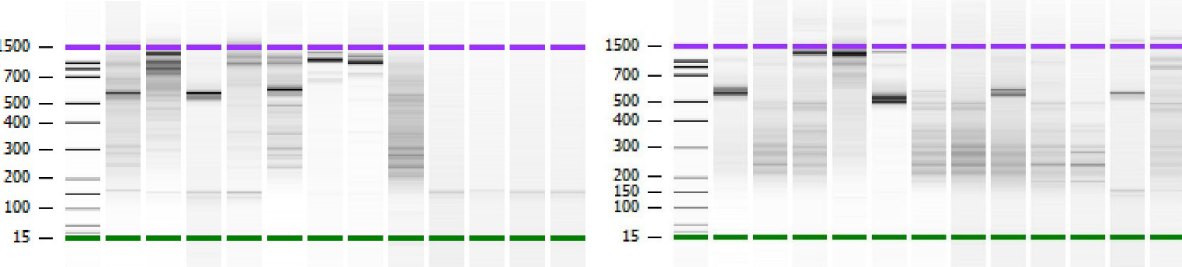
Indeks-PCR for ITS-2



Indeks-PCR for rpoB-ent



Indeks-PCR for rpoB-ess



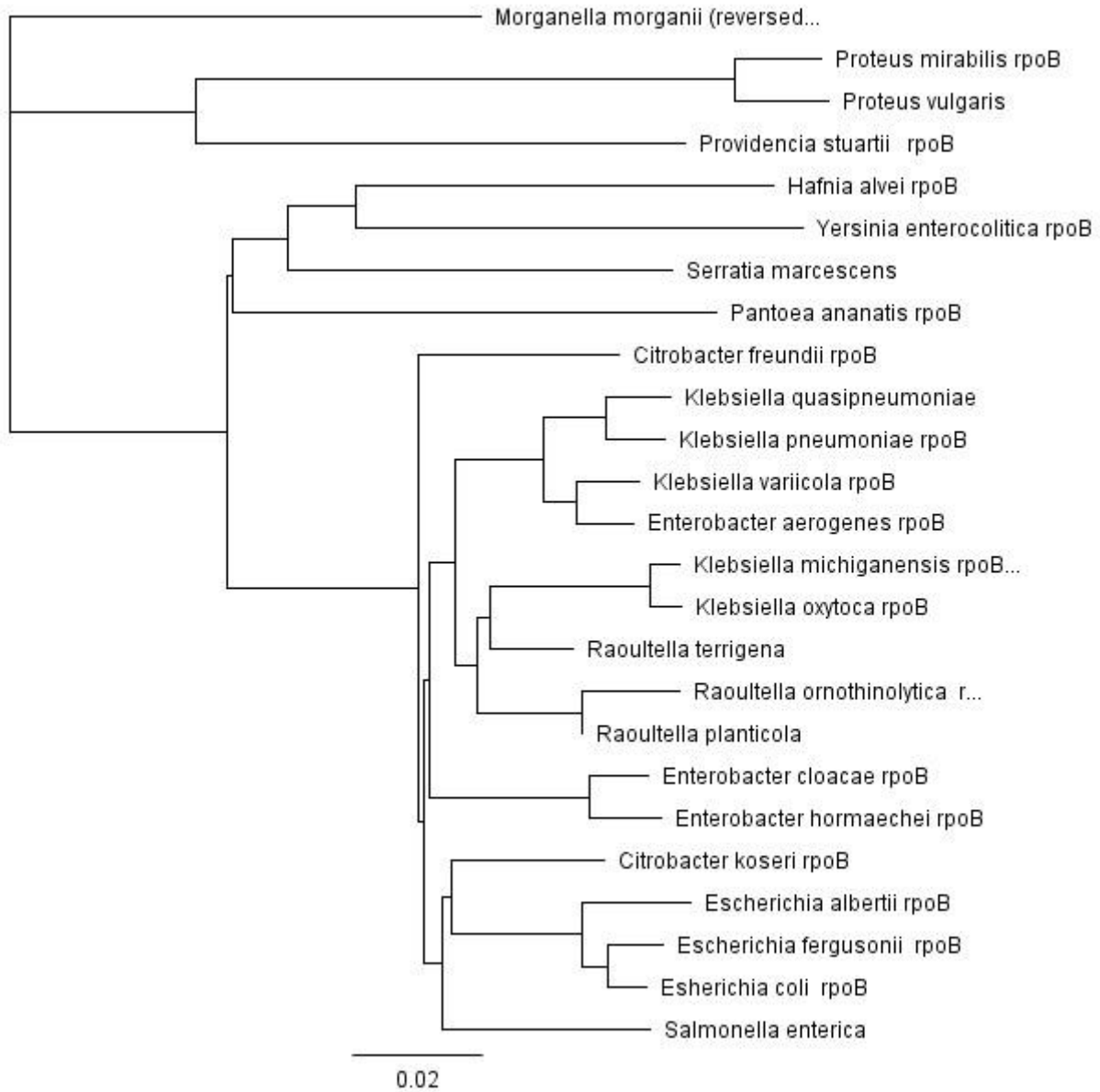
VEDLEGG B

Sanntids-PCR utført på galleprøvene i forkant av oppsettet med massiv parallell sekvensreing. 16S rRNA ble utført for å kontrollere tilstedeværelse av DNA i prøvene.

oppga	labnr	dyrkningsresultat	16S	16S blar	RpoBen	RpoBen	RpoBssL	RpoBss	ITS-2	ITS-2 l
9	64331435	<i>S.salivarius</i> R	15,4	33,4	36,37	36,88	19,56	38,06	42,9	35,8
10	64499812	<i>K.pneumoniae</i> R	20,5	33,5	22,34	35,79	40,28	43,96	39,9	40,3
1	64407850	<i>E.coli</i> R	15,4	31,6	19,32	38,11	37,76	33,33	0	37,2
2	64474458	<i>K.oxytoca</i> R, <i>E.coli</i> R, <i>Raoultella ornithinolytica</i> R	15,9	32,6	19,14	35,76	29,31	36,79	38,9	38,6
3	64669440	<i>E.faecalis</i> R, <i>K.pneumoniae</i> S	13,1	33,2	19,46	37,46	15,97	39,35	36,6	38,4
4	64536829	<i>E.coli</i> R, <i>S.massiliensis</i> M	16,9	32,9	23,66	36,17	27,51	42,74	0	37,8
5	64619507	<i>K.pneumoniae</i> R	18,4	33,1	22,15	0	43,57	37,58	0	36,2
6	64646907	<i>Streptococcus salivarius</i> S	16,4	33,8	36,84	35,4	19,64	43,06	38,7	36,5
