

TARTU ÜLIKOOL

LOODUS- JA TÄPPISTEADUSTE VALDKOND  
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT  
BIOINFORMAATIKA ÕPPETOOL

**Patogeensete mikroorganismide tuvastamine rakuvabast DNA-st**

Bakalaureusetöö

12 EAP

Pilleriin Paidla

Juhendajad

PhD Mairo Remm

TARTU 2023

## INFOLEHT

Rakuvaba DNA (cfDNA) on rakuväline DNA, mis koosneb veres ja teistes bioloogilistes vedelikes leiduvatest inimese ja mikroobide nukleiinhappe fragmentidest. cfDNA sisaldab endas nakkuseid puudutavat informatsiooni ning seda sekveneerides saab leida infot üle terve organismi haigusttekitavate patogeenide kohta.

Käesoleva töö eesmärgiks on anda referatiivne ülevaade seni avaldatud teadustöödest mikroobide tuvastamisel cfDNA-st ning tutvustada erinevaid meetodeid patogeensete mikroorganismide identifitseerimiseks. Rakuvälise DNA sekveneerimine annab võimaluse haiguste varajaseks diagnoosimiseks ning efektiivse ravikava määramiseks.

Märksõnad: cfDNA, sekveneerimine, mikroorganismid, tuvastamine

CERCS: B110 (Bioinformaatika, meditsiiniinformaatika, biomatemaatika, biomeetrika)

Cell-free circulating DNA (cfDNA) consists of human and microbial fragments of nucleic acid that are found from blood and other body fluids. cfDNA integrates information about infections from all over the body. Sequencing of cell-free plasma can help detect infection-causing pathogens.

The purpose of this study is to present a review of published scientific articles concerning detection of microbes from cfDNA thus far and introduce different methods for identifying pathogenic microorganisms. Sequencing of cfDNA gives an opportunity for early diagnosis of diseases and the determination of an effective treatment plan.

Keywords: cfDNA, sequencing, microorganisms, detection, identification

CERCS: B110 (Bioinformatics, medical informatics, biomathematics, biometrics)

## SISUKORD

<b>INFOLEHT</b> .....	<b>2</b>
<b>SISUKORD</b> .....	<b>3</b>
<b>SISSEJUHATUS</b> .....	<b>5</b>
<b>KASUTATUD LÜHENDID</b> .....	<b>6</b>
<b>KIRJANDUSE ÜLEVAADE</b> .....	<b>8</b>
<b>1. PATOGEENSETE MIKROOBIDE TUVASTAMINE</b> .....	<b>8</b>
1.1. Kultiveerimine.....	9
1.1.3. Kliiniline tähtsus.....	9
Verekuultuur.....	10
1.2. MALDI-TOF.....	10
1.2.1. Tööpõhimõte.....	10
1.2.2. Kliiniline tähtsus.....	11
Verekuultuuris kasutamine.....	12
1.3. PCR.....	12
1.3.1. Reaalaja PCR (real time).....	13
1.3.2. Multiplex PCR.....	13
1.4. Metagenoomne sekveneerimine 16S piirkonnast (eelneva amplifitseerimisega)..	14
1.5. Mikroobide tuvastamine isolaadi genoomi sekveneerimisel (WGS).....	15
1.6. Mikroobide tuvastamine sekveneerimisel ilma eelneva amplifitseerimiseta (WMS).....	16
<b>2. MIKROOBIDE TUVASTAMINE VEREPLASMAST</b> .....	<b>18</b>
2.1. Millest koosneb cfDNA, millised on tema omadused.....	18
2.1.1. cfDNA kliiniline tähtsus.....	19
2.2. Näiteid bakterite tuvastamisest cfDNAST.....	20
2.2.1. Sepsis.....	20
2.2.2. Tuberkuloos.....	22
2.2.3. Kopsupõletik.....	22
2.2.4. Kuseteede haigused.....	24
2.2.5. Haruldased infektsioonid.....	25
<i>Mycobacterium chimaera</i> .....	25
<i>Propionibacterium acnes</i> .....	26
Nakkused peale tüvirakkude siirdamist.....	26
2.3. Näiteid patogeensete seente/pärmide tuvastamisest cfDNAST.....	27
2.4. Näiteid viiruste tuvastamisest cfDNAST.....	28
<b>3. PROGRAMMID MIKROOBIDE TUVASTAMISEKS</b> .....	<b>30</b>
3.1. Homoloogiaotsing BLAST.....	30
3.2. Lugemite paigutamine referentsgenoomile.....	30

3.3. Joendusvabad meetodid.....	32
3.3.1. Kraken.....	32
3.3.2. StrainSeeker.....	33
<b>KOKKUVÕTE.....</b>	<b>37</b>
<b>SUMMARY.....</b>	<b>38</b>
<b>TÄNUSÕNAD.....</b>	<b>39</b>
<b>KASUTATUD KIRJANDUS.....</b>	<b>40</b>
KASUTATUD VEEBIAADRESSID.....	47
<b>LIHTLITSENTS.....</b>	<b>48</b>

## SISSEJUHATUS

Rakuvaba DNA (cfDNA) koosneb väikestest nukleiinhappe fragmentidest, mis on leitav verest ja teistest kehavedelikest. See võib sisaldada nii inimese enda DNA molekule kui ka mikroobset DNA-d. Olenevalt inimese tervislikust seisundist võib cfDNA koostis ja kontsentratsioon vereplasmas muutuda. cfDNA-d analüüsid on võimalik tuvastada seal leiduvaid patogeenseid mikroorganisme. Konventsionaalsed haiguste diagnoosimismeetodid nõuavad tihti invasiivseid proovivõtmise meetodeid ja on ajakulukad. cfDNA koostise analüüsimine sekveneerimise abil on potentsiaalne alternatiiv mitteinvasiivseks ja kiireks haigusttekitavate patogeenide tuvastamiseks. Varajane efektiivse ravikava rakendamine patsiendile võib vähendada infektsioonidest tingitud suremust.

Käesoleva töö eesmärgiks on anda ülevaade mikroobide tuvastamise meetoditest ning koostada kokkuvõtte seni avaldatud teadustöödest, mis puudutavad cfDNAst mikroobide määramist ja selleks kasutatavatest arvutiprogrammidest.

## KASUTATUD LÜHENDID

AMR	<i>antimicrobial resistance</i>	antimikroobne resistentsus
AFB	<i>acid fast bacillus</i>	happekindel bakter
BALF	<i>Bronchoalveolar lavage fluid</i>	bronhoalveolaarne loputusvedelik
BC/GS	<i>blood culture/ Gram stain analysis</i>	verekultuur/ Grami meetodil värvimine
BKV		BK polüoomiviirus
bp	<i>base pair</i>	aluspaar
BSI	<i>bloodstream infection</i>	vereringe infektsioon
cfDNA	<i>cell-free DNA</i>	rakuvaba DNA
ctDNA	<i>circulating tumor DNA</i>	ringlev kasvaja DNA
EBV	<i>Epstein-Barr virus</i>	Epstein-Barri viirus
HSV	<i>Herpes simplex virus</i>	Herpes simplex viirus
IFD	<i>invasive fungal disease</i>	invasiivne seennakkus
mcfDNA	<i>microbial cell free DNA</i>	mikroobne rakuvaba DNA
mNGS	<i>metagenomic next generation sequencing</i>	metagenoomne järgmise põlvkonna sekveneerimine
MS	<i>mass spectrometry</i>	massispektromeetria
MTB	<i>Mycobacterium tuberculosis</i>	Kochi kepike ehk tuberkuloosibakter (kitsalt liigi tähenduses)

MTC	<i>Mycobacterium tuberculosis complex</i>	tuberkuloosi mükobakteri kompleks, rühm geneetiliselt lähedasi mükobakteri liike
MRSA	<i>Methicillin-resistant Staphylococcus aureus</i>	metitsilliinresistentne stafülokokk
M/Z	<i>mass/charge</i>	massi ja laengu suhe
NSG	<i>next generation sequencing</i>	järgmise põlvkonna sekveneerimine
NPC	<i>nasopharyngeal carcinoma</i>	nasofarüingeaalne kartsinoom
ONT	<i>Oxford Nanopore Technologies</i>	
PD	<i>primer-dimer</i>	praimeri dimeer
PMF	<i>peptide mass fingerprint</i>	peptiidide massi sõrmejalg
SIQ	<i>sepsis indicating quantifier</i>	sepsist näitav kvantor
TB		tuberkuloos
TOF	<i>time of flight</i>	lennuaeg
UTI	<i>urinary tract infection</i>	kuseteede nakkus?
VAP	<i>ventilator-associated pneumonia</i>	ventilaatorpneumoonia
WGS	<i>whole genome sequencing</i>	terve genoomi sekveneerimine (mikroobi isolaadi genoomi sekveneerimine)
WMS	<i>whole metagenome sequencing</i>	kogu metagenoomi sekveneerimine
qPCR	<i>quantitative real-time PCR</i>	kvantitatiivne reaalaaja PCR

## KIRJANDUSE ÜLEVAADE

### 1. PATOGEENSETE MIKROOBIDE TUVASTAMINE

Inimhaiguste põhjustajateks on seni teadaolevalt üle 1000 erineva mikroorganismi, mis kollektiivselt põhjustavad ainuüksi Ameerika Ühendriikides ligikaudu 5 miljonil inimesel haiglaravi vajava seisundi (Blauwkamp et al., 2019). Varane patogeensete mikroobide tuvastamine on efektiivse ravi rakendamise võtmeks (Kumar et al., 2009). Paljudel juhtudel on mikroobide tuvastamine aeg-kriitiline, näiteks sepsise puhul on õigete antibiootikumide manustamine elulise tähtsusega (Grumaz et al., 2016). Sellistes olukordades võib juhtuda, et traditsioonilised patogeeni ning antimikroobse resistentsuse (AMR) tuvastamismeetodid, mis võivad mõjutada patsiendi ravikava määramist, ei ole piisavalt kiired (Ramanan et al., 2018). AMR on mikroobide (viiruste, bakterite, seente ja algloomade) võime muutuda vastupanuvõimeliseks seni toimunud antimikroobsetele preparaatidele (nt. antibiootikumid) (2023, <https://www.sm.ee/amr>). Teatavatel juhtudel on haigusttekitavate patogeeni identifitseerimine raskendatud ka seetõttu, et neid on kas liiga keeruline või lausa võimatu väljaspool organismi kultiveerida (Fenollar and Raoult, 2007). See aitab omakorda kaasa haigestumise, suremuse ning teiste ebasoodsate tulemuste (nt. ravi maksumuse) kasvule (Ramanan et al., 2018).

Kõiki bakteriaalseid nakkusi kahjuks ainult sümptomite põhjal alati tuvastada ei saa, seega on levinud empiirilise antibiootikumravi kasutamine kuni patsiendi mikrobioloogiliste proovide tulemused selguvad (Schreckenberger ja McAdam, 2015). Empiirilist antibiootikumiravi saab keskmiselt iga viies patsient USAs ning seda seostatakse antibiootikumidele resistentsete patogeeni põhjustatud nakkustega (Sameer et al., 2021). Marquet jt. viisid läbi uurimuse, kus Medline vahendusel (2004-2014 jooksul) otsiti süstemaatiliselt ebasobivaid empiirilise antibiootikumiravi juhtumeid tõsiste nakkustega (nagu vereringeinfektsioonid, kopsupõletik, sepsis jm.) patsientidel. Koostati 27 artikli põhjal metaanalüüs, kus selgus, et ebasobiv empiiriline antibiootikumiravi tõstab oluliselt haiglapatsientide suremust raskesse infektsiooni 30 päeva jooksul, sest ei ole suunatud õigele haigustekitajale (Marquet et al., 2015).



Antibiootikumide liigtarbimisega seotud riskid hõlmavad endas ka antibiootikumresistentsete mikroorganismide, tõsiste haiguste, haiguse kulu pikkuse, tüsistuste tekkimise, suremuse, haiglaravi maksumuse, (elu)ohtlike kõrvalmõjude ning iseparanevate nakkuste medikaliseerimise kasvu (Llor ja Bjerrum, 2014). Varane efektiivne mikroobidevastane ravi on nakkustest tervenemise ja AMR ärahoidmise puhul kriitilise tähtsusega. Antimikroobse resistentsuse vastu testimine aitab valida optimaalse antibiootikumi ravi (Weis et al., 2022).

## **1.1. Kultiveerimine**

Bakterite kultiveerimine on meetod laboratoorsetes tingimustes bakterirakkude kasvatamiseks söötme peal või sees. Bakterid peavad kasvama söötmel isoleeritult ning uurimiseks tuleb eraldada puhaskultuurid. Puhaste bakterikultuuride kultiveerimine on vajalik peale liigi identifitseerimise ka selleks, et uurida nende virulentsust ja vastuvõtlikkust antibiootikumidele. Peale kultiveerimist on võimalik sekveneerida mikroorganismide genoom, mis aitab mõista ja ravida tema põhjustatud haiguseid (Lagier et al., 2015).

### **1.1.3. Kliiniline tähtsus**

Bakteriaalsete infektsioonide efektiivne ravimine nõuab tihti kiiret ja täpset bakterite tuvastamist steriilsetes kehavedelikes nagu veri ja tserebrospinaalvedelik (CSF). Selle jaoks kõige levinum meetod on kultuuride kasvatamine söötmel, mis nõuab vähemalt 8 tunnist inkubatsiooniaega, millele järgnevad biokeemilised ja/või immunoloogilised testid, et bakter identifitseerida. Vajaminev aeg bakterite identifitseerimiseks võib olla isegi pikem, kui tegemist on aeglaselt kasvavate organismide või bakteri populatsioonidega (Greisen et al., 1994). Miinimumaeg organismi tuvastamiseks ja identifitseerimiseks on 24-48 tundi (Nieman et al., 2016), kusjuures testi tundlikkus on kriitilises seisus patsientidele umbes 70% ning isegi väiksem nõudlike mikroorganismide puhul (Beutz et al., 2003).

Suureks väljakutseks on olnud patogeensete mikroorganismide puhaskultuuri isoleerimine mikrobiotast (Chapin ja Lauderdale, 2007), sest enamus uuringud ja testid on usaldusväärse tulemusega vaid puhaskultuuris (Kumar, 2012).

