Culture-Independent Metagenomics Characterisation of Infection

Solomon Mwaigwisya

Doctor of Philosophy, February 2018

Culture-Independent Metagenomics Characterisation of Infection

Solomon Mwaigwisya

Norwich Medical School

Faculty of Medicine and Health Sciences

University of East Anglia, Norwich, UK

Submitted in partial fulfilment of the requirements of the Degree of Doctor of Philosophy, February 2018.

Statement of Innovation

I, Solomon Bupe Mwaigwisya, confirm that the work presented in this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Where information has been derived from other sources, I confirm that this had been indicated in the thesis.

This thesis is the result of three and half years of research conducted at the Medical School Faculty of Health Sciences at the University of East Anglia in Norwich, UK. Throughout the period of the research, some of the contents concerning Nanopore sequencing have been published in the following articles:

Schmidt K., Mwaigwisya S., Crossman L., Doumith M., Munroe D., Pires C., Khan A., Woodford N., Saunders N., Wain J., O'Grady J., Livermore D. *Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore- based metagenomic sequencing*. J Antimicrob Chemother (2017) 72 (1): 104-114.

Ashton, P. M., Nair, S., Dallman, T., Rubino, S., Rabsch, W., Mwaigwisya, S., Wain, J. & O'grady, J. 2015. *MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island*. Nature biotechnology, 33**,** 296-300.

Ip, C. L., Loose, M., Tyson, J. R., De Cesare, M., Brown, B. L., Jain, M., Leggett, R. M., Eccles, D. A., Zalunin, V., Urban, J. M., Piazza, P., Bowden, R. J., Paten, B., Mwaigwisya, S., Batty, E. M., Simpson, J. T., Snutch, T. P., Birney, E., Buck, D., Goodwin, S., Jansen, H. J., O'grady, J. & Olsen, H. E. 2015b. *MinION Analysis and Reference Consortium: Phase 1 data release and analysis*. F1000Res, 4**,** 1075.

Nair, S., Ashton, P., Doumith, M., Connell, S., Painset, A., Mwaigwisya, S., Langridge, G., De Pinna, E., Godbole, G. & Day, M. 2016. *WGS for surveillance of antimicrobial resistance: a pilot study to detect the prevalence and mechanism of resistance to azithromycin in a UK population of non-typhoidal Salmonella*. J Antimicrob Chemother, 71**,** 3400-3408.

Therefore, as one of the main authors, I have used some parts of the published article in my thesis.

Funding

This project was funded by University of East Anglia - International Student Scholarship.

Additional funding was provided by Rosetrees Trust as PhD student medical research award.

Abstract

Next-generation sequencing (NGS) technologies are revolutionising our ability to study and characterise microorganisms and investigate infectious diseases. The potential of metagenomics sequencing for use as a single, all-inclusive diagnostic test for comprehensive detection of pathogens, resistance genes and virulence markers directly from clinical samples has been discussed at length in the literature in recent years. However, implementation has been slow as there are several challenges associated with applying metagenomics sequencing to clinical microbiology. These include the large number of human cells, the often low proportion of pathogen cells/DNA and, in some cases, the high background of normal microbiological flora present in clinical samples. Here we report rapid, culture-independent metagenomics workflows that overcome these challenges.

Metagenomics pipelines were developed and evaluated in three model samples: i) blood, for the diagnosis of sepsis, ii) urine, for the diagnosis of urinary tract infections, and iii) stool, for the diagnosis of *Clostridioides difficile* infection. Developed workflows comprised of rapid depletion of unwanted cells/DNA (human and normal flora (in stool)), genomic DNA extraction from remaining microorganisms, whole genome amplification (in blood), rapid nanopore library preparation and real-time metagenomics analysis.

These pipelines enabled comprehensive detection of pathogens and resistance genes in clinical blood samples within eight hours and in clinical urine samples within four hours. The *C. difficile* pipeline could enrich for and sequence the pathogen directly from stool within 24 hours. However, further optimisation of this pipeline is required to increase genome coverage before it can be utilised for typing *C. difficile* directly from stool.

The rapid host depletion and metagenomics sequencing pipelines developed here demonstrate that this technology can provide clinicians with the necessary information to tailor antibiotic therapy for the specific infecting pathogen before second dose of empiric therapy is administered (usually eight-hour intervals), thereby improving patient outcomes and antibiotic stewardship.

Acknowledgements

First, I would like thank my supervisors, Dr. Justin O'Grady, Professor John Wain and Professor David Livermore. I would also like to thank the University of East Anglia and Rosetress Trust for the funding.

Dr. Justin, I would like to especially thank you for being there with me from the very beginning of this journey right to the end. It would have been impossible for me to achieve this accomplishment without your advice, insights, and your constant communication. Thank you for sharing the joys and troubles of science with me.

I would also like to thank Professor John Wain for your insights, ideas and optimism and Professor David Livermore for challenging and encouraging my research, it has all helped me to grow as a scientist.

Dr. Geoff Warhurst from The University of Manchester, Division of Infection, Immunity and Respiratory Medicine and Dr. Tim Leary from Norfolk and Norwich University NHS Foundation Trust, thank you for the clinical samples. Dr. Samir Dervisevic, Dr. Elumogo Ngozi, Dr. Jennifer Moore and Dr. Emma Meader from Clinical Microbiology laboratory of Norfolk and Norwich University hospital, my deepest gratitude for taking your time provide me clinical samples whenever I needed them. My work would have been extremely difficult without your help.
To Dr. Gemma Kay, thank you very much for the data analysis, and taking your time to read my thesis, I am very grateful.

I could not have achieved this goal without the support of my friends and family. To my father, thank you very much for supporting my academic journey, you are pillar of the family, the best father. To my mother, thank your understanding and allowing me to express myself knowing you will always be there for me. Nothing will inspire my life like your pride, confidence, determination and hard work. To my siblings, Ntuli, Paul, Tuntufye, Nsume, Ntwa and Subilaga, thank you for being there for me in good and bad times, and for supporting my decisions.

My friends Ayoub Mushi, Dr. Annapaula Correia and Dr. Katarzyna Schmidt you are one of the best and nicest people I ever met and worked with, I am glad to call you friends.

Dedication

I dedicate this work to my father who has sacrificed a lot for the education of my siblings and mine.

Abbreviations

ARMA	Antimicrobial Resistance Mapping Application
BLAST	Basic Local Alignment Search tool
BSI	Blood Stream Infection
CARD	Comprehensive Antibiotic Resistance Database
CDI	*Clostridioides difficile* Infection
CU	Clinical Urine
Cq	Quantitative Cycle
DH	Department of Health
EI	Earlham Institute
EUCAST	European Committee on Antimicrobial Susceptibility Testing
IFR	Institute of Food Research
JIC	John Innes Centre
MALDI-TOF MS	Matrix-Assisted Laser-Desorption/Ionization Time of Flight Mass Spectrometry
MAP	MinION Earl-Access Program
MRSA	Methicillin-Resistant Staphylococcus aureus
NGS	Next Generation Sequecning
NNUH	Norfolk and Norwich University Hospital
ONT	Oxford Nanopore Technology
PBS	Phosphate-buffered saline
PLC	Phospolipidase C
PMA	Propidium Monoarzide
PHE	Public Health England
qPCR	Quantitative PCR
UTI	Urinary Tract infection
SQK	Sequencing Kit
WHO	World Health Organisation
WIMP	‘What’s in my pot’ software
WGA	Whole genome amplification
WGS	Whole genome sequencing

Table of Contents

1: Introduction

1.1: Infectious Diseases

Infectious diseases caused by pathogenic microorganisms such as bacteria, viruses, parasites or fungi are still the major cause of morbidity and mortality worldwide. According to the World Health Organisation (WHO), approximately 12.5% of all deaths globally are due to infectious diseases (WHO 2014). In the UK, the Chief Medical Officer has estimated that, in 2010, infectious diseases accounted for 7% of all deaths. The total economic burden from infectious diseases in England is estimated at £30bn annually (this includes costs to the health service, the labour market and to individuals) (House of Parliament, 2017).

New and re-emerging infections are on the rise as global travel becomes increasingly common, the human population grows, people live closer to wildlife and population density increases. This has been demonstrated by new threats (e.g. Zika outbreak in Brazil 2016) and re-emerging infections (e.g. Ebola outbreak in West Africa 2014). Drug resistance mechanisms such as the New Delhi Metallo-beta-lactamase-1 (NDM-1) in *Enterobacteriaceae*, mobilized colistin resistance (mcr-1) in Gram-negative bacteria and drug resistant HIV are becoming increasingly common. Furthermore, opportunistic and hospital acquired infections are becoming more of a concern, often causing serious illness and sometimes fatality (Melzer & Welch, 2013). In the face of rapidly changing global epidemiology, increased invasive techniques, transplantation and surgical procedures, increased use of immunosuppressive drugs and chemotherapy, and increasing numbers of multidrug-resistant microorganisms, rapid and improved methods for diagnosis and surveillance of infectious diseases become increasingly important in reducing global mortality, morbidity and healthcare associated costs.

1.2: Current Methods for Infection Diagnosis

Medical microbiology is a field that focuses on characterising pathogens in clinical samples for the direct management of patients (diagnostic microbiology) and infectious disease monitoring for transmission and epidemiology (public health microbiology). Over the last century, the principles behind diagnostic bacteriology have not changed much, except for the use of rapid molecular methods such as PCR, which has had the biggest impact on virology and sexually transmitted infections. Diagnostic bacteriology is mainly dependent on a large repertoire of culture methods to determine the microbial composition of a sample (Baquero, 2017; P.-E. Fournier et al., 2013).

Different culture methods and media types are used for diagnosis, based on the suspected pathogen and different clinical sample type. Samples that are normally sterile (e.g. blood, cerebrospinal fluid) are often cultured in a rich medium that will support the growth of any culturable organism present in the sample. Whereas samples that are contaminated with colonizing flora, such as stool and respiratory samples are often cultured on selective media to favour the growth of the suspected pathogen (Didelot, Bowden, Wilson, Peto, & Crook, 2012). These methods have been improved over the years by using liquid cultures, automated culture systems, matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) (P. Fournier, Drancourt, …, & 2013, n.d.; Tille, 2013). However, despite these advances, culture techniques are still hampered by low sensitivity, in some conditions, and long turn-around times. The sensitivity of culture is particularly low in sterile specimens such as blood and prosthetic joint infection samples. This sensitivity can be further diminished if antimicrobial treatment has been initiated in a particular patient (Jonathan Cohen et al., 2015). Furthermore, it has been estimated that ~10% of all bacteria are thought to be unculturable or difficult to grow (Didelot et al., 2012; He et al., 2013; Stewart, 2012). Methods for detecting such organisms are largely dependent on serological, antigen and nucleic acid amplification tests.

Following microbial culture, species identification can be performed by colony staining, colony growth and morphology depending on the media used and as judged by the biomedical/clinical scientist (Washington, 2012). Identification can also

be done by molecular hybridization tests and rapid biochemical reactions. Species characterization using 16S ribosomal RNA (rRNA) gene sequencing has been tried but had a number of drawbacks including complicated data interpretation, only genus level identification for some bacteria, long turn-around time and was labour intensive (Mwaigwisya, Assiri, & O'Grady, 2015; J. B. Patel, 2001; Poretsky, Rodriguez-R, Luo, Tsementzi, & Konstantinidis, 2014). Recently, there has been increasing implementation of MALDI–TOF for culture identification in clinical laboratories. MALDI-TOF works by analysing the spectrum of ionised molecules (e.g. ribosomal proteins) in the gas phase. Each spectrum from a species is unique and is separated and analysed based on molecular weight after migration in an electric field. Each molecule detected is characterized against a large database based on molecular mass, charge, mass/charge ratio, and relative signal intensity to identify bacterial and fungal species (Bizzini & Greub, 2010; Burillo et al., 2014).

The cost of MALDI-TOF analysis per sample is low, € 1.35 -1.44 (Seng et al., 2013), the sample preparation method is simple, turnaround time is rapid (less than 10 minutes (Seng et al., 2013)), and the technology has a large database for species and genus identification (Bizzini & Greub, 2010)(Bader et al., 2011). Currently, there are two commercially available MALDI-TOF systems, the VITEK MS (bioMérieux Clinical Diagnostics) and MALDI Biotyper (Brüker, Daltonics GmbH). The major difference between the systems is the number of taxa in the reference database. The VITEK MS system contained >25,000 spectra covering 586 species, consisting of 508 bacterial and 78 fungal, while the Brüker Biotyper contained >80,000 spectra covering 2048 species and 385 genera (Bilecen, Yaman, Ciftci, & Laleli, 2015). Both systems demonstrate similar efficiency, sensitivity of over 99% and specificity over 98% compared to the traditional methods (Mellmann et al., 2008), workflow robustness (Bilecen et al., 2015), and provide results in less than an hour. These features make this technology an attractive alternative to traditional culture identification methods by increasing the speed and accuracy.

After pathogen identification, the susceptibility profile is determined for an isolate. Phenotypic methods for antibiotic susceptibility testing (AST) are almost exclusively based on the inhibition of bacterial growth when exposed to test antimicrobials (Minimum Inhibitory Concentration) (Tille, 2013). To speed up AST and standardize

reading of results, several automated systems that combine both classical identification and AST have been introduced. The systems include:

- Vitek® (bioMérieux Clinical Diagnostics);
- The MicroScan WalkAway system (Beckman Counter, Brea, CA, USA);
- The BD Phoenix system (BD Diagnostics, Sparks MD, USA; and
- Sensititre ARIS 2X (Trek Diagnostic Systems, Cleveland, OH, USA)

The Vitek® (bioMérieux Clinical Diagnostics) is a fluorometric test with biochemical substrates for bacterial growth during an abbreviated incubation period in microwells. The system can identify Gram-negative fermenting and non-fermenting bacilli, Gram-positive cocci and spore-/non-spore forming bacilli and yeasts to the species level along with susceptibility testing within 8-10 hours (Putnam, Howard, Pfaller, Koontz, & Jones, 1997). The MicroScan WalkAway system (Beckman Coulter, Brea, CA, USA) contains 40–96 microdilution trays for identification of Gram negative and Gram positive bacteria, and susceptibility testing by periodically measuring colour change or increases in turbidity (Burns et al., 2001). The system provides results within 16-18 h (Burns et al., 2001). The BD Phoenix system (BD Diagnostics, Sparks MD, USA) has the capacity to process 100 test panels for pathogen identification and AST. The system monitors each panel every 20 min for up to 16 h (Menozzi et al., 2006). Finally, Sensititre ARIS 2X (Trek Diagnostic Systems, Cleveland, OH, USA), a bench-top incubating and reading system with a 64-panel capacity for both identification and AST. Presumptive identification of Gram-negative pathogens can be obtained within 5 h and comprehensive identification to species level after 18 h (Jorgensen & Ferraro, 2009). Molecular tests have also been used to determine the resistance profiles of pathogens using well characterized markers such as the *mecA* gene in *Staphylococcus aureus* (Shrestha, Tuohy, Hall, Isada, & Procop, 2002) or *rpoB* gene and *inhA* gene to predict rifampicin and isoniazid resistance in *Mycobacterium tuberculosis* isolates (Telenti et al., 1997).

Routine characterisation of virulence determinants of isolates has not been a major priority in the clinical laboratory. Exceptions to this, include the identification of virulence determinants in *Streptococcus pneumoniae*, *Corynebacterium diphtheria, Clostridioides difficile* (formerly known as *Clostridium difficile*) and

foodborne/waterborne pathogens (*Salmonella, Shigella, Campylobacter*, Shiga toxin-producing *Escherichia coli* (STEC), pathogenic *Enterobacteriaceae*, *Listeria*, *Clostridium perfringens, Vibrio* and *Cronobacter*) (Didelot et al., 2012; Gerner-Smidt et al., 2006). Presence of toxin producing genes in these specific isolates is important in determining antitoxin administration to a patient or if treatment is required. Currently, virulence tests are performed by serotyping (e.g. pneumococcal capsule) or EIA (enzyme immuno-assays) e.g. *C. difficile*. For public health purposes, virulence determinants, such as capsule type, are important in vaccine studies for pathogens including *Haemophilus influenzae*, *Streptococcus pneumoniae* and *Neisseria meningitidis* (Didelot et al., 2012).

When a pathogen with the potential to cause hospital or public outbreak such as meticillin-resistant *S. aureus* (MRSA), *Salmonella spp* or *C. difficile* is identified, further typing tests are performed by public health specialists (PHE 2017) (Sabat et al., 2013). Isolated pathogens are chosen for typing based on the infection, epidemiological criteria and other public health guidelines, due to this complex process, many outbreaks are likely to be missed (Didelot et al., 2012). Current typing methods can take a few weeks and up to months for outbreak characterization. Also, some of the tests are faced with several draw-backs such as being species-specific, labour intensive and dependent on the level of biomass produced by culture methods, which takes 24 to 48 hours (Sabat et al., 2013). Typing tests are performed in reference/specialised laboratories and only a small number of clinical laboratories can carry out these tests routinely (Köser et al., 2012).

In summary, the current methods in the clinical laboratory are largely dependent on growing pathogens and are limited by complex procedures that have many contingencies to classify species and their susceptibility profile. In outbreak cases, further tests are required, which demands extensive knowledge, have long turnaround times, are often complex and labour-intensive (P. Fournier et al., n.d.; Sabat et al., 2013; Washington, 2012). Overall, the described multiple-step process may take from days for rapidly growing bacteria such as *E. coli,* one to two weeks for bacteria such as *C. difficile* to a month for slow-growers such as *M. tuberculosis*.

1.3: Methods for Blood Stream Infection, Urinary Tract Infection and Clostridioides difficile Infection Diagnosis

1.3.1: Blood Stream Infections

The presence of infectious organisms in the blood, including bacteria, viruses, or fungi is generally known as a blood stream infection (BSI). BSIs may be associated with infections, such as endocarditis, and procedures, such as central venous catheter, colonoscopy or dental extraction. BSIs can also be caused by secondary focal infections, such as pneumonia, urinary tract infections (UTIs) or abscesses. BSI often develops to sepsis, defined as a life-threatening organ dysfunction caused by a dysregulated host response to infection (Singer et al., 2016). This happens when the innate immune response becomes amplified and dysregulated, leading to an imbalance between pro- and anti-inflammatory responses with the excessive release of cytokines and other inflammatory regulators (Schulte, Bernhagen, & Bucala, 2013). Septic shock, a subset of sepsis in which underlying circulatory and cellular metabolism abnormalities are profound enough to substantially increase mortality, is a more severe illness with a much higher likelihood of death than sepsis (J Cohen, 2002; Rhodes et al., 2017; Singer et al., 2016). In Europe and the USA combined, sepsis causes more than 400,000 deaths every year, costing more than $20 billion in the USA in 2011 (Daniels, 2011; McPherson et al., 2013; Torio, 2013).

High mortality rate associated with sepsis can be influenced by a combination of factors, including the characteristics of the infecting pathogen and the host response to infection. Studies suggest that every hour delay in administration of effective antimicrobial therapy in patients with septic shock is associated with a 7% decrease in survival rate (Ferrer et al., 2014; Kumar et al., 2006, 2009). The Surviving Sepsis Campaign recommends confirming sepsis by objective evidence of infection, which is achieved by identifying grown pathogens by blood culture methods. Culture analysis, including antibiotic susceptibility testing (AST), takes at least 48 h for positive and fast growing pathogens and five days to identify specimens as negative (Dellinger et al., 2013; Kumar et al., 2006).

Due to temporal separation between initial clinical suspicion of sepsis and laboratory confirmation of BSI, a ‘safety first’ strategy of using potent broad-spectrum antibiotics

within 1 h of sepsis onset is employed, aimed at covering the most probable pathogens (Dellinger et al., 2013; Kumar et al., 2006). Although effective (Kumar et al., 2009), rapid empirical treatment of septic patients has several limitations:

- under-treatment of a few patients with resistant pathogens, with contingent mortality (Kollef, 2000);
- over-treatment of considerably more patients with susceptible pathogens, with antibiotic overuse and increased cost and occurrence of adverse drug effects (Kollef, 2000; Kumar et al., 2009);
- clinical utility of blood samples taken for blood culture after antibiotic therapy is administered can be diminished due to sterilization of the sample (Ferrer et al., 2014).

Also, international guideline for management of sepsis and states that antibiotic treatment should be tailored following microbiological culture results (Dellinger et al., 2013). Benefits of tailored treatment include: use of less toxic agents; use of antibiotics that achieve higher concentrations; the option to select more effective agents if inherent or acquired resistance is detected; and it is associated with a more favorable clinical outcome (Garnacho-Montero et al., 2014; Stoneking et al., 2013). However, a study by Garnacho-Montero et al. (2014) showed that 12% of patients die before microbiology results are available to clinicians. In addition, blood culture is, at best, only 60% sensitive (Afshari, Schrenzel, Ieven, & Harbarth, 2012; Ellepola & Morrison, 2005; Garnacho-Montero et al., 2014; Rivers et al., 2001; Vincent et al., 2006). The sensitivity is particularly poor for slow-growing and fastidious organisms responsible for community-acquired pneumonia such as *Mycoplasma pneumoniae*, *Legionella pneumophila*, and *Chlamydia pneumoniae* (Fenollar & Raoult, 2007; Gadsby et al., 2016). If no pathogens grow, there is no scope to guide the refinement of treatment. The slow turnaround time of culture methods means prolonged use of broad-spectrum antibiotics even if a pathogen can be grown and identified. Extended use of broad-spectrum antibiotics is associated with increased antibiotic resistance and collateral damage to the normal gut flora, resulting in colonization by opportunistic pathogens, such as *C. difficile* and/or the selection of *Enterobacteriaceae* with transferable drug resistance (Cooper & Shlaes, 2011; Dark et al., 2011).

A number of new technologies have been introduced that speed up phenotypic diagnosis. While progress has been made, none of the methods have succeeded in changing clinical practice (P.-E. Fournier et al., 2013; Raoult, Fournier, & Drancourt, 2004). The speed of molecular diagnostics assays have the potential to improve the current situation by providing results to clinicians before a second dose of broad-spectrum treatment is administered (8 h), resulting in improved antibiotic stewardship and reduced patient morbidity and mortality. Currently, the detection of pathogen DNA in blood by molecular methods relies on nucleic acid amplification rests (NAATs) such as real- time PCR, with or without a period of incubation in blood culture. Because of the large and diverse range of pathogens that cause sepsis, PCR assays are often multiplexed. Multiplexed and broad range assays (e.g. 16S rRNA and IST assays) have been designed to cover most likely pathogens and, in some cases, important antibiotic resistance markers. However, some of these tests are hindered by low sensitivity, sub-optimal positive predictive values and are reliant on a pre-set range of targets meaning less common pathogens will be missed. Consequently, this and other NAAT diagnostics are used as an adjunct to shorten turn-around time of microbiological techniques and not as a replacement (Mwaigwisya et al., 2015).

Rapid BSI diagnosis

To speed up the diagnosis of BSI, several manufacturers have developed technologies for pathogen identification either directly from whole blood or from positive blood culture bottles.

Verigene® (Nanosphere, Chicago, IL, USA) is a microarray based test for pathogen detection in positive blood cultures. There are two Verigene® assays: Gram-Positive Blood Culture (BC-GP) and Gram-Negative Blood Culture (BC-GN) assays (Nanosphere). The BC-GP identifies 13 species and genera, and three resistance markers (*mecA*, *vanA,* and *vanB*) (Wojewoda et al., 2013), whereas the BC-GN detects 8 species and genera bacteria and six antimicrobial resistance genes (bla_{CTX-M} genes encoding extended-spectrum β- lactamases (ESBLs) and bla_{IMP}, bla_{KPC}, bla_{NDM}, bla_{OXA}, and bla_{VIM} for the detection of carbapenemases) (Hill, Tran, Barton, Labreche, & Sharp, 2014). The range of evaluated sensitivity and specificity

of the assay was 80.4-94.3% and 95-100% respectively (Dodémont, De Mendonça, Nonhoff, Roisin, & Denis, 2014; Hill et al., 2014; Mwaigwisya et al., 2015). FilmArray® (Biofire Diagnostics, Salt Lake City, UT, USA) is a multiplex PCR based test for the identification of 24 bacterial targets, five fungal targets, and four antibiotic resistance genes including *mecA, vanA/B*, and *Klebsiella pneumoniae* carbapenemase (bla_{KPC}) from blood cultures (Guido et al., 2016). The reported range of specificity and sensitivity was 80.4-94.3% and 95-100% respectively(Mwaigwisya et al., 2015).

Prove-it™ Sepsis (Mobidiag, Espoo, Finland) is a commercially available technology for pathogen identification in positive blood cultures (Tissari et al., 2010). The Prove-it™ system is a broad-range PCR and microarray-based assay, which identifies BSI-causing bacteria and fungi from positive blood culture in three hours. This test panel covers 80 species of Gram-negative and Gram-positive bacteria, 13 fungal species and antibiotic resistance marker *mecA* for MRSA (Tissari et al., 2010). However, the test panel misses important pathogens such as *Streptococcus viridans*, *Candida spp*., and coagulase-negative Staphylococci (Guido et al., 2016). Three diagnostic studies have shown the sensitivity and specificity of the assay to range between 96-99% and 98-100% respectively (Mwaigwisya et al., 2015).