7	64683165	<i>K.pneumoniae</i> , to ulike R og M	17,5	33,2	19,08	35,39	39,33	40,26	0	33,2
8	64673435	<i>E.coli</i> R	15,7	33,1	19,73	36,29	0	0	0	
11	64231023	<i>E.faecium</i> R, <i>C.albicans</i> S, <i>E.coli</i> S	19,5	32,4	35,97	36,11	22,06	44,2	34,2	36,2
12	64298652	<i>E.coli</i> R, <i>S.sanguinis</i> U	12,5	29,7	15,97	37,65	22,59	35,46	31,2	37,1
13	64417287	<i>E.faecium</i> R	16,3	31,3	25,53	38,69	20,17	36,82	39,7	36,1
14	64442523	<i>E.coli</i> R, <i>Granulicatella adiacens</i> R	12,9	34,5	16,11	36,82	34,15	41,09	37,8	37,4
15	64507869	<i>Cellulosimicobium cellulans</i> R, <i>Lactobacillus paracasei</i> M	17,6	33,9		36,04	20,64	43,4	39,7	34,2
16	64536826	<i>E.faecium</i> S	27,9	32,9	36,74	38,93	31,54	38,26	34,8	35,2
17	64581242	<i>S.aureus</i> R	14,7	32,6	0	34,76	17,11	41,07	0	38,4
18	64619719	<i>E.coli</i> R	15,1	33,1	16,33	0	35,04	37,58	37,2	36,2
19	64395188	<i>C.perfringens</i> R, <i>S.marcescens</i> R, <i>E.coli</i> R	12,9	32,1	15,42	35,49	28,21	0	38,3	37
20	64401301	<i>C.perfringens</i> R, <i>E.faecium</i> R, <i>S.marcescens</i> M, <i>E.coli</i> M	14,8	28,8	24,69	0	23,36	0	36,8	35,8