## Verekultuur

Verekultuur koos Grami meetodil värvimisega (BC/GS) on meditsiiniline laborianalüüs, mis on bakteriaalsete ja seenhaiguste tuvastamise üks peamisi viise, kuid see nõuab suurt proovi mahtu ja teiste meetoditega võrreldes pikemat inkubatsiooniaega, et maksimeerida proovi tundlikkust (Tsalik et al., 2010). See on nt. sepsise diagnoosimise standardmeetod, kuid paljudel juhtudel pole kultiveerimise piisavalt tundlik ega spetsiifiline (Grumaz et al., 2019). Verekultuuri tundlikkus väheneb, kui antimikroobse raviga alustatakse enne kultiveerimist (Ramanan et al., 2018). Enamus kliinilisi patogeene saab tuvastada peale 24 - 48 tunnilist inkubeerimist, kuid rutiinseks kontrolliks tuleb proove inkubeerida vähemalt 5 päeva (Chapin, 2007). Näiteks *Helicobacter pylori*'t, mis põhjustab seedetrakti haavandeid, saab tuvastada alles peale viiepäevast inkubeerimist (Marshall ja Warren, 1984). Lisaks on organisme, kes ei kasva rutiinsetes vereseerumites (nt. *Coxiella burnetii*, *Tropheryma whipplei*, ja *Rickettsia* liigid) ning jäävad seetõttu tuvastamata (Ramanan et al., 2018).

## 1.2. MALDI-TOF

MALDI-TOF MS on kerkinud esile kui paljulubav uus meetod mikroobide identifitseerimiseks. MALDI-TOF MS ehk maatriksi assisteeritud laser desorptsioon-ionisatsioon/ lennuaeg massispektromeetria (*matrix-assisted laser desorption-ionization/time-of-flight mass spectrometry*) on kiire, tundlik ja võrdlemisi ökonoomne tehnoloogia oma hinna ning tööjõu kulu poolest (Singhal et al., 2015).

### 1.2.1. Tööpõhimõte

MALDI-TOF läbiviimiseks asetatakse agarplaadil kasvatatud proov MALDI taldrikule, kuhu lisatakse maatriks (Giuliano et al., 2019). Maatriks on energiat neelav orgaaniline ühend, mis kuivades kristalliseerub. Tänu maatriksi kristalliseerumisele kristalliseerub ka kasutatav proov (proovi-maatriksi segu). Seda proovi ioniseeritakse laseriga (Singhal et al., 2015). Maatriks aitab laseril kuumuseks muutuda, tänu millele eemaldub proov taldrikult ning moodustab seejärel ioniseeritud molekuli. Molekulid muutuvad gaasiks, mis lendavad vastavalt oma suurusele ja laengule TOF torusse (Giuliano et al., 2019). TOF-i põhimõte seisneb selles, et antud ioniseeritud molekulid kiirendatakse kindla potentsiaalini, kus nad vastavalt m/z suhtele

üksteisest eralduma hakkavad (Singhal et al., 2015). Väiksema suuruse-laengu suhtega ( $m/z$ ) molekulid lendavad kiiremini ja see loob molekulide suurusel ning laengul põhineva massi-spektri (Giuliano et al., 2019). MS ongi analüütiline tehnika, kus keemilised osakesed ioniseeritakse laetud molekulideks ning mõõdetakse nende massi-laengu suhet ( $m/z$ ) (Singhal et al., 2015). Igal organismil on unikaalne massi-spekter, tänu millele saab neid identifitseerida kasutades raamatukogudes leitavaid standardseid spektri referentsväärtuseid (*a library of standard reference spectra*) (Paolucci et al., 2010). Vastavalt TOF-ilt saadud informatsioonile, moodustatakse uuritavatele analüütidele peptiidide massi sõrmejälj (PMF) (Singhal et al., 2015).

### 1.2.2. Kliiniline tähtsus

Peamine MALDI-TOF eelis on kiire (alla 5 tunni) kliiniliste patogeenide (seened, bakterid, viirused) tuvastamine (Giuliano et al., 2019). Seng jt. analüüsisid MALDI-TOF MS võimekust bakterite identifitseerimisel ning leidsid, et 1660-st bakteri isolaadist suutis MALDI-TOF MS õigesti tuvastada 95%, millest 84% liigi tasandil ning 11% perekonna tasandil. Keskmiselt võttis ühe isolaadi identifitseerimine aega kuus minutit (Seng et al., 2009).

Küll aga on üheks limiteerivaks faktoriks uute isolaatide tuvastamisel tõsiasi, et see on võimalik vaid siis, kui spektri andmebaasis on olemas antud perekonna/ liigi/ alaliigi/ tüve PMF (Singhal et al., 2015). Hiljuti on avalikkusele kättesaadavaks tehtud uus massispektri profiile sisaldav andmebaas, mis sisaldab kliiniliselt kõige olulisemaid isolaate ja nende antibiootikumi tundlikkuse fenotüüpe. See andmestik hõlmab endas neljast meditsiiniastutusest kogutud üle 300 000 massi-spektri, koos rohkem kui 750 000 AMR fenotüübiga. (Weis et al., 2022)

MALDI-TOF MS laialdane kasutamine on kiirendanud isolaatide identifitseerimist võrreldes traditsiooniliste biokeemiliste meetoditega, kuid see lähenemine hõlmab tüüpiliselt endas ka mikroorganismide esmast kultiveerimist, mis tähendab potentsiaalset ajakulu (Ramanan et al., 2018). Perez ja kolleegid leidsid, et kasutades MALDI-TOF tulemuste põhjal peaaegu koheselt antimikroobseid ravimeid, lühenes vereringeinfektsioonide puhul antimikroobse ravi kohandamise aeg 46 tunni võrra (Perez et al., 2013). Huang ja kolleegid leidsid, et kui kombineerida MALDI-TOF analüüs kandidoosi ja baktereemia korral antimikroobse raviga,

lüheneb optimaalse ravi rakendamise aeg 43 tundi ning antibiootikumiravi aeg 9,7 tundi (Huang et al., 2013).

#### Verekultuuris kasutamine

Kui tüüpiliselt rakendatakse MALDI-TOF meetodit isoleeritud kolooniatele agarplaadil, on mitmed uurimused analüüsinud ka MALDI-TOF meetodi rakendamist otse verekuultuurile (Kothari et al., 2014). Verekuultuuri kasutamise korral on vajalik proovi eelnev prepareerimine, et eemaldada bakter erütrotsüütidest jt. verekuultuuri koostisosadest (Croxatto et al., 2012). Antud meetodil identifitseeritakse bakter umbes 75% juhtudest ning tulemus saadakse keskmiselt ühe tunniga (Kothari et al., 2014).

Ramanan jt. testisid MALDI-TOF MS meetodit verekuultuuril (peale 4 h inkubatsioonieaga), et identifitseerida Gram(-) baktereid, mis Grami meetodil värvimise järgi proovis leidsid. MALDI-TOF MS suutis identifitseerida 92% organismidest. See tähendab, et võrreldes konventsionaalse testimisviisiga, on võimalik tulemused saada palju kiiremini ja ka ilma lisakuludeta, kui seadmed juba laboris olemas on (Ramanan et al., 2018).

### 1.3. PCR

PCR ehk polümeraasi ahelreaktsiooni kasutuselevõtt, tõi diagnostilisse mikrobioloogiasse uue ajastu. Tema toimemehhanismiks on DNA segmentide paljundamine, mida saab kasutada bakterite identifitseerimiseks (Fenollar and Raoult, 2004).

PCR analüüsi plussideks on tema kiirus ja võimekus tuvastada baktereid ka juba antibiootikumravi saaval patsiendil (Scerbo et al., 2016). PCR ei nõua otseselt eelnevat bakterite kultiveerimist (Schmidt et al., 2016). Võrreldes tavapärase BC/GS (verekuultuur koos Grami meetodil värvimisega) analüüsiga, piisab PCR testi jaoks väiksemast proovikogusest (5 ml vs 20-30 ml). Lisaks raporteeritakse PCR testidel suuremat tundlikkust (vähem vale-negatiivseid vastuseid) võrreldes kultiveerimisega (Scerbo et al., 2016).

Küll aga leidub testil ka limiteerivaid tegureid. PCR võib tuvastada baktereid, mis ei ole peale ravi saamist enam elujõulised. Vale-positiivsed tulemused võivad tulla ka peremeesorganismi mikrofloora või saaste tõttu (Scerbo et al., 2016). Saaste võib tekkida laboriinstrumentidest

(tuubid, pipetid, pinnad, tehniku riided) või kanduda üle ühelt proovilt teisele (Fenollar and Raoult, 2004). Esineda võivad ebamäärased, raskesti tõlgendatavad tulemused. Tavalise PCRi abil on raskendatud patsiendi proovis esinevate bakterite koguse ja antibiootikumresistentsuse markerite kvantitatiivne mõõtmine (Grumaz et al., 2016). Keerukas on olukord, kus proovis esineb palju erinevaid patogeene, millest testida on vaja vaid ühte (Scerbo et al., 2016). PCR ei suuda katta kõiki esinevaid organisme ning potentsiaalseid resistentsuse määrajaid (Schmidt et al., 2016).

### **1.3.1. Reaalaja PCR (*real time*)**

Algse DNA koguse kvantiseerimine on keeruline, sest PCR annab küllastusfaasi jõudes sama koguse produkti olenemata algsest DNA kogusest. Selle piirangu lahendas 1992 aastal Higuchi ja tema kolleegid, kes arendasid välja reaalaja PCR meetodi (Higuchi et al., 1992).

Kvantitatiivne reaalaja PCR (qPCR) on PCR tehnoloogia edasiarendus, kus amplifitseerimine ja amplifitseeritud produktide tuvastamine toimuvad koos ühes reaktsiooni anumal (Fenollar ja Raoult, 2004). Produktide tuvastamine võib toimuda kahel viisil, kas fluorestseeruvate ja interkaleeruvate värvidega, mis amplifitseerimise käigus seonduvad mitte-spetsiifiliselt DNA-ga või fluorestsentsmärgisega DNA sond, mis seonduvad spetsiifiliselt sihtmärki piironda (Mege et al., 1997).

Mõlema märki puhul emiteerub fluorestsentsvärv, kui segus on olemas uuritav DNA kaheahelalises vormis. Värv intensiivsus kasvab proportsionaalselt produktide koguse suurenemisega (Fenollar ja Raoult, 2004). Arvestades amplifitseerimise efektiivsust (tavaliselt kahekordistub molekulide arv ühes tsükli), on võimalik arvutada algsete DNA molekulide hulk proovis (Kubista et al., 2006).

Reaalaja PCR-l on tavapärase PCR ees mitmeid eeliseid, sh. kiirus, lihtsus, reprodutseeritavus, kvantitatiivne võimekus ning fluorestsentsmärgisega sondi kasutamisel ka madalam saastuse tuvastamise oht (Fenollar and Raoult, 2004).

### **1.3.2. Multiplex PCR**

Multiplex PCR toimib traditsioonilise PCR-ga samal meetodil, kuid ühes reaktsioonituubis töötavad korraga mitmed praimeripaarid (Scerbo et al., 2016). See tähendab, et korraga

amplifitseeritakse mitut sihtmärk järjestust ning ühes reaktsioonituubis saab tuvastada korraga nii viiruseid, baktereid ja/ või teisi nakkuslikke patogeene (Elnifro et al., 2000). Testiga saab tuvastada ka antibiootikumidele resistentseid geene (Giuliano et al., 2019). Multiplex PCR test aitab alla tuua analüüsi hinda ja laiendada testi diagnostilist võimekust (Elnifro et al., 2000).

Kuigi multiplex PCR testid on kiired, vajavad enne testi tegemist eeldatavat diagnoosi (Blauwkamp et al., 2019). Multiplex analüüsi on ka raskem disainida, sest kõik produktid, mis korraga reaktsioonis osalevad, võistlevad reagentide üle. Et vähendada võistlust, tuleb kasutada piiratud arvu praimereid (Kubista et al., 2006). Rohkem kui ühe praimeri olemasolu reaktsioonis tõstab võimalust, et moodustuvad praimeri dimeerid, mis takistavad praimeritel sihtmärgile seonduda (Brownie et al., 1997).

Praimerite valikust sõltub suuresti amplifitseerimise edukus (Suzuki ja Giovannoni, 1996). Praimerid peavad omavahel sobima nii, et nende vahel ei toimuks sekkumist. Samas peavad praimeripaarid kaasama nii palju sihtmärk patogeeni tüvesid kui võimalik ning need peavad olema lihtsasti tuvastatavad näiteks geel-elektroforeesil (Elnifro et al., 2000). Praimerite disainimine on ka keerulisem, sest praimerite vahel ei tohi olla komplementaarsust (Kubista et al., 2006).

2008. aastal kiitis USA Toidu- ja Raviamet (FDA) heaks esimese PCR multiplex paneeli (hingamisteede patogeenidele). Sellele on järgnenud veel mitme hingamisteede ja seedetrakti patogeeni paneeli vastu võtmine. Need analüüsid tuvastavad korraga 12-20 patogeeni (Schreckenberger ja McAdam, 2015).

#### **1.4. Metagenoomne sekveneerimine 16S piirkonnast (eelneva amplifitseerimisega)**

16S rRNA geen on peaaegu kõikidel bakteritel olemasolev, umbes 1550 aluspaari pikk geen, mis koosneb nii varieeruvatest kui ka konserveerunud aladest (Ranjan et al., 2016). 16S rRNA geenijärjestuste võrdlemine lubab perekonnatasemel eristada erinevaid organisme ning klassifitseerida tüvesid liikide ja alamliikide tasemel (Clarridge, 2004). Disainides praimereid, mis on komplementaarsed varieeruvate aladega, on teoorias võimalik tuvastada ükskõik mis bakteri olemasolu, mille DNA järjestus on olemas andmebaasis (Fenollar ja Raoult, 2004). Mignard ja Flandrois viisid läbi uurimuse, kus 30 kuu vältel proovisid kliinilistelt proovidelt

identifitseerida 683 bakteri isolaati. 83% juhtudest said nad vaste liigitasandil, 15% jäi vaste perekonnatasemele ning vaid 1% uuritud tulemustest jäid tuvastamatuks. Autorite hinnangul on 16S rRNA geeni amplifitseerimisel ja sekveneerimisel oluline roll bakteriaalsete patogeenide tuvastamisel, eriti aeglaselt kasvavate, ebatavaliste ja nõudlike liikide puhul, kui ka bakterite, keda konventsionaalsete meetoditega on raske eristada (Mignard ja Flandrois, 2006).