The most widely deployed rapid molecular test for pathogen identification in positive blood cultures is PNA-FISH (Fluorescent *in Situ* Hybridization, FISH), a DNA hybridization based test (Perry-O'Keefe et al., 2001; Wagner, Horn, & Daims, 2003). The test is based on fluorescently labeled oligonucleotide probes specific for rRNA, which have been used to detect 95% of bacteria and fungi in blood culture in 2.5–3 hours . The most commonly used target in prokaryotes is the 16S rRNA gene (Perry-O'Keefe et al., 2001; Wagner et al., 2003). The test is quick, easy to interpret and requires only basic laboratory equipment but is limited with a requirement of an organism concentration of at least 10^5 CFU/mL for detection which may be problematic for identification of slow-growing, or fastidious organisms (Guido et al., 2016; Harris & Hata, 2013).

Blood culture dependent molecular tests based on microarrays, hybridization, and PCR are rapid and sensitive enough for pathogen identification. However, the

diagnostic advantage of these tests is challenged by MALDI-TOF methods for identifying pathogens in positive blood cultures (Mwaigwisya et al., 2015; Salipante et al., 2013). Molecular assays applied directly to clinical samples, without culture, have the potential to significantly shorten the time required for microbial identification and, thereby, ensure patients receive the appropriate antibiotic treatment faster. This would improve the care of individual patients and reduce the unnecessarily prolonged use of broad-spectrum antibiotics, which presently are given until a pathogen is grown and characterised by conventional methods (taking a total of 48-72 h).

The LightCycler® SeptiFast assay (Roche Diagnostics, Basel, Switzerland) is a real-time PCR multi-pathogen probe-based test with a detection panel consisting of 25 different sepsis-causing pathogens and the *mecA* gene. The assay can be completed within six hours (Vince et al., 2008).The range of sensitivity and specificity of the LightCycler SeptiFast test has been reported to be between 68–75% and 86–92% respectively (Dark et al., 2015; Mwaigwisya et al., 2015).

The SepsiTest (Molzym, Bremen, Germany) is another commercially available real-time PCR test, which targets conserved regions of the 16S and 18S rRNA genes of bacteria and fungi respectively. This assay involves selective degradation of human DNA, which at high concentrations can inhibit pathogen PCR assays, while leaving pathogen cells intact. This technology can detect more than 345 bacteria and fungi with a limit of detection (LoD) of 20-460 CFU/ml (Muhl, Kochem, Disque, & Sakka, 2010). The first part of this assay can be completed within four hours, however, positive samples require sequencing, which extends the turnaround time by at least an extra 8 to 12 hours (Guido et al., 2016) (new sequencing technology, such as nanopore sequencing, could reduce this turnaround time). Studies have shown the range of sensitivity and specificity to be 86–87% and 83–85% respectively (Haag, Locher, & Nolte, 2013; Mwaigwisya et al., 2015; Stevenson et al., 2016).

The PLEX-ID (Abbott Molecular, Ibis Biosciences, Carlsbad, CA, USA) test involves automated DNA extraction and PCR which targets 16S rRNA, 23S rRNA, and four housekeeping genes. The PCR amplicon is analysed by ESI-MS to identify a broad range of pathogens and four resistance markers (*mecA, vanA/B*, and *bla*KPC) from

whole blood (Simner et al., 2013). The newer version of the PLEX ID is reported to have a sensitivity of 83% and a specificity of 94% (Guido et al., 2016).

The Magicplex™ sepsis real-time assay (SeeGene, Seoul, Korea) has a multistep approach that involves conventional PCR and real-time PCR associates to screen for 90 pathogens (73 Gram-positive, 12 Gram-negative, and six fungi) as well as three drug resistance markers (*mecA* and *vanA/B*) from whole blood samples (Opota, Jaton, & Greub, 2015). The turnaround time of the test is six hours with 65% sensitivity and 92% specificity (Carrara et al., 2013; Guido et al., 2016).

One instrument that has gained CE (European Conformity) and FDA (US Food and Drug Administration) approval for diagnosis of BSI is T2Dx® Instrument (T2Biosystems, Lexington, Massachusetts). The instrument enables nanoparticle-mediated rapid detection of five species of *Candida* (*C. albicans, C. tropicalis, C. parapsilosis, C. glabrata,* and *C. krusei*) and six species of bacteria (*Enterococcus faecium, S. aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa* and *E. coli*) (T2Biosystems, Lexington, Massachusetts).

The assays, T2Bacteria and T2Candida, use specific primers to PCR amplify candida and bacteria DNA, which then binds to paramagnetic nanoparticles coated with complementary probe. For each target, two nanoparticles are designed, each bearing a different complementary probe; one hybridizing to the 5' end of the single-stranded DNA target and one to the 3' end. Upon hybridization, the nanoparticles cluster and the extent of clustering increases with target DNA concentration (Neely et al., 2013). This clustering alters the magnetic resonance signal in the sample, indicating the presence of the target. The change in signal (T2 relaxation) is measured by the T2Dx® Instrument. According to the manufacturer and other studies, test panels have a limit of detection of 1 CFU/ml and can be performed within 3 h for candida and 5 h for bacteria (De Angelis et al., 2018; Neely et al., 2013).

Evaluation study of T2Bacteria has reported the overall sensitivity and specificity of 83.3% and 97.6% respectively, in comparison with blood culture (De Angelis et al., 2018). For T2Candida test, the reported sensitivity was 100% and specificity was

98% (De Angelis et al., 2018). The technology is now of considerable interest for clinical use for microbial diagnosis of BSI internationally.

Despite higher sensitivity and quick turn around-times compared to culture assays, these technologies have failed to replace culture due to limited detection panel, failure to detect drug resistance and higher cost.

1.3.2: Urinary Tract Infections

UTIs are one of the most common bacterial infections at all levels of healthcare settings (Foxman, 2010; Wilson & Gaido, 2004). UTI account for over 8 million physician visits per annum in the USA (Schappert & Rechtsteiner, 2011). Most are trivial, but, in severe cases, infection may ascend to the kidneys, with overspill to the bloodstream precipitating bacteraemia and urosepsis. In the UK, UTI was the condition with the highest emergency admissions rate in 2012/13 with 67 admissions per 100,000 population per quarter on average (NHS 2014). Laboratory investigation of urine samples in the UK commonly relies on:
Triage on a screening system;
Culture on chromogenic agar;
Disc testing by BSAC (British Society for Antimicrobial Chemotherapy) or EUCAST (European Committee on Antimicrobial Susceptibility Testing) methodology (Hay et al., 2016).

Primary urine analysis is based on the visual inspection of colour, turbidity and odour. In recent years, the diagnostic technology for UTIs has improved significantly, with moves to automate traditional manual methods and to adopt fully automated systems. The pressure to increase laboratory efficiency, reduce costs, and allow clinicians to optimise therapy earlier prompts the use of rapid and innovative technologies. Although microscopy and quantitative culture are still the gold standard, non-culture techniques with urine analyzers are increasingly used as the first triage to predict the presence of infections (Roberts, 2011; Schmiemann, Kniehl, Gebhardt, Matejczyk, & Hummers-Pradier, 2010).

The Public Health England (PHE) guideline for UTI management in adults in primary care recommends initiation of empirical treatment without routine culture for patients

with uncomplicated UTIs with three or more clinical symptoms (i.e., dysuria, frequency, polyuria, haematuria, suprapubic tenderness) without vaginal discharge or irritation (PHE, 2014). The diagnosis of mild UTIs with two or fewer clinical symptoms requires collecting a urine specimen and performing a point-of-care dip-stick test to identify nitrites and leukocytes as UTI markers. If both results are negative the GP would consider other diagnoses; negative nitrite and positive leukocytes or positive nitrite and negative leukocytes indicates a likely UTI, therefore the patient would be treated with a first-line antibiotic, and a simultaneous urine sent for culture; if both markers are positive, treatment with first-line agents is required. The dipstick tests are routinely used to identify pathological changes in urine (PHE, 2014).

Analysis of the dipstick test is based on the detection of colour changes contingent on the presence of compounds, enzymes or cell types that ordinarily are absent from urine e.g. nitrite, leukocytes esterase, erythrocytes, protein (albumin), glucose etc (H. D. Patel, Livsey, Swann, & Bukhari, 2005). Presence of nitrile is associated with the presence of nitrite Gram-negative rods, e.g. *Enterobacteriaceae* nitrate reductase. Nitrile is not detected if the causative uropathogen is not nitrate-reducing (or if it reduces nitrate to nitrogen gas) as with e.g. *Enterococcus spp., S. saprophyticus*, *Pseudomonas spp*., or *Acinetobacter spp* (Semeniuk & Church, 1999). However, the leukocytes esterase and nitrite tests are less reliable in diabetic and elderly individuals (Ipe, Sundac, Benjamin Jr, Moore, & Ulett, 2013). The mean sensitivity of dipstick is 48% (Putnam et al., 1997) for the detection of clinically significant bacteriuria of $>10^5$ CFU/mL to 68% (Gangaram, Ojwang, Moodley, & Maharaj, 2005).

Microscopy observation for red and white blood cells (RBC/WBC), epithelial cells, bacteria, casts (hyaline, cellular, granular) and crystals, is another rapid point-of-care test for UTI. However, the method is limited by low sensitivity for detecting bacteria in uncentrifuged and unstained urine, which is around 10^4 CFU/mL (Goswitz et al., 1993; Pezzlo, 1988). Sensitivity and specificity of the test can also vary greatly between experienced and inexperienced workers (Pezzlo, 1988).

Rapid semi-automated urine analysers are used as an alternative to manual laboratory screening methods. As for microscopy observation, the systems analyse

the presence of RBCs, WBCs, crystals, casts, epithelial cells, mucus, bacteria and yeasts. These analysers include:

- iQ® 200 (Iris Diagnostics, Chatsworth, CA, USA) based on flow-cell digital imaging;
- SediMax (Menarini Diagnostics, Florence, Italy) based on sediment analysis; and UF- 1000i (Sysmex Corporation, Kobe, Japan) based on fluorescence flow cytometry.

Semi-automated urine analysers have been variously reported to reduce the number of urine cultures performed by more than 40% (Broeren, Bahçeci, Vader, & Arents, 2011; van der Zwet, Hessels, Canbolat, & Deckers, 2010), thereby reducing both costs and workload. However, the sensitivity and specificity of the analysers vary depending on the parameters and cut-offs employed. There has also been contradictory reports of the performance of analysers. One study (Sterry-Blunt, Randall, Doughton, Aliyu, & Enoch, 2015) reported that the SediMax analyser was not suitable to screen routine urines prior to culture due to a low negative predictive value (87.5%); but other studies showed that the system could be reliably applied to urine screening (Falbo et al., 2012; Tessari, Osti, & Scarin, 2015). None of the automated screening analysers identify the bacterial species present or test their antibiotic susceptibility.

Culture based methods remain the gold standard for pathogen identification and AST. The guidelines recommend sending specimens for culture following a triage of screening analysis and if there are two or more signs of infection (PHE 2014). Quantitative urine culture is commonly performed on non-selective solid media including blood and nutrient agar, or selective media such as MacConkey, Cysteine Lactose Electrolyte deficient (CLED), and chromogenic agar where the colony colour varies with the organism species (Hay et al., 2016). Culture methods require 24h to grow a pathogen followed by species identification by MALDI-TOF; and a further 24h for species identification using other automated systems and AST. Susceptibility testing is performed by either manually or automated systems (discussed in section 1.2). Results of the tests are classified as either susceptible, intermediate, or resistant based on the criteria provided by EUCAST (European Committee on Antimicrobial Susceptibility Testing) or CLSI (Clinical and Laboratory Standards

Institute) (PHE, 2014) (Hay et al., 2016).

Rapid UTI diagnosis

Global concern about antibiotic resistance and the lack of new classes of antibiotics has driven the development of more rapid, comprehensive, sensitive, and specific diagnostics. A fast pathogen identification and AST for UTI would have an important clinical impact given high number of UTI cases and high frequency of resistance to primary antibiotics (Stamm & Norrby, 2001). New approaches based upon PCR, metagenomics sequencing, and rapid phenotypic methods have been developed and tested. Also, advances has been made on phenotypic methods including microfluidics and lab on chip technologies approaches for ART (Murray, Adeyiga, Owsley, & Di Carlo, 2015; Reece et al., 2016), laser scattering technologies (Hayden et al., 2016) and digital microscopy (Zaman et al., 2010). These technologies are mainly focused on either rapid screening or AST by measuring early bacterial growth curves after exposure to antimicrobial agents.

Multiple in-house PCR methods have been developed targeting the most common bacterial pathogens and resistances involved in UTIs. These tests have shown to be more rapid and sensitive than culture based methods (Felt et al., 2017; Shigemura et al., 2005; Zhang, Niu, & Zhang, 2014). Currently commercially available real-time PCR tests include:

- Cepheid Xpert MRSA test (Cepheid Sunnyvale, CA, USA)
- BD GeneOhm MRSA (Becton Dickenson, Heidelberg, Germany)
- Check-Direct CPE kit and Check-MDR ESBL kit (Check-Points, Aageningen, Netherlands)
- Easyplex® SuperBug CPE kit (Amplex Diagnostics GmbH, Mark Gars, Germany).

Although used for urine testing, these tests are largely applied to nasal or rectal swabs to screen for carriers of resistant strains for the implementation of infection control measures rather than to guide therapy (Bradley et al., 2015; Dolinger & Jacobs, 2015; Emmadi et al., 2011; Huang et al., 2015; Lucke, Hombach, Hug, & Pfyffer, 2010). These tests screen for resistance such as;

- *mecA* and *mecC* genes for MRSA (Cepheid Xpert/Gene Expert);
- *vanA, vanB* and *vanC* genes for glycopeptide resistance (e.g. Xpert/Gene Expert, Cepheid; LightCycler® VRE Detection Kit, (Roche);
- four carbapenemase genes (*bla*KPC-like, *bla*OXA-48-like, *bla*NDM-like, *bla*VIM-like) (Check-Direct CPE kit); and
- two ESBL genes (*bla*CTX-M-1 and *bla*CTX-M-9 groups) (Easyplex® SuperBug CPE) (Dolinger & Jacobs, 2015; Lutz E Lehmann et al., 2011).

Microfluidics approaches, pathogen growth-based antimicrobial resistance testing, followed by single-cell imaging (Boedicker, Li, Kline, & Ismagilov, 2008) and digital PCR (Schoepp et al., 2016) have been developed. Considerable advances have been reported regarding in microfluidic devices including microchambers (Choi et al., 2014), microchannels (Chen et al., 2010), and microdroplets (Baraban et al., 2011). In these devices small populations of bacteria are cultured in small volumes (nanolitre or picolitre) under different antibiotic conditions and cell growth. Cell divisions are then closely monitored via methods such as microscopy and fluorescence. Small volumes shortens the detection time to one to three hours (B. Li et al., 2014). A big challenge in using microfluidic devices is contamination and polymicrobial infections which can be observed in complicated UTIs (Baltekin, Boucharin, Tano, Andersson, & Elf, 2017).

Another approach is based on digital PCR of nucleic acid quantification, (Schoepp et al., 2016), the method uses DNA markers to perform a phenotypic AST after short antibiotic exposure. Digital methods partition bacterial DNA into thousands of compartments and then use targeted amplification to determine the number of "positive" compartments containing DNA carrying one or more copies of the target gene. Relative DNA copy number increase in antibiotic-treated aganist reference cultures is quantified using digital PCR, results suggest that a biological response can be detected within 15 min after exposure to an antibiotic (Schoepp et al., 2016).

1.3.3: *C. difficile* infection

Clostridioides difficile (formerly known as *Clostridium difficile*) is a Gram-positive toxin producing and spore forming anaerobic bacillus that infects the gastrointestinal

tract. Symptoms of infection range from mild diarrhoea to life-threatening colitis, and is associated with substantial morbidity and mortality mainly in elderly, hospitalised patients particularly those receiving antibiotics, and in long-term care facilities (Kelly, MD & LaMont, MD, 1998; Leffler & Lamont, 2015). In the USA, CDI is the leading cause of hospital-associated gastrointestinal illness (Surawicz et al., 2013); a surveillance study in 2011 identified 453,000 cases of CDI and 29,000 associated deaths, approximately a quarter of those infections were community-acquired (Lessa et al., 2015). Nosocomial *C. difficile* infection more than quadruples the cost of hospitalizations (Lofgren, Cole, Weber, Anderson, & Moehring, 2014), increasing annual expenditures by approximately $1.5 billion (Zimlichman et al., 2013).

C. difficile can also be found in healthy people's intestines where it causes no symptoms (up to 3% of adults and 66% of babies) (Mark H Wilcox et al., 2017). The disease pathogenesis is largely considered to be due to the production of two potent toxins by the bacteria, TcdA and TcdB (Leffler & Lamont, 2015). Most of the *C. difficile* infections are believed to be acquired from the patients' immediate surrounding such as health care settings, care homes, and environmental contamination and is the major cause of hospital acquired diarrhoea. Once CDI cases have been identified, immediate measures to stop transmission are implemented, following isolation and cleaning guidelines from the Department of Health UK (DH 2012) (S. H. Cohen et al., 2010).

Due to its physical properties and strict anaerobic requirements, the dormant spore is the infectious and transmissible morphotype of *C. difficile.* These spores are responsible for the persistence of CDI in patients, hospital transmission and environmental contamination (Britton & Young, 2014; Deakin et al., 2012; Paredes-Sabja, Shen, & Sorg, 2014). Bacterial spores have structural layers that give them unique resistance properties. The core, which contains the spore DNA, RNA and most enzymes, is surrounded by a compressed inner protein membrane, which has a similar phospholipid composition to growing bacteria but exhibits very low permeability to small molecules such as water (Paredes-Sabja et al., 2014). The germ wall, which becomes the cell wall of the outgrowing bacterium, surrounds the core/inner membrane (Moir, Corfe, & Behravan, 2002). Surrounding the germ cell wall, is a thick peptidoglycan layer, the cortexis, surrounded in turn by an outer

membrane derived from the mother cell, which is essential for spore formation (**Figure 1-1**). A proteinaceous coat surrounds the outer membrane, the coat contains unique proteins to spores, and is essential for spore resistance to commonly used decontaminants. The outermost layer of the *C. difficile* spore is exosporium (Driks, 1999; Paredes-Sabja et al., 2014). Structure and stability of exosporium is thought to be species dependent, while *C. difficile* strain 630 spores have an electron-dense, compact exosporium layer, strains R20291, M120, TL176 and TL178 have a hair-like exosporium layer (Paredes-Sabja et al., 2014). Apart from thick and compressed layers, other factors contributing to spore stability and resistance properties are the low water content (25–60% of wet weight), elevated levels of dipicolonic acid (DPA) (25% of core dry weight), and the saturation of DNA with α/β-type small acid soluble proteins (SASP) (Paredes-Sabja et al., 2014).

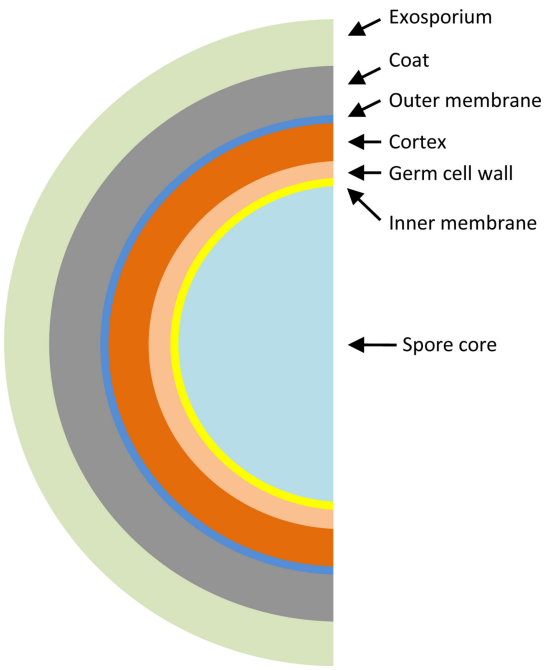


Figure 1- 1 Structural layers of bacterial spores. The main layers of bacterial spore structure are shown and not drawn to scale. Presence of exosporium layer is species dependent (Paredes-Sabja et al., 2014b)

These structural properties make *C. difficile* spores metabolically dormant and intrinsically resistant to antibiotics, attacks from the host's immune system, and resistant to bleach-free disinfectants commonly used in hospital settings. When shed into the environment they spread easily because of their structure and morphology (Paredes-Sabja et al., 2014). Once a person is infected, *C. difficile* spores germinate

in the colon to form vegetative cells that initiate CDI. In fact, strains that are unable to form spores during CDI are unable to persist in the colonic tract of the host and be transmitted (Leffler & Lamont, 2015; Paredes-Sabja et al., 2014).

In the UK, laboratory investigation for CDI is requested for patients with diarrhoea (Bristol Stool Chart types 5-7) that is not attributable to an underlying condition such as inflammatory colitis, overflow or therapy (e.g., laxatives and enteral feeding). Initially, testing for CDI was commonly performed by EIAs as a standalone testing method. However, suboptimal performance reported by several studies (S D Goldenberg & French, 2011; Simon D Goldenberg, Price, Tucker, Wade, & French, 2011; Longtin et al., 2012) led to the amendment of the guidelines for CDI testing by the Department of Health and National Health Services (NHS) in 2012 (DH 2012). The current guidelines recommend performing two tests, glutamate dehydrogenase (GDH) EIA *or* PCR *and* toxin EIA. CDI is likely to be present if GDH EIA (or PCR) positive, and toxin EIA positive. If EIA toxin is negative but GDH EIA (or PCR) is positive, CDI could be present and the patient may have potential to transmit spores/organism. CDI is unlikely to be present if GDH EIA (or PCR) and toxin EIA are negative (DH 2012). Once CDI has been confirmed, the Health Protection Agency (HPA) is notified and the samples are sent away for culture and further investigation including molecular typing at the reference laboratory.

Generally, CDI transmission and epidemiology is investigated by the following molecular typing methods:

- multilocus variable-number tandem-repeat analysis (MLVA) targeting DNA repeat units;
- amplified fragment length polymorphism (AFLP) targeting *PstI* and *MseI* restriction sites;
- multilocus sequence typing (MLST) utilising housekeeping loci;
- pulsed-field gel electrophoresis (PFGE) targeting *SmaI* restriction sites;
- PCR ribotyping targeting 16S-23S spacer region; and
- whole-genome sequencing (WGS) (Huber, Foster, Riley, & Paterson, 2013; Killgore et al., 2008).

Restriction endonucleases methods (AFLP and PFGE) MLST are limited by being laborious and results cannot easily be compared between laboratories (Huber et al., 2013; Killgore et al., 2008; Sabat et al., 2013). PCR ribotyping, is the main method for studying CDI transmission and outbreaks in the UK (PHE, 2010), Europe, Australia and North America (Clements, Magalhães, Tatem, Paterson, & Riley, 2010; Kachrimanidou & Malisiovas, 2011; Richards et al., 2011). The assay targets the 16S-23S rRNA intergenic spacer region and *C. difficile* has considerable interspecies diversity in the ITS regions (Kachrimanidou & Malisiovas, 2011; Killgore et al., 2008), making it suitable for PCR ribotyping. The idea of making use of the variability of the intergenic 16S-23S rRNA spacer regions in *C. difficile* was first described in 1993 by (Gurtler, 1993). This approach has since been improved with the use of new primers. However, the technology is still faced with drawbacks, including the dependence on biomass produced by culture methods, increasing the turn-around time (currently 10 days to 2 weeks) (Bidet et al., 2000; Morris, Wilson, & Wilcox, 2012). Furthermore, culturing *C. difficile* is difficult because of anaerobic conditions required and sporulation (Marler et al., 1992). The other challenge is the poor resolution of the technology compared to WGS (Huber et al., 2013; M H Wilcox et al., 2012).

Molecular C. difficile diagnosis

Commercial molecular assays for the detection of toxigenic strains include: Illumigene *C. difficile* assay (Meridian Bioscience Inc., Cincinnati, OH), Xpert *C. difficile* (Cepheid), BD GeneOhm Cdiff (BD Diagnostics) and proGastro Cd (Gen-Probe Prodesse, Inc., Waukesha, WI).

The Illumigene *C. difficile* assay uses loop-mediated isothermal amplification technology to detect toxin A gene (*tcd*A) in approximately one hour. The reported range of sensitivity of Illumigene *C. difficile* assay is 91.8% - 95.2% and specificity of 95.3% - 100%(Emmadi et al., 2011; Pancholi, Kelly, Raczkowski, & Balada-Llasat, 2012). Xpert *C. difficile,* BD GeneOhm Cdiff (BD Diagnostics), and proGastro Cd are real-time PCR tests detection of *C. difficile* and the toxin gene *tcd*B (Emmadi et al., 2011). The Xpert *C. difficile* test detects *C. difficile* 027/NAP1/BI strains in approximately 45 minutes with sensitivity and specificity of 93.50% and 94 - 100% respectively (Emmadi et al., 2011). The BD GeneOhm Cdiff assay detects *C. difficile*

and the toxin gene *tcd*B in <2 hours; and has sensitivity and specificity ranges from 93.8- 95.5% and 95.5- 97.9% respectively (Emmadi et al., 2011; Viala et al., 2012). The sensitivity and specificity range of the proGastro Cd test is 77. 3 - 91.7% and 94.7 – 99.2 % respectively (Emmadi et al., 2011; Stamper et al., 2009).