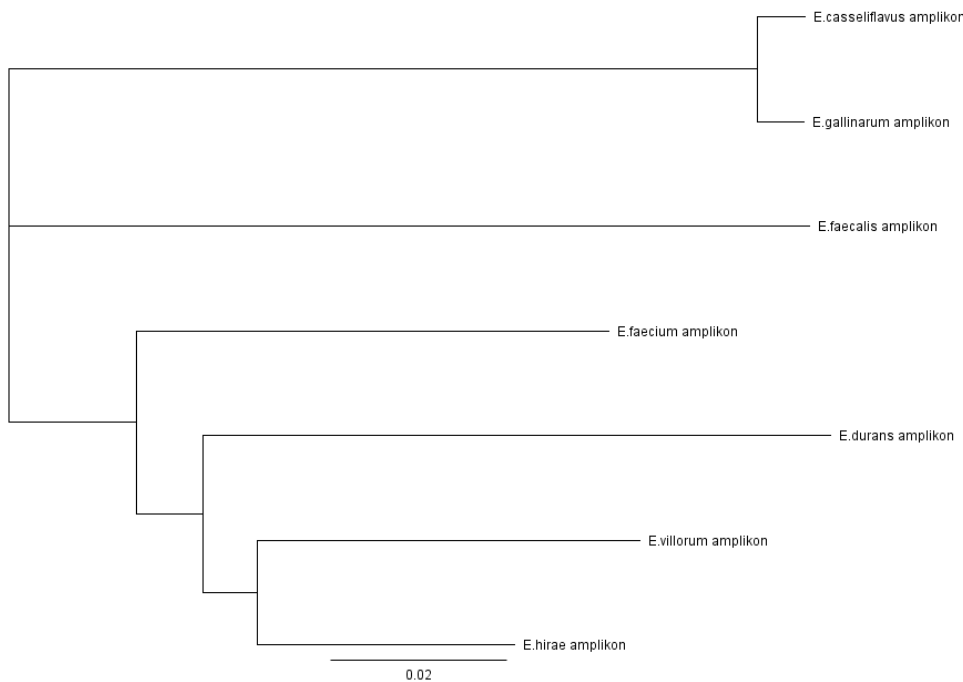
VEDLEGG C

Fylogenetisk tre, hele *rpoB*-genet hos *Enterobacteriaceae*

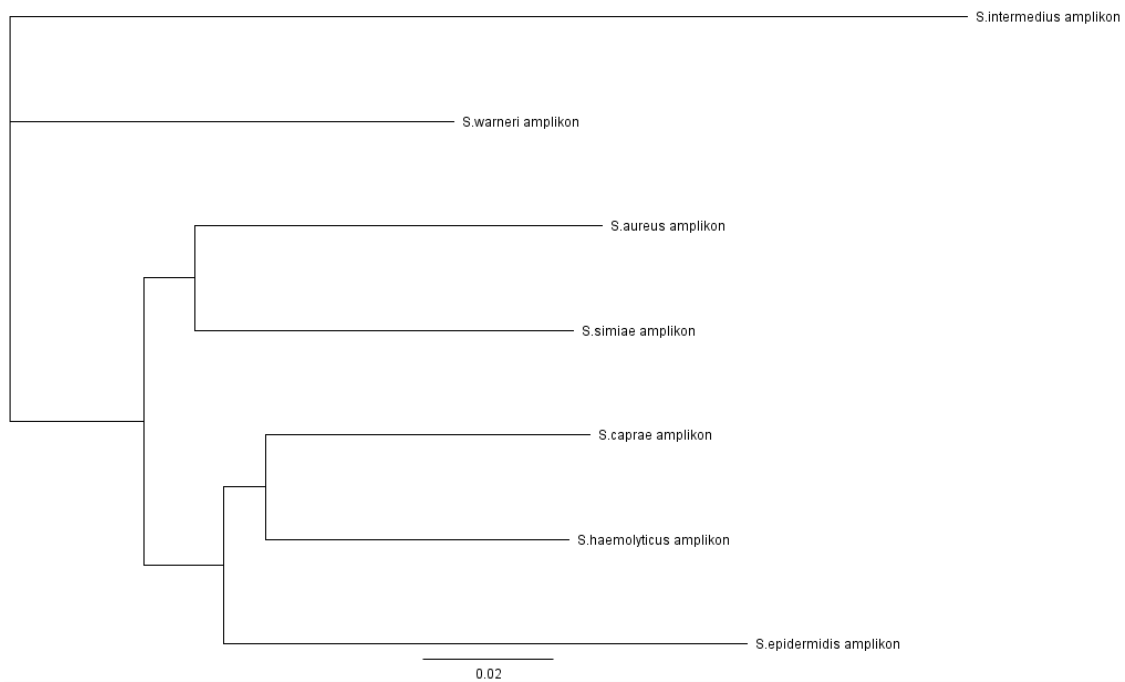


VEDLEGG D

Fylogenetisk tre, rpoB-amplikon Enterococcus




Fylogenetisk tre, rpoB-amplikon Staphylococcus



VEDLEGG E

Valideringsrapport, ITS-2 PCR

	Valideringsrapport, ITS-2 PCR og sekvensering	R08653
---	--	--------

Plan godkjent

Relaterte dokumenter

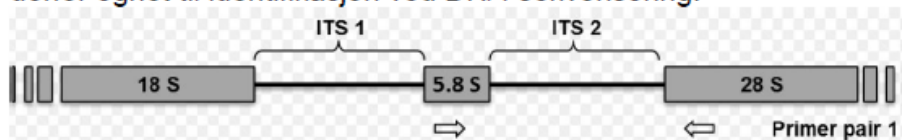
	DokID	Tittel
Plan for validering/verifisering	R07428	Plan for validering av ITS-2 PCR
Verifisering av onlinekobling		Ikke aktuelt
Sjekkliste for etablering og endring	R08664	
Risikovurdering		Ikke aktuelt
Prosedyrer	D27978 D27990 D35276	Analyseprosedyrer
Annet		

Gjennomføring, resultater og konklusjon

Gjennomføring

Primere ble valgt ved å studere resultatene i artikkelen «*Sequencing and Analysis of Fungal rRNA Operons for Development of Broad-Range Fungal PCR Assays*»
Prasanna D. Khot,¹ Daisy L. Ko,¹ and David N. Fredricks^{1,2}

Primerne binder til konserverte områder på 5,8S og 28S. *Internal transcribed spacer 2* (ITS-2) ligger mellom disse genene. ITS-2 varierer i lengde og sammensetning og er derfor egnet til identifikasjon ved DNA-sekvensering.



ATCC stamme 10231 (*C.albicans*) er brukt til innkjøring. To ulike enzymløsninger ble testet, TAKARA og Roche SYBR. Ulike temperaturprofiler ble testet før annealingtemperatur ble valgt. Annealingtemperatur som er 2°C lavere enn den høyeste temperaturen som fungerer ble valgt, slik at det også kan være mulig å detektere typer som har mutasjon i primerområdet. Annealingtid som er 15 sekunder lengre enn den korteste tiden som fungerer ble valgt, slik at det også er mulig å detektere lengre amplicon enn kontrollstammens.