Kui traditsiooniliselt sooritati 16S rRNA geeni sekveneerimist Sangeri meetodil, siis järgmise põlvkonna sekveneerimise (NGS) kasutuselevõtmine on seda protsessi oluliselt lihtsustanud. Nüüd võimaldavad nõ "kapipealsed NGS masinad" laboritel 16S rRNA geeni paari päevaga põhjalikult sekveneerida (Sanschagrin ja Yergeau, 2014).

Tema eelisteks on kuluefektiivsus, juba kindel andmeanalüüsimise meetod ning suur olemasolev referentsjärjestuste andmebaas (Ranjan et al., 2016). Samas võib 16S rRNA geeni sekveneerimine tuvastamata jätta haigustekitaja, kui tegemist on seene, parasiidi või viirusega. Ta ei pruugi näidata ka kõiki baktereid ning ei kuva informatsiooni virulentsusfaktorite, antibiootikumresistentsuse või tüve tüüpide kohta (Afshinnekoo et al., 2017). Seetõttu ei ole antud metoodika kliinilises praktikas praeguseks eriti levinud.

### **1.5. Mikroobide tuvastamine isolaadi genoomi sekveneerimisel (WGS)**

Erinevalt 16S geeni sekveneerimisest, identifitseerib terve genoomi sekveneerimine (WGS) mikroorganismid kõikidest taksonoomilistest kuningriikidest: nii bakterid, viirused, seened kui algloomad ning teeb kindlaks nende nukleotiidsed järjestused (Ranjan et al., 2016). WGS annab võimaluse iseloomustada nii geneetilist ja genoomset mitmekesisust, kui ka infot potentsiaalsete ning uute funktsioonide kohta (läbi geeni poolt kodeeritavate valkude funktsiooni ennustamise) analüüsitud keskkonnas (Pérez-Cobas et al., 2020). WGS-i saab kasutada mikroobide taksonoomiliseks liigitamiseks, potentsiaalsete funktsioonide kirjeldamiseks ning terve genoomi järjestuse taastamiseks. (Quince et al., 2017).

Kõigepealt kultiveeritakse huvipakkuvad organismid, seejärel eemaldatakse need taldrikult ning eraldatakse DNA. Edasi tuleb koostada raamatukogu ja sekveneerimine väljavalitud platvormil. Peale sekveneerimist assambleeritakse lugemid bioinformaatilise programmiga

ning kasutades referents andmebaase, tuvastatakse huvipakkuvad organismi (Hilt ja Ferrieri, 2022).

Illumina ja Thermo Fisher pakuvad lühikese lugemiga (100-400 aluspaari (bp)) platvorme. Pacific Biosciences ja Oxford Nanopore Technologies (ONT) pakuvad pika lugemiga sekveneerimisi (>500 bp) (Shendure ja Ji, 2008). Igal platvormil on enda eelised ja puudused, arvestades täpsust, efektiivsust ja hinda. Olenemata kasutatud tehnoloogiast, on vaja bioinformaatilisi teadmiseid, et andmeid töödelda ja analüüsida (Petersen et al., 2019). Bioinformaatiline analüüs nõuab professionaalseid töötajaid ja pädevaid analüüsimise instrumente (k.a. referents andmebaas, arvuti infrastruktuur, standardsete protseduuride loomine) (Gu et al., 2019).

Esialgne proovide ettevalmistus, kus pärilikkusaine tuleb isoleerida, kvaliteedinäitajad kontrollida ning raamatukogu ettevalmistamise protokoll paika panna, võib võtta aega tunde kuni päevi. Sekveneerimine ise kestab 1-6 päeva, olenevalt kasutatavast platvormist, lugemite pikkusest ja genereeritud andmete kogusest (Petersen et al., 2019).

#### **1.6. Mikroobide tuvastamine sekveneerimisel ilma eelneva amplifitseerimiseta (WMS)**

Metagenoomne sekveneerimine (WMS) ehk *shotgun* (i.e. ilma sihtmärgita) sekveneerimine on kõigi ("meta") proovis leiduvate mikroobide genoomide suure läbilaskevõimega sekveneerimine, mis võimaldab ühes proovis teostada genoomseid analüüse kõikidele mikroobidele (Quince et al., 2017). Mitmed uurimused on näidanud, et metagenoomse sekveneerimise tehnoloogiad on paljulubav viis identifitseerida ja jälgida ka nakkushaiguste etioloogiaid ning tuvastada AMR-i tekke (Rasko et al., 2011; Köser et al., 2014; ).

Metagenoomne teise põlvkonna sekveneerimine ehk mNGS võimaldab otse proovist sekveneerida samaaegselt kõiki seal esinevaid nukleiinhappeid, võimaldades tuvastada suurema osa proovis leiduvatest mikroobidest (Wu et al., 2022; Wang et al., 2019). Tänu sellele on mNGS abil võimalik tuvastada ka uusi ja haruldasi mikroorganisme (Diao et al., 2022). mNGS-i saab rakendada suurele hulgale erinevatele proovidele (rõga, kõri tampooniproov, veri, alveoolide loputusvedelik, pleuravedelik, tserebrospinaalvedelik, mäda, koeproovid jne.), kust siis otse nukleiinhappeid tuvastada (Zheng et al., 2021). Tulemused on väiksema tõenäosusega mõjutatud eelnevast antibiootikumiravist kui traditsioonilise kultiveerimismeetodi puhul (Miao et al., 2018). Antud meetodiga on võimalik analüüsida



hingamisteede mikrobioomi, inimorganismi vastuseid erinevatele teguritele (antibiootikumid, vaktsiinid jm.) ning ennustada ravimiresistentsust (Diao et al., 2022).

mNGS hindab proovis leiduvate mikroorganismide tüüpe lähtudes sekveneerimisel saadud informatsioonist ning võrreldes ja analüüsides neid tuntud mikroobijärjestuste andmebaasiga (Zinter et al., 2019). Peaaegu kogu pärilikkusaine enamustes kliinilistes proovides on pärit inimeselt, samas on mNGS puhul huvipakkuvad mikroobse päritoluga nukleiinhapped. Seega on patogeenide leidmine võrdne nõela heinakuhjast otsimisega (Gu et al., 2019). Kuna mNGS on äärmiselt tundlik, võivad ilmned ka vale-positiivsed tulemused. Nukleiinhapete saastumine võib toimuda alates proovi korjamisest kuni selle töötlemiseni, kusjuures saastumine saab tulla ka ümbritsevast keskkonnast, nt. proovi hoidmise nõust (Zheng et al., 2021). Kuigi saastumist lõplikult elimineerida on võimatu ning keeruline on eristada saastumise käigus ilmnenu mikroobset järjestust õigest järjestusest, on võimalik kasutada ettevaatusabinõusid, et saastumisohtu minimeerida. See hõlmab endas nt. tihedamat materjalide ja pindade puhastamist 10% naatriumhüpokloritiga ning ühesuunalise töövoo ja füüsilise eraldatuse hoidmist (Diao et al., 2021).

Erinevad uurimused on patogeenide tuvastamisel hinnanud mNGS tundlikkuseks 70%-94% ning spetsiifilisuseks 63%-90%, kasutades referentseteks standarditeks konventsionaalsete meetodite kasutamist (Long et al., 2016; Blauwkamp et al., 2019). Huang ja kolleegid leidsid, et 240 kopsupõletiku kahtlusega patsiendi seast tuvastas mNGS nakkuse 89%-l, samas kui konventsionaalne meetod tuvastas nakkuse vaid 26%-l (Huang et al., 2020).

Laialdast rakendamist hoiab tagasi kõrge hind ning vastuste saamiseks kulub, olenevalt tehnoloogia kättesaadavusest, parimal juhul 24-48 tundi (Petersen et al., 2019). Küll aga on enamuse mNGS testides saadaval vaid kolmandate osapoolte laboratuurimites, mis tähendab, et transpordi peale kulub palju aega ning proovi võtmisest lõppdiagnoosi saamiseni läheb keskmiselt 98 tundi (48 - 243 tundi) (Farnaes et al., 2019). Hinna langemisel muutub tehnoloogia kättesaadavamaks ning omakorda lüheneb diagnoosimiseks kuluv aeg (Han et al., 2020).

## **2. MIKROOBIDE TUVASTAMINE VEREPLASMAST**

### **2.1. Millest koosneb cfDNA, millised on tema omadused**

Rakuvaba ringlev DNA (cfDNA) avastati esmalt 1948. aastal (Mandel ja Metais, 1948). Ta koosneb nukleiinhappe fragmentidest, mis leiduvad veres ja teistes bioloogilistes vedelikes (Weerakoon ja McManus, 2016). Lisaks inimese DNA-le leidub plasmas ka mikroorganismide nukleiinhappeid (mcfDNA). Usutakse, et nukleiinhapped pärinevad kehas leiduvatelt lagunevatelt mikroorganismidelt, kes surres enda sisu verre vabastavad (Fernández-Carballo et al., 2019). Mikroobne DNA võib vereringesse sattuda kahel viisil: kas läbi välismaailmaga suhtlevate organite epiteeli limaskestast (seedetrakt, suu, reproduktiivorganid jm.) või kui kudede limaskestast on lokaalse nakkuse või füüsilise kahjustuse tõttu kahjustunud. Vereringes lagundatakse mikroobsed nukleiinhapped nukleaaside poolt, mis seejärel moodustavad väikesed DNA fragmendid ehk mcfDNA (Han et al., 2020).

Enamus rakuvälisest DNA-st (kuni >99%) moodustab inimese DNA ning ülejäänud mcfDNA, millest 0,08%-4,85% pärineb bakteritelt, 0,00%-0,01% seentelt ning 0,00%-0,16% viirustelt/faagidelt (Huang et al., 2018). Grumaz jt. näitasid, et haiguslikus seisundis võib see osakaal muutuda. Nt. sepsise korral oli mikroobse DNA osakaal plasmas keskmiselt 9,8%, kui tervetel patsientidel oli see keskmiselt 3,5% (Grumaz et al., 2020).

Tervete inimeste vereplasmas on DNA kontsentratsioon võrdlemisi madal: 3,6-5,0 ng/ml, kusjuures enamus DNA fragmente on 180 aluspaari pikad (Suzuki et al., 2008). Ilmselt tänu enda lühikesele pikkusele, suudab osa DNA-st ka neerud ületada ning uriini jõuda (Lichtenstein et al., 2001). Uuringud on näidanud, et cfDNA omadused ja kontsentratsioon võivad muutuda erinevates füsioloogilistes või haiguslikes tingimustes (Weerakoon ja McManus, 2016). Näiteks sepsise puhul on täheldatud cfDNA kontsentratsiooni suurt kasvu, eriti infektsiooni algusjärgus on cfDNA kontsentratsioon keskmiselt 197 ng/ml - 377 ng/ml, mis viitab suuremale bakterite koormusele. Suurenenud cfDNA tasemeid on kirjeldatud ka teistel füsioloogilistel või patoloogilistel tingimustel, nt. raseduse, põletike, traumade, vähi ja operatsioonide tõttu (Grumaz et al., 2016).

### 2.1.1. cfDNA kliiniline tähtsus

Üks levinumaid cfDNA analüüsi kasutusalasid on loote varane ja mitteinvasiivne geneetiline analüüs (NIPT), tänu sellele, et loote DNA ringleb ema veres (Sayres ja Cho, 2011). Paljulubav on cfDNA analüüs ka onkoloogias, kus “vedela biopsia” abil saab kiirelt, kuluefektiivselt ja mitteinvasiivselt analüüsida apoptoosis või nekroosis kasvajakude eritatud ringlevaid DNA fragmente (ctDNA), et kasvajakke monitoorida või kirjeldada (Qin et al., 2016). Kolmas paljulubav ala cfDNA kasutamiseks on siirdamiste valdkond, kus rakuvälise DNA kogust saab kasutada organi äratõukereaktsiooni markerina (Moreira et al., 2009).

Hong jt. näitasid, et puhastatud plasmast saab sekveneerimisega identifitseerida patogeene, mis on pärit erinevatest organitest. See tähendab, et cfDNA integreerib infektsioone puudutavat informatsiooni üle keha (Hong et al., 2018). NGS meetodiga on kerkinud võimalus mitteinvasiivselt mikroobset cfDNA-d sekveneerides tuvastada ühe analüüsiga infektsioone üle terve keha, tänu millele saab cfDNA sekveneerimist kasutada varajaseks diagnoosimiseks (Blauwkamp et al., 2019; Weerakoon ja McManus, 2016).

Huvi nakkushaiguste diagnoosimiseks cfDNA-st aina kasvab, eriti haiguste puhul, millel pole lihtsasti kasutatavaid vere või uriini teste (Fernández-Carballo et al., 2019). mcfDNA sekveneerimine suudab identifitseerida patogeene ka siis, kui tavapärased testid on negatiivsed või diagnostilisi proove on liiga keeruline võtta (O’Grady, 2019). See toimib ka juhtumite korral, kus enne proovi võtmist on juba alustatud antibiootikumikuuriga või kui bakterid on muidu raskesti kultiveeritavad (Han et al., 2020).

Oluline mure mcfDNA analüüsimisel on nukleiinhappe saastumine, kas siis inimese DNA või proovi töötlemisel välise mikroobse DNAGA. Kuna mcfDNA-d leidub plasmas niivõrd vähe, vähendab patsiendi DNA foon oluliselt testi tundlikkust ning võib esineda vale-negatiivsust (Han et al., 2020). Välise mikroobse DNAGA saastumine tõstab tulemuste tõlgendamise keerukust ning võib esineda vale-positiivne tulemus (Eisenhofer et al., 2019).