Due to the higher sensitivity of molecular tests and the potential for healthy individuals to be *C. difficile* carriers, testing for CDI is limited to only patients with the symptoms (Emmadi et al., 2011). However, these tests are not used as standalone diagnostics. In the UK, the Department of Health recommends a two test algorithms to diagnose CDI as described above (DH.2012). Furthermore, mutations or polymorphisms in primer or probe binding regions may affect the detection of *C. difficile tcd*A or t*cd*B variants, resulting in false-negative results (Emmadi et al., 2011).

Overall, the clinical utility of culture independent molecular tests is limited by the predefined pathogen detection panels and limited antimicrobial susceptibility information, hence they cannot replace culture (Liesenfeld, Lehman, Hunfeld, & Kost, 2014; Mwaigwisya et al., 2015).

1.4: The Potential for Metagenomics for Diagnosis of Infectious Diseases

Originally the term metagenomics was strictly used for the characterisation of a collection of microbial genomes (sequences) in a multi-organism sample including unculturable organisms obtained by the use of shotgun sequence libraries and second generation sequencing (Mardis, 2008). The term has now been adapted for sequencing without a priori knowledge of sequence targets to identify all constituents (e.g., infectious agent(s) or marker(s) of interest, novel, emerging agent(s), microbiota, human background, and contaminants) in a biological samples (Greninger et al., 2015; Gyarmati et al., 2016). The evolution of sequencing technologies for metagenomics applications in terms of cost and turnaround time has made it feasible to apply this approach to infectious diseases diagnostics. Metagenomics has the potential to address some of the limitations of current diagnostic methods. Availability of rapid, inexpensive and user-friendly genome

sequencing technologies hold the key in replacing the complex multistep culture based procedures for characterising pathogens in clinical samples. Using metagenomics sequencing, all the tests can be combined in one. Furthermore, culture methods can only identify bacteria that can be grown, whereas metagenomics sequencing can identify all the bacteria, dead and non-culturable such as *Francisella tularensis, Bartonella* spp*., Rickettsia* spp*.,* and *Nocardia* spp (Lagier et al., 2015), parasitic, viral, and fungal pathogens (Didelot et al., 2012; Dunne Jr., Westblade, & Ford, 2012; J Wain & Mavrogiorgou, 2013). There are barriers, both perceived and real, to the routine application of metagenomics sequencing for infectious disease diagnostics. Some of the perceived barriers includes prohibitive cost, turnaround time to results (considered to be equivalent to culture), and complicated data analysis. New Next Generation Sequencing (NGS) technologies such as single molecule sequencing technologies and semiconductor sequencing are being designed to address some of the shortcomings of the current technologies. MinION, single molecule sequencing platform (Oxford Nanopore Technologies Ltd, ONT) for example, has unique features of being portable, producing long-reads (1 kilobase – 1 megabase) and providing rapid turnaround (rapid library preparation (15 min – 3 hours) and real-time sequencing and data analysis). The technology has addressed many of the weaknesses of current platforms for clinical diagnosis, including short sequence reads, time consuming library preparation procedures, long sequencing run times (48 to 72 hours) and end point data analysis (Fox, 2014).

The real barrier, in our own and others' experience, is trying to detect bacterial pathogens without culture in samples that contains large numbers of human cells (e.g. in blood) coupled with often very low numbers of pathogen or a sample highly contaminated with normal flora (e.g. stool). Human DNA sequences can be removed using bioinformatics (Cho & Blaser, 2012; Kuczynski et al., 2012), then analysing the remaining reads or alternatively, the reads can be mapped to a reference genome from the hypothesized pathogen (Naccache et al., 2014; Schmieder & Edwards, 2011). These approaches are restricted by the availability of sufficient data to overcome the low proportion of pathogen DNA in a clinical sample. In many sterile samples, such as blood, pathogen cells causing infection can be present at concentrations as low as 1–100 CFU/mL (Lutz Eric Lehmann et al., 2010). In

samples containing host microbiota, such as stool, pathogen cells/DNA can be overwhelmed by human cells/DNA and an enormous amount of normal flora (estimated to be 10^{11} per gram) (Sender, Fuchs, & Milo, 2016). Recovering complete pathogen genome sequences in such samples requires longer sequencing time, hence sequencing cost and the volume of data generated increase. Obtaining sufficient genome coverage for analysis of virulence markers and resistance genes using this approach is challenging. In these situations, to be able to detect pathogen DNA by metagenomics sequencing, sample preparation methods capable of reducing the huge dilution effects of contaminating DNA need to be developed. These include methods for pathogen nucleic acid enrichment or depletion of host DNA (Mwaigwisya et al., 2015).

In this thesis, three specimen types, blood, urine and stool, were chosen as model clinical samples to address different challenges of metagenomic detection of pathogen clinical samples independent of culture:

- Blood – 'Sterile' sample that contains low numbers of pathogen cells coupled with large amounts of human DNA/cells.
- Urine – Contains a high number of bacterial cells and a variable amount of host cells/DNA.
- Stool - Variable pathogen numbers coupled with variable host cell numbers and high commensal floral numbers.

Blood

Using metagenomics to identify pathogens directly from whole blood could enable the identification and genotyping of viable, dead and viable but non-culturable bacteria, fungi, and viruses present as well as detecting antibiotic resistance markers, virulence determinants, and novel pathogens (Didelot et al., 2012; Livermore & Wain, 2013). All the limitations of NAATs could be overcome by developing real-time metagenomics based diagnostics, which would offer non-restricted rapid pathogen detection. Implementation of NGS technologies in clinical microbiology for the diagnosis of BSI is faced with a number of challenges. One challenge is the development of bioinformatics tools that will identify clinically relevant sequence information and communicate the results to microbiologists and

clinicians in a timely fashion (Naccache et al., 2014; Rhoads, Sintchenko, Rauch, & Pantanowitz, 2014). However, perhaps the biggest barrier for the microbiologists is detecting tiny amounts of pathogen nucleic acid in amongst the vast amount of human nucleic acid in the blood sample. The number of leukocytes in one ml blood is estimated to be 10^6-10^7, each cell containing 6.6pg of DNA, whereas there might be 1 to 10 pathogen cells (Kellogg, Manzella, & Bankert, 2000; John Wain et al., 1998; Yagupsky & Nolte, 1990) containing approximately 3 to 6 fg of DNA per cell. Hence, the ratio of human DNA to pathogen is estimated to be as high as 10^9:1 per ml blood. This suggests to get sufficient clinically actionable pathogen sequence from blood would require terabases of sequence data per sample, which clearly is not feasible. The solution is either host DNA depletion – removing as much human DNA (or cells) as possible whilst maintaining sufficient pathogen DNA for library preparation or pathogen DNA enrichment – or separation of pathogen DNA (or cells) from host DNA (or cells) i.e. pathogen DNA enrichment (Mwaigwisya et al., 2015).

In this thesis, we chose the depletion of host DNA rather than pathogen DNA capture. Benefits of depleting human DNA in clinical samples include increased pathogen sequence depth, decreased cost of sequencing, shorten turnaround time, simplified bioinformatics analysis and reduced data volumes (and the subsequent data storage). This has previously been demonstrated in a study by (Oyola et al., 2013), in which host DNA depletion from blood samples (80% reduction), increased the amount of Plasmodium sequences 9-fold (Oyola et al., 2013).

Several molecular kits have been developed aiming to remove human DNA and improve the detection of pathogens in clinical samples. Most of the kits which have been designed for use in blood samples with PCR for identification rather than metagenomics sequencing (Hansen, Bruggeman, & Wolffs, 2009). Commercially available PCR dependent tests include SepsiTest (Molzym GmbH & Co. KG, Bremen, Germany) which uses the MolYsis™ kit to deplete human DNA; and VYOO (SIRS-Lab GmbH, Jena German) which uses SIRS-Lab's Looxster kit to isolate and purify pathogen DNA. The Molzym's MolYsis kit, is the most widely tested human DNA depletion method in blood (Mwaigwisya et al., 2015; Springer et al., 2011; Thoendel et al., 2016). The method is based on the differential lysis of leukocytes and other host cells (without lysing bacteria or fungi) using chaotropic buffer. This is

followed by DNase treatment of the released human DNA using chaotropic-resistant *Mol*DNase B – the DNase is then inactivated and the pathogen DNA extracted and analysed (MolYsis™ Basic 5; Molzym GmbH & Co. KG, Bremen, Germany) (Horz et al., 2008). The method removes about 99% (100 fold) of human DNA in 1mL of blood sample (Loonen et al., 2013).

The LOOXSTER® kit by SIRS-Lab, separates pathogen DNA from the host's by specific binding of prokaryotic DNA using enrichment matrix and separation columns (Mancini et al., 2010; Mwaigwisya et al., 2015). The technology is based on differences in the frequency of methylated and unmethylated CpG-dinucleotides occurring within each DNA species. In separating pathogen DNA, the unmethylated CpG motifs bind to proteins immobilized in enrichment matrix and separation columns and then eluted using elution agent gradient (Osterloh and Felsmann, 2013). The manufacturer claims the kit improves pathogen detection using PCR by removing 95% of human DNA. However, a study by (Glassing et al., 2015) showed the kit reduced the amount of human DNA in the samples by only 40–70%, resulting in a 3.5-fold increase in the number of 16S rRNA bacterial gene sequences detected using the Illumina MiSeq platform (Glassing et al., 2015).

Another commercially available test designed to deplete human DNA is the NEBNext Microbiome DNA Enrichment kit (New England Biolabs, Ipswich, Massachusetts, United States). This kit removes human DNA by capturing eukaryotic DNA at the CpG methylation sites (Methylcytosine in CpG dinucleotides), which are rare in microbial DNA (Feehery et al., 2013). The selective binding and removal of the CpG-methylated host DNA is performed by MBD2-Fc protein immobilized on magnetic beads. The manufacturer claims 90% of human DNA is depleted using this method; similar results have been observed by other researchers (Feehery et al., 2013; Thoendel et al., 2016; Yigit, Hernandez, Trujillo, Dimalanta, & Bailey, 2014).

None of the described methods deplete more than 99% (100 fold) of host DNA from the blood sample. We estimated that the ratio of human to pathogen DNA should be reduced from the approximately $10^{9:}$ 1 to $10^{3:}$ 1 (10^6 fold) to reliably detect pathogens in blood by metagenomics sequencing (requirements would be lower for other clinical sample types). The need for efficient human DNA depletion methods for

detecting pathogens in blood and other clinical samples was the basis of this PhD project.

Urine

Real-time sequencing using MinION technology and real-time data analysis has potential to identify pathogens and resistance genes from clinical urine samples within clinical relevant timeframe to allow early de-escalation and refinement of antimicrobial treatment. We have investigated application of MinION in UTI because of high number of hospital visits per annum and high frequency of resistance to primary antibiotics. Most of the cases are trivial, but, in severe cases, infection may ascend to the kidneys, with overspill to the bloodstream precipitating bacteraemia and urosepsis (Ackermann & Monroe, 1996). Complicated UTIs are a growing cause of hospitalisation, mostly of elderly patients (Ackermann & Monroe, 1996). Increased bacteria resistance to antibiotics, particularly to primary antibiotics in severe and bacteraemic infections, drives the use of previously reserved antibiotics, including 'empirical' therapy (Fair & Tor, 2014). With proliferating bacterial resistance, and few alternative therapies in reserve, rapid diagnostics becomes increasingly important to guide early targeted therapy.

Stool

WGS has the potential to be the method of choice for the investigation of *C. difficile* strains because of accurate and reproducible data with very high discriminatory power. This has been demonstrated in studies that have used Illumina NGS technology to identify various sources of *C. difficile* infection (David W Eyre et al., 2013). However, the method has not been used as a typing method on a regular basis by public health agencies. One of the hurdles for application of WGS is the long turn-around time to results. The current WGS methods depend on challenging culture methods which take up to 48hrs, after that DNA and library preparation and sequencing which can take up to 72 hrs, followed by additional time for data analysis and interpretation. Therefore, the overall turnaround time may take longer than the current 14 days needed for the PCR ribotyping in the UK (PHE 2010).

To address this problem, we have investigated the use of metagenomics sequencing to study *C. difficile* transmission directly from clinical stool samples independent of culture. We used the differences in physical features/properties between *C. difficile* spores, normal flora and human cells to enrich for spores and sequence directly from stool samples. A metagenomics approach to CDI directly from stool (resulting in strain level identification of *C. difficile*) would enable diagnosis, management, infection control and epidemiology all in one test. This would be a dramatic improvement on the current methods.

1.5: Project Hypothesis.

Rapid metagenomics sequencing based characterisation of pathogens and antimicrobial resistance directly from clinical samples can perform comparably to culture but in a dramatically reduced timeframe (from 48h to <8 hours).

1.6: Project Aims.

Current methods of detecting infectious pathogens are culture based and hindered by long turnaround times and poor sensitivity. We propose the use of metagenomics sequencing for the rapid diagnosis of infectious diseases to better guide treatment of patients and enable infection control decisions in a rapid manner. Metagenomics sequencing has the potential to detect pathogens, predict their resistance profiles, detect toxin genes and presence of virulence determinants. To enable metagenomics detection without culture enrichment, methods for depleting human and normal flora DNA need to be developed and evaluated on clinical samples.

1.7: Objectives.

The objectives of this research were to:

- Develop novel methods for depleting human DNA in blood, urine and stool samples to enable pathogen identification by metagenomics sequencing.

- Develop a workflow(s) for metagenomics based diagnosis of BSI, UTI and CDI.

- Evaluate and build capacity of the new Nanopore based real-time sequencing

- Integrate sample preparation methods, library preparation, sequencing and data analysis in rapid metagenomics sequencing workflows to allow unbiased diagnosis of pathogens in blood, urine and stool samples.

Chapter 2: Material and Methods

2.1: Ethical Approval

Ethical approval for collecting blood samples was granted by the Faculty of Medicine and Health Sciences of the University of East Anglia (UEA) and the Norwich and Norfolk University Hospital (NNUH) research and development office (reference number 2012/2013-44HT). This approval allowed us to collect a blood sample for metagenomics analysis whenever a sample was being taken for routine blood culture of suspected septic patient. Healthy volunteer blood was collected for method development after verbal consent was provided, there was no signs or recent history of infection or recent antibiotic treatment.

Retrospectively collected blood samples were frozen at -80°C before being shipped from the University of Manchester on dry ice to the University of East Anglia for processing. These samples were collected as part of a sepsis diagnostics evaluation study conducted by the University of Manchester and funded by the National Institute for Health Research. Ethics approval from North West 6 Research Ethics Committee (reference number 09/ H1003/109) was in place for the retrospective molecular analysis of these blood samples (Warhurst et al., 2015).

For urine and stool samples, the UEA ethical review board approved the use of excess clinical samples which would routinely be discarded (reference number 08/H0305/85), informed consent was not required. Excess stool and urine samples were collected from routine clinical samples submitted to the NNUH Clinical Microbiology Laboratory, no patient information was collected.

2.2: Bacterial Growth Conditions

Bacterial strains *Staphylococcus aureus* NCTC 6571, *Streptococcus pneumoniae* ATCC 49619 and *Escherichia coli* H141480453 were used for spiking experiments. There strains were obtained from the Clinical Microbiology laboratory at the NNUH. All bacterial strains were cultured at 37°C with shaking aerobically at 180 rpm in 10 mL Luria-Bertini (LB) broth (OXOID, Hampshire, England) except for *S. pneumoniae*

where Brain Heart Infusion (BHI) broth (OXOID) was used. The strains were incubated for various time periods (8 - 18 hours).

The non-toxigenic bacterial strain *Clostridioides difficile* NCTC 12726 was used to optimise methods for the DNA extraction and enrichment of *C. difficile* spores in stool samples. *C. difficile* cultures were grown anaerobically on Brazier's CCEY Agar (BioConnections, Knypersley, UK) at 37 °C for 48 hrs. For spore formation, CCEY agar plates were exposed to aerobic conditions in a cabinet for 48 hrs. Thereafter, spores were harvested by scooping using inoculation loop into a PBS (Gibco, ThermoFisher Scientific) and stored at 2 to 8°C for up to three weeks.

Sequencing experiments for optimisation and assessment of MinION™ sequencing technologies (Oxford Nanopore Technologies, Oxford, UK) in sequencing bacterial genomes were performed using *E. coli* strain K-12 substr. MG1655, *Bifidobacterium longum* 8809, *Salmonella* Typhi H58, *Salmonella* Blockey and *Pseudomonas fluorescens* 28E. Culturing of *E. coli* was done by inoculating 10 mL LB broth and then incubating overnight at 37 °C with shaking aerobically at 180 rpm. *P. fluorescens* was cultured by incubating overnight at room temperature (RT) in 10 mL LB broth. *B. longum* was cultured anaerobically in 50 mL of Difco™ Lactobacilli MRS broth and cysteine (50 mg/L) for 48 hrs. Following incubation, broth cultures were centrifuged at 5,000 × g for 5 min in a benchtop centrifuge to collect bacterial cells for DNA extraction as described in section **2.7.** Genomic DNA from *Salmonella* strains (*Salmonella* Blockely and *Salmonella* Typhi H58) was extracted from cultures at the Public Health England (PHE) and shipped to UEA.

2.3: Methods for Host DNA Depletion in Blood

2.3.1: Molzym MolYsis™ Complete5 (MolYsis™)

MolYsis™ technology utilizes the selective lysis of human cells to deplete human DNA and enable bacterial DNA extraction from remaining intact bacterial cells (Horz et al., 2008). The MolYsis™Complete5 DNA extraction kit (Cat number D-321-100; Molzym GmbH & Co. KG, Bremen, Germany) was used for bacterial DNA extraction from whole blood and urine samples. One mL of the whole blood collected in a K2 EDTA tube (BD Vacutainer) was used for depletion experiments.

Following the manufacturer's instructions, the procedure was performed as follows; chaotropic buffer CM (250 µL) was added to 1 mL of clinical sample, vortexed for 10 s, and incubated for 5 min at room temperature to lyse human cells. After incubation, buffer DB1 (250 µL) and *Mol*DNase B (10 µL) were added to the lysate, vortexed for 10 s, and incubated for 15 min at room temperature to degrade nucleic acids released from host cells. Bacterial cells were pelleted by centrifugation at ≥12,000 × g for 10 min and the supernatant discarded. The pellet was resuspended in buffer RS (1 mL), vortexed, and centrifuged at ≥12,000 × g for 5 min and the supernatant discarded to remove any residual chaotropic or DNase activity.

To lyse bacterial cells, the pellet was resuspended in buffer RL (80 µL) and vortexed, followed by adding 20 µL of Buglysis enzyme cocktail and vortexed for 10 s. The sample was incubated in a thermomixer (Eppendorf AG, Hamburg, Germany) at 37°C at 1000 rpm for 30 min to digest bacterial cell walls. For additional bacterial lysis and to denature proteins, 180 µL of MagNA Pure bacterial lysis buffer (Roche Diagnostics, Basel, Switzerland) and 20 µL Proteinase K (Roche Diagnostics) was added, vortexed and incubated in a thermomixer at 65 °C for 5 min. DNA from the lysate was then isolated as described in section **2.6.**

2.3.2: Dynabeads® CD45 Immunomagnetic Separation

Anti-CD45 coated Dynabeads® (ThermoFisher Scientific, Massachusetts, United States) are uniform, superparamagnetic beads (4.5 µm diameter) coated with a primary monoclonal mouse IgG_{2a} antibody specific for a CD45 membrane antigen common to all known isoforms of CD45. This method was used to remove human leukocytes (hence host DNA) by Immunomagenetic separation, binding anti-CD45 coated Dynabeads to the CD45 antigen on the cell membrane of leukocytes.

Briefly, Dynabeads CD45 were resuspended and transferred to a new Eppendorf tube (250 µL per 1 mL blood sample). The beads were then washed by adding 750 µL sterile isolation buffer 1 into an Eppendorf tube containing 250 µL of Dynabeads CD45 and resuspend by slowly inverting the tube six times. Isolation buffer-1 was made in a falcon tube by adding 25 mL PBS, 25 mg bovine serum albumin (BSA) and 100 µL of EDTA (1M, pH 8). The Eppendorf tube containing resuspended beads was placed on a magnetic separation rack for 2 min before discarding the

supernatant. Washed beads were resuspended in 200 µL of Isolation Buffer-1 in a round bottomed tube and placed on a thermomixer (Eppendorf AG, Hamburg, Germany) at 2-8 °C with 600 rpm shaking, 1 mL of blood was then slowly added to the beads and incubated for 30 min with gentle mixing and rotation on a Hulamixer® (Life Technologies, CA, USA). The tube was then placed on the magnetic separation rack for 3 min to separate the bead-bound leukocytes. While the tube was on the magnetic separation rack, the supernatant was transferred to a graduated sterile 2 mL screw cap polypropylene tube. The sample (supernatant) was then centrifuged at 12,000 × g for 10 min, the supernatant was discarded and the pellet resuspended in 1 mL PBS.

2.3.3: Dynabeads® CD45 Immunomagnetic Separation Combined with Molzym MolYsis™ Complete5 (MolYsis™).

The two methods were combined with the aim of achieving higher levels of depletion of human DNA. CD45 imunomagnetic separation (IMS) was first performed as described in section **2.3.2.** Thereafter, residual blood cells (and pathogens) were collected by centrifuging at 12000 × g for 10 min. The supernatant, plasma and isolation buffer 1 (added during IMS procedure), was discarded and cells were resuspended in PBS to make up total volume of 1mL. Thereafter, chaotropic buffer CM (Molzym) for lysing host cells and the rest of MolYsis procedure was followed as described in section **2.3.1**.

2.3.3.1: Limit of Detection of CD45 IMS combined with MolYsis

The LoD of human DNA depletion method (CD45 IMS combined with MolYsis) was determined repeatedly (three times) by processing the blood samples spiked with known dilution of *S. aureus* and *E. coli* (ranging from 10 to 100 CFU), as described below.

Bacterial strains *S. aureus* NTCTC 6571 for Gram positive and *E. coli* H141480453 for Gram negative bacteria were used for spiking blood samples. Bacterial cultures were grown as described in section **2.2**, except incubation time was 6 hours for *E. coli* and 18 hours for *S. aureus*. Serial dilutions in estimated number in CFU were made in sterile PBS and plated on nutrient agar plate (in triplicate) for counting. The

same dilutions (in µL) were spiked in 1 mL of blood samples (from healthy volunteers) to make the estimated final concentration of 10 CFU/ mL, 50 CFU/ mL and 100 CFU/ mL. Human DNA in spiked samples were thereafter depleted as described in **2.3.3.** After pathogen DNA enrichment and DNA extraction, samples were analysed by real-time PCR (**2.9**) and by metagenomic using MiSeq platform (Illumina) (**2.11.2**).

2.3.4: A novel Human DNA Depletion Method using Cytolysins for Differential Lysis of Leukocytes

2.3.4.1: Endonucleases

To digest released human DNA directly in blood samples, different DNase enzymes were tested:

- HL-SAN DNase at 25 U/µl (heat labile, salt active nuclease) (ArcticZymes, Tromsø, Norway);
- HL-double stranded DNase Enzymes at 5 U/µl (ArcticZymes, Tromsø, Norway); and
- Nuclease micrococcal from *S. aureus* at 10 U/µL (Sigma Aldrich, Dorset, UK).

Endonuclease activity was tested by adding equal amount of each enzymes (2 µL) to the 200 µL blood sample with lysed human cells and incubated at 37 °C with and without shaking on a Thermomixer (Eppendorf) for 15 -30 min. Human cells were lysed by treating blood sample through three cycles of freeze at -80ºC for about 10 minutes and thawing at 37 ºC. After incubation, DNA was extracted as explained in section **2.6.1** from the sample and an undigested control and was quantified by Qubit® dsDNA broad range assay (Thermo Fisher Scientific. Massachusetts, United States) and qPCR as described in sections **2.8.3 and 2.9.1** respectively.

2.3.4.2: Bacterial Cytolysins

All cytolysins were supplied by Sigma Aldrich (Dorset, UK). The cytolysins were suspended in water at the following concentrations:

- Phospholipase C (PLC) from *Bacillus cereus* at 40 µg/µl;

- PLC from *Clostridium perfringens* at 40 μg/μl;
- Streptolysin O from *Streptococcus pyogenes* at 2 mg/μL; and
- Alpha-hemolysin from *Staphylococcus aureus* at 82 units/mg/μL.