Fortynningsrekke ble laget for å vurdere PCR-effektivitet.

Vevsprøver som er rike på humant DNA ble testet for å se etter kryssreagering.

Negative kontroller (vann) ble tatt med under valideringen for å se etter eventuell kontaminasjon av reagenser/utstyr.

SLP-panel med 3 prøver fra NEQAS mykopanel 2016 ble analysert.

Valideringsrapport, ITS-2 PCR og sekvensering

Resultater

Resultatene finnes her: F:\Metodevalidering\Molekylærbiologisk seksjon\ITS\ITS-2\validering av ITS-2.xlsx

Testen oppfyller alle kravene som ble definert under planlegging av valideringen. Metoden har god sensitivitet og spesifisitet, og det er tatt hensyn til at metoden skal være robust mot mutasjoner og ulike lengder på amplicon.

Fortynningsrekke viste bra PCR-effektivitet, i alle fall til området 32-34 sykluser. Resultatene av denne har sammen med analyseresultater fra positive og negative pasientprøver vært brukt til å sette cut-off grense for hvilke prøver som skal sekvenseres.

Prøver med $ct \leq (ct_{blank} - 3)$ og samtidig $ct \leq 32$ vurderes som mulig positive og sendes til sekvensering. Det har vært vanskelig å sette en klar cut-off grense nå, så denne må trolig justeres når metoden har vært i bruk en periode, og når flere reagenslot er inkludert.

Terskelen for å sende prøver til sekvensering bør være lav i starten, og det er mulig grensen på 32 sykluser oppjusteres dersom sekvensering av prøver med høy ct gir negative resultater. Dersom lege ønsker sekvensering til tross for at disse kriteriene ikke er oppfylt er dette naturligvis mulig.

SLP-prøvene ble identifisert som *Paecilomyces variotii*, *Geothricum candidum* og *Neoscytalidium*, noe som stemmer med fasiten. Kontrollstammen vår *C.albicans*, ATCC 10231, ble også identifisert med 100% treff i Genbank

Konklusjon

Valideringen, sammen med dokumentasjon fra den grundige uttestingen som er gjort av Prasanna D. Khot,¹ Daisy L. Ko,¹ and David N. Fredricks^{1,2*}
Fred Hutchinson Cancer Research Center, Seattle, Washington,¹ and University of Washington, Seattle, Washington²
viser at metoden kan erstatte dagens metode (påvisning og sekvensering av ITS-1)

Dato for oppstart av analysen

august 2016 (se sjekklisten R08664 for detaljer)

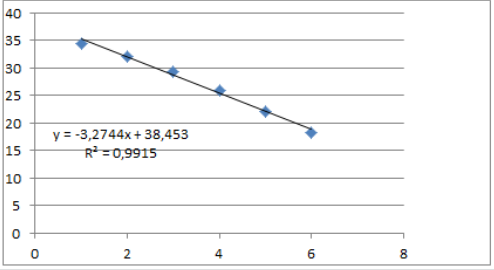
Møteprotokoll

Planleggingsmøte 25.09.2015: Øyvind, Birgitte, Rutt-Lovise og Randi
Gjennomgang av valideringsrapport 11.05.2016: Øyvind, Birgitte, Solveig, Randi

Rapporten er godkjent *Elektronisk signatur*

Birgitte Bargård Madsen, Kristinsson, Solveig Gunn, Paul Christoffer Lindemann, Rebecca Irene Breistein, Rutt-Lovise Strand, Øyvind Kommedal