## 2.2. Näiteid bakterite tuvastamisest cfDNAst

### 2.2.1. Sepsis

Sepsist defineeritakse kui eluohtlikku organite talitlushäiret, mis on põhjustatud peremeesorganismi reguleerimata vastusest infektsioonile (Purcarea ja Sovaila, 2020). Sepsise juhtumid on viimaste aastakümnetega kahekordistunud ning ennustatakse jätkuvat kasvu. Seda vananeva populatsiooni, siirdamiste, kemoteeraapia, invasiivsete protseduuride ning immunosupressiivse ravi ja krooniliste tervisemurede kasvu tõttu (Mayr, 2014). Sepisega kaasneb kõrge haigestumus ja suremus; et tulemusi parandada, on vaja rakendada kohest ravi (Purcarea ja Sovaila, 2020).

Blauwkamp jt. hindasid Kariuse testi võimekust tuvastada sepsist põhjustav patogeen 348 patsiendi hulgas. Kariuse test on kvantitatiivne mikroobse cfDNA sekveneerimise test, mis identifitseerib inimese verest 1250 inimpatogeeni nende genoomse DNA järgi. Kõikidele patsientidele teostati nii cfDNA sekveneerimine kui ka verekuultuur. Võrreldes verekuultuuriga, oli cfDNA sekveneerimise tundlikkus 93,7%. Kariuse test andis positiivseid vasteteid ka patsientidel, kes olid eelnevalt saanud antibiootikume või kellel esines viiruseid, eukarüootseid patogeene ja kultiveerimatuid baktereid. Nendel verekuultuur positiivset vastet ei näidanud.

Leiti, et sepsiseohus patsientide testimisel oli cfDNA sekveneerimisel suurem patogeene tuvastamise määr (169/348) kui verekuultuuril ning teistel mikrobioloogilistel testidel (koe ja vedelike kultiveerimine, seroloogia, nukleiinhapete testimine) kokku (132/348). Kariuse testiga oli keskmine vastuse saamise aeg 53,0 tundi, traditsioonilistel tuvastamismeetoditel oli see 92,4 tundi.

27,7% patsientidest olid saanud antimikroobset ravi kahe nädala jooksul enne proovi andmist ning nendest suutis cfDNA sekveneerimine tuvastada sepsist põhjustavaid patogeene 47,9%-l. Verekuultuuril oli see võimekus vaid 19,6%.

Organismid, keda cfDNA sekveneerimine identifitseeris kui võimalikke või ebatõenäolisi sepsise põhjustajaid, hõlmasid endas aktiveerunud herpesviiruseid ja kroonilisi nakkuseid nagu *Helicobacter pylori* ning inimese papilloomiviirus. Asümptomaatiliste patsientide proovidest leiti mikroobset cfDNA-d 22,8% juhtudes, tuvastades kõige tihedamini *H. pylori*, *Klebsiella pneumoniae* ja *Haemophilus influenzae* bakterid (Blauwkamp et al., 2019).

Grumaz jt. analüüsisid ühe aasta jooksul 120 patsiendi vereplasmat, et sekveneerimise abil tuvastada haigust põhjustavad mikroobid rakuvabast DNA-st. Nad näitasid, et proovi võtmisest tulemusteni läks aega kuni 30 tundi. Tuvastati, et cfDNA sekveneerimise tulemused olid heas korrelatsioonis verekultuuri jt. bioloogiliste proovidega, kuigi leidis paar erandit:

1) Ühel patsiendil ei näidanud sepsise jooksul verekultuur kordagi positiivset vastet, küll aga haavalapid ning kõhukoopa loputus andsid *Escherichia coli* ja *Enterococcus faecium* vasted. Sepsise alguses võetud cfDNA proovid näitasid *E. coli*, *E. faecium*, ja *Bacteroides fragilis* baktereid. Nakkuse jooksul jäi patsient ka kopsupõletikku, mida põhjustasid traditsioonilistel meetoditel tuvastatud *E. coli*, *Stenotrophomonas maltophilia* ja *Klebsiella pneumoniae*. cfDNA NGS näitas vasteid *K. pneumoniae* ja *B. fragilis* bakteritele. Seitsmendal päeval tuvastas qPCR HSV1 viiruse, mida kinnitas cfDNA sekveneerimine plasmast. Antud patsiendi juhtum näitab, et NGS lähenemine võib olla tundlikum ja spetsiifilisem kui verekultuur.

2) Teisel patsiendil tuvastas verekultuur koos qPCR-iga *Staphylococcus aureus*'e. Seda kinnitas plasma cfDNA sekveneerimine. Seitsmendal päeval tuvastas vereplasma sekveneerimine HSV1 viiruse, mille tuvastas ka qPCR 14-ndal päeval. Kultiveerimine näitas ka väikest kogust *Candida albicans* seent, mis sekveneerimisel ei tulnud välja kui oluline organism.

Uurimistulemused näitasid, et cfDNA sekveneerimine suudab suhteliselt väikse kallutatusega analüüsida vereringe infektsioone ning võib olla eriti kasulik kui klassikaline mikrobioloogia või molekulaardiagnostilised meetodid ei õnnestu (Grumaz et al., 2016).

Abril et al. raporteerisid juhtumit, kus 60-aastane patsient saabus haiglasse palaviku, kõrgenenud vererõhu ning vaimse seisundi muutustega. Verekultuur ei andnud tulemusi, misjärel viidi läbi mcfDNA sekveneerimine ning leiti, et sümptomid on põhjustanud *Capnocytophaga canimorsus* põhjustatud sepsis. See näitab, et rakuvaba DNA sekveneerimine on paljulubav meetod raskesti kultiveeritavate patogeenide tuvastamisel (Abril et al., 2016).

### 2.2.2. Tuberkuloos

Tuberkuloos (TB) on hea näide haigusest, mille puhul cfDNA sekveneerimine on eriti kasulik (Han et al., 2020). Tuberkuloosi põhjustab *Mycobacterium tuberculosis* (MTB), mille kliiniline tuvastamine on raskendatud tema pika peiteaja ja mittespetsiifiliste sümptomite tõttu. Tüüpiliselt diagnoositakse TB happekindla värvingu, kultiveerimise ja rögast nukleiinhappe amplifitseerimise kaudu (Ushio et al., 2016). See võib aega võtta kuni paar nädalat ning sügavalt juurdunud infektsioonide puhul on tihti vaja teostada invasiivset biopsiat (Nomura et al., 2019). Varajaseks tuberkuloosi diagnoosimiseks on loodud mitu PCR meetodil põhinevat cfDNA testi nii verest kui uriinist, näidates, et mcfDNA on kasulik biomarker tuberkuloosi tuvastamiseks ja ravi jälgimiseks (Fernández-Carballo et al., 2019).

Click jt. uurisid 40 patsiendi seas plasma cfDNA võimekust diagnoosida tuberkuloosi võrreldes traditsiooniliste meetoditega ning leidsid, et isegi kui ei esine mükobaktereemiat, suudab qPCR tuvastada vereplasmast *Mycobacterium tuberculosis complex* (MTC) DNA-d 45% haigetest. Uuringu läbiviijate seisukoht on, et cfDNA uurimine qPCRi abil on paljulubav meetod tuberkuloosi diagnoosimiseks, mida tuleb veel optimeerida (Click et al., 2018).

### 2.2.3. Kopsupõletik

Kopsupõletik on levinud nakkus, millele tihti ei suudeta diagnoosi panna. Paljudele patsientidele määratakse antibiootikumid, mis võivad vähendada kultiveerimise edukust. Kopsupõletiku diagnoosimisel muudab sekveneerimistulemuste analüüsi keerulisemaks suufloora mikroobne foon, kus esinevad bakterid võivad teatavatel juhtudel olla ka patogeeneid. Seega võib kvantitatiivsetest või semikvantitatiivsetest statistilistest analüüsides olla abi, et eristada infektsiooni kolonisatsioonist (Gu et al., 2019).

Langelier viis koos kolleegidega läbi uurimuse, mille eesmärgiks oli näidata, et plasma cfDNA sekveneerimine suudab patsientide verest tuvastada kopsupõletikku põhjustavate patogeene DNA, kaasa arvatud patsientidel, kelle verekuultuur on negatiivne. Valimi

suurus oli 25 kriitiliselt haiget patsienti, kellest 18-l oli kopsupõletik ning seitsmel akuutne hingamisteede haigus. Teiseks eesmärgiks oli võrrelda plasma cfDNA ja hingamisteede vedeliku mNGS vasteid patogeenidele (viimast peetakse täpseks kopsupõletiku diagnoosimise meetodiks). Uurimusele eelnevalt olid 83% patsientidest saanud empiirilist laia spektriga antibiootikumide ravi.

Esmalt tehti mNGS kontrolltestid kultiveerimise käigus tõestatud baktereemiale (*Cytomegalovirus viremia*) ja leiti, et plasma cfDNA sekveneerimine identifitseeris patogeene igal juhtumil.

Uurimuse käigus said kaks patsienti ventilaatorpneumoonia (VAP) diagnoosi. Nendel juhtudel oli plasma cfDNA sekveneerimine esimene, mis tuvastas patogeensed bakterid. Kultiveerimine andis sama tulemuse 4 päeva hiljem. Nendeks bakteriteks olid esimese patsiendi puhul *Enterobacter cloacae* ja *Staphylococcus aureus* ning teise patsiendi puhul *Serratia marcescens*. Teine patsient sai juba ravi (tsesfriaksoon) *Escherichia coli* ja *Klebsiella pneumoniae* vastu, kuid mNGS tuvastas ka SRT-1  $\beta$ -laktamaas geeni, mis põhjustab tsesfriaksooni resistentsust ning põhjendab, miks VAP võis tekkida.

Sekveneeriti ka seitse kriitiliselt haiget, kuid negatiivse kopsupõletiku tulemusega patsienti, kellel esines hingamispuudulikkus. Kahel juhtumil ei leitud mikroobe, kuid viiel patsiendil ilmnis teiste patogeene seas Inimese Herpesviirus 6, mis võib esineda tervetel indiviididel ja muutuda aktiivseks sepsise korral. Sellistel juhtumitel võivad tuvastatud mikroobid olla kopsupõletiku tekitajad või hoopis sattunud sinna soolestikust, hingamisteedest, nahast või mujalt. Patogeene leidmine näitab, et cfDNA sekveneerimine on kasulik täiendav diagnoosimeetod kui verekuultuur on negatiivne.

Näidati, et plasma mNGS ei ole nii täpne kui hingamisteede mNGS (67% vs 100%), kuid suudab tuvastada kultiveerimisega tõestatud hingamisteede patogeene cfDNA-st ning võib olla väärtuslik abivahend, kui hingamisteede proove ei saa võtta. Tulemused näitasid ka, et kuigi verekuultuur on negatiivne, võib toimuda siiski patogeense DNA vabanemine vereringesse (Langelier et al., 2020).

Wu jt. teostasid 502 kopsupõletiku kahtlustusega patsiendi seas mNGS bronhoalveolaarsest loputusvedelikust (BALF), et leida haigusttekitav patogeen. Tulemused võrreldi konventsionaalsete mikrobioloogiliste testidega. Leiti, et mNGS tundlikkus oli 72,5% ning täpsus 74,9%. Konventsionaalsete testide seas olid need tulemused vastavalt 25,4% ja 36,9%. Nende patsientide seast diagnoositi kopsupõletik 84,1%-le ning 15,9%-l esines mittenakkav haigus. 355-l patsiendil suudeti tuvastada põhjustav patogeen, kuid

67-l patsiendil jäi see teadmata. Ühesugused haigustekitajad suutsid mNGS ja konventsionaalsed meetodid tuvastada 120-l juhul. 239 patogeeni suutis tuvastada mNGS üksi ning 12 patogeeni vaid konventsionaalsed meetodid.

Kokkuvõtteks võib öelda, et hingamisteede mNGS näitas paremat diagnostilist võimekust kui konventsionaalsed meetodid, eriti mükoplasma, klamüüdia ja viiruste korral. *Acinetobacter baumannii*, *Nocardia*, *Staphylococcus aureus*, *Escherichia coli* ja *Legionella* puhul ei leitud selle valimi juures kahe diagnoosimismeetodi vahel statistiliselt olulist erinevust (Wu et al., 2022).

#### **2.2.4. Kuseteede haigused**

Kuseteede haigused (UTI) on ühed levinumad inimeste nakkused. Semikvantitatiivne bakterite kultiveerimine on UTI-de diagnoosimise standardmeetod, kuid see on aeganõudev ning ei pruugi olla usaldusväärne, kui patsient saab antibiootikume. Metagenoomne uriini cfDNA sekveneerimine võib olla väärtuslik abivahend nakkustekitaja identifitseerimiseks (Janes et al., 2022).

Burnham jt. testisid uriinis leiduva rakuvaba DNA kasulikkust bakteriaalsete ja viraalsete kuseteede haiguste monitoorimiseks. Nad testisid NGS meetodil 141 uriiniproovi ning leidsid, et ühe analüüsiga on võimalik infot saada nii uriini cfDNA koostise, antibiootikumitundlikkuse, bakterikasvu dünaamika, siirdatud neeru vigastuste kui ka peremeesorganismi nakkuse vastuse kohta. Peaaegu kõikides proovides, kus leidis kliiniliselt tõestatud viraalseid või bateriaalseid kuseteede infektsioone, suutis cfDNA sekveneerimine identifitseerida põhjustava patogeeni. Ühes tuvastas kultiveerimine bakteri, mis cfDNA-s ei kajastunud: *Raoultella ornithinolytica*. Lisaks tuli cfDNA-st välja baktereid, keda praegune kliiniline praktika tuvastada ei suuda, nt. *Haemophilus influenzae*, mis põhjustas patsiendil baktereemia ning selgus lõpuks ka kultiveerimise käigus, kui antud proov uuesti šokolaadi agarile külvati. Uuring leidis, et uriini cfDNA sekveneerimine on väga mitmekülgne vahend UTI haiguste monitoorimiseks (Burnham et al., 2018).



Janes jt. võrdlesid metagenoomse uriini cfDNA sekveneerimise tulemusi traditsioonilise bakterite kultiveerimisega. Nad analüüsisid 86 uriiniproovi, millest 43 olid kultiveerimisel saanud positiivse tulemuse ning 43 negatiivse.