To lyse human cells, cytolysins were added directly to blood samples spiked with bacteria (as described in section ***2.2***), then incubated for 15 min at 37 °C with shaking at 500 rpm using a Thermomixer (Eppendorf). After host cell lysis, samples were treated with an endonuclease for digestion of the released DNA (as described in section **2.4.2**).

2.3.4.3: A novel PLC Based Method for Human Nucleic Acid Depletion

PLC (4 mg - from *Clostridium perfringens*) was reconstituted in 100 µl of molecular grade water (40 μg/μl) and 1 mL blood samples were spiked with *E. coli* and *S. aureus* overnight cultures. The PLC was added to the blood samples (1 mL blood, 100 µl of PLC) and incubated at 37 °C for 20 min with slow mixing using a Hulamixer®. After incubation, 500 µl of HL-SAN buffer (3M NaCl, 100mM $MgCl_2$, pH 8.5) was added followed by 10 µl of HL-SAN DNase, mixed briefly by vortexing then incubated at 37 °C for 15 min. The samples were centrifuged for 10 min at 12,000 x g, the supernatant was carefully discarded and the pellet was washed by resuspended in 1.5 mL PBS. Bacterial cells were then collected by centrifuging at 12000 × g for 5 min. The bacterial cell pellet was then used for DNA extraction as described in **2.6.1** followed by DNA quality control assays (section **2.9**).

2.3.5: Host DNA Modification by using Propidium monoazide (PMA™)

To lyse human cells in 1 mL of spiked blood, 250 µl chaotropic buffer CM (from the MolYsis kit) was added and incubated for 15 min at room temperature as recommended. Thereafter, bacterial cells were collected by centrifugation at top speed (≥12,000 × g) for 10 min. The pellet was washed with saline, then resuspended in 200 µl of saline (PBS was avoided because of interaction with PMA). To the resuspended sample/lysate, the released DNA was degraded by adding 10 µl of 1.25 mM PMA followed photo-activation by 15 min incubation on PMA-Lite™

LED Photolysis device (Biotium Inc). This was followed by extraction of bacteria DNA and PCR of human and bacterial DNA, sections **2.6.1 and 2.9.**

2.4: Methods for Human DNA Depletion in Urine

2.4.1: NEBNext® Microbiome DNA Enrichment Kit

The NEBNext Microbiome DNA Enrichment Kit (New England Biolabs, MA, USA) was used to enrich for pathogen DNA in urine samples, following the manufacturer's instructions. First, 2 mL of the urine sample was centrifuged at 12000 ×g for 5 min using a benchtop centrifuge (Stuart, Staffordshire, UK), supernatant was discarded and total DNA from the pellet was extracted as explained in section **2.6.2**. Thereafter human DNA was separated from the pathogen DNA as explained in the steps below.

Binding of MBD-Fc Protein to Protein A Paramagnetic Beads

Protein A paramagnetic beads were uniformly suspended in bind/wash buffer by gentle pipetting. One mL of the suspension was transferred to an Eppendorf tube, and 160 µL of MBD-Fc protein solution was added. The paramagnetic beads and MBD-Fc protein mixture was gently rotated for 10 min at room temperature. The tube was placed on a magnetic separation rack until the supernatant was clear and beads were collected on the wall of the tube. The supernatant was removed and discarded using a pipette without disturbing the beads. The beads were then washed three times by removing the tube and resuspending in 1 mL of 1 × wash/bind buffer. After the final wash, beads were resuspended in 1 mL of ice cold 1 × wash/bind buffer and kept in the fridge for no more than seven days.

Enrichment of Microbial DNA

The total DNA from the sample (prepared as described above) was mixed with MBD-Fc protein A beads in a ratio of 1 µg of sample DNA to 160 µL of beads. The sample DNA was directly added to the bead slurry and incubated for 15 min at room temperature with gentle rotation. The incubated mixture was placed on a magnetic separation rack until the supernatant was clear and beads were collected on the wall of the tube (2–5 min). The supernatant, containing enriched microbial DNA, was

carefully removed with a pipette without disturbing the beads then washed using 1.8 × volume of Agencourt AMPure XP beads (Beckman Coulter) following manufacturer's protocol. The washed DNA was eluted in 50 µL of 10 mM Tris-HCl (pH8.5).

2.4.2: Molzym MolYsis™ Complete5 (MolYsis™)

Pathogen DNA enrichment using MolYsis™ basic 5 kit was performed on urine samples according to the manufacturer's instructions as detailed in section **2.3.2.** Urine samples collected from NNUH clinical microbiology laboratory were aliquoted (2 mL) into 2 mL Eppendorf tubes and centrifuged at 13,000 × g for 5 min using a benchtop centrifuge (Stuart). The collected cells were resupended in 1 mL of PBS, this was followed by depletion of human DNA using MolYsis, following manufacturer's instructions. Thereafter DNA was extracted as explained in section **2.6.2** followed by qPCR (section **2.9**).

2.4.3: Differential Centrifugation Combined with MolYsis Basic 5 kit

In this method, host cells and pathogens were separated based on the difference in size and density by differential centrifugation. Differential centrifugation was combined with the MolYsis assay to deplete human DNA in 2 mL of urine. The combination of methods was performed as follows; human cells were pelleted from 2 mL urine samples by differential centrifugation at relative centrifugal force (RCF) of 300 ×g for 2 min on benchtop mini centrifuge (Stuart) leaving most of the bacteria in suspension. The supernatant was removed by pipetting into a new Eppendorf tube and the pellet which contained mainly leukocytes was discarded. All the cells in the supernatant were collected by centrifugation at ≥12,000 × g for 10min. The supernatant was discarded and the pellet which contained the remaining leukocytes and pathogens was resuspended in 1 mL of PBS and host DNA was depleted by MolYsis method as described in **2.3.1.** Thereafter DNA was extracted as explained in section **2.6.2** (Schmidt et al., 2017), followed by qPCR as described in **2.9.**

2.5: Methods for Host DNA Depletion in Stool

2.5.1: Microbiota Separation by Nycodenz Medium

Bacterial cells from stool samples were separated as follows: Nycodenz (Sigma Aldrich, Dorset, UK) was reconstituted by dissolving 1.3 g/mL in 10 mM Tris-HCl (pH 8), 1 mM EDTA to make equivalent to 56% (w/v) solution. Stool sample was homogenised by vigorously vortexing 1 g of stool sample in 1.5 mL PBS. Large particles were pelleted by centrifuging the homogenised sample at 300 × g for 1 min. Supernatant was transferred to a new sterile tube and pelleted large particles were rinsed three times with a single volume of 500 µl PBS and pooled with recovered supernatant (total of 2 mL). To separate bacterial cells from faeces, 1 mL aliquots of the faecal slurry were layered onto 650 µl of Nycodenz (1.3 g/mL in 10 mM Tris-HCl (pH 8), 1 mM EDTA), in a sterile 2 mL screw top micro centrifuge tube and centrifuged at 16,000 × g for 6 min at 4 °C. The upper aqueous layer was discarded and the microbiota layer (the whitish layer between the soluble layer and insoluble faecal debris) collected, as shown in **Figure 2-1** (Hevia, Delgado, Margolles, & Sanchez, 2015). The collected microbiota cells were washed in 500 µl PBS and centrifuged at 16,000 × g for 6 min.

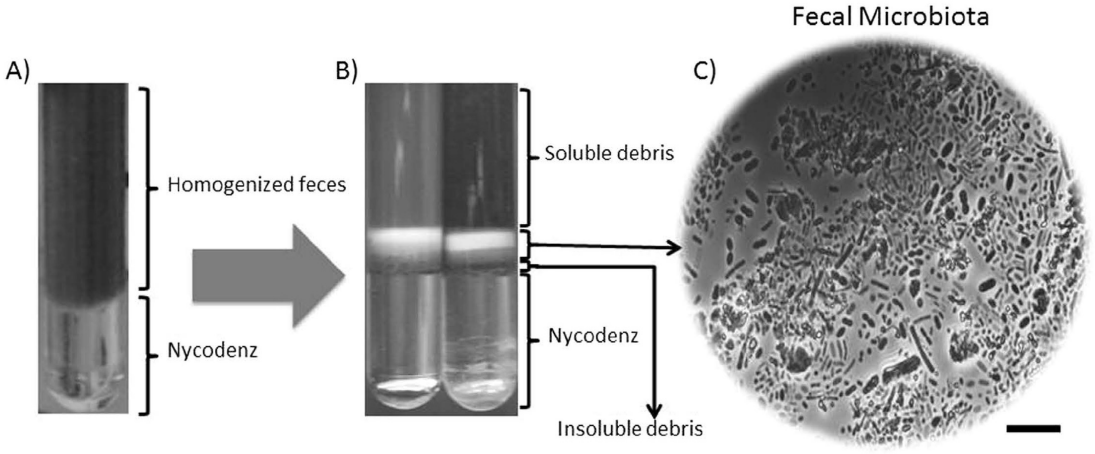


Figure 2- 1 Figure 2-1 (A) before centrifugation and (B) after centrifugation. After centrifugation four layers are formed. The layer corresponding to the faecal microbiota is in between two layers containing soluble (upper) and insoluble (lower) faecal debris, all above the Nycodenz medium (bottom) (Hevia et al., 2015).

The collected cells were depleted of human DNA using the MolYsis kit, microbial DNA was extracted as explained in section **2.6.1** for further analysis by NGS using Illumina MiSeq, section **2.11.2**.

2.5.2: C. difficile Spore Separation by Gastrografin Medium

Preliminary experiments were performed to separate spores from vegetative cells using Gastrografin® medium (Bayer AG, Leverkusen, German). The method was optimised as follows: 1 mL of overnight bacterial cultures, *E. coli* and *S. aureus*, were centrifuged to collect cells. Bacterial pellets were mixed and stained with Crystal violet, stained cells were washed three times with PBS to remove residual dyes whilst changing the Eppendorf tubes. Stained and washed cells were then mixed with a prepared suspension of *C. difficile* spores section **2.2**. The mixed suspension of cells was layered onto a 50% (v/v) Gastrografin solution in an Eppendorf tube and centrifuged at 15,000 × g at 4 ºC for an hour to separate the spores from vegetative cells. After centrifugation, spores (the whitish layer) were collected at the bottom of centrifuge tube whereas vegetative cells (crystal violet) remained at the top (**Figure 2-2**).

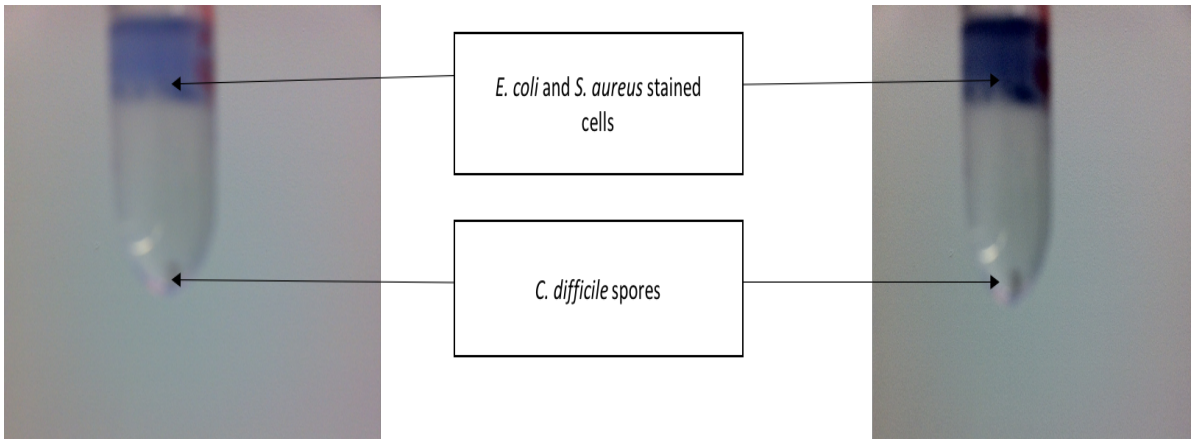


Figure 2- 2 Eppendorf tube showing Gastrografin spore purification, spore as seen settling at the bottom and vegetative cells remained at the top

Clinical samples were processed as follows: about 200 mg of stool was homogenised by vigorous vortexing and, whenever necessary, sterile PBS was

added to liquefy the sample. Good vortexing was necessary to release bacteria and spores into suspension and achieve higher DNA yields. Large particles were separated by pulse centrifuging at no more than 300 × g.

The 200 µl of homogenized stool sample was carefully layered onto the 1 mL of Gastrografin 50% (v/v) in the centrifuge tube. Differential centrifugation was carried out at 16,000 × g (Eppendorf 5424) at 4 ºC for 1 hour. For quantification analysis, the top fraction assumed to contain mostly vegetative and human cells was separated to a new tube, the aliquot was pelleted 12,000 × g for 5 minutes and washed using 1 mL of PBS. The middle fraction which mainly contained Gastrografin medium was discarded. The pellet (which was expected to contain most of the C. difficile spores) was washed by resuspending in 1 mL of PBS and cells pelleted at top speed for 10 min. DNA from the separated pellet and upper layer was extracted as described in section **2.6.3** for further analysis by qPCR.

2.5.3: C. difficile Spores Enrichment by Chemical Differential lysis

Approximately 200 mg of stool sample was placed in a screw cap tube, 1 mL of PBS was added then homogenised vigorously to release all the cells from the solid waste. The tube was left to stand for 30 s to allow large particles to settle. Thereafter, the supernatant was transferred to a new screw cap tube and centrifuged at 15,000 ×g (Eppendorf 5424) for 5 min and the supernatant was discarded. To lyse vegetative and human cells, the pellet was resuspended in buffer FL from the GeneAll Exgene Stool gDNA purification kit (Seoul, South Korea) by pipetting and thereafter incubated for 5 min at room temperature. The samples were then centrifuged for 5 min at 15,000 × g to collect the remaining cells, supernatant was discarded and the pellet was washed twice in 1 mL PBS and pelleting at 15,000 × g.

To digest DNA from the lysed cells, the washed pellet was resuspended in 1 mL of PBS followed by adding 250 µL of buffer DB1 (Molzym) then 10 µL of *Mol*DNase B (Molzym), mixed well and incubated for 15 min at room temperature. After incubation, remaining cells were pelleted by centrifuging at 15,000 × g and then washed in 1.5 mL PBS and pelleted again at 15,000 × g. The supernatant was

discarded and pellet enriched for *C. difficile* spores was used for DNA extraction as explained in section **2.6.2.**

2.6: Bacterial DNA Extraction from Collected Samples

Extraction, isolation and purification of bacterial DNA from all clinical samples was performed on the automated MagNA Pure Compact System (Roche, Basel, Switzerland). The MagNA Pure Compact System is an automated benchtop instrument that uses ready-to-use pre-filled cartridges to extract, isolate and purify DNA. The MagNA Pure cartridges are filled with chaotropic lysis buffers, binding buffers and magnetic glass particles (https://shop.roche.com/shop/products/magna-pure-compact-instrument). Bacterial DNA from all samples was purified using the MagNA Pure Compact Nucleic Acid Isolation Kit 1 cartridges (Roche) and the DNA Bacteria v3_2 protocol. Prior to loading samples onto the MagNa Pure, dependent on the sample type there were a number of pre-lysis steps performed (as detailed in sections **2.6.1** to **2.6.3**).

2.6.1: Blood Samples

DNA was extracted from spiked blood of healthy volunteers, after human cell/DNA depletion, as follows: bacterial cell pellets depleted of human DNA were resuspended in 380 µl of MagNA Pure bacterial lysis buffer (Roche Diagnostics) followed by 20 µl Proteinase K (Roche Diagnostics). For all positive control samples (200 µl of spiked blood which was not depleted), 180 µl of MagNA Pure bacterial lysis buffer was added followed by 20 µl of Proteinase K. After mixing by vortexing, all the samples were incubated for 5 min at 65 °C followed by MagNA Pure DNA extraction.

Clinical blood samples for the studies were processed through the established pathogen DNA enrichment workflow to remove human DNA (sections **2.3.1** and **2.3.2)**. After human DNA depletion, pathogens were lysed as per MolYsis kit instructions. Briefly, to the processed sample, 80 µl of buffer RS (Molzym) and 20 µl of Buglysis enzyme cocktail (Molzym) were added, mixed and then incubated for 30 min at 37°C with shaking at 1000 rpm on a Thermomixer (Eppendorf). Afterwards, 280 µl of MagNA Pure bacterial lysis buffer (Roche) and 20 µl of proteinase K

(Roche) were added, incubated for 5 min at 65 °C before loading to the MagNA Pure.

For positive controls from clinical samples, 200 µl of blood sample was aliquoted and treated as follows: 80 µl of MagNA Pure bacterial lysis buffer was added followed by 80 µl of buffer RS and 20 µl of Buglysis cocktail. The mixture was incubated for 30 min at 37 °C with shaking at 1000 rpm on a Thermomixer. Thereafter, 20µl of proteinase K was added and incubated for 5 min at 65 °C then loaded on the MagNA Pure.

2.6.2: Urine Samples

This method was mainly used for DNA extraction from bacterial pallet from urine samples but it was also tested for stool samples (section **2.5.1**). Human DNA was depleted from urine samples (section **2.4** to **2.6**) and DNA extracted was extracted from the bacterial pellet. Cells from unenriched samples (2 mL) were pelleted at 12,000 × g for 5 min, supernatant was discarded. The pellets (both enriched and unenriched) were resuspended in 180 µl of buffer RS (MolYsis™) and 20 µl of lysis enzymes (MolYsis™). The mixture was incubated for 30 min at 37 °C with shaking at 1000 rpm. Afterwards, 20 µl of proteinase K (Roche) and 200 µL bacterial lysis buffer (Roche) were added and incubated for 5 min at 65 °C then loaded on the MagNA Pure for DNA extraction.

2.6.3: Stool Samples

To make DNA extraction from spores easier, spores were stimulated to germinate by adding 200 µl of germination matrix to the *C. difficile* spore enriched samples and positive controls and then incubated for 1 hr at 37 °C in a Thermomixer. Germination matrix was made by adding 59.6 g/L Thioglycollate medium (Sigma) and 7.4 g/L Sodium taurocholate (sigma) to the100 µL distilled water, followed by autoclaving for 15 min at 121 °C.

After incubation, bacterial cells were lysed by adding 400 µl of MagNA Pure bacterial lysis buffer, followed by 50 µl of lysozyme (100 mg/mL, Sigma) and 5 µl of RNase A (100 mg/mL, QIAGEN) then incubated for 1 hr at 37 °C with shaking at 500 rpm in a

Thermomixer. After incubation, 50 µl of Proteinase K was added then another incubation at 65 °C for a another 1 hour with shaking at 500 rpm in a Thermomixer or alternatively overnight (without shaking). The gDNA from the lysed cells were isolated and purified using MagNA Pure.

2.7: DNA extraction from Bacterial Cultures for MinION Sequencing

Genomic DNA from bacterial cultures for MinION sequencing was extracted manually using the QIAGEN Genomic-Tip kit, with alterations depending on the bacterial species as explained below.

2.7.1: DNA Extraction from E. coli Cultures

E. coli DNA was extracted using the QIAGEN Genomic-tip 20/G (QIAGEN) as follows: *E. coli* cells were pelleted by centrifuging 1.5 mL of overnight culture at 5,000 × g for 10 min. The bacterial cell pellet was resuspended in 1 mL of buffer B1 containing 2 µl of RNase A solution (100 mg/mL, QIAGEN), thereafter, 20 µl of lysozyme stock solution (100 mg/mL, Sigma) and 45 µl of Proteinase K (Roche) were added then incubated for 30 min at 37 °C. After incubation, 350 µl of buffer B2 was added and mixed by inverting the tube several times then further incubation at 50 °C for 30 min. The Genomic tip was placed over the tube using a tip holder. The tip was equilibrated using 1 mL of buffer QBT and left to drain by gravity. The lysed sample was mixed by vortexing for 10 s at maximum speed and loaded onto the equilibrated tip and again allowed to drain by gravity. The tip was washed three times with 1 mL of buffer QC and allowed to flow by gravity.

Genomic DNA was eluted with 2 × 1 mL of QF elution buffer over a 10 mL collection tube. Precipitation was done by adding 1.4 mL of room temperature isopropanol and immediately inverting the tube several times until the DNA precipitate was visible. Genomic DNA was isolated by spooling into a new Eppendorf tube containing 100 µl of 10 mM TrisHCl, pH 8.5. The DNA was dissolved overnight by shaking at 500 rpm on a Thermomixer.

2.7.2: DNA Extraction from B. longum Cultures

Bacterial DNA from *B. longum* 8809 was extracted using QIAGEN Genomic-tip 100/G (QIAGEN, Hilden, Germany). After incubation at 37 °C for 48 hrs in anaerobic conditions, the bacterial cells were pelleted by centrifuging 20 mL of culture at 3,000 × g for 10 min. The bacterial cell pellet was resuspended in 3.5 mL of buffer B1 containing 14 µl of RNase A solution (100 mg/mL, QIAGEN), thereafter, 160 µl of lysozyme stock solution (100 mg/mL, Sigma) and 200 µl of Proteinase K (Roche) were added then incubated for 8 hrs at 37 °C. After incubation, 1.2 mL of buffer B2 was added and mixed by inverting the tube several times before incubating at 50 °C overnight. The genomic tip was placed over the tube using a tip holder. The tip was equilibrated using 4 mL of buffer QBT and left to drain by gravity. The lysed sample was mixed by vortexing for 10 s at maximum speed and loaded onto the equilibrated tip and allowed to drain by gravity. The tip was washed three times with 7.5 mL buffer QC, allowed to flow by gravity.

Genomic DNA was eluted with 2 × 5 mL of QF elution buffer over a 10 mL collection tube. DNA was precipitated by adding 3.5 mL of room temperature isopropanol and inverting the tube several times until the DNA precipitate was visible. Genomic DNA was isolated by spooling/centrifugation and resuspending the DNA in 50 to 500 µl of 10 mM TrisHCl, pH 8.5. DNA was dissolved overnight by shaking at 500 rpm on a Thermomixer (Eppendorf).

2.8: DNA Quality Control

After DNA extraction or Whole Genome Amplification (WGA) and before sequencing library preparation, the quality and quantity of DNA was checked as follows.

2.8.1: DNA Purity

The purity of DNA was measured by absorbance using the NanoDrop 2000c (Thermo Scientific) UV spectrophotometer. DNA has a maximal absorbance near 260 nm, therefore when this wavelength is passed through the sample, the level of absorbance indicates the amount of DNA present in the sample. Absorbance at 280 nm was measured to determine the level of protein and phenol present in the

sample. The ratio of absorbance at 260 nm to 280 nm was used to assess the purity of DNA, a ratio of approximately 1.8 was generally accepted as pure DNA and if lower indicated the presence of protein, phenol or other contaminants that absorbed strongly at 280 nm. Also, for pure nucleic acids, 260/230 values were expected to range between 2.0 and 2.2. A lower ratio indicated the presence of contaminants which absorbed at 230 nm.

2.8.2: DNA Integrity and Fragment Size Analysis

For the purposes of sequencing (using MinION and Illumina platforms), the DNA integrity and average fragment size was assessed using the 2200 TapeStation instrument (Agilent Technologies). The system estimates the fragment size of DNA based on electrophoresis and digital analysis. Results were presented as images, estimated values of fragments in Kilobase pairs and quality of the sample in DNA Integrity Number (DIN). DNA analysis was performed according to the genomic DNA ScreenTape quick guide, in brief; 1 µL of DNA was mixed with 10 µL of sample buffer (Agilent Technologies). The DNA ladder (1 µL) (Agilent Technologies) was mixed with 10 µL of sample buffer then placed in the first tube of an 8-tube strip followed by the samples. The prepared strip was vortexed at high speed for a few seconds, spun down and placed in the TapeStation instrument together with DNA screening tape (Agilent Technologies) for analysis.

2.8.3: DNA Quantification

In this project, DNA was quantified using the Qubit® dsDNA broad range and the high sensitivity assays (ThermoFisher Scientific). Briefly, a working solution was made by preparing a 1:200 dilution of the Qubit reagents. Two standards were then made by adding 10 µL of standard to 190 µL of working solution, and samples were prepared by adding 2 µL of sample to 198 µL of working solution. The samples were vortexed and incubated for two minutes at room temperature in the dark. The DNA concentration was then measured on the Qubit® 3.0 Fluorometer (ThermoFisher Scientific).

2.8.4: DNA Purification

Whenever necessary, DNA was purified using AMPure XP beads (Beckman Coulter, Brea, CA) at various concentrations dependent on the required DNA fragment lengths (0.4 × for long fragments, 1× for general purification and 1.8 × for amplified PCR products). AMPure XP beads were added to DNA samples and incubated at room temperature for 5 min, before being placed on a magnet separation rack. The supernatant was discarded, the bead pellet was washed twice with freshly made 80 % ethanol for a minimum of 30s followed by 5-10 min room temperature -drying or 1 min at 37 °C. The bead pellet was then resuspended in 10 mM Tris-HCl pH8.5 and incubated at room temperature for 5 min to release DNA before being placed back on the magnetic separation rack. Washed DNA was stored at 4 °C.