Rådata, validering av ITS-2 PCR

Tester ulike PCR-betingelser/reagenser	Dato	denat	Annealing	ekstensjo	komment	pos ktr	ktr:10	vann	tm ktr
SmartCycler, TaKaRa	12.09.2015	95, 15s	60C, 30s	-	neg ktr kommer før 1:10 fort. Ktr	28,4	32,43	30,21	
SmartCycler, TaKaRa	13.09.2015	94, 30s	60C, 30s	72, 10s		37,32	41,38	37,38	
SmartCycler, Roche SYBR	13.09.2015	94, 30s	60C, 30s	72, 10s		36,12	IU	IU	
LightCycler, Roche SYBR	13.09.2015	95C, 10s	60C, 10s	72C, 10s	Fine kurve	28,63	31,84	>40	85,39
LightCycler, Roche SYBR	13.09.2015	95C, 10s	55C, 10s	72C, 15s	Fine kurve	28,95	31,85	>40	IU
SmartCycler, Roche SYBR	14.09.2015	95C, 10s	60C, 10s	-	Fine kurve	42,12	43,27	-	
SmartCycler, TaKaRa	14.09.2015	95C, 5s	60C, 30s	-					
SmartCycler, Roche SYBR		95, 15s	60, 20s	-	Fine kurve	30,91	35,97	39,07	
LightCycler, Roche SYBR	19.09.2015	95, 15s	60, 20s	-	Fine kurve	29,03	35,11	neg	
SmartCycler, Roche SYBR	20.09.2015	95C, 10s	60C, 10s	72C, 15s		IU	30,82	neg	
SmartCycler, Roche SYBR	20.09.2015	95C, 10s	60C, 10s	72C, 15s		IU	31,45	39,56	
SmartCycler, Roche UTEN PRIMERE	20.09.2015	95C, 10s	60C, 10s	72C, 10s		-	-	neg	
Oppsummert: Ønsker å bruke smartcycler av praktiske årsaker (enkler å få ut templatet til sekvensering). Beste protokoll er da 95C i 10 sek, 60C i 10 sek og 72C i 15 sek.									
Lager forfynningsrekke. Må forfynne utgangskontrollen noe pga lite materiale.									
	ufort	1/10	1/100	1/1000	1/10000	1/100000	1/1000000	vann	
	6	5	4	3	2	1			
gjennomsnittsverdi	18,22	22,05	25,9	29,225	32,135	34,425	35,895	35,19	
1.parallell	18,22	21,98	25,42	29,25	32,05	34,12	36,69	36,21	
2.parallell	18,22	22,12	26,38	29,2	32,22	34,73	35,1	34,17	
									
	Effektivitet	2,020224							

VEDLEGG F

Rådata fra optimalisering av protokoll etter innføring av adapter i primersekvensene

16S

Uten adapter: ca.20

	Pos ktr. 160124			vann		
ada st.prot. fl	36,71	36,18		0		
10+30+30, 60C	25,51	24,98		35,44		
10+30+30, 58C	26,51	26,6		38,6		
halv primerkons	27,23	26,39		39,87		
10+60+30, 60C	24,28	25,35		36,56		
10+60+60, 60C						
10+30+30, 61C	26,23	25,91				
10+30+30, 62C	26,17	25,56				
10+30+30, 55C						
	Pos ktr. 160512					
10+60+30	23,37	23,53				
10+30+30	24,38	24,28	24,27	38,20	38,35	38,32
Illumina F+R 10+30+30	22,45	22,63	22,83	34,69	34,28	35,87
III F og okom R 10+30+30	23,46	23,42	23,78	35,68	35,37	36,73
III F og okom R 10+60+30	21,15	21,81	22,63	34,18	33,39	34,17
Illumina F+R 10+60+30	20,83	22,56	21,68	33,65	34,89	32,64
III F og okom R 30+30+30	23,42	23,41	23,67	36,82	36,88	36,93
Ny test av våre 10+60+30	23,65	23,76	23,69	35,18	33,22	33,35
III F og okom R, 55C	22,11	22,2	22,31	33,86	33,15	33,83

Lang annealingtid var avgjørende for lav ct-verdi (god PCR-sensitivitet)

Negative kontroller ble inkludert i protokollene som ga lav nok ct-verdi for positiv kontroll til at protokollen var aktuell for bruk.

Når PCR-protokoll var valgt for 16S ble denne testet på de andre genene.

Annealingtemperaturen var ulik for genene, mens resten ble gjort likt