86st proovist identifitseeriti 200 bakteriliiki, millest kõige rohkem esinesid *Escherichia coli*, *Pseudomonas aeruginosa*, *Enterococcus faecalis*, *Proteus mirabilis*, *Klebsiella pneumoniae*, *Aerococcus urinae*, *Bifidobacterium longum*, *Serratia marcescens*, *Morganella morganii* ja *Actinotignum schaalii*. cfDNA sekveneerimine tuvastas positiivseteks 39 proovi, millest 87% olid õiged positiivsed. Viit vale-positiivset tulemust tõlgendas kultiveerimine kui fekaalne saaste. 47st negatiivsest proovist olid päriselt negatiivsed 38 ehk 81%. Üheksast valenegatiivsest tulemusest kaheksa olid Gram-positiivseid baktereid, millest võib järeldada, et cfDNA sekveneerimine tuvastab paremini Gram-negatiivseid baktereid kui Gram-positiivseid.

Kõikidel teistel proovidel suutis uriini cfDNA sekveneerimine tuvastada rohkem liike kui kultiveerimine, kaasa arvatud patogeene, keda kultiveerimine tihti tuvastada ei suuda, nt. *A. schaalii*. Uurimine näitas, et uriinist mcfDNA sekveneerimine võib olla väärtuslik diagnoosimeetod (Janes et al., 2022).

### 2.2.5. Haruldased infektsioonid

#### *Mycobacterium chimaera*

*M. chimaera* on *Mycobacterium avium* kompleksi kuuluv bakter, mida kirjeldati esmalt 2004 aastal. Ta põhjustab hingamisteede haiguseid, eelkõige immuunpuudulikkusega patsientidel. Viimasel aastakümnel raporteeritakse üha rohkem tema põhjustatud infektsioone, eriti avatud südameoperatsiooni järgselt (Riccardi et al., 2020). *M. chimaera* poolt põhjustatud sümptomid on mitte-spetsiifilised, hõlmates endas väsimust, higistamist, palavikku, köha, kaalulangust ja hingeldamist ning need võivad esineda kuni 6 aastat peale operatsiooni (Marra et al., 2017). Standardne bakterite tuvastamine teostatakse happekindla bakteri (AFB) kultiveerimise ja värvinguga, mis nõuab invasiivset proovivõtmist, on limiteeritud tundlikkusega ning mükobakteri kasvamine võib võtta aega kuni 8 nädalat (Kohler et al., 2015).

Nomura ja kolleegid testisid ühe aasta jooksul südameoperatsiooni järgselt kümnet *M. chimaera* bakteriga kokkupuutunud nakkuse kahtlustusega patsienti, teostades neile

cfDNA sekveneerimise. Plasma cfDNA sekveneerimine tuvastas *M. chimaera* üheksal patsiendil kümnest, AFB värving vaid neljal, kusjuures AFB kultiveerimiseks oli vaja invasiivset proovi võtta 90% patsientidest. Proovi võtmisest tulemuseni läks NGS meetodiga aega keskmiselt 4 päeva, AFB kultiveerimisel saadi tulemus keskmiselt 41 päeva peale proovi võtmist.

Antud valimil oli suremus 60%. Neljal patsiendil, kes infektsiooni üle elasid, tuvastas NGS esimesena haigustekitaja, mis näitab, et varasel diagnoosimisel on potentsiaal vähendada suremust (Nomura et al., 2019).

### *Propionibacterium acnes*

Ye et al. raporteerisid juhtumit, kus haruldase juveniilse müelomonotsütaarse leukeemiaga 2-aastasel poisil esines teadmata päritoluga lööve ning kõrge palavik. Traditsiooniline kliiniline diagnostika ei suutnud tuvastada ühtki põhjustavat patogeeni. Vereplasmale teostatud NGS suutis tuvastada *Propionibacterium acnes* bakteri. Alustades kohest spetsiifilist ravi taganesid sümptomid täielikult viie päevaga (Ye et al., 2016).

### Nakkused peale tüvirakkude siirdamist

Fung et al. kirjeldasid kolmel allogeensete vereloome tüvirakkude siirdamisega patsienti, kellel oleks vereplasma mcfDNA sekveneerimine hõlbustanud ja kiirendanud nakkust põhjustava bakteri kindlaks tegemist.

Esimesel patsiendil oli palsma NGS positiivne *Chlamydia trachomatis* bakterile 30 päeva enne, kui klassikaline testimismeetod (invasiivne ureteroskoopia stendi paigaldusega) seda näitas ning oleks lühendanud haiglasoleku aega 10 päeva võrra.

Teisel patsiendil teostati juba kahe nädalane metitsilliinresistentse *Staphylococcus aureus* (MRSA) baktereemia ravi ning verekuultuur näitas bakteri taganemist. Sellegipoolest tema näidud halvenesid. Plasma NGS näitas peale esialgse ravi manustamist veel MRSA esinemist, kuid verekuultuur tuvastas selle alles 10 päeva hiljem, peale mida patsient suri. Plasma NGS suudab tuvastada patogeene 2-5 päeva enne korduva baktereemia tuvastamist ning jääb positiivseks keskmiselt 3 päeva kauem kui verekuultuur. Plasma NGS oleks varasemalt näidanud baktereemia taas tekkimist ning potentsiaalselt päästnud patsiendi elu.

Ka kolmandal patsiendil esines korduv MRSA baktireemia, kuid nakkust põhjustavaks patogeeniks tuvastati ekslikult *S. epidermidis*. Sümptomid süvenesid ja alles kolm päeva hiljem tuvastas verekuultuur *S. aureus* bakteri. Plasma NGS oleks tuvastanud juba varem püsiva varjatud infektsiooni, sest peale esialgse ravi saamist jäi MRSA näit veel positiivseks (Fung et al., 2018).

### 2.3. Näiteid patogeensete seente/pärmide tuvastamisest cfDNAst

Immuunsupresseeriv ravi ja seenevastasele ravile resistentsete organismide kasv on toonud invasiivsete seennakkuste (IFD) suure kasvu, mis põhjustab immuunpuudulikkusega patsientide seas suure haigestumise ja suremuse (Perlin et al., 2017). Praegune diagnostiline meetod jagab IFD kolme klassi: tõestatud, tõenäoline (riskikrupis olev peremeesorganism vastas kliinilistele ja mükoloogilistele kriteeriumitele) ja võimalik (riskikrupis olev peremeesorganism vastab kliinilistele kriteeriumitele kuid mitte mükoloogilistele) haigus (De Pauw et al., 2008). Võttes arvesse patogeensete seente suurt mitmekesisust, on kerkinud esile kriitiline vajadus kiire ja mitteinvasiivse meetodi järele. Ideaalis peaks see suutma haigustekitajaid identifitseerida liigi tasemeni, et rakendada spetsiifilist seenevastast ravi (Han et al., 2020).

Hong ja kolleegid analüüsisid 2018 aastal IFD patsientide plasma cfDNA-d ja leidsid, et sekveneerides on võimalik tuvastada nakkust põhjustavad patogeenid, mida on raske kultiveerida, nende hulgas *Aspergillus*, *Rhizomucor* ja *Scedosporium* liike.

Seitsmel juhul üheksast tuvastas plasma NGS sama seene, mis nakatunud koe (kopsu, siinuse, aordi, aju jt.) biopsia. Mõnel juhul tuvastas konventsionaalne mikrobioloogia hallituse vaid perekonnatasemel (nt. *Rhizopus* liigi). Ühel juhul tuvastas plasma NGS *Aspergillus lentuluse*, kuid konventsionaalne mikrobioloogia tuvastas *Aspergillus fumigatus* liigi. *A. lentulus* ja *A. fumigatus* on morfoloogiliselt identsed ning jagavad 91% järjestuses homoloogiat. Mõnel juhul ei leidnud plasma NGS patogeeni, kuid see võib olla põhjustatud proovi võtmise hilinemisest. Ühel juhul võeti plasma proov 20 päeva hiljem kui teostati biopsia. Teistel juhtudel esinesid biopsiaga tõestatud *Aspergillus fumigatus* järjestused proovis, kuid mitte piisavalt suurel hulgal, et kvalifitseeruks positiivse proovi alla. Kuuel juhul üheksast identifitseeris plasma NGS lisaks seentele proovist veel viiruseid ja baktereid, mida kultiveerimine ei näidanud.

Uurimus näitas, et nakkuseid põhjustavaid mikroobe saab mitteinvasiivelt kiiremini diagnoosida, kui seda võimaldab nakatunud koe biopsia ning tõestas, et cfDNA sekveneerimine võib olla hea alternatiiv (Hong et al., 2018).

Armstrong jt. võtsid ühe aasta vältel korduvaid vereproove 40-lt patsiendilt, et uurida cfDNA NGS võimekust identifitseerida patogeenseid seeni. Kuuel patsiendil esines mikroskoopiliselt tõestatud IFD, ühel oli tõenäoline, 11-l võimalik IFD, ühel oli kõrgenenud 1,3- $\beta$ -d-glukaani tase, kuid ei esinenud kliinilisi leide ning 21 patsienti ei vastanud IFD kriteeriumitele.

Neljal juhul kuuest tõestatud IFD juhtumist identifitseeris cfDNA sekveneerimise sama patogeeni, mis invasiivne biopsia (kas nahalt, kopsust või pseudotsüsti vedeliku äravoolust) ja verekuultuur. Ühel juhul esines kultiveerimisel *Rhizopus oryzae*, kuid cfDNA sekveneerimine andis negatiivse vaste. See võib tuleneda lokaliseeritud ja levinud haiguse erinevusest, sest *R. oryzae* tuvastati pealmisest nahabiopsiast. Teisel juhul näitas kultiveerimine *Aspergillus fumigatus* kohalolu, kuid cfDNA NGS tuvastas hoopis *Pneumocystis jirovecii* mikroobi, mis võib seletada temal esinenud normist kõrvalekaldunud rinna kompuutertomograafia tulemust koos teiste sümptomitega ning võib viidata koinfektsioonile.

Tõenäolise IFD-ga patsiendil ei andnud kultiveerimine vastet, kuid cfDNA NGS tuvastas *P. jirovecii*.

Kõikidel teistel juhtumitel kattusid cfDNA NGS ja konventsionaalsete testimismeetoditega tehtud tulemused (Armstrong et al., 2019).

#### **2.4. Näiteid viiruste tuvastamisest cfDNAst**

Viirused mängivad inimese mikroobioomis suurt rolli. Täiskasvanud terve inimene kannab endas väga palju erinevaid viirusosakesi, mis varieeruvad indiviiditi. Neid inimeste viiruslikke mustreid veel suures jaos ei mõisteta ning kuna viirustel ei esine ühte universaalselt geeni, ei saa teostada ka laia amplikonide põhise uuringut (Kuczynski et al., 2011).

Burnham jt. analüüsisid uriinis leiduvat cfDNA-d, et tuvastada UTI tekkepõhjused ning peaaegu pooltel patsientidest (66/141) tuvastas cfDNA sekveneerimine ka viiruseid, mis

võivad olla kliiniliselt olulised, kuid mida rutiinsed analüüsid ei tuvasta. 5-8% neerusiirdamise patsientidest kannatavad kolme aasta jooksul nefropaatia (igasugune mittepõletikuline neeruhaigus) all, mida põhjustab BK polüoomiviirus (BKV). Teised viirused, mis tihti neerusiirdamise patsientidele komplikatsioone tekitavad, on adenoviirused, JC polüoomiviirus, tsütomegaloviirus (CMV) ja parvoviirus, mida cfDNA sekveneerimine suutis edukalt tuvastada.

Kahel proovil, kus patsientidel traditsiooniliste meetoditega parvoviirus B19 infektsioon diagnoositi, tuvastas cfDNA sekveneerimine antud viiruse esimesel juhul 8 päeva enne kliinilist diagnoosi ja teisel juhul 80 päeva enne diagnoosi. Kolmandal patsiendil tuvastas cfDNA skriinimine viiruse 25 päeva peale ametlikku diagnoosi. Kusjuures tema uriiniproovis leidis rohkelt ka BKV-d, mis korreleerus talle teostatud viirus-spetsiifilise PCR testiga. Ühe patsiendi cfDNA näitas suures koguses inimese adenoviirus B DNA-d juba 15 päeva enne ning 9 päeva peale kliinilist diagnoosi.

Need andmed näitavad, et uriini cfDNA-d saab kasutada nii levinud kui ebaharilike viiruste tuvastamiseks (Burnham et al., 2018).

### 3. PROGRAMMID MIKROOBIDE TUVASTAMISEKS

Metagenoomsete proovide sekveneerimisel on seal leiduvad liigid, perekonnad ja isegi hõimkonnad teadmata ning eesmärgiks on nende tuvastamine saadud järjestuste abil nii täpselt kui võimalik. Laias laastus on kolm põhimõtteliselt erinevat viisi DNA järjestuste abil organismide tuvastamiseks:

- 1) järjestuste homoloogiaotsingul baseeruv tuvastamine
- 2) lugemite paigutamine (mäppimine) referentsgenoomile ja
- 3) joondusvabad meetodid

#### 3.1. Homoloogiaotsing BLAST

Järjestuste homoloogial baseeruv otsinguviiis identifitseerib sarnased valgud või geenid ning statistiliselt olulise sarnasuse järgi tuvastab nende ühise eellase. Tavaliselt kasutatakse selleks BLASTi. BLAST (*Basic Local Alignment Search Tool*) on järjestuse homoloogial baseeruv tööriist, mis on kõige laialdaselt kasutatud ja usaldusväärsem nii DNA kui valgu järjestuste homoloogiaotsinguks (Altschul, 1997). BLAST klassifitseerib tundmatu järjestuse, leides sellele parima joonduse genoomijärjestuse andmebaasist. Kuigi BLASTi ei disainitud spetsiaalselt metagenoomsete järjestuste jaoks, saab seda hõlpsalt selleks otstarbeks rakendada (Brady ja Salzberg, 2009). NGS metagenoomsete andmete analüüsiks on ta liiga aeglane (Niu et al. 2011). Suure andmestiku protsessimise aeg võib võtta mitmest päevast nädalateni.