2.9: PCR Assays

All the PCR assays were performed on the Light Cycler 480 (Roche).

2.9.1: Hydrolysis Probe based qPCR Assays

Probe based qPCR was performed on samples to quantify human and bacterial (*E. coli*, *S. aureus, S. pneumoniae* and *C. difficile*) DNA. The qPCR master mix consisted of 10 µl LightCycler 480 probe master (2 ×), 0.5 µl of each reverse and forward primer (20 µM), 0.4 µl probe (10 µM) and 3.6 µl PCR grade water. To this master mix, 5 µl of DNA template was added to make a total volume of 20 µl. The qPCR conditions were as follows: [95 °C for 5 min], 45 cycles [95 °C for 10 s, 60 °C 30 s], and a final hold at 40 °C. Details of primer sequences can be found in **Table 2-1.**

2.9.2: SYBR Green qPCR Assay

The 16S rRNA V3-V4 region SYBR green qPCR assay was set up using primers with Illumina adapters (Fadrosh et al., 2014). The assay was performed on all clinical samples with unknown pathogens. The master mix consisted of 10 µl of LightCycler 480 SYBR green I master (2 ×), 2.5 µl each of forward and reverse primers (20 µM). To the master mix, 5 µl of template DNA was added to make a total volume of 20 µl

of qPCR reaction. qPCR conditions were as follows: [95 °C for 5 min pre-incubation], followed by 40 cycles [95 °C for 30 s, 55 °C 30 s and 72 °C 30 s] and a final hold at

Table 2- 1 Details of primers and probes used in this thesis

Primer	**Sequence (5' to 3')**	**Gene target**
Human forward primer	TGAAGCCGTGCGGAAGG	RNA polymerase II
Human reverse primer	ACAAGAGAGCCAAGTGTCG	
Human probe	TACCACGTCATCTCCTTTGATGGCTCCTAT	
S. aureus forward primer	ACTGTAACTTTGGCACTGG	*eap*
S. aureus reverse primer	GCAGATACCTCATTACCTGC	
S. aureus probe	ATCGCAACGACTGGCGCTA	
E. coli forward primer	CGATAATCGCCAGATGGC	*cyaA*
E. coli reverse primer	CCTAAGTTGCAGGAGATGG	
E. coli probe	TAGAGCGCCTTCGGTGTCGGT	
S. pneumoniae forward primer	TGCAGAGCGTCCTTTGGTCTAT	*ply*
S. pneumoniae reverse primer	CTCTTACTCGTGGTTTCCAACTTGA	
S. pneumoniae probe	TGGCGCCCATAAGCAACACTCGAA	
16s rRNA forward primer	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGC CTACGGGNGGCWGCAG	V3-V4 region
16s rRNA reverse primer	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACA GGACTACHVGGGTATCTAATCC	

2.10: Whole Genome Amplification

Isothermal Whole Genome Amplification (WGA) was performed by Multiple Displacement Amplification (MDA) using Phi 29 polymerase. WGA was carried out on blood samples depleted of human DNA using REPLI-g® single cell kit (QIAGEN GmbH) and/or Illustra™ Single cell GenomiPhi™ kit (GE Healthcare).

2.10.1. Repli-g Single Cell Kit (Qiagen)

Using the Repli-g Single Cell Kit, 4 µl of DNA was denatured by adding 3 µl of buffer D2 then 10 min incubation at 65 °C. Afterwards, 3 µl of stop solution was added to the denatured DNA and stored on ice. The master mix was prepared in a total volume of 40 µL, which consisted of 29 µL REPLI-g SC Reaction buffer, 2 µl of Phi 29 DNA polymerase and 9 µL of nuclease free water. The denatured DNA was then added to the master mix and incubated on a thermocycler (Applied Biosystems) at 30 °C for 3-8 hrs then enzyme inactivation performed at 65 °C for 3 min and held at 4 °C.

Alternative protocol of increased template volume was also used. WGA using increased volume protocol was performed under the same conditions as above but in larger volume. The template volume was 15 µL, denatured using 2 µL Buffer DLB for 3 minutes at room temperature, and the reaction was stopped by 3 µL Stop Solution. The master constituted of 29 µL REPLI-g SC Reaction buffer and 2 µl of Phi 29 DNA polymerase. Denatured DNA was added to the master mix and incubated and inactivated as described above.

2.10.2: Illustra™ Single cell GenomiPhi™ DNA Amplification Kit (GE Healthcare)

When using Illustra Single Cell GenomiPhi DNA amplification kit, 1 µl of DNA was denatured using 1 µl of lysis buffer followed by incubation at 65 °C for 10 min. Denaturation was stopped by adding 1 µl of neutralization buffer and denatured DNA was stored on ice. The master mix was prepared by adding 11 µL of reaction buffer, 1 µl of Phi 29 DNA polymerase, 1 µl of amplification mix and 4 µl sterile water. The denatured DNA was then added and mixed to the master mix to make a 20 µl total

reaction volume. The reaction mix was incubated on a thermocycler (Applied Biosystems) at 30 °C for 2 hrs and then enzyme inactivation at 65 °C for 3 min and held at 4 °C.

2.11: Sequencing

2.11.1: Sanger Sequencing

All 16S rRNA gene PCR products were sent to Source Bioscience Genomics Service (Cambridge, UK) for Sanger sequencing with the forward and reverse primers. Sequence chromatograms were trimmed and identified using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch).

2.11.2: Illumina - MiSeq

DNA (1ng) quantified by Qubit HS kit (section **2.9.3**) was converted into Nextera XT libraries according to the library preparation guide (Illumina). In brief, the sequencing library was prepared by random tagmentantion using Nextera transposome. Indexes were then ligated to the tagmented DNA and amplified. The amplified library was washed, quantified by Qubit HS assay and tapestation analysis (Agencourt Technologies) before being normalized to 4 nM and then pooled in equimolar amounts and denatured using 0.1 M NaOH. 12pM of the denatured library was loaded to the cartridge for sequencing**.** Sequencing was performed on the Illumina MiSeq platform using V3 2 × 300bp chemistry.

2.11.3: MinION Library Preparation Using Sequencing Kits

Five sequencing kits, SQK MAP-002 -006, were used to generate library for MinION sequencing. There was either minor or major variations in the protocols between the kits, but generally, library for each kit was generated by following the following steps: genomic DNA (1 µg – 1.5 µg) was sheared for 8 kb fragments in a G-tube (Covaris, Brighton, UK) by spinning at 6000rpm (Eppendorf 5424) for 1 minute before inverting the tube and centrifuging again with the same conditions. Whenever necessary, the sample was pulse-centrifuged at 4,000 rpm if DNA had not completely passed through the ruby orifice. Fragmented DNA was repaired by either NEBNext FFPE

DNA Repair Mix (NEB) or PreCR treatment (NEB) depending on the protocol. The ends of repaired DNA was tailed with nucleotides Ts (end repair) and As (dA-tailed) or both simultaneously (end prep), followed by ligation of adapters and tethers. Adapted DNA was then purified by magnetic beads for loading in the flow cells and sequencing. Experiments variations has been highlighted in **Table 2-2** and explained in details in **Chapter 4**. In this thesis, the kits were used as follows: (i) Sequencing kit SQK-MAP-002 was used to sequence blood sample P4, urine samples, *Salmonella* typhi H58 (Ashton et al., 2015), and *Bifidobacterium longum* 8809 following ONT procedure as described by (Ashton et al., 2015); (ii) Kit SQK-MAP- 003 was used to sequence urine samples and *Pseudomonas fluorescens* following the protocol as described by (J Quick, Quinlan, & Loman, 2014); (iii) Kit SQK-MAP-004 following ONT protocol as described by (Urban et al., 2015) was used to sequence urine samples, *P. fluorescens* and *B. longum* 8809 genomes; (iv) Kit SQK- MAP-005 was used to sequence urine samples and *E. coli* strain K-12 substrain MG1655 (Ip et al., 2015); (V) Sequencing kit SQK- MAP-006 was used to sequence *B. longum* 8809 and *Salmonella* Blockley (Nair et al., 2016).

Table 2-2 shows library preparation steps using SQK-MAP-002-006. To generate library using SQK-MAP-002-005, genomic DNA (1 µg-1.5µg) was fragmented to the average size of 8 kb using G-tube and Eppendorf 5424 centrifuge at 6000 rpm for 1 minute in both directions. Fragmented DNA in 80 µL nuclease free water was washed with 1 × by volume Agencourt AMPure XP Beads (Beckman Coulter, High Wycombe, UK). The beads were incubated for 5 minutes to allow DNA binding followed by supernatant aspirations using magnetic rack (Invitrogen MagnaRack) washing twice with 200 µL of freshly prepared 80% ethanol while still on the magnet, and then eluted into 80 µL of 10 mM Tris-HCl pH 8.5.

The optimal step of DNA repair was performed by PreCR mix (NEB) (SQK-MAP-002-005) or NEBNext FFPE DNA Repair Mix (NEB) for SQK-MAP006 following manufacturer's protocol. An end-repair step was then performed using the NEBNext End Repair Module (NEB) according to manufacturer's instruction, the resulting blunt-ended DNA (100 µL) was cleaned using 1 × Agencourt AMPure XP Beads as described above, and eluted in 26 µL of 10 mM Tris-HCl pH 8.5.

The end repaired DNA was then dA tailed using the NEBNext dA-Tailing Module (NEB), following manufacturer protocol. Using SQK MAP006, end repair and dA tail was performed simultaneously using NEBNext Ultra II End-repair/dA-tailing Module (NEB). Adapters and Tethers provided by ONT were used to generate a MinION sequencing library. Briefly, to the adenylated DNA, adapters and tethers were added then ligated by Blunt/TA ligase master mix (T4 DNA ligase, NEB); reaction was left to stand for 10 minutes at room temperature. AMPure XP Beads (SQK-MAP002 - 003) or His-Tag Dynabeads (SQK-MAP004 - 005) or MyOne C1 Streptavidin (Life Technology, Paisley, UK) (SQK-MAP006) was used to clean the adapter-ligated DNA followed by elution in 25 µL of Elution Buffer (Oxford Nanopore) **Table 2-2**. After each clean-up,1 µL of the library was quantified HS Qubit assay before sequencing.

Table 2- 2 Library preparation procedures using early versions of MinION sequencing kits.

	Step	SQK-MAP002	SQK-MAP003	SQK-MAP004	SQK-MAP005	SQK-MAP006
1	Fragmentation	Shear genomic DNA in a Covaris g-TUBE in a total volume of 80 µL, at 6000rpm (Eppendorf 5424) for 1 minute on each side of the tube.	Shear genomic DNA in a Covaris g-TUBE in a total volume of 80 µL, at 6000rpm (Eppendorf 5424) for 1 minute on each side of the tube.	Shear genomic DNA in a Covaris g-TUBE in a total volume of 80 µL, at 6000rpm (Eppendorf 5424) for 1 minute on each side of the tube.	Shear genomic DNA in a Covaris g-TUBE in a total volume of 81 µL, at 6000rpm (Eppendorf 5424) for 1 minute on each side of the tube.	Genomic DNA in 46 µL NFW sheared for 8 kb fragments in a Covaris g-TUBE using Eppendorf 5424; 6000rpm. 2 x 1 minute
2	DNA Repair	PreCR treatment (NEB) following manufacturer's protocol. 37 °C for 30 minutes.	PreCR treatment following manufacturer's protocol. 37 °C for 30 minutes.	PreCR treatment following manufacturer's protocol. 37 °C for 30 minutes.	PreCR treatment following manufacturer's protocol. 37 °C for 30 minutes.	DNA was repaired by NEBNext FFPE DNA Repair Mix (NEB), following manufacturer's protocol. 15 minutes at room temperature.
3	DNA Clean up	Clean up the reaction products with 1x Agencourt AMPure XP. beads, eluting in 80 µL nuclease-free water (NFW)	Clean up the reaction products with 1x Agencourt AMPure XP, eluting in 80 µL NFW.	Clean up the reaction products with 1x Agencourt AMPure XP beads by volume, eluting in 80 µL NFW.	Clean up the reaction products with 1x Agencourt AMPure XP beads, eluting in 81 µL NFW.	Clean up the reaction products with 1x Agencourt AMPure XP, eluting in 41 µL nuclease-free water
4	End-Repair	End-repair using NEBNext End Repair module (NEB), following manufacturer's protocol. 30 minutes at RT.	End-repair using NEBNext End Repair module (NEB). 30 minutes at RT.	End-repair using NEBNext End Repair module (NEB). 30 minutes at RT	End-repair using NEBNext End Repair module (NEB). 30 minutes at RT	DNA was repaired with NEBNext Ultra II End-repair/dA-tailing Module. 20°C for 5 minutes and 65°C for 5 minutes
5	DNA Clean up	Clean up the reaction products with 1x Agencourt AMPure XP, eluting in 80 µL nuclease-free water (NFW)	Clean up the reaction products with 1x Agencourt AMPure XP, eluting in 80 µL NFW.	Clean up the reaction with 1x Agencourt AMPure XP, eluting in 80 µL NFW.	Clean up the with 1x Agencourt AMPure XP. beads by volume, following manufacturer's protocol, eluting in 81 µL NFW	Clean up with 1x Agencourt AMPure XP. beads by volume, following manufacturer's protocol, eluting in 46 µL NFW

6	dA tailing	NEBNext dA-Tailing Module (NEB) following manufacturer's protocol. 37 °C for 15 minutes.	NEBNext dA-Tailing Module (NEB). 37 °C for 15 minutes.	NEBNext dA-Tailing Module (NEB). 37 °C for 15 minutes.	NEBNext dA-Tailing Module (NEB). 37 °C for 10 minutes.	Omitted from ONT the protocol
7	DNA clean up	No clean-up performed	No clean-up performed	No clean-up performed	No clean-up performed	1x Agencourt AMPure XP. beads by volume.
8	Ligation	Ligation reaction consist of: dA-tailed DNA -30 µL Adapter Mix (ONT)- 10 µL HP adapter (ONT) - 10 µL Blunt/TA Ligase Master Mix (NEB)- 50 µL 10 minutes incubation at RT.	Ligation reaction consist of: dA-tailed DNA -30 µL Adapter Mix (ONT)- 10 µL HP adapter (ONT) - 10 µL Blunt/TA Ligase Master Mix - 50 µL 10 minutes at RT.	Ligation reaction consists of: Nuclease free water 8µL dA-tailed 30µL Adapter mix 10 µL HP Adapter 2µL Blunt TA Ligase Master Mix 50µL 10 minutes at RT	Ligation reaction consists of: Nuclease free water dA-tailed 30µL Adapter mix 10 µL HP Adapter 10 µL Blunt TA Ligase Master Mix 50µL 10 minutes at RT	Ligation reaction consists of: NFW8 µL Adapter Mix10 µL HPA 2 µL 50 µL NEB Blunt / TA Master Mix 10 minutes at RT
9	DNA Clean up	Clean up the reaction products with 1x Agencourt AMPure XP. beads by volume, following manufacturer's but using 150 µL ONT's wash buffer. Elute in 25 µL eluting buffer (ONT).	Clean up with 42.8 µl of Agencourt AMPure XP beads containing BSA*, wash using 150 µL wash buffer (ONT) containing BSA 10 mins incubation at RT.	Library purification using His-Tag beads and wash buffer.	Library purification using His-Tag beads and wash buffer.	
10	Tether annealing	Reaction consist of: Ligated DNA 25 µL Tether 10 µL 10 mins at RT	Omitted from ONT the protocol	Omitted from ONT the protocol	Omitted from ONT the protocol	Add 1 µL HP Tether Incubate at RT for 10 minutes

11	Library conditioning	Reaction consists of Adapted and tethered DNA fragments 35 µL HP Motor 25 µL.	Omitted from ONT the protocol	Omitted from ONT the protocol	Omitted from ONT the protocol	
12	Library Elution	Conditioned library/ Pre-sequencing mix is ready for sequencing	Pre-sequencing mix eluted in 25 µL eluting buffer	Pre-sequencing mix eluted in 25 µL eluting buffer	Pre-sequencing mix eluted in 25 µL eluting buffer	Library was purified by MyOne C1 beads and BBB wash buffer.

MinION sequencing was variously performed using R7.0, R7.3 (Goodwin et al., 2015) and R9 flow cells (Jain et al., 2017) (Oxford Nanopore Technology). A QC run was firstly performed to assess the flow cell's number of active pores followed priming a flow cell for sequencing using a mixture of Fuel Mix, Running Buffer/ EP Buffer and water (Oxford Nanopore Technology). The Pre-Sequencing Library was then diluted in nuclease free water, EP Buffer/Running Buffer, and Fuel Mix (Oxford Nanopore Technologies) and loaded in a flow cell for sequencing using MAP_48Hr_Sequencing_Run script. Most libraries were run for less than 48 h, the library was reloaded after 12 and 24 hours of sequencing.

2.12: Data Analysis

2.12.1: NGS Data Analysis - Kraken

To quickly classify and assign taxonomic labels to the Illumina sequences, we used rapid classification method which uses an exact alignment of *k-mers* to assign taxonomic labels to metagenomics DNA sequences (Norling, Karlsson-Lindsjö, Gourlé, Bongcam-Rudloff, & Hayer, 2016), Kraken BaseSpace application (Illumina). All the bacteria, fungi or virus sequences were identified to species or genus level, and species that did not have k-mers in the database (including human) or if insufficient evidence exists were left as unclassified by Kraken (Rose et al., 2015).

Kraken's sensitivity is about 90% and precision over 99% (Wood & Salzberg, 2014). Kraken results were presented either as in excel file or in a graphical representation in Krona taxonomic classification as an interactive pie chart.

Kraken was mainly used to characterize metagenomics sample and identify contaminant sequences rapidly. Further bioinformatics analysis were required in addition to Kraken.

2.12.2: Illumina and MinION Data Analysis

The Fast5 files for *Salmonella* Typhi H58 were analysed and the sequence extracted in fasta and fastq format using the Poretools (Nicholas J Loman & Quinlan, 2014) v 0.3.0 library, commands 'poretools fasta' and 'poretools fastq'. The reads were

mapped to the *de novo* assembly of the Illumina data using the LAST aligner (Frith, Hamada, & Horton, 2010) with the paraments 'lastal –s 2-T 0 –Q 0 –a 1'(Ashton et al., 2015). Miscalled bases (assuming no SNPs and perfect alignment) were determined by analysing the alignment output. If two alignments over- lapped by >75%, the alignment with the lower E value was discarded (Ashton et al., 2015).

A hybrid Illumina MinION assembly was generated using SPAdes (Bankevich et al., 2012) (v3.1.1) with the '—careful' flag, k-mers of 21, 33, 55 and 77) with the concatenated MinION reads for that isolate input under the–PacBio flag (Ashton et al., 2015). The reads from other genomes (*B. longum* 8809 and *P. fluorescens*) were mapped to the *de novo* assembly of the Illumina data using the LAST aligner (Frith et al., 2010) with the paraments 'lastal –a 2'.

For *Salmonella* Blockley, SPAdes v.3.7.0 hybrid assembly (Bankevich et al., 2012) was used to combine the MinION reads with the Illumina reads. MinION reads were mapped back to the hybrid assembly and this mapping was used to confirm the contiguity of key parts of the hybrid assembly. These processed reads were mapped to a de novo assembled *Salmonella* Blockley strain (73626) using BWA-MEM (Nair et al., 2016a).

The FAST5 files for the *E. coli* K-12 substr. MG1655 were aligned to the *E. coli K-12* reference genome (NCBI RefSeq Accession NC_000913.1) using BWA-MEM (H. Li, 2013)(H. Li et al., 2009) version 0.7.12-41044 with the nanopore data parameters '-× ont2d' and LAST version 460. Both the BWA-MEM and the LAST alignment (Kiełbasa, Wan, Sato, Horton, & Frith, 2011) were post-processed using marginAlign (Jain et al., 2015) version 0.1(Ip et al., 2015).

The FAST 5 files for urine sample were mapped to the reference genome of particular strain using BLAST alignment tools. Statistics on alignments for all the genome studies using MinION sequencing were computed by SAMtools (H. Li et al., 2009).

All the reads for the blood samples were analysed using default BWA- MEM settings, version 0.7.13-r1126 (minION reads used –x ont2d) to map against the human genome (index –a bwtsw hg19). Reads aligning to the hg19 genome were removed

using samtools –f 4 option and PCR duplicates removed with samtools rmdup -S. Kraken version 0.10.5-beta was used to identify the most abundant organism within the samples using the standard database. Genomes were then chosen from PubMed to map the reads against specific pathogens, including *S. aureus* (NCTC8325), *S. haemolyticus* (JCSC1435) and *E. faecalis* (V583). Single genomes were assembled using SPAdes version 3.8.0 for Illumina data and metagenomes were assembled with MEGAHIT version 1.0.3 using default settings. Assembled contigs were uploaded to the CARD (Comprehensive Antibiotic Resistance Database) website with the aid of the Resistance Gene Identifier (RGI) analysis tool (McArthur et al., 2013) to identify antibiotic resistance genes (discovery criteria – perfect, strict and loose hits); positive hits were defined as ≥95% identity.
Reads for *C. difficile* were counted in every fastq file using: cat input.fq | echo $((`wc –l `4/)) and aligned to *C. difficile* genome 630: bwa mem ref.fa input.fq | samtools view –bS - | samtools sort - -o output.bam.

Also, the ONT 'What's in my pot?' (WIMP) (Juul et al., 2015) Metrichor application was used to classify MinION metagenomics data in blood and urine samples in real-time, using a reference database and Kraken.

All the alignments were viewed in Qualimap (García-Alcalde et al., 2012) for genome coverage.

2.12.3: PCR Cq Calculation

Raw Cq values of qPCR assays were translated by estimating number of human and bacteria (*S. aureus* NCTC 6571) cells using DNA/cell conversion factors. The amount DNA in each human cell is estimated to be 6.6pg, hence conversion factor of 6.6 pg DNA/cell (Saiki et al., 1988) was used to make standards and develop standard curve. Most of the human positive control were detected at around Cq 22 which is equivalent to approximately 10^5 human cells (**Figure 2-3**). For PCR analysis, positive control used one tenth of the DNA from 200 μL of blood sample. Depletion of human DNA was considered successful when human DNA equivalent of one cell was left in a sample (one cell equals to approximately Cq 37) (**Figure 2-3** and **2-4**). One tenth of the extracted DNA from 1 mL blood sample was used for PCR analysis.

Inc	Pos	Name	Type	CP	Concentration	Standard	Status
☑	A1	10^5 cells	Standard	22.91	8.85E4	1.00E5	
☑	A2	10^4 cells	Standard	25.63	1.03E4	1.00E4	
☑	A3	10^3 cells	Standard	28.61	9.72E2	1.00E3	
☑	A4	10^2 cells	Standard	31.60	9.10E1	1.00E2	
☑	A5	10^1 cells	Standard	35.14	5.75E0	1.00E1	
☑	A6	10^0 cells	Standard	37.86	1.42E0	1.00E0	
☑	A7	Sample 7	Negative Control				
☑	B1	10^5 cells	Standard	22.77	9.90E4	1.00E5	
☑	B2	10^4 cells	Standard	25.50	1.14E4	1.00E4	
☑	B3	10^3 cells	Standard	28.51	1.05E3	1.00E3	
☑	B4	10^2 cells	Standard	31.50	9.89E1	1.00E2	
☑	B5	10^1 cells	Standard	34.21	1.16E1	1.00E1	
☑	B6	10^0 cells	Standard	39.36	1.00E0	1.00E0	
☑	B7	Sample 19	Negative Control				
☑	C1	10^5 cells	Standard	22.80	9.71E4	1.00E5	
☑	C2	10^4 cells	Standard	25.59	1.06E4	1.00E4	
☑	C3	10^3 cells	Standard	28.61	9.77E2	1.00E3	
☑	C4	10^2 cells	Standard	31.44	1.04E2	1.00E2	
☑	C5	10^1 cells	Standard	34.38	1.01E1	1.00E1	
☑	C6	10^0 cells	Standard	40.00	9.41E-1	1.00E0	>
☑	C7	Sample 31	Negative Control				

> - Late Cp call (last five cycles) has higher uncertainty

Figure 2- 3 Human Cq values and cell equivalents in human standard curve

For *S. aureus*, conversion factor of 2.9fg DNA/cell was used. For method development purposes, bacterial cells equivalents of $10^{2\text{-}3}$ (appox. Cq 27 to 32) were spiked into blood samples (**Figure 2-4**).

Inc	Pos	Name	Type	CP	Concentration	Standard	Status
☑	A1	10^6	Standard	17.95	9.74E5	1.00E6	
☑	A2	10^5	Standard	21.28	9.71E4	1.00E5	
☑	A3	10^4	Standard	24.50	1.05E4	1.00E4	
☑	A4	10^3	Standard	27.85	1.03E3	1.00E3	
☑	A5	10^2	Standard	31.13	1.07E2	1.00E2	
☑	A6	10^1	Standard	34.01	1.25E1	1.00E1	
☑	A7	10^0	Standard	36.48	1.52E0	1.00E0	
☑	A8	-ve H2O	Negative Control				
☑	B1	10^6 2	Standard	18.02	9.26E5	1.00E6	
☑	B2	10^5 2	Standard	21.27	9.79E4	1.00E5	
☑	B3	10^4 2	Standard	24.21	1.28E4	1.00E4	
☑	B4	10^3 2	Standard	27.99	9.33E2	1.00E3	
☑	B5	10^2 2	Standard	31.45	8.51E1	1.00E2	
☑	B6	10^1 2	Standard	34.16	1.11E1	1.00E1	
☑	B7	10^0 2	Standard	37.49	6.01E-1	1.00E0	
☑	B8	-Ve cont H2O	Negative Control				

Figure 2- 4 *S. aureus* Cq values and cell equivalents in *S. aureus* standard curve

The volume of positive control used for blood samples was 200 μL, 5-fold less of the sample volume (1 mL). The 5-fold difference in volume was estimated to be equal to 2.2 Cq.

Chapter 3. Results: Development of Host DNA Depletion Methods in Clinical Samples

3.1: Chapter Introduction

The aim of this chapter was to develop host DNA depletion methods to enable unbiased pathogen identification in clinical samples (blood, urine and stool) using metagenomics sequencing.

A number of studies have used the MolYsis kit (Molysis, Germany) to deplete human DNA in blood samples with the aim of improving the sensitivity of PCR in large volumes of blood (5-10mL) and more recently, for application in metagenomics in human samples (Gyarmati et al., 2016; Hansen et al., 2009). The dearth of studies applying metagenomics to infection diagnosis is partly because of the overwhelming amount of host DNA in clinical samples, particularly blood (Mwaigwisya et al., 2015), and the inefficient human DNA depletion observed using MolYsis and other commercially available methods (Mwaigwisya et al., 2015; Springer et al., 2011; Thoendel et al., 2016).

We estimated that, for reliable pathogen identification directly from clinical blood samples using metagenomics, the ratio of human to pathogen DNA should be reduced from approx. 10^9:1 to $<10^3$:1 (a 10^6-fold reduction). As none of the commercially available or published methods could achieve this level of human DNA depletion, improvements were made to commercial methods (mainly MolYsis) and novel depletion methods were developed to enable metagenomics based pathogen and antibiotic resistance profiling directly from clinical samples.

For blood samples, the MolYsis kit was chosen as this method was the most efficient of the commercially available kits according to the literature (Hansen et al., 2009). This kit is based on the combination of chemical lysis and enzymatic digestion of human DNA, which makes the process relatively simple and rapid. However, MolYsis by itself did not provide sufficient DNA depletion in blood (approx. 99% or 10^2-fold depletion), therefore another simple method was used to complement MolYsis depletion. The concept was to remove leukocytes in blood before MolYsis

treatment, thereby improving overall depletion efficiency. This was achieved by looking for a common surface marker to all classes of leukocytes that could be targeted for depletion by immunomagnetic separation (IMS). Leukocytes cell surface marker CD45 was identified as a potential target and we found commercially available anti-CD45 coated Dynabeads® (ThermoFisher Scientific) that could be used directly in blood for leukocytes depletion by IMS. These methods were optimised and tested in blood, and subsequently, in clinical urine samples. Urine was a less challenging sample type compared to blood, due to the higher numbers of bacteria present in a UTI - human:pathogen DNA ratio can be as high as 1:10 (up to 10^9 CFU/mL pathogen) (Ferreira et al., 2010).

Additionally, in this chapter a novel method of depleting host DNA in blood based on differential lysis of leukocytes using cytolysins and degradation of released DNA by nucleases or propidium monoazide (PMA) was developed. A cytolysin (also known as a cytolytic toxin) is a protein secreted by a microorganism, plant, fungus or animal which is specifically toxic to a heterologous cell type(s), particularly promoting lysis of target cells. Cytolysins secreted by bacteria (also known as bacteriocins) are particularly lethal for a broad range of eukaryotic cells, killing them by forming pores in their membranes thereby releasing DNA. Several cytolysins were investigated, including streptolysin O produced by *Streptococcus pyogenes*, alpha-hemolysin by *Staphylococcus aureus* and phospholipase C that that digests phospholipids by *Clostridium perfringens.* There are four types, A, B, C and D, we chose phospholipase C (PLC) (i.e. a phospholipase that cleaves before the phosphate, releasing diacylglycerol and a phosphate-containing head group) as it appeared to possess the activity necessary for efficient human cell lysis and it was commercially available from Sigma. Bacterial PLC enzymes are classified in four groups as follows:

Group 1 – Zinc metallophospholipases
Group 2 – Sphingomyelinases (e.g. sphingomyelinase C)
Group 3 – Phosphatidylinositol
Group 4 – Pseudomonad PLC

We have investigated the ability of these cytolysins to differentially lyse human leukocytes, followed by DNA depletion using nucleases or PMA.

PMA is a high affinity photoreactive DNA binding dye. The dye is weakly fluorescent by itself but increases in fluorescence after binding to nucleic acids. It preferentially binds to dsDNA with high affinity. Upon photolysis, the photoreactive azido group on the dye converts into a highly reactive nitrene radical, which readily reacts with any hydrocarbon moiety at the binding site to form a stable covalent nitrogen-carbon bond, thus resulting in permanent DNA modification (Nocker, Sossa-Fernandez et al. 2007). The dye is nearly completely cell membrane-impermeable, and thus can be selectively used to modify only exposed DNA from dead cells while leaving DNA from viable cells intact (Biotium Inc, Fremont, CA, USA) (Nocker, Cheung et al. 2006). This feature was useful for our approach since the cytolysins selectively lyse human cells, leaving bacterial cells intact. PMA was used to degrade the released human DNA.

The third clinical sample type studied was stool, which presented different challenges to blood and urine because of the tremendous amount of normal microbial flora present coupled with variable numbers of host cells. In these samples, methods for enriching *C. difficile* spores from *C. difficile* infected patient samples were developed for genotypic typing of *C. difficile* directly from stool. Initially, Nycodenz® (Sigma) medium was used to enrich *C. difficile* spores from stool (Hevia et al., 2015) (Akiyoshi et al., 2003). Nycodenz has very low toxicity to cells, low osmolarity and a high density; these features make it a potentially suitable media to separate macromolecules and viable cells from biological samples by gradient centrifugation. Nycodenz is a standard media for enriching microbiota from various samples in metagenomics studies (Delmont, Robe, Clark, Simonet, & Vogel, 2011; Manichanh et al., 2006). The media was tested for enriching *C. difficile* spores.

Another gradient centrifugation media, Gastrographin® (Bayer AG, Leverkusen, German), was also tested for enriching *C. difficile* spores from stool samples. Gastrografin (Diatrizoate Meglumine and Diatrizoate Sodium Solution) is a water-soluble iodinated radiopaque contrast medium made for gastrointestinal tract radiological examination via oral or rectal administration. Each mL of Gastrografin contains 660 mg diatrizoate meglumine and 100 mg diatrizoate sodium (Bayer AG, Leverkusen Germany). Gastrografin has attributes of a density gradient medium including high density, high stability, low viscosity and low toxicity. These properties

make the medium suitable for separating macromolecules and viable cells by differential centrifugation. Gastrografin has been used as the matrix for density gradient to separate *Bacillus subtilis* and *C. difficile* spores from vegetative/culture cell suspensions (Fichtel, Köster, Rullkötter, & Sass, 2007; Sixt et al., 2013) but not from metagenomic samples.

The last method tested on stool samples was the in-house developed method based on chemical differential lysis of human and vegetative cells using buffer FL (GeneAll, Seleou Korea), leaving *C. difficile* spores intact. This was followed by enzymatic digestion of the released DNA and DNA extraction from the spore enriched sample.

The described depletion/enrichment strategies for various clinical samples were developed over the course of my PhD to improve the metagenomics detection of pathogens.

The qPCR results presented below are representative of the biological replicates performed. Variation between treatments did not affect the conclusions of the experiments.

3.2: Host DNA Depletion Method Development for Blood Samples

3.2.1. Method 1: Host DNA Depletion using Dynabeads® CD45 and MolYsis™ Basic 5 kit

Initially, human DNA depletion was performed using the MolYsis kit following the manufacturer's instructions as described in section **2.3.1.** Human DNA depletion was quantified using a human qPCR assay as described in section **2.9.1.** The impact of the depletion method on spiked bacteria (*S. pneumonia*e) was also assessed by qPCR. Human DNA depleted blood samples were compared to the respective un-depleted sample (positive control – PC). NTC is a qPCR control with no template DNA. Taking into account the difference in volume, less than 50% of the human DNA was depleted (∆ Cq 1.64 and 5 fold bigger volume in depleted sample), with ~75% loss of bacterial signal (∆ Cq -0.19). (**Table 3.1,** and **Figure 3-1** and **3-2**).

Table 3- 1 qPCR results showing Cq of human and *S. pneumoniae* DNA before (PC) and after depletion by MolYsis. NTC is the no template control.

	Sample ID	**Human qPCR (Cq)**	**S. pneumo qPCR (Cq)**
1	MolYsis_1mL	23.22	25.85
2	PC_ 200µL	21.58*	26.04
3	NTC	_**	_

*Cq value suggests <10^5 cell (<10^6 cells in total input)
**No DNA was detected

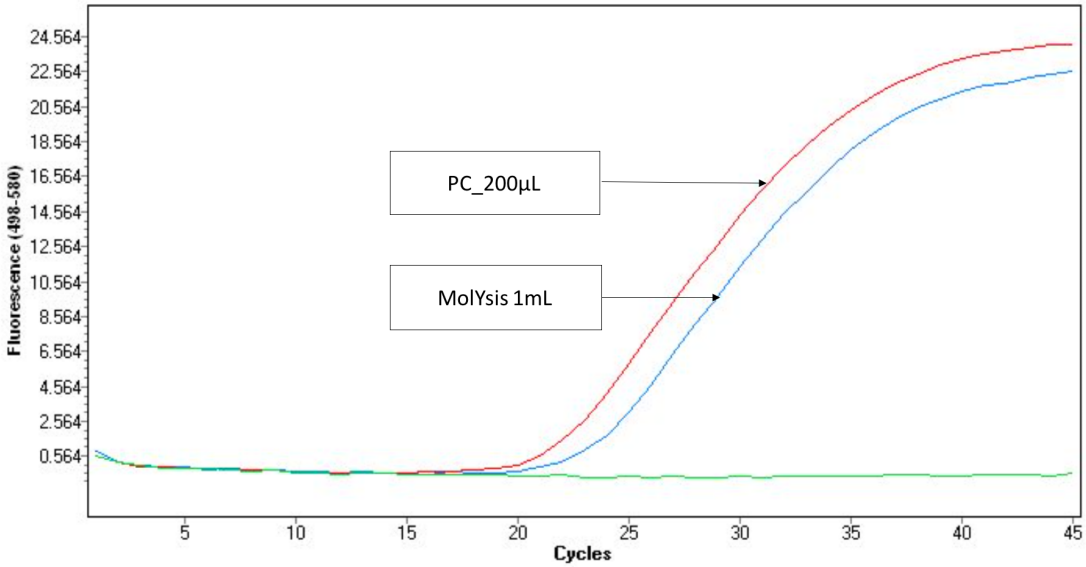


Figure 3- 1 real-time amplification curve obtained using human qPCR showing 1.64 quantitative cycles host DNA depletion using MolYsis Basic 5 kit. The X-axis denotes the cycle number of the quantitative qPCR reaction and the Y-axis denotes the fluorescence.

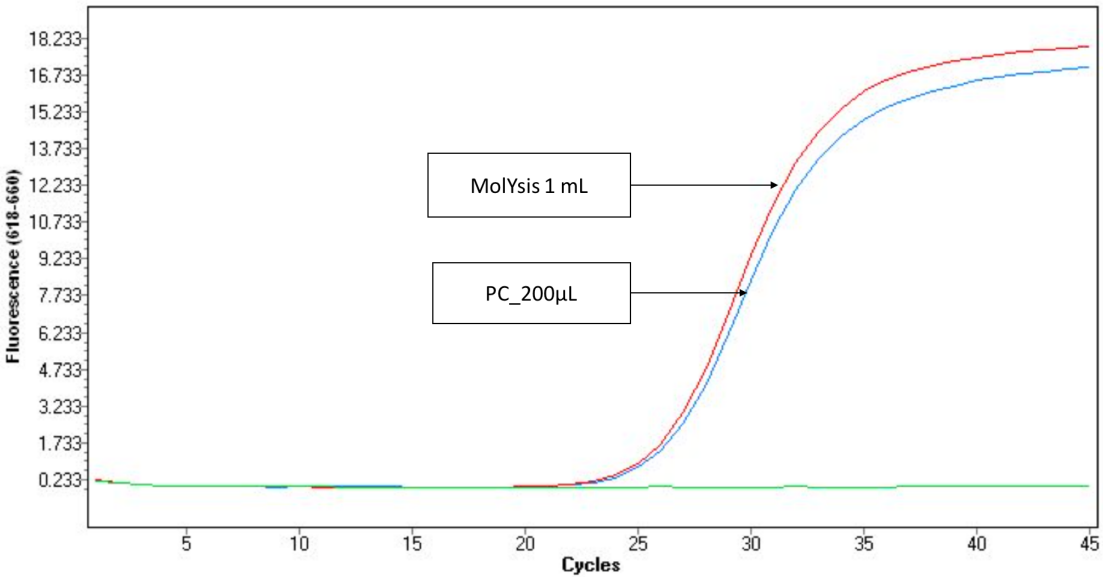


Figure 3- 2 qPCR amplification curves using *S. pneumoniae* assay showing 80% loss of the bacterial DNA.

The MolYsis method was repeated several times (by different personnel) but only 10 to 90% of human DNA was depleted. The level of host DNA depletion using MolYsis kit was insufficient for metagenomics identification of pathogens in blood. As an alternative, Dynabeads CD45 were used to remove white cells as the CD45 cell surface marker is expressed on all human leukocytes. Dynabeads® CD45 are uniform, superparamagnetic beads (4.5 µm diameter) coated with a primary monoclonal mouse IgG_{2a} antibody specific for a CD45 membrane antigen common to all known isoforms of CD45. Due to the bead size, Dynabeads® CD45 can efficiently isolate and remove leukocytes from viscous samples such as blood by IMS as seen in **Figure 3-3.** Leukocytes depletion using Dynabeads could be completed in 30 minutes.

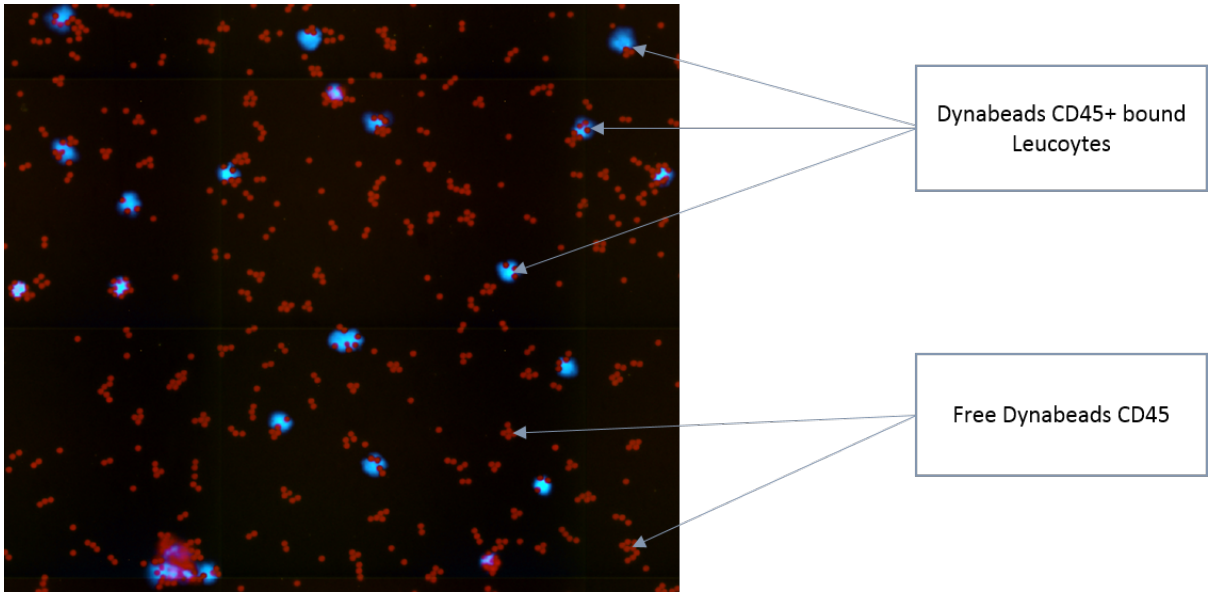


Figure 3- 3 DAPI stained slide showing Dynabeads CD45 (red) bound to leukocytes (blue) before depletion by IMS.

The beads were used to remove leukocytes from blood, hence deplete the human DNA. After human DNA depletion with beads, qPCR results showed >95% human DNA was depleted, approximately five qPCR cycles (Δ Cq 4.82) between the CD45 IMS depleted blood and the un-depleted sample (positive control). **Table 3-2** and **Figure 3-4.**

Table 3- 2 qPCR results showing Cq of human DNA before (PC) and after depletion (CD45 IMS blood) using 200µL blood sample. NTC is the no template control.

	Sample ID	**Human qPCR (Cq)**
1	CD45 IMS_200µL	27.69
2	PC_ 200µL	22.87
3	NTC	_*

*No DNA was detected

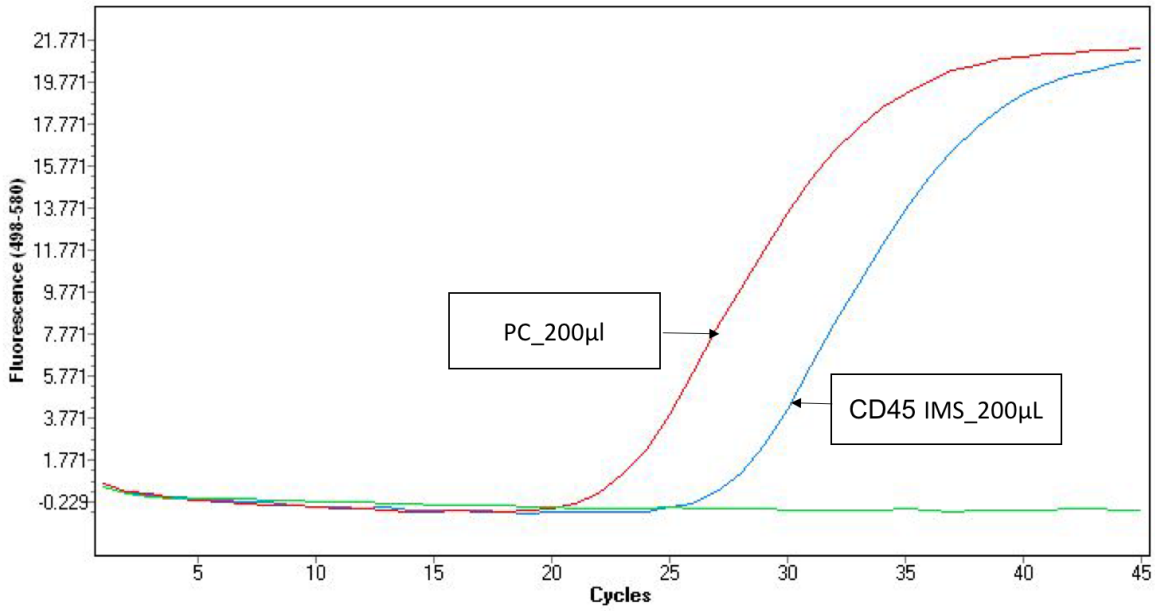


Figure 3- 4 real-time qPCR amplification curve obtained using human qPCR showing 4.8 quantitative cycles host DNA depletion using CD45 immunomagnetic separation.

During human DNA depletion by CD45 IMS there was no loss of bacterial signal as observed by qPCR for spiked *S. pneumoniae* (∆ Cq 0.04) (**Table 3-3** and **Figure 3-5**).

Table 3- 3 qqPCR Cq results showing no loss *S. pneumoniae* DNA of before (PC) and after depletion (CD45 depleted blood) using of spiked 200µL blood sample.

	Sample ID	**S. pneumo qPCR (Cq)**
1	CD45_S. pneumo	25.09
2	PC_ S. pneumo	25.05
3	NTC	_

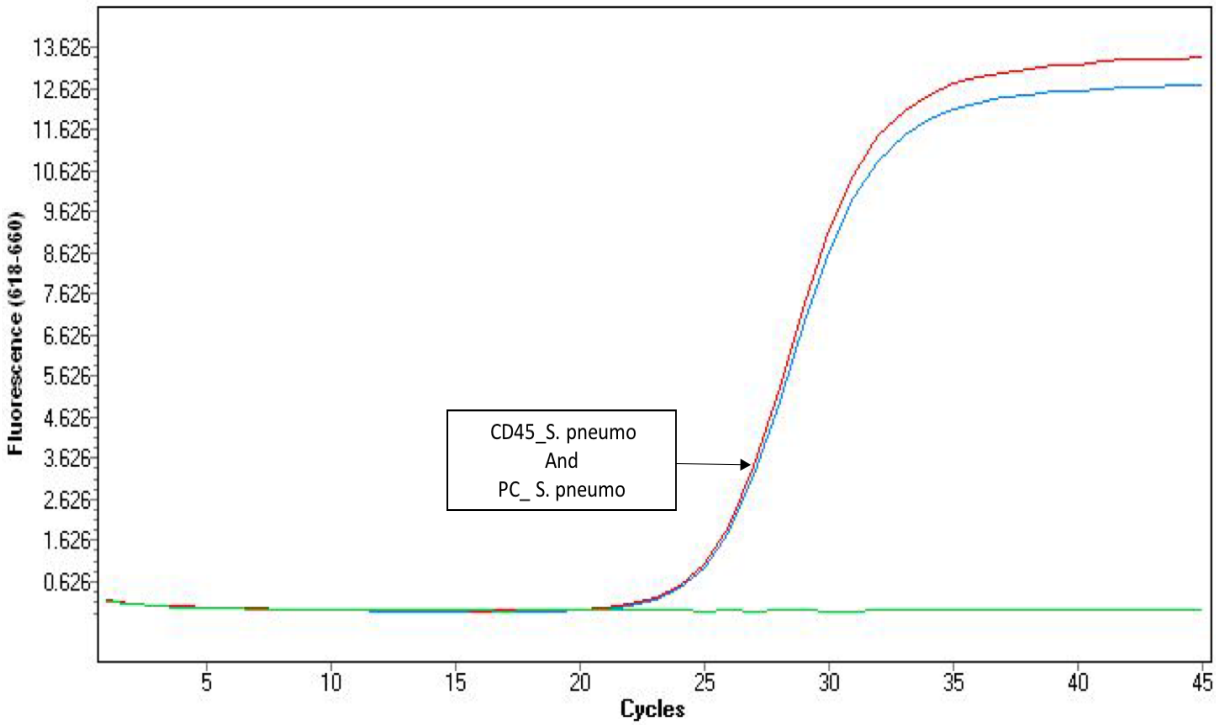


Figure 3- 5 amplification curves obtained using qPCR for *S. pneumoniae* showing no loss of bacterial DNA during CD45 IMS.

We estimated 10^6-fold depletion was necessary for cost and time efficient metagenomics detection of pathogens in blood – reducing the ratio of human: pathogen DNA from 10^9:1 to 10^3:1. Hence, the depletion levels achieved using IMS alone were not sufficient. Therefore, we decided to combine two complementary depletion methods (CD45 IMS as described in **2.3.2** and MolYsis in **2.3.1**), removing >90% of the leukocytes by IMS followed by differential lysis and DNase removal of

the remaining human DNA using MolYsis. When the two methods were combined, this resulted in 99.3% depletion of human DNA (Δ Cq 6.03), taking into account the difference in volume (5-fold difference or 2.2 Cq), loss of bacterial signal was still around 80% (Δ Cq 0.01) (**Table 3-4** and **Figure 3-6** and **3-7**)

Table 3- 4 qPCR results (Cq) showing human and *S. pneumoniae* DNA after depletion using CD45 IMS followed by MolYsis compared to the respective un-depleted (PC) 200µL blood sample.