#### 3.2. Lugemite paigutamine referentsgenoomile

Peale terve metagenoomi *shotgun* sekveneerimist, on kogu proovis leiduv DNA analüüsiks kasutatav miljonite lühikeste nukleotiidijärjestuste lugemitena (Petersen et al., 2017). Kasutatav programm peab olema suuteline antud FASTQ vormingus järjestusi mäppima referents andmebaasidesse. Teiseks võimaluseks on nende lühikeste järjestuste kokkupanek kontiigidesse, et neid seejärel referentsgenoomidega võrrelda (Menzel et al., 2016).

Üheks probleemiks metagenoomsete järjestuste võrdlemisel on referentsgenoomide kättesaadavus. Teatavad laialt levinud mudelorganismid või patogeenid on andmebaasides üle-esindatud, teised raskesti kultiveeritavad organismid ala-esindatud, mis teeb mõnede äärmuslikest keskkondadest võetud proovide klassifitseerimise keeruliseks (Menzel et al., 2016). Mida suurem on sihtmärk andmebaas, seda suurem on ka tõenäosus suvalise vaste saamiseks, mis tõstatab vale-positiivsete tulemuste probleemi (Petersen et al., 2017). Kolmandaks probleemiks on mikroobide (eriti viiruste) kiire muutumine evolutsiooni käigus. Seega jääb metagenoomsetel uuringutel tihti suur osa järjestusi klassifitseerimata või näitavad mõned järjestused vaid väikest sarnasust teadaolevate liikidega (Menzel et al., 2016).

Üks efektiivsemad määpimisel põhinevaid klassifitseerijaid on MGmapper (Petersen et al., 2017). See võimaldab paralleelselt kasutada mitmeid terve genoomi referentsjärjestuste andmebaase: bakterite, viiruste, seente, taimede, selgroogsete ja selgrootute organismide, samuti ka näiteks AMR geenide ja 16S rRNA või ükskõik mis kohandatud andmebaasi, mis põhineb FASTA vormingul.

Teine laialt kasutatav liikide määramise tarkvara, mis põhineb lugemite määpimisel, on MetaPhlAn (*Metagenomic Phylogenetic Analysis*). See on tööriist, millega on võimalik minutitega töödelda miljoneid metagenoomseid lugemeid ja neist täpsed mikroobikoosluste profiilid määrata. Saadud lugemitest hindab MetaPhlAn keskmist mikroobirakkude arvu ja paigutab need klaadi markerjärjestuste järgi vähemalt liigitasemele.

MetaPhlAn võrdleb proovist saadud metagenoomseid lugemeid markerjärjestuste kataloogiga, mis sisaldab ~4% sekveneeritud mikroobide geenidest. Igal lugemil on tänu markeri unikaalsusele vähemalt üks kõrge usaldusväärsusega vaste. Metagenoomsele DNA-le ei ole tänu markerjärjestusele tarvis teha ka eeltöötlust (vigade tuvastust, assambleerimist, geeni annotatsiooni), sest valedele lugemitele ei teki suure tõenäosusega vastet. MetaPhlAn2 sisaldab sekveneeritud alasid üle 7500-lt liigilt (iga bakteriliigi kohta 184±45 regiooni) (Segata et al., 2012). Tüvespetsiifilised järjestused võimaldavad tüvede määramist, samuti on võimalik viiruste ja eukarüootsete mikroobide klassifitseerimine (Tin Truong et al., 2015).

### 3.3. Joendusvabad meetodid

Joendusvaba taksonoomiline klassifitseerimine on eelnevatest kiirem ja lihtsam, sest pole vajadust joendamise ega metagenoomse assambleerimise järele. Joendusvaba klassifitseerimine põhineb taksoni-spetsiifiliste  $k$ -meeride tuvastamisel järjestustest.

$K$ -meer on fikseeritud pikkusega ( $k$ ) lühike nukleotiidne järjestus (Ounit et al., 2015). Kuna see meetod ei nõua geenide leidmist ega nende võrdlemist, on tal suur eelis lühikeste järjestuste (näiteks mcfDNA fragmentide) identifitseerimisel. Nii toimivad näiteks Kraken (Wood ja Salzberg, 2014), CLARK (Ounit et al., 2015), StrainSeeker (Roosaare et al., 2017) ja Phymm (Brady ja Salzberg, 2009). Tavaliselt hoiavad need programmid enda indekseeritud andmebaasis andmeid kõikide uuritavatele organismidele iseloomulike (spetsiifiliste)  $k$ -meeride kohta. Tänu antud andmebaasile saab  $k$ -meeride tuvastamisel iga uuritava järjestuse kiirelt sobiva taksonoomilise rühma alla lahterdada. Järjestused, mille  $k$ -meere andmebaasis pole, jäävad tuvastamata (Wood ja Salzberg, 2014).  $K$  suurus määrab otsingu tundlikkuse ja täpsuse. Liiga suure  $k$  valimisel ei pruugi vasteid leida ning liiga väikese  $k$  korral võib ilmneda liiga palju vale-positiivseid vasteid (Menzel et al., 2016). Tavaliselt kasutatakse  $k$ -meere kuni 31-aluspaarise pikkusega.

#### 3.3.1. Kraken

Kraken on kiire ja kõrge täpsusega programm metagenoomsete DNA järjestuste taksonoomilisteks rühmateks jagamisel  $k$ -meeride kaudu. Tema täpsus on võrreldav BLAST programmiga ning kõige kiirema režiimi peal suudab Kraken klassifitseerida üle 4,1 miljoni lugemi minutis, mis on 909 korda kiirem kui Megablast ja 11 korda kiirem kui MetPhlAn. Kraken suudab kiirelt identifitseerida taustamüra järjestused enne kui neid järjestuse kokkupanekusse kaasatakse (Wood ja Salzberg, 2014).

Kuna Kraken vajab järjestuste klassifitseerimiseks suurt muutmälu mahtu (72,4 GB), loodi Kraken2, mis vajab klassifitseerimiseks vaid 10,6 GB mälu, säilitades kõrge täpsuse ning tõstes kiirust viiendiku võrra. Kraken2 pakub ka suurenenud tundlikkust viiruste metagenoomseks analüüsiks tänu uuele transleeritud otsingu programmile (Kraken 2X), mis kasutab vähendatud aminohapete tähestikku valkude andmebaasiga võrdlemiseks (Wood et al., 2019).



Wool ja Salzberg võrdlesid 10 000 järjestuse klassifitseerimist erinevate meetoditega (Kraken, PhymmBL, NBC, Megablast, MetaPhlAn) ning mõõtsid perekonnatasemel tundlikkust (kui paljud järjestused määrati õige perekonna alla) ja täpsust (*positive predictive value*, näitab õigete klassifikatsioonide hulka kogu mahust). Selgus, et Megablasti tundlikkus on kõige kõrgem ning Kraken jäi sellele alla vaid 2,5%. Krakeni täpsus ja klassifitseerimise kiirus olid kõige kõrgemad, järgmisel kohal Megablast (150-240 korda aeglasem) (Wool ja Salzberg, 2014).

### 3.3.2. StrainSeeker

StrainSeeker on programm, mis suudab otse WGS-ga saadud lugemitest klassifitseerida bakterite isolaate klonaalsetesse rühmadesse või klaadidesse. Klassifitseerimine toimub erinevatel taksonoomilistel tasemetel jagatud *k*-meeride arvu järgi. Juhtpuu (fülogeneesipuu, mille saab kasutaja vajadusel ise ette anda) põhjal hindab StrainSeeker proovi fülogeneetilist sugulust referentsgenoomidega. Tulemused on näha visuaalselt kasutatud juhtpuul koos uuritavale isolaadile määratud kõige sarnasema taksoniga. StrainSeeker on võimeline assambleeritud NGS andmetest baktereid identifitseerima 1-2 minutiga.

Roosaare jt. võrdlesid erinevate programmide võimekust 100 *E. coli* tüve identifitseerimisel ning leidsid, et StrainSeekeri antud andmestikul oli taksoni määramise täpsus 99% ning Krakeni täpsus 69% (Roosaare et al., 2017).

## ARUTELU

cfDNA sekveneerimine on paljulubav meetod mikroorganismide tuvastamiseks. Järgmise põlvkonna sekveneerimisega on kerkinud võimalus mitteinvasiivselt mikroobset cfDNA-d sekveneerides tuvastada ühe analüüsiga infektsioone üle terve keha, tuvastades nii baktereid, seeni kui ka viiruseid. See annab infot nii cfDNA koostise, antibiootikumitundlikkuse ja bakterikasvu dünaamika kohta, kui ka võimaldab hinnata haiguse kulgu. cfDNA analüüsi saab kasutada veel loote varajaseks geneetiliseks analüüsimiseks, kasvajakude tuvastamiseks ning organi äratõukereaktsiooni ennustamiseks.

Paljude haiguste korral on korrektne diagnoosimine aeg-kriitiline, kus iga möödunud päev võib põhjustada patsiendi seisundi kiiret halvenemist. Traditsioonilised haiguste tuvastamise meetodid on tihti ajakulukad - bakterite tuvastamiseks kultiveerimisega läheb aega vähemalt kaks päeva, kusjuures paljud bakteritüved ja seened nõuavad veel pikemat kultiveerimise aega (Greisen et al., 1994). Leidub ka baktereid, keda kultiveerimine ei suuda tuvastada ning baktereid, keda on keeruline kultiveerida, näiteks *Haemophilus influenzae* tuleb külvata spetsiaalsele šokolaadi agarile, kuid seda standardmeetodid ette ei näe (Burnham et al., 2018). Tavaliselt määratakse antibiootikume enne mikrobioloogilise testi tulemusi empiirilisel, vastavalt kliinilistele ettekirjutustele ja epidemioloogiale. Empiiriline antibiootikumravi toob endaga kaasa antibiootikumresistentsete mikroobide kasvu ning võib tõsta oluliselt patsientide suremust raskesse infektsiooni, sest ei ole suunatud õigele haigustekitajale (Marquet et al., 2015). Seega on oluline leida kiirem viis patogeensete mikroorganismide tuvastamiseks, et rakendada haigusttekitavale patogeenile suunatud efektiivset ravi.

Mikroobe on võimalik tuvastada erinevate meetoditega, igal ühel enda plussid ja miinused. Kultiveerimine on üpris odav ning hästi uuritud mikroobide tuvastamise viise, küll aga on tema tundlikkus piiratud antibiootikumide ja seennakkuste vastaste ravimite tõttu ning viiruseid kultiveerimisega tuvastada ei saa. PCR testid on head, kui potentsiaalne haigustekitaja on juba eelnevalt diagnoositud. Multiplex PCR testid on küll kiired, kuid madala spetsiifilisusega ning esineb kõrge vale-positiivsete tulemuste oht. MALDI-TOF on kõrge spetsiifilisusega ja kiire, kuid nõuab eelnevat kultiveerimist. mNGS lähenemine kombineerib endas molekulaarbioloogilised tehnikad ja informaatika, et filtreerida inimese DNA järjestused ja identifitseerida patogeenide järjestused otse patsiendi vereplasmast.

cfDNA sekveneerimine on suuteline väikese kallutatusega tuvastama igat proovis esinevat mikroorganismi. Ta on suure läbilaskevõimega ja annab vastused kliiniliselt olulises ajaaknas. NGS ei nõua eelnevat kultiveerimist ning suudab tuvastada ka haruldasemad organismid. Küll aga võtab kõigi proovis leiduvate järjestuste sekveneerimine aega ning on kallim kui teised meetodid. NGS meetodil ei ole standardset bioinformaatilist lähenemist, mistõttu võib andmete analüüsimisel esineda erisusi olenevalt andmete töötlemisel tehtud valikutest, näiteks lugemite joondamisel ja andmete filtreerimise kriteeriumitest. Tuvastatav organism peab olema leitav ka genoomide andmebaasist. Täpseks diagnoosimiseks on oluline mikroobi tuvastamine vähemalt liigi tasemele.

cfDNA sekveneerimise tõhusus sõltub mitmest faktorist: mcfDNA kontsentratsioon, sekveneerimismeetod, bioinformaatiline töövoog ning proovi kvaliteet. Erinevad uuringud on hinnanud cfDNA sekveneerimise tundlikkuseks 70%-94% (Long et al., 2016; Blauwkamp et al., 2019). Kuna mcfDNA-d leidub plasmas niivõrd vähe, vähendab nii peremeesorganismi DNA foon kui ka välise mikroobse DNAGA saastumine testi tundlikkust. cfDNA sekveneerimine ei pruugi vahet teha bakteriliigi tüvedel, kui nende järjestuses esineb suur homoloogia (Hong et al., 2018). Vastus ei pruugi ilmnedagi ka siis, kui nakkus on lokaliseerunud väiksele alale kehas (Armstrong et al., 2019). See võib olla tingitud cfDNA sekveneerimise kallutatusest, sest mcfDNA on pärit erinevatest kehaosadest ning mikroobse DNA osakaal erinevatest allikatest võib varieeruda, mis omakorda võib põhjustada selle tuvastamata jäämise. cfDNA sekveneerimise kallutus võib olla tingitud veel mitmest faktorist. Näiteks raamatukogu ettevalmistamine mõjutab analüüsi täpsust. PCR amplifitseerimisel võivad ühed DNA fragmendid olla rohkem esindatud kui teised. See mõjutab järjestuste tuvastamist ja kvantifitseerimist. Kallutus võib sõltuda ka GC osakaalust DNA järjestuses. Väga kõrge või madala GC osakaaluga piirkondi võib olla raskem täpselt sekveneerida ja ka amplifitseerida raamatukogu valmistamise käigus.

Kliinilises praktikas on oluline analüüsi peale kuluv aeg ja selle täpsus. Olenevalt patsiendi seisundist võib vajaminevate invasiivsete proovide võtmine olla võimatu. Sellistel juhtudel on diagnoosimiseks vaja leida mitteinvasiivne lähenemine. Nende kriteeriumite järgi on cfDNA sekveneerimisel palju potentsiaali töötada kui standardne diagnoosimismeetod, kuid selle eelduseks on analüüsi kättesaadavus ning professionaalsed töötajad bioinformaatilise analüüsi teostamiseks. Pädeva tehnika soetamine laboritesse

tähendab suurt väljaminekut. Teiseks variandiks on proovide saatmine laborisse, kus vajaminevad instrumendid juba olemas on. See toob kaasa lisakulu transpordile ning potentsiaalselt pikemat aega tulemuste kättesaamiseni.