	Sample ID	**Human qPCR (Cq)**	**S. pneumo qPCR (Cq)**
1	CD45 IMS + MolYsis_1mL	28.61*	29.00
2	PC_ 200µL	22.58	29.01
3	NTC	_	_

*Cq value suggests <10^3 cells

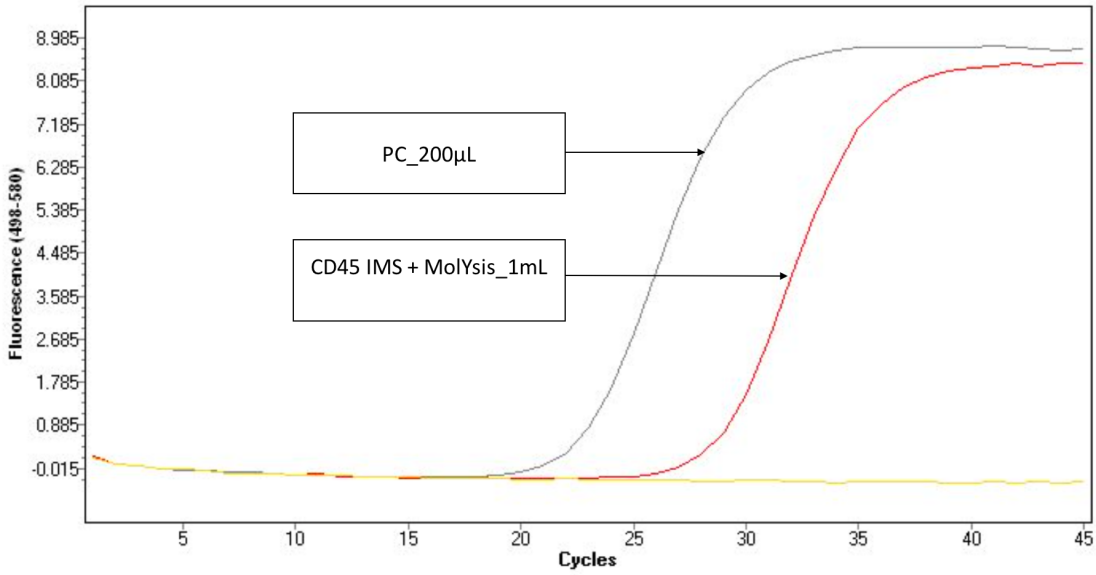


Figure 3- 6 qPCR amplification curve obtained using human showing host DNA depletion (6.03 Cq) using combined method of CD45 IMS and MolYsis.

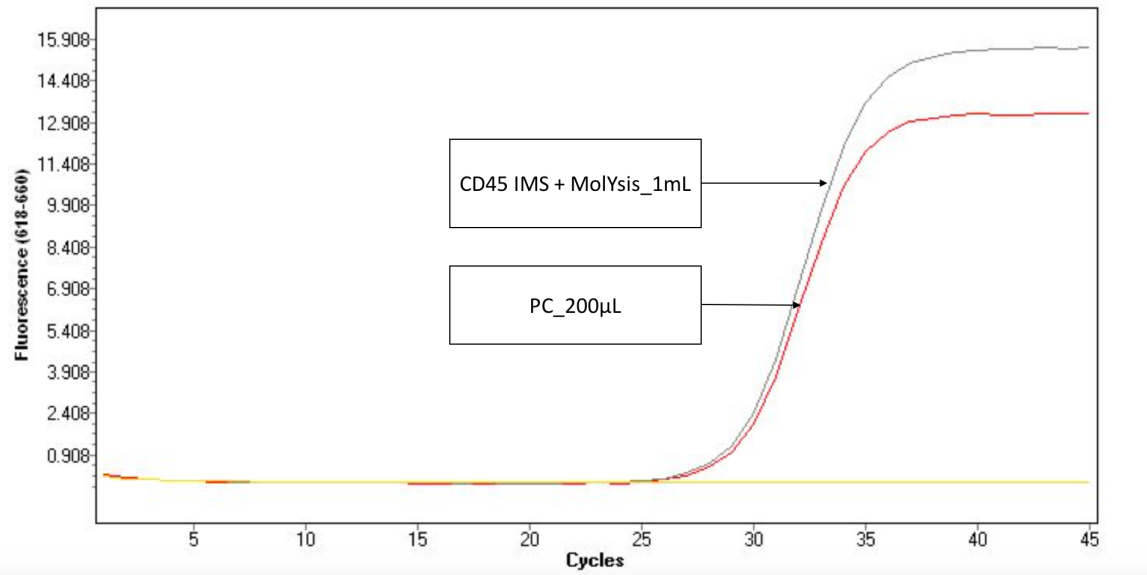


Figure 3- 7 qPCR amplification curve showing loss 80% (∆ Cq 0.01) of *S. pneumoniae* after host DNA depletion using CD45 IMS and MolYsis.

Although the combined method depleted more 99% of human DNA, higher levels of depletion were expected. This is because 90% depletion by CD45 IMS would have reduced the initial amount of leukocytes and DNA considerably, making the sample matrix less viscous for the MolYsis buffers and enzyme to work. It was therefore hypothesised that the buffer used for CD45 IMS (Isolation buffer 1) was interacting with MolYsis buffer (buffer CM), which reduced the efficiency of the MolYsis method. Therefore, when combining the two methods, plasma and isolation buffer 1 (added during CD45 IMS procedure) was removed before adding chaotropic buffer CM for lysing host cells (section **2.3.3**). The combined method was tested on blood samples spiked with *S. pneumoniae* and depletion was quantified by qPCR assays. After depletion, 1 mL of sample was compared to 200 µL of PC. qPCR results showed ∆Cq 17.32, which equates to an estimated 99.9999% (10^6 fold) depletion of human DNA (when the 5-fold lower volume of the PC is taken into account) (Table **3-5** and Figure **3-8**).

Table 3- 5 qPCR results (Cq) showing human depletion using CD45 IMS and MolYsis. In 1 ml samples a depletion of approx. 10^6 -fold was observed, meaning approximately 1 cell equivalent of human DNA was remained.

	Sample ID	**Human qPCR (Cq)**
1	CD45 IMS +MolYsis _1mL	40.00*
2	PC_ 200µL	22.68
3	NTC	_

* Cq value suggests <1 cell (<10 cells in total input)

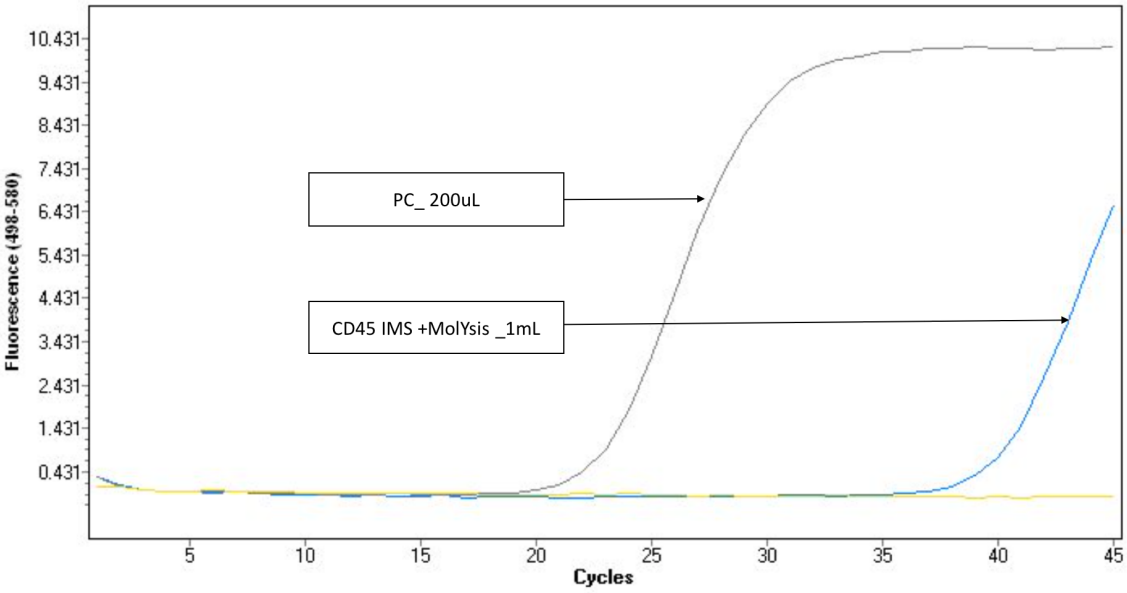


Figure 3- 8 amplification curves obtained using qPCR for human showing host DNA depletion (ΔCq 17.32) using CD45 IMS and MolYsis.

Table 3-6 and **Figure 3-9** below show corresponding qPCR results for *S. pneumoniae* after depletion of host DNA. Cq (Δ Cq -0.62) results of two samples were comparable despite the difference in volume, suggesting approximately 70% loss of bacterial DNA during depletion of human DNA.

Table 3- 6 qPCR results showing loss *S. pneumoniae* DNA in 200µL positive (PC) compared to 1mL of depleted blood sample.

	Sample ID	**S. pneumo qPCR (Cq)**
1	CD45 IMS + MolYsis_1mL	28.95
2	PC_200µL	29.57
3	NTC	_

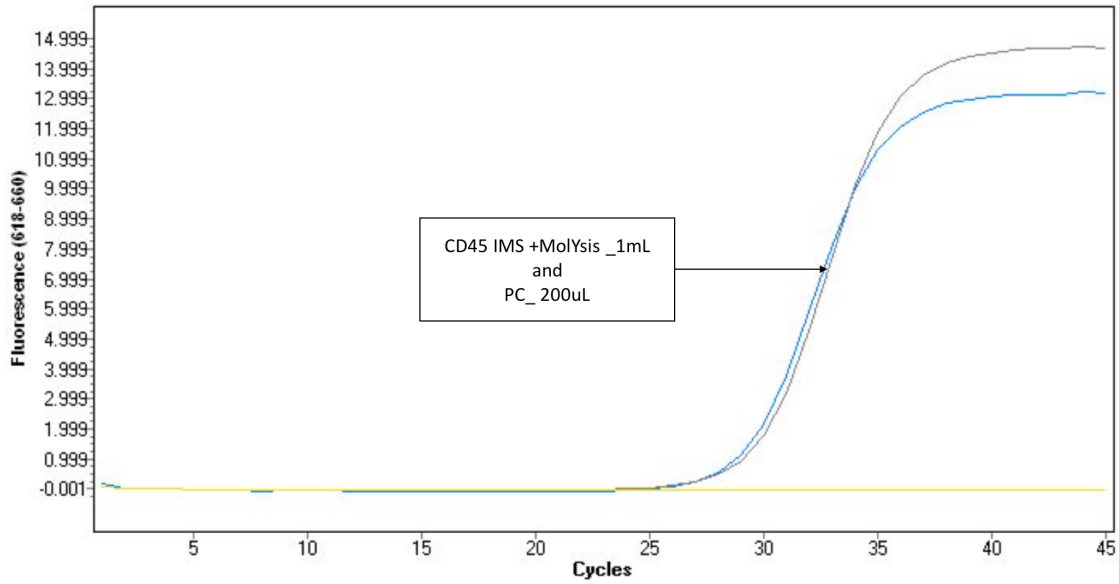


Figure 3- 9 qPCR results showing approximately 70% loss of *S. pneumoniae* DNA in 1mL of depleted blood compared to 200µL positive (PC).

This combination of methods resulted in 99.9999% (10^6 fold) depletion of human cells in a 1 ml blood sample spiked with *S. pneumoniae* without significant loss of bacterial DNA (approx. 4-fold). The MolYisis procedure appeared to be working more efficiently using this approach and we postulated that this was related to the removal of excess EDTA (contained in the in the blood sample collection tubes) that may interfere with the activity of the chaotropic lysis buffer in the MolYsis kit. To confirm this, the MolYsis procedure was modified to remove excess EDTA by spinning down the cells for 10 minutes at ≥ 12000 × g, discarding the supernatant, resuspending the cell pellet in sterile PBS (to make 1mL) and then performing the standard MolYsis procedure as explained in section **2.3.1.** Modifying the MolYsis method

using this approach improved human DNA depletion from 90% to 99.95% (Δ Cq 9.83 plus 5x bigger volume used), with approximately 75% loss of bacterial signal, 4 fold (Δ Cq -0.51 but 5x bigger volume used) (**Table 3-7** and **Figure 3-10** and **3-11**).

Table 3- 7 qPCR results showing human and *S. pneumoniae* DNA after depletion by modified MolYsis compared to the respective 200µL positive (PC).

	Sample ID	**Human qPCR (Cq)**	**S. pneumo qPCR (Cq)**
1	Modified MolYsis_1mL	31.78	26.28
2	PC_ 200µL	21.95	26.79
3	NTC	_	_

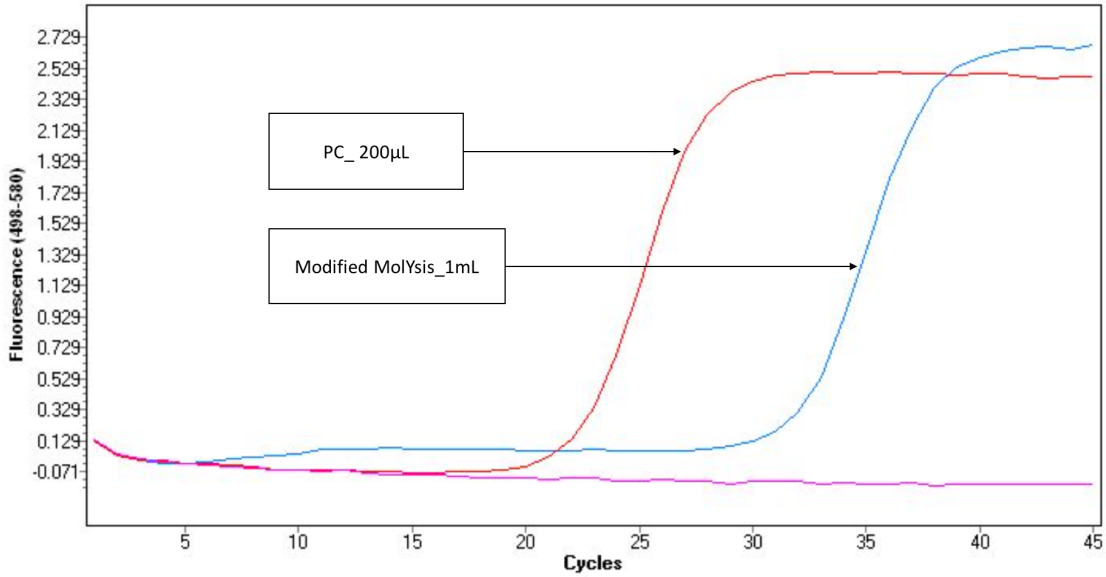


Figure 3- 10 qPCR amplification curves showing (Δ Cq 9.83) human DNA depletion by modified MolYsis procedure.

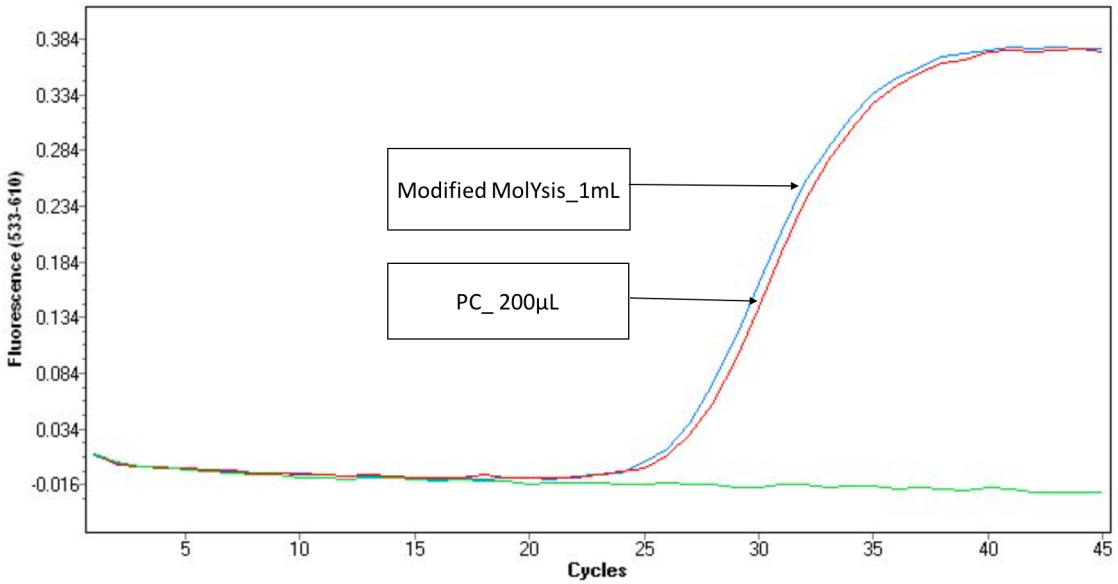


Figure 3- 11 qPCR amplification curves showing *S. pneumoniae* DNA depletion by modified MolYsis protocol

Despite improvement, the modified MolYsis alone was not as efficient at depleting human DNA compared to CD45 IMS with MolYsis. Therefore, it was concluded that CD45 IMS in combination with MolYsis would provide the level of human DNA depletion necessary for metagenomic pathogen identification.

3.2.1.1: Development of Metagenomics Based Pathogen Detection Workflow

After development of an efficient host DNA depletion method, the next step was to integrate this into a metagenomics based workflow for pathogen and antibiotic resistance profiling that could be performed in a clinically relevant timeframe. We started by sequencing human DNA depleted blood samples spiked with approximately 1000 CFU/mL *S. pneumoniae* using Nextera XT libraries on the Illumina MiSeq. This resulted in very few reads (n=2), all of which were human. The respective positive control produced 309,253 reads but only 2 were *S. pneumoniae,* the rest were human (**Table 3-8**).

Table 3- 8 Illumina reads showing total number of reads, and number of human reads and S. pneumoniae reads in human depleted sample and respective positive control.

Sample ID	Total reads	Human reads	S. pneumo reads
CD45 IMS + MolYsis_1mL	2	2	-*
PC_200µL	309253	309251	2

*No *S. pneumoniae* read was detected

Although Illumina Nextera XT library preparation required 1 ng of input DNA, samples depleted by CD45 IMS and MolYsis contained less than 1 ng (estimated to be in femtogram range by qPCR). The low amount of input DNA resulted in few sequence reads and poor sensitivity with Illumina sequencing. This result suggested the need for whole genome amplification (WGA) after human DNA depletion in order to:

1. have enough DNA input for Illumina and/or nanopore based sequencing using MinION;
2. improve the sensitivity of the assay;
3. generate enough reads to cover the pathogen genome for accurate pathogen identification and antibiotic resistance genes.

WGA using multiple displacement amplification (MDA) technology generated up to 10,000-fold amplification of DNA from host DNA depleted samples. WGA reactions were performed as described in section **2.10**. Quality assurance of amplified DNA was done by Qubit dsDNA broad range assay, NanoDrop 2000c and 2200 TapeStation instrument (Methods section **2.8**) and qPCR using probe based primers and universal primers as explained in Methods section **2.9**. The downstream applications of whole genome amplified DNA were Illumina sequencing (Nextera XT) and nanopore sequencing as described in the workflow below (**Figure 3-12).**

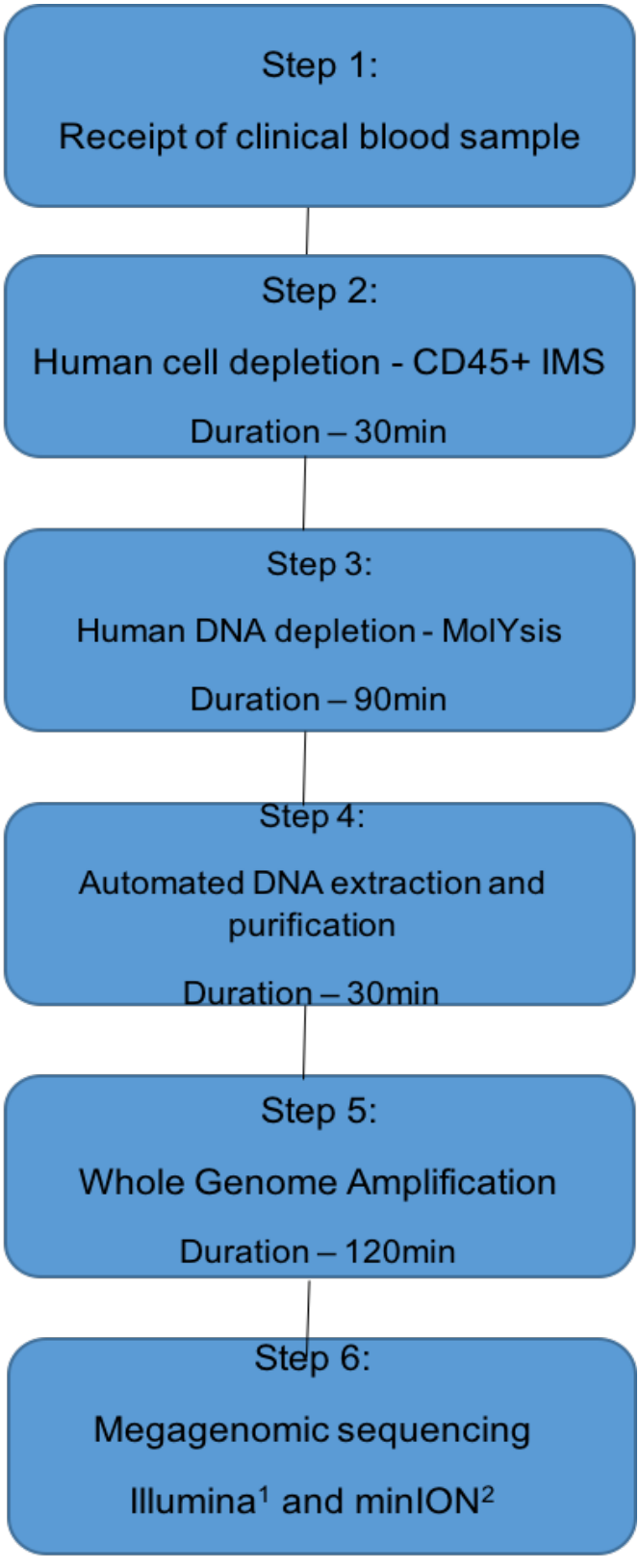


Figure 3- 12 Overview of sample processing workflow for detecting pathogens in clinical blood samples Turnaround time is 7 hours or 3 days depending on sequencing platform.

3.2.2. Method 2: A novel Human DNA Depletion Method using Cytolysins for Differential Lysis of Leukocytes

3.2.2.1: Human DNA Depletion using Endonucleases

Several endonucleases were tested to find the most efficient enzyme in digesting dsDNA in raw blood. Tested endonucleases included HL-SAN DNase (heat labile, salt active nuclease, ArcticZymes, Tromsø, Norway), HL-double stranded DNase (ArcticZymes, Tromsø, Norway) and nuclease micrococcal from *S. aureus* (Sigma Aldrich, Dorset, UK) (see Methods section **2.3.4**). There was no evidence available to suggest that these enzymes have previously been tested for DNA depletion in blood.

Experiments to find the most efficient enzyme were performed in blood which had been though three cycles of freezing at -80ºC for 20 minutes or more and thawing at 37 ºC to lyse leukocytes and release host DNA. The enzymes were tested on lysed blood as described in section **2.3.4.,** except for *Mol*DNase B (Molzym). The *Mol*DNase B from MolYsis kit which had proven to work in blood was used as a control enzyme with its respective buffer, DB1, following the manufacturer's protocol. Of all the enzymes tested on the lysed blood samples only HL-SAN DNase degraded human DNA in blood. The HL-SAN enzyme reduced human DNA 4-fold (∆ Cq 2), whereas DS-DNase and Nuclease did not show any activity (**Table 3-9** and **Figure 3-13**)**.** The *Mol*DNase B control worked best, showing a ∆ Cq of 7. Enzyme buffers were not used except for the control *Mol*DNase B.

Table 3- 9 Human qPCR results showing activity of endonucleases on freeze-thawed human blood without addition of buffers except for the control, *Mol*DNase B.