Kliinilises diagnostikas mängib suurt rolli ka sekveneerimise hind. Kuigi laialdasem NGS sekveneerimise kasutuselevõtmine aitab langetada hinda, on see ikka võrdlemisi kallis. Kallim hinnaklass võib potentsiaalselt piirata kriitiliselt haigete patsientide diagnoosimist, kelle üldine ravikulu päevas on juba võrdlemisi kõrge. Hinna tasuvusel tuleb kaaluda konkreetseid vajadusi - kui täpset, tundlikku ja sügavamahulist informatsiooni vaja on. Juba laialdaselt kasutatava mitteinvasiivse loote geneetilise uuringu (NIPT) pealt on näha, et hindade alanemine aitab seda laialdasemalt kasutusele võtta ning teeb teenuse kättesaadavamaks.

Antud uurimistöö näitas, et cfDNA analüüsimine võimaldab edukalt tuvastada patogeenseid mikroorganisme ka siis, kui standardsed mikrobioloogilised uurimismeetodid seda ei suuda. Tulemused on võimalik saada mitteinvasiivselt kliiniliselt olulises ajaaknas kõrge täpsusega. Kuigi cfDNA mNGS on kallim kui alternatiivsed meetodid, pakub ta väärtuslikku infot proovi mikroobse koosseisu ja geneetiliste omaduste kohta. Küll aga on esitletud uurimised tehtud võrdlemisi väikeste valimitega ning oleks vaja koostada suuri uuringuid, et hinnata cfDNA sekveneerimise kliinilist kättesaadavust.

## KOKKUVÕTE

Käesolevas töös tutvustatakse mikroobide tuvastamise erinevaid meetodeid ning antakse ülevaate seni avaldatud teadustöödest, mis puudutavad cfDNAst mikroobide määramist ja selleks kasutatavatest arvutiprogrammidest. cfDNA sekveneerimine on suure tundlikkuse ja täpsusega viis mitteinvasiivselt patogeensete mikroobide ja viiruste tuvastamiseks. cfDNA analüüsimine võimaldab edukalt tuvastada patogeenseid mikroorganisme ka siis, kui standardsed mikrobioloogilised uurimismeetodid seda ei võimalda. Kasutades cfDNA järgmise põlvkonna sekveneerimist on võimalik ühe analüüsiga saada informatsiooni proovis leiduva mikroobse koosluse, antibiootikumitundlikkuse, bakterikasvu dünaamika ning haiguse kulgemise kohta. Tulemused on võimalik saada kliiniliselt olulises ajaaknas kõrge täpsusega. Kuigi cfDNA sekveneerimine on kallim kui alternatiivsed meetodid, pakub ta laia spektriga informatsiooni, mis on teatavatel haigusjuhtudel elulise tähtsusega. Patogeensed mikroorganismid on erinevaid bioinformaatilisi programme kasutades võimalik tuvastada paljudel juhtudel vähemalt liigi tasemel, tänu millele saab rakendada efektiivset ravikava. See näitab cfDNA sekveneerimise võimekust täiendada traditsioonilisi diagnoosimismeetodeid.

## SUMMARY

### **Detection and Identification of Pathogenic Microbes from Cell-Free DNA**

Pilleriin Paidla

Cell-free circulating DNA (cfDNA) consists of human and microbial fragments of nucleic acid that are found from blood and other body fluids. cfDNA integrates information about infections from all over the body. The purpose of this study was to present a review of published scientific articles concerning detection of pathogenic microbes from cfDNA thus far and introduce different methods for identifying microorganisms.

There are many ways to identify infectious agents, all have their pros and cons. Culture-based methods are often considered the standard for bacterial identification, but they are time-consuming, labor-intensive and might not detect fastidious organisms. PCR is a targeted amplification technique that can detect specific DNA sequences or genetic markers but needs a previous diagnosis. MALDI-TOF is a rapid and cost-effective technique used for bacterial identification but requires culture-positive isolates. cfDNA metagenomic next-generation sequencing is a rapid high-throughput DNA sequencing technology that allows for the identification of known and novel microorganisms, determination of their genetic composition, and assessment of potential drug resistance or virulence factors. mNGS has shown great promise in the field of infectious disease diagnosis and surveillance, enabling rapid and accurate detection of a wide range of pathogens, especially in cases where the specific pathogen or causative agent is unknown or difficult to culture using traditional methods.

Although it has higher cost than alternative methods, cfDNA mNGS has high accuracy, sensitivity, and depth of information provided. Sequencing of cfDNA gives an opportunity to non-invasively give early diagnosis of infectious diseases and thus the determination of an effective treatment plan, potentially reducing mortality. It makes cfDNA a great tool for clinical practices with a potential to become a standard diagnostic method.

## **TÄNUSÕNAD**

Ma soovin tänada kõiki, kes aitasid kaasa minu töö edukasse valmimisse. Margus Leppik, kes aitas leida uurimisrühma ja Mairo Remm, kes oli lahkelt valmis võtma mind enda käe alla ning suure vilumusega juhendas ja suunas töö kirjutamisel.

## KASUTATUD KIRJANDUS

- Abril, M. K., Barnett, A. S., Wegermann, K., ... Kraft, B. D. (2016). Diagnosis of *Capnocytophaga canimorsus* Sepsis by Whole-Genome Next-Generation Sequencing. *Open Forum Infect Dis.* 3(3):ofw144.
- Afshinnekoo, E., Chou, C., Alexander, N., Ahsanuddin, S., Schuetz, A. N., Mason, C. E. (2017). Precision Metagenomics: Rapid Metagenomic Analyses for Infectious Disease Diagnostics and Public Health Surveillance. *J Biomol Tech.* 28(1):40-45.
- Altschul, S. F., Gish, W., Miller, W., Myers, W. E., Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology.* 215(3):403-410.
- Armstrong, A. E., Rossoff, J., Hollemon, D., Hong, D. K., Muller, W. J. and Chaudhury, S. (2019). Cell-free DNA next-generation sequencing successfully detects infectious pathogens in pediatric oncology and hematopoietic stem cell transplant patients at risk for invasive fungal disease. *Pediatric Blood & Cancer.* 66:e27734.
- Beutz, M., Sherman, G., Mayfield, J., Fraser, V. J., Kollef, M. H. (2003). Clinical utility of blood cultures drawn from central vein catheters and peripheral venipuncture in critically ill medical patients. *Chest.* 123(3):854-61.
- Blauwkamp, T. A., Thair, S., Rosen, M.J., ... Yang, S. (2019). Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nature Microbiology.* 4:663–674.
- Brownie, J., Shawcross, S., Theaker, J., Whitcombe, D., Ferrie, R., Newton, C., Little S. (1997). The elimination of primer-dimer accumulation in PCR. *Nucleic Acids Res.* 25(16):3235-41.
- Burnham, P., Dadhania, D., Heyang, M., Chen, F., Westblade, L. F., Suthanthiran, M., Lee, J. R., De Vlaminck, I. (2018). Urinary cell-free DNA is a versatile analyte for monitoring infections of the urinary tract. *Nat Commun.* 9(1):2412.
- Chapin, K. C. 2007. Principles of stains and media, p 182–191. In P. R. Murray, E. J. Baron, J. H. Jorgensen, M. L. Landry, M. A. Pfaller (ed.), *Manual of clinical microbiology*, 9th ed ASM Press, Washington, DC.
- Chapin, K. C., Lauderdale, T.L. 2007. Reagents, stains, and media: bacteriology, p 334–364. In P. R. Murray, E. J. Baron, J. H. Jorgensen, M. L. Landry, M. A. Pfaller (ed.), *Manual of clinical microbiology*, 9th ed., ASM Press, Washington, DC.
- Clarridge, J. E. 3rd. (2004). Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin Microbiol Rev.* 17(4):840-62
- Click, E. S., Murithi, W., Ouma, G. S., McCarthy, K., Willby, M., Musau, S., Alexander, H., Pevzner, E., Posey, J., Cain, K. P. (2018). Detection of Apparent Cell-free *M. tuberculosis* DNA from Plasma. *Sci Rep.* 8(1):645.



- Croxatto, A., Prod'hom, G., Greub, G. (2012). Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiology Reviews*. 36:380–407
- De Pauw, B., Walsh, T. J., Donnelly, P., ... Bennet, J. E. (2008). Revised Definitions of Invasive Fungal Disease from the European Organization for Research and Treatment of Cancer/Invasive Fungal Infections Cooperative Group and the National Institute of Allergy and Infectious Diseases Mycoses Study Group (EORTC/MSG) Consensus Group. *Clinical Infectious Diseases*. 46(12):1813–1821.
- Diao, Z., Han, D., Zhang, R., Li, J. (2021). Metagenomics next-generation sequencing tests take the stage in the diagnosis of lower respiratory tract infections. *J Adv Res*. 38:201-212.
- Elnifro, E. M., Ashshi, A. M., Cooper, R. J., Klapper, P. E. (2000). Multiplex PCR: optimization and application in diagnostic virology. *Clin Microbiol Rev*. 13(4):559-70.
- Farnaes, L., Wilke, J., Ryan Loker, K., Bradley, J. S., Cannavino, C. R., Hong, D. K., Pong, A., Foley, J., Coufal, N.G. (2019). Community-acquired pneumonia in children: cell-free plasma sequencing for diagnosis and management. *Diagn Micr Infec Dis*. 94:188–91.
- Fenollar, F., Raoult, D. (2004) Molecular genetic methods for the diagnosis of fastidious microorganisms. *APMIS*. 112(11-12):785-807.
- Fenollar, F., Raoult, D. (2007) Molecular diagnosis of bloodstream infections caused by non-cultivable bacteria. *International Journal of Antimicrobial Agents*. 30:7-15.
- Fernández-Carballo, L. B., Broger, T., Wyss, R., Banaei, N., Denking C. M. (2019). Toward the Development of a Circulating Free DNA-Based In Vitro Diagnostic Test for Infectious Diseases: a Review of Evidence for Tuberculosis. *Journal of Clinical Microbiology*. 57:e01234-18
- Fung, M., Zompi, S., Seng, H., ... Chin-Hong, P. (2018). Plasma Cell-Free DNA Next-Generation Sequencing to Diagnose and Monitor Infections in Allogeneic Hematopoietic Stem Cell Transplant Patients. *Open Forum Infect Dis*. 5(12):ofy301.
- Giuliano, C., Patel, C. R., Kale-Pradhan, P. B. (2019) A Guide to Bacterial Culture Identification And Results Interpretation. *Pharmacy and Therapeutics*. 44(4):192-200.
- Greisen, K., Loeffelholz, M., Purohit, A., Leong, D. (1994). PCR primers and probes for the 16S rRNA gene of most species of pathogenic bacteria, including bacteria found in cerebrospinal fluid. *Journal of Clinical Microbiology*. 32:335–351.
- Grumaz, C., Hoffmann, A., Vainshtein, Y., ... Sohn, K. (2020). Rapid Next-Generation Sequencing Based Diagnostics of Bacteremia in Septic Patients. *The Journal of Molecular Diagnostics*. 22:405-418.
- Grumaz, S., Grumaz, C., Vainshtein, Y., Stevens, P., Glanz, K., Decker S. O., Hofer, S., Weigand, M. A., Brenner, Sohn, K. (2019). Enhanced Performance of Next-Generation

Sequencing Diagnostics Compared With Standard of Care Microbiological Diagnostics in Patients Suffering From Septic Shock. *Critical Care Medicine*. 47:e394–e402.

Grumaz, S., Stevens, P., Grumaz, C., Decker S. O., Weigand, M. A., Hofer, S., Brenner, T., Haeseler, A., Sohn, K. (2016). Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Medicine*. 8:73.

Gu, W., Deng, X., Lee, M., ... Chiu, C. Y. (2021). Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nature Medicine*. 27:115–124.

Gu, W., Miller, S., Chiu, C. Y. (2019). Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection. *Annu Rev Pathol*. 14:319-338.

Han, D., Li, R., Shi, J., Tan, P., Zhang, R., Li, J. (2020). Liquid biopsy for infectious diseases: a focus on microbial cell-free DNA sequencing. *Theranostics*. 10(12):5501-5513.

Higuchi, R., Dollinger, G., Walsh, P. S., Griffith, R. (1992). Simultaneous amplification and detection of specific DNA-sequences. *Bio-Technology*, 10(4):413-417.

Hilt, E. E., Ferrieri. P. (2022). Next Generation and Other Sequencing Technologies in Diagnostic Microbiology and Infectious Diseases. *Genes (Basel)*. 13(9):1566.

Hong, D. K., Blauwkamp, T. A., Kertesz, M., Bercovici, S., Truong, C., Banaei, N. (2018). Liquid biopsy for infectious diseases: sequencing of cell-free plasma to detect pathogen DNA in patients with invasive fungal disease. *Diagnostic Microbiology and Infectious Disease*. 92:210-213.

Huang, A. M., Newton, D., Kunapuli, A., Gandhi, T. N., Washer, L. L., Isip, J., Collins, C. D., Nagel, J. L. (2013). Impact of rapid organism identification via matrix-assisted laser desorption/ionization time-of-flight combined with antimicrobial stewardship team intervention in adult patients with bacteremia and candidemia. *Clin Infect Dis*. 57(9):1237-45.

Huang, J., Jiang, E., Yang, D., Wei, J., Zhao, M., Feng, J., Cao, J. (2020). Metagenomic Next-Generation Sequencing versus Traditional Pathogen Detection in the Diagnosis of Peripheral Pulmonary Infectious Lesions. *Infect Drug Resist*. 13:567-576.

Huang, Y. F., Chen, Y. J., Fan, T. C., ... Chiu, K. P. (2018). Analysis of microbial sequences in plasma cell-free DNA for early-onset breast cancer patients and healthy females. *BMC Med Genomics*. 11:16.

Janes, V. A., Matamoros, S., Munk, P., ... Schultsz, C. (2022). Metagenomic DNA sequencing for semi-quantitative pathogen detection from urine: a prospective, laboratory-based, proof-of-concept study. *Lancet Microbe*. 3:e588–97.

Kohler, P., Kuster, S. P., Bloemberg, G., ... Hasse, B. (2015). Healthcare-associated prosthetic heart valve, aortic vascular graft, and disseminated *Mycobacterium chimaera* infections subsequent to open heart surgery. *Eur Heart J*. 36:2745–2753.

- Kothari, A., Morgan, M., Haake, D. A. (2014). Emerging Technologies for Rapid Identification of Bloodstream Pathogens. *Clinical Infectious Diseases*. 59: 272–278.
- Kuczynski, J., Lauber, C. L., Walters, W. A., Parfrey, L. W., Clemente, J. C., Gevers, D., Knight, R. (2011). Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet*. 13(1):47-58.
- Kubista, M., Andrade, J. M., Bengtsson, M., ... Zoric, N. (2006).
- Kumar, A., Ellis, P., Arabi, Y., ... Chateau, D. (2009) Initiation of Inappropriate Antimicrobial Therapy Results in a Fivefold Reduction of Survival in Human Septic Shock. *Original Research Critical Care Medicine*. 136(5):1237-1248.
- Kumar, S. 2012. *Textbook of microbiology*. JP Medical Ltd, New Delhi.
- Köser, C. U., Ellington, M. J., Peacock, S. J. (2014). Whole-genome sequencing to control antimicrobial resistance. *Trends Genet*. 30(9):401-7.
- Lagier, J. C., Edouard, S., Pagnier, I., Mediannikov, O., Drancourt, M., Raoult, D. (2015) Current and Past Strategies for Bacterial Culture in Clinical Microbiology. *Clinical Microbiology Reviews*. 28(1): 208–236.
- Lichtenstein, A. V., Melkonyan, H. S., Tomei, L. D., Umansky, S. R. (2001). Circulating nucleic acids and apoptosis. *Ann N Y Acad Sci* 945:239–249.
- Llor, C., Bjerrum, L., (2014). Antimicrobial resistance: risk associated with antibiotic overuse and initiatives to reduce the problem. *Therapeutic Advances in Drug Safety*. 5(6):229-41
- Long, Y., Zhang, Y., Gong, Y., ... Tan, H. (2016). Diagnosis of sepsis with cell-free DNA by next-generation sequencing technology in ICU patients. *Arch Med Res*. 47:365–71.
- Mandel, P., Metais, P. 1948. Les acides nucléiques du plasma sanguin chez l'homme. *C R Seances Soc Biol Fil*. 142:241–243.
- Marshall, B. J., Warren, J. R. (1984). Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet*. 1(8390):1311-5.
- Marquet, K., Liesenborgs, A., Bergs, J., Vleugels, A., Claes, N. (2015) Incidence and outcome of inappropriate in-hospital empiric antibiotics for severe infection: a systematic review and meta-analysis. *Critical Care*. 19(1): 63
- Marra, A. R., Diekema, D. J., Edmond, M. B. (2017). *Mycobacterium chimaera* Infections Associated With Contaminated Heater-Cooler Devices for Cardiac Surgery: Outbreak Management. *Clin Infect Dis*. 65(4):669-674.
- Mayr, F. B., Yende, S., Angus, D. C. (2014). Epidemiology of severe sepsis. *Virulence*. 5(1):4-11.

- Mege, J. L., Maurin, M., Capo, C., Raoult, D. (1997). *Coxiella burnetii*: the ‘query’ fever bacterium. A model of immune subversion by a strictly intracellular microorganism. *FEMS Microbiol Rev.* 19:209–17.
- Menzel, P., Ng, K. L., Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications.* 7:11257.
- Miao, Q., Ma, Y., Wang, Q., ... Hu, B. (2018). Microbiological Diagnostic Performance of Metagenomic Next-generation Sequencing When Applied to Clinical Practice. *Clin Infect Dis.* 67:S231-S240.
- Mignard, S. ja Flandrois, J. P. (2006). 16S rRNA sequencing in routine bacterial identification: A 30-month experiment. *Journal of Microbiological Methods.* 67(3):574-581.
- Gao, Q., Zeng, Q., Wang, Z., ... Fan, J. (2022). Circulating cell-free DNA for cancer early detection. *Innovation (Camb).* 3(4):100259.
- García Moreira, V., Prieto García, B., Baltar Martín, J. M., Ortega Suárez, F., Alvarez, F. V. (2009). Cell-free DNA as a noninvasive acute rejection marker in renal transplantation. *Clin Chem.* 55(11):1958-66.
- Nieman, A. E., Savelkoul, P. H. M., Beishuizen, A., Henrich, B., Lamik, B., MacKenzie, C. R., Kindgen-Milles, D., Helmers, A., Diaz, C., Sakka, S. G., Schade, R. P. (2016). A prospective multicenter evaluation of direct molecular detection of blood stream infection from a clinical perspective. *BMC Infect Dis.* 16:314.
- Niu, B., Zhu, Z., Fu, L., Wu, S., Li, W. (2011). FR-HIT, a very fast program to recruit metagenomic reads to homologous reference genomes. *Bioinformatics* 27:1704–1705.
- Nomura, J., Rieg, G., Bluestone, G., Tsai, T., Lai, A., Terashita, D., Bercovici, S., Hong, D. K., Lee, B. P. (2019). Rapid detection of invasive *Mycobacterium chimaera* disease via a novel plasma-based next-generation sequencing test. *BMC Infect Dis.* 19(1):371.
- Ounit, R., Wanamaker, S., Close, T.J., and Lonardi, S. (2015). CLARK: fast and accurate classification of metagenomic and genomic sequences using discriminative k-mers. *BMC Genomics* 16:236.
- O’Grady, J. (2019) A powerful, non-invasive test to rule out infection. *Nature Microbiology.* 4:554–555.
- Pan, W., Ngo, T. T. M., Camunas-Soler, J., Song, C. X., Kowarsky, M., Blumenfeld, Y. J., Wong, R. J., Shaw, G. M., Stevenson, D. K., Quake, S. R. (2017). Simultaneously monitoring immune response and microbial infections during pregnancy through plasma cfRNA sequencing. *Clin Chem.* 63:1695–704.
- Paolucci, M., Landini, M. P., Sambri, V. (2010). Conventional and molecular techniques for the early diagnosis of bacteraemia. *Int J Antimicrob Agents.* 2:S6-16.

- Perez, K. K., Olsen, R. J., Musick, W. L., ... Musser, J. M. (2013). Integrating rapid pathogen identification and antimicrobial stewardship significantly decreases hospital costs. *Arch Pathol Lab Med.* 137:1247–54.
- Pérez-Cobas, A. E., Gomez-Valero, L., Buchrieser, C. (2020). Metagenomic approaches in microbial ecology: an update on whole-genome and marker gene sequencing analyses. *Microb Genom.* 6(8):mgen000409.
- Petersen, L. M., Martin, I. W., Moschetti, W. E., Kershaw, C. M., Tsongalis, G. J. (2019). Third-Generation Sequencing in the Clinical Laboratory: Exploring the Advantages and Challenges of Nanopore Sequencing. *J Clin Microbiol.* 58(1):e01315-19.
- Petersen, T. N., Lukjancenko, O., Thomsen, M. C. F., Sperotto, M. M., Lund, O., Møller Aarestrup, F., Sicheritz-Pontén, T. (2017). MGmapper: Reference based mapping and taxonomy annotation of metagenomics sequence reads. *PLoS One.* 12(6):e0179778.
- Purcarea, A., Sovaila, S. (2020). Sepsis, a 2020 review for the internist. *Rom J Intern Med.* 58(3):129-137.
- Qin, Z., Ljubimov, V. A., Zhou, C., Tong, Y., Liang, J. (2016). Cell-free circulating tumor DNA in cancer. *Cancer Communications.* 35:1-9.
- Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology.* 35:833–844
- Ramanan, P., Bryson, A. L., Binnicker, M. J., Pritt, B. S., Patel, R. (2018). Syndromic Panel-Based Testing in Clinical Microbiology. *Clinical Microbiology Reviews.* 31: e00024-17.
- Ranjan, R., Rani, A., Metwally, A., McGee, H. S., Perkins, D. L. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochem Biophys Res Commun.* 469(4):967-77.
- Rasko, D. A., Webster, D. R., Sahl, J. W., ... Waldor, M. K. (2011). Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. *N Engl J Med.* 365(8):709-17.
- Riccardi, N., Monticelli, J., Antonello, R. M., Luzzati, R., Gabrielli, M., Ferrarese, M., Codecasa, L., Di Bella, S., Giacobbe, D. R. (2020). *Mycobacterium chimaera* infections: An update. *J Infect Chemother.* 26(3):199-205.
- Roosaare, M., Vaheer, M., Kaplinski, L., Möls, M., Andreson, R., Lepamets, M., Kõressaar, T., Naaber, P., Kõljalg, S., Remm, M. (2017). StrainSeeker: fast identification of bacterial strains from raw sequencing reads using user-provided guide trees. *PeerJ.* 5:e3353.
- Sameer, K. S., Lai, Y. L., Warner, S., ... Adjemian, J. (2021) Inappropriate empirical antibiotic therapy for bloodstream infections based on discordant in-vitro susceptibilities: a retrospective cohort analysis of prevalence, predictors, and mortality risk in US hospitals. *The Lancet Infectious Diseases.* 21:241-251.

Sanschagrín, S., Yergeau, E. (2014). Next-generation sequencing of 16S ribosomal RNA gene amplicons. *J Vis Exp.* 29;(90):51709.

Sayres, L. C., Cho, M. K. (2011). Cell-free fetal nucleic acid testing: a review of the technology and its applications. *Obstet. Gynecol. Surv.* 66:431-442.

Scerbo, M. H., Kaplan, H. B., Dua, A., Litwin, D. B., Ambrose, C. G., Moore, L. J., Murray, C. C., Wade, C. E., Holcomb, J. B. (2016). Beyond Blood Culture and Gram Stain Analysis: A Review of Molecular Techniques for the Early Detection of Bacteremia in Surgical Patients. *Surg Infect (Larchmt)*. (3):294-302.

Schmidt, K., Mwaigwisya, S., Crossman, L. C., ... Livermore, D. M. (2016) Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *Journal of Antimicrobial Chemotherapy.* 72: 104–114.

Schreckenberger, P. C., McAdam, A. J. (2015) Point-Counterpoint: Large Multiplex PCR Panels Should Be First-Line Tests for Detection of Respiratory and Intestinal Pathogens. *Journal of Clinical Microbiology.* 53(10): 3110–3115.

Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O., Huttenhower, C. (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. *Nature Methods.* 9:811–814.

Seng, P., Drancourt, M., Gouriet, F., La Scola, B., Fournier, P. E., Rolain, J. M., Raoult, D. (2009). Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Infect Dis.* 49(4):543-51.

Shendure, J., Ji, H. (2008). Next-generation DNA sequencing. *Nat Biotechnol.* 26(10):1135-45.

Singhal, N., Kumar, M., Kanaujia, P. K., Viridi, J. S. (2015). MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology.* 6.

Suzuki, N., Kamataki, A., Yamaki, J., Homma, Y. (2008). Characterization of circulating DNA in healthy human plasma. *Clinica Chimica Acta.* 387:55-58.

Zheng, Y., Qiu, X., Wang, T., Zhang, J. (2021). The Diagnostic Value of Metagenomic Next-Generation Sequencing in Lower Respiratory Tract Infection. *Front Cell Infect Microbiol.* 11:694756.

Zinter, M. S., Dvorak, C. C., Mayday, M. Y., ... DeRisi, J. L. (2019). Pulmonary Metagenomic Sequencing Suggests Missed Infections in Immunocompromised Children. *Clin Infect Dis.* 68(11):1847-1855.

Tin Truong, D., Franzosa, E. A., Tickle, T. L., Scholz, M., Weingart, G., Pasolli, E., Tett, A., Huttenhower, C., Segata, N. (2015). MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nature Methods.* 12:902–903

Tsalik, E. L., Jones, D., Nicholson, B., ... Woods, C. W. (2010). Multiplex PCR To Diagnose Bloodstream Infections in Patients Admitted from the Emergency Department with Sepsis. *Journal of Clinical Microbiology*. 48: 26–33.

Ushio, R., Yamamoto, M., Nakashima, K., ... Kaneko, T. (2016). Digital PCR assay detection of circulating *Mycobacterium tuberculosis* DNA in pulmonary tuberculosis patient plasma. *Tuberculosis (Edinb)* 99:47–53.

Wagner, J. (2012). Free DNA--new potential analyte in clinical laboratory diagnostics? *Biochem Med (Zagreb)*. 22(1):24-38.

Wang, J., Han, Y., Feng, J. (2019). Metagenomic next-generation sequencing for mixed pulmonary infection diagnosis. *BMC Pulm Med*. 19(1):252.

Weerakoon, K. G., McManus, D. P. (2016). Cell-free DNA as a diagnostic tool for human parasitic infections. *Trends Parasitol* 32:378–391.

Weis, C., Cuénod, A., Rieck, B., ... Egli, A. (2022). Direct antimicrobial resistance prediction from clinical MALDI-TOF mass spectra using machine learning. *Nature Medicine*. 28(1):164-174.

Wood, D. E., Lu, J., Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*. 20:257.

Wu, D., Wang, W., Xun, Q., Wang, H., Liu, J., Zhong, Z., Ouyang, C., Yang, Q. (2022). Metagenomic next-generation sequencing indicates more precise pathogens in patients with pulmonary infection: A retrospective study. *Front Cell Infect Microbiol*. 12:977591.

Ye, M, Wei, W, Yang, Z, ... Jiang, H. (2016). Rapid diagnosis of *Propionibacterium acnes* infection in patient with hyperpyrexia after hematopoietic stem cell transplantation by next-generation sequencing: a case report. *BMC Infect Dis*. 16:5.

## **KASUTATUD VEEBIAADDRESSID**

2023, <https://www.sm.ee/amr>

## LIHTLITSENTS

### **Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks**

Mina, Pilleriin Paidla (05.07.1995)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose

**Patogeensete mikroorganismide tuvastamine rakuvabast DNA-st,**

mille juhendaja on Maido Remm,

1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-i lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 29.05.2023