	Sample ID	**Human qPCR (Cq)**
1	DS-DNase	22.80
2	Nuclease	22.35
3	HL-SAN	24.77
4	*Mol*DNase B	29.47
5	PC	22.58
6	NTC	_

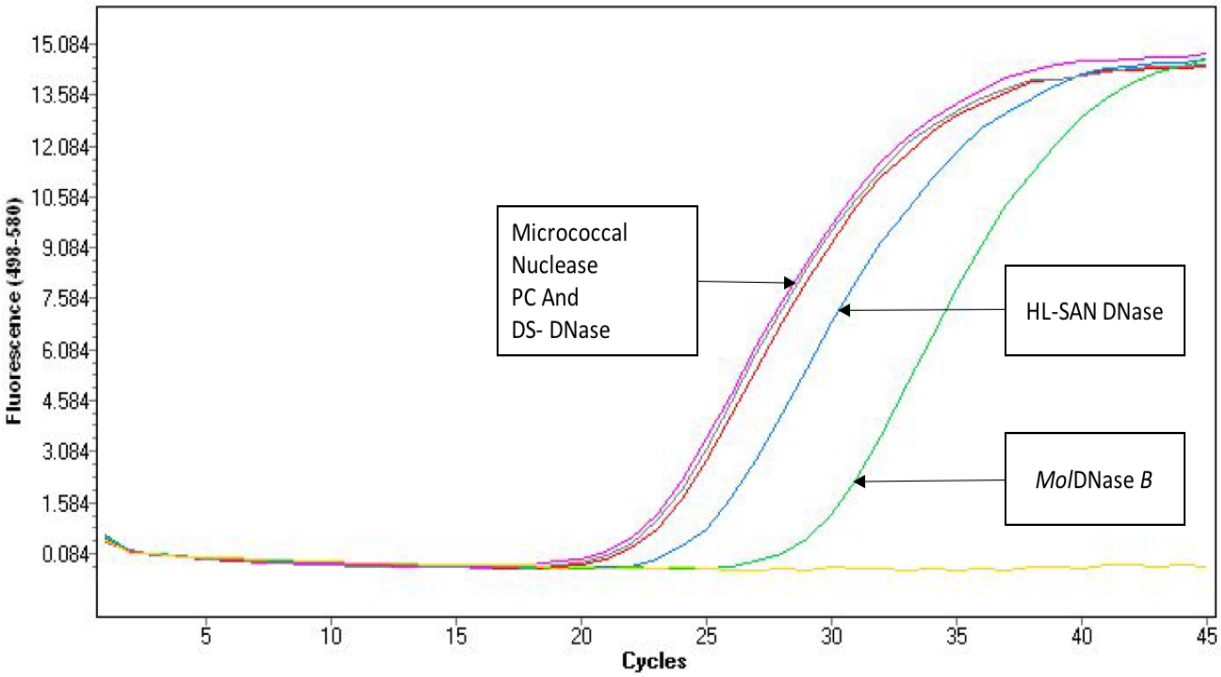


Figure 3- 13 human qPCR amplification curves showing activity of endonucleases on freeze-thawed human blood

After demonstrating the potential of HL-SAN to work in blood, a suitable buffer was made to increase its activity to match or improve on that of the *Mol*DNase B. The HL-SAN enzyme required a high concentration of Mg^{2+} and NaCl (section **2.3.6**). The same amount of HL-SAN enzyme (2 µL) was tested using different volumes of HL-SAN buffer **(Table 3-10).** qPCR results showed buffering HL-SAN DNase with Mg^{2+}

and NaCl resulted in the increased efficiency of HL-SAN to digest human DNA present in the blood samples. Using this combination (HL-SAN buffer and HL-SAN DNase) there was approx. 25-fold increase of activity observed, with human qPCR showing ΔCq 7 depletion in the sample with the highest volume of the buffer (180µL). Almost the same level of human DNA depletion was observed as the known control (*Mol*DNase *B*) buffered with DB1 buffer (Molzym) (**Table 3-10** and **Figure 3-14**).

Table 3- 10 Human qPCR results showing activity of HL-SAN DNase on freeze-thawed human blood, *Mol*DNase B is the control. HL-SAN DNase was buffered with different volumes of HL-SAN buffer.

	Sample ID	**Human qPCR (Cq)**
1	20µL HL-SAN Buffer	22.27
2	100µL HL-SAN Buffer	24.64
3	180µL HL-SAN Buffer	27.02
4	*Mol*DNase B	27.63
5	PC	22.32
6	NTC	_

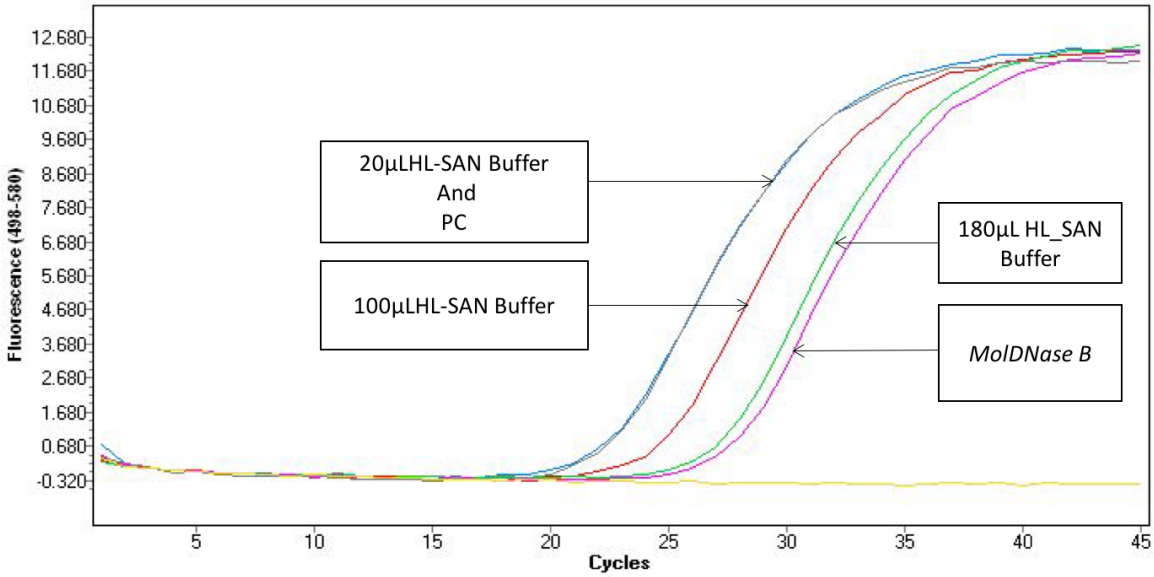


Figure 3- 14 Human qPCR results showing activity buffered HL-SAN DNase on freeze-thawed human blood. *Mol*DNase B was used as the control.

To test the robustness of the optimized HL-SAN DNase method, the experiment was repeated in 200 µL blood sample with 2 µL of the enzyme and an optimised volume of HL-SAN buffer (180 µL), against *Mol*DNase B (2 µL) with DB1 buffer. This experiment confirmed that, under optimized buffer conditions, approximately 100-fold reduction of human DNA could be achieved by HL-SAN DNase digestion. The qPCR results showed HL-SAN DNase could work as well, if not more effectively, than *Mol*DNase B in blood to deplete human DNA (**Table 3-11** and **Figure 3-15**).

Table 3- 11 Human qPCR results showing activity HL-SAN DNase under optimized conditions on freeze-thawed human blood compared to *Mol*DNase B.

	Sample ID	**Human qPCR (Cq)**
1	HL-SAN	30.20
2	*Mol*DNase B	28.83
3	PC	24.45
4	NTC	_

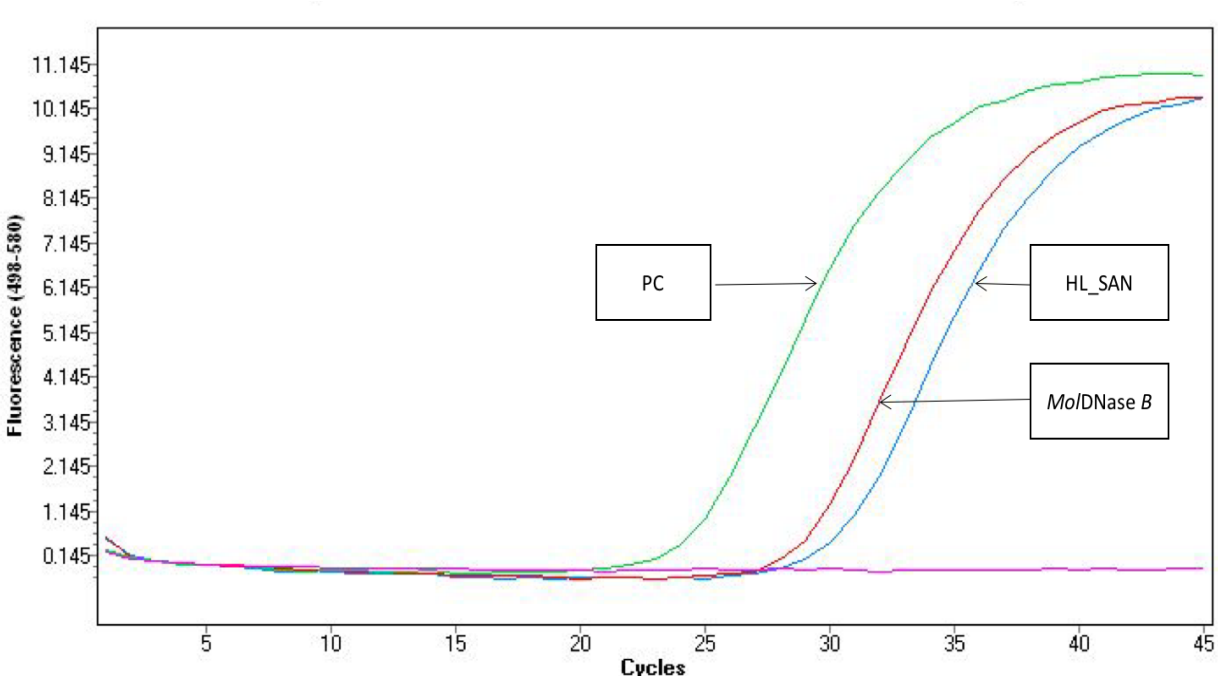


Figure 3- 15 Human qPCR results showing HL-SAN DNase activity under optimized conditions compared to *Mol*DNase B.

After this experiment, HL-SAN DNase and *Mol*DNase B were used interchangeably as the endonucleases of choice for the digestion of released DNA from blood cells in investigating the effectiveness of cytolysin on cell lysis.

3.2.2.2: Differential Lysis Human Leukocytes using Cytolysins

Cytolysins can have a detergent effect on the target cell membrane (e.g., *S. aureus* delta-toxin, a 26 amino acid delta toxin produced by *Staphylococcus*) or form pores in the target cell membrane (e.g. alpha-hemolysin from *S. aureus*, streptolysin O from *S. pyogenes*, and perfringiolysin O produced by *C. perfringens*). We investigated whether cytolysins could be used to differentially lyse human cells, leaving bacteria cells intact, followed by human DNA digestion with endonucleases. A simple combination of enzymatic human cell lysis with enzymatic human DNA digestion was an attractive depletion strategy due to simplicity, specificity and speed.

The first pore forming cytolysins to be investigated for host DNA depletion in human blood samples were streptolysin O (*S. pyogenes*) and alpha-hemolysin (*S. aureus*). To *S. aureus* spiked blood samples, cytolysins were added (individually and in combination) to lyse host cells as explained in Methods section **2.3.4.2.** Samples were then incubated and released DNA from lysed cells was digested with buffered *Mol*DNase B and then a DNase inactivation reagent was added to stop the reaction.

When used alone streptolysin O and alpha-hemolysin showed almost the same level of human DNA depletion, approximately 10^3-fold reduction, and 10-fold increase (10^4) when used in combination. (**Table 3-12** and **Figure 3-16**).

Table 3- 12 Human qPCR results showing lysis activity of cytolysins on human blood.

	Sample ID	**Human qPCR (Cq)**
1	MolYsis	32.04
2	Streptolysin O	31.96
3	α-hemolysin	31.32
4	α-hemolysin and Streptolysin	35.31
5	PC	21.79
6	NTC	_

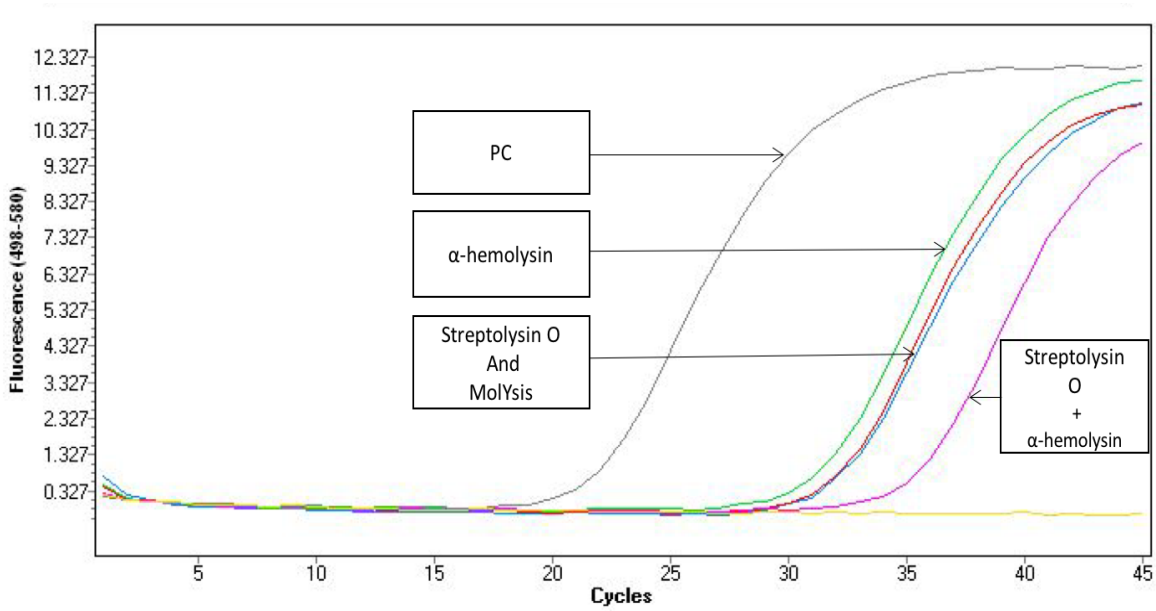


Figure 3- 16 Human qPCR results showing lysis activity of cytolysins on human blood

However, during the host depletion process with cytolysins, a large proportion of spiked *S. aureus* DNA (99%) was also lost (**Table 3-13** and **Figure 3-17**Error! Reference source not found.). This was unexpected as there was no documented evidence suggesting cytolysins produced by bacteria would lyse bacteria protected by a cell wall. The loss of *S. aureus* was therefore assumed not to be due to the lysis activity by cytolysins but rather because of inefficient inactivation of *Mol*DNase B before bacterial DNA extraction. The assumption was supported by the fact that the

loss was the same across all treated samples despite including the MolYsis control (**Table 3-13** and **Figure 3-17**).

Table 3- 13 *S. aureus* qPCR results after cytolysin treatment in human blood

	Sample ID	***S. aureus* qPCR (Cq)**
1	MolYsis	29.68
2	Streptolysin O	29.58
3	α-hemolysin	29.28
4	α-hemolysin + streptolysin O	29.22
5	PC	22.88
6	NTC	> 40

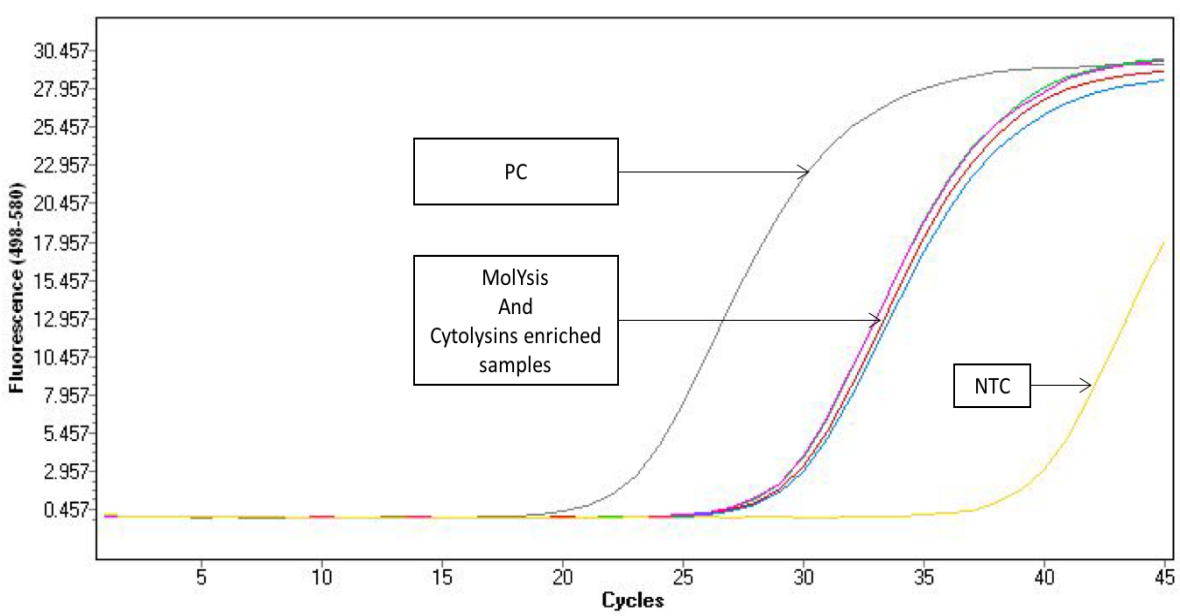


Figure 3- 17 *S. aureus* qPCR results after cytolysin treatment in human blood

These initial results showed that membrane pore forming cytolysins were able to target human cells and lyse them to enable host DNA depletion. The combination of the two cytolysins that produced the greatest human DNA depletion.

This experiment was repeated using combination of enzymes (α-hemolysin and streptolysin O) in 200 µL blood sample. Results showed lower level of human DNA

depletion and similar levels of loss of bacterial signal. Less than 10^2 fold (ΔCq 4.85 – 5.8) depletion of human was observed across all three samples (**Table 3-14** and **Figure 3-18**).

Table 3- 14 Human qPCR results showing lysis activity of cytolysins on human blood

	Sample ID	**Human qPCR (Cq)**
1	α-hemolysin and Streptolysin followed by HL-SAN	26.99
2	α-hemolysin and Streptolysin with HL-SAN	27.94
4	α-hemolysin and Streptolysin followed by *Mol*DNAse	27. 68
5	PC	22.14
6	NTC	–

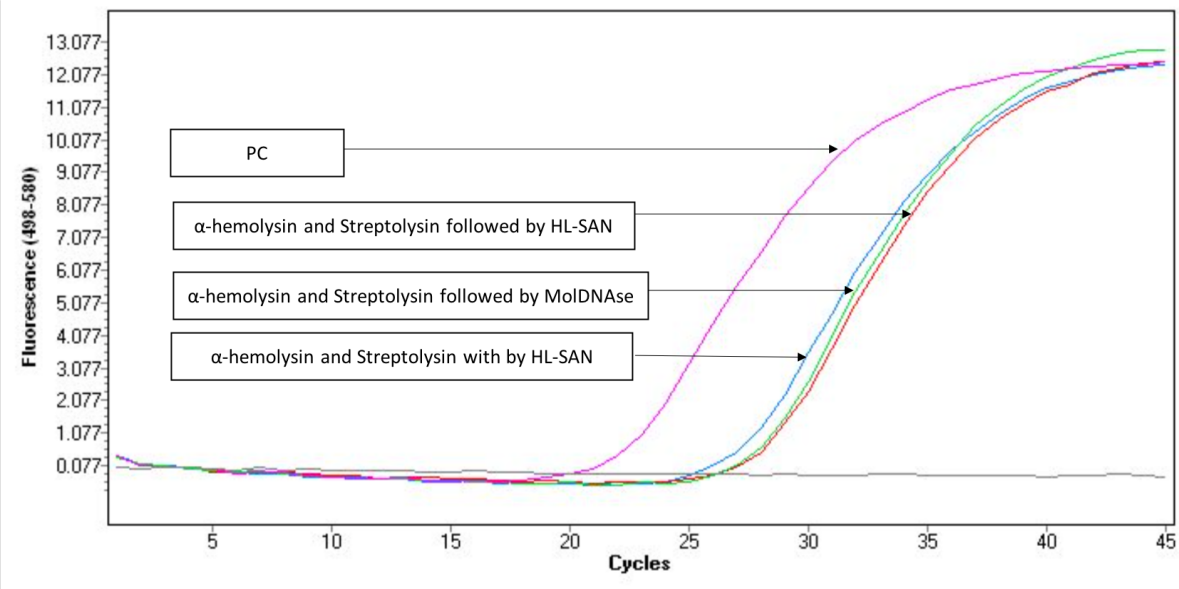


Figure 3- 18 Human qPCR results showing combined lysis activity of α-hemolysin and Streptolysin on human blood

Also, loss of bacteria DNA was the same across all three samples. Less than 2 folds (ΔCq 1.47) loss of *S. aureus* DNA was observed in the samples **Table 3-15, Figure 3-19**

Table 3- 15 *S. aureus* qPCR results after lysis of human cells using α-hemolysin and Streptolysin in human blood

	Sample ID	***S. aureus* qPCR (Cq)**
1	α-hemolysin and Streptolysin followed by HL-SAN	40.00*
2	α-hemolysin and Streptolysin with HL-SAN	40.00
4	α-hemolysin and Streptolysin followed by *Mol*DNAse	40.00
5	PC	38.53**
6	NTC	–

* Cq value suggests <10 cell (<100 cells in total input)
* Cq value suggests <1 cell (<10 cells in total input)

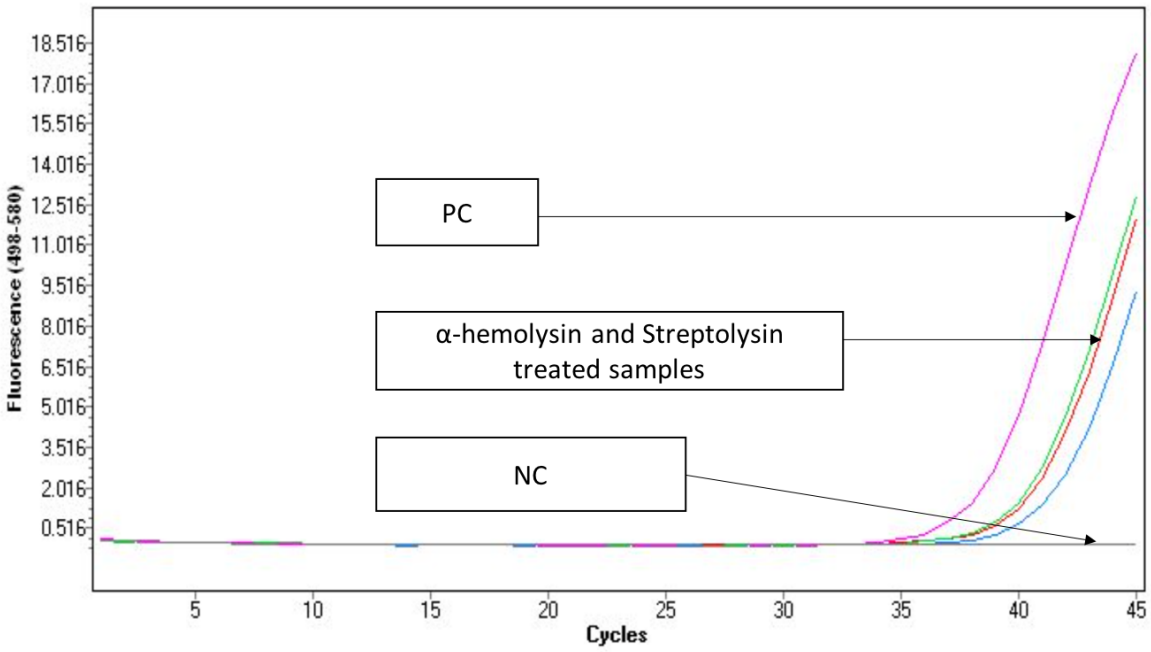


Figure 3- 19 *S. aureus* qPCR results showing effect of combined host lysis activity of α-hemolysin and Streptolysin bacteria signal.

Experiments using streptolysin O and alpha-hemolysin were not very reproducible in our hands (depletion levels achieved were very variable), hence, we changed our focus to another member of the cytolysin family, namely, phospholipase C (PLC) from *C. perfringens.*

PLC was tested for specific host cell lysis and subsequent host DNA digestion using HL-SAN DNase. PLC is a zinc metallophospholipase and requires the presence of zinc for activity (Titball, Leslie, Harvey, & Kelly, 1991); it was however unknown whether the concentrations of zinc in human blood would be sufficient for PLC to work. Also required for PLC activity, are calcium and magnesium ions (Bernardo et al., 2000)(Stonehouse et al., 2002). With these experiments using blood collected with EDTA preservative, there was a concern that EDTA would chelate the required calcium and metal ions necessary for PLC activity. Therefore, PLC was tested on blood with no preservative, blood containing EDTA and on blood in the presence of a metal ion containing buffer.

PLC was added to the various blood sample types and incubated with shaking for host cell lysis. HL-SAN DNase (with HL-SAN buffer) was then added and incubated for host DNA digestion followed by heat inactivation of HL-SAN DNase. PC and NTC samples were included, and DNA was extracted from all samples followed by human qPCR.

Interestingly, after qPCR, results showed the level of human DNA depletion when PLC was tested on blood with no preservative was poor (sample 3). Furthermore, results showed there was no improved PLC activity when buffer was added (Components PLC buffer 0.1M $ZnCl_2$ and 0.1M $MgCl_2$) (sample 2), (**Table 3-16** and **Figure 3-20**). The highest level of host DNA depletion was observed in sample 1, the blood sample preserved with EDTA (with no addition of buffer or metal ions), with approx. 9 Cq difference compared to the positive control, suggesting a 1000-fold depletion of human DNA. Compared to sample 3, without any preservative, there was approximately 5 Cq difference between -EDTA and +EDTA blood as shown below.

Table 3- 16 Human qPCR results showing lysis activity of PLC on blood +/- EDTA and PLC buffer.

	Sample ID	**Human qPCR (Cq)**
1	PLC + EDTA blood	31.04
2	PLC+ PLC buffer + EDTA blood	25.80
3	PLC – EDTA blood	26.12
4	PC	22.85
5	NTC	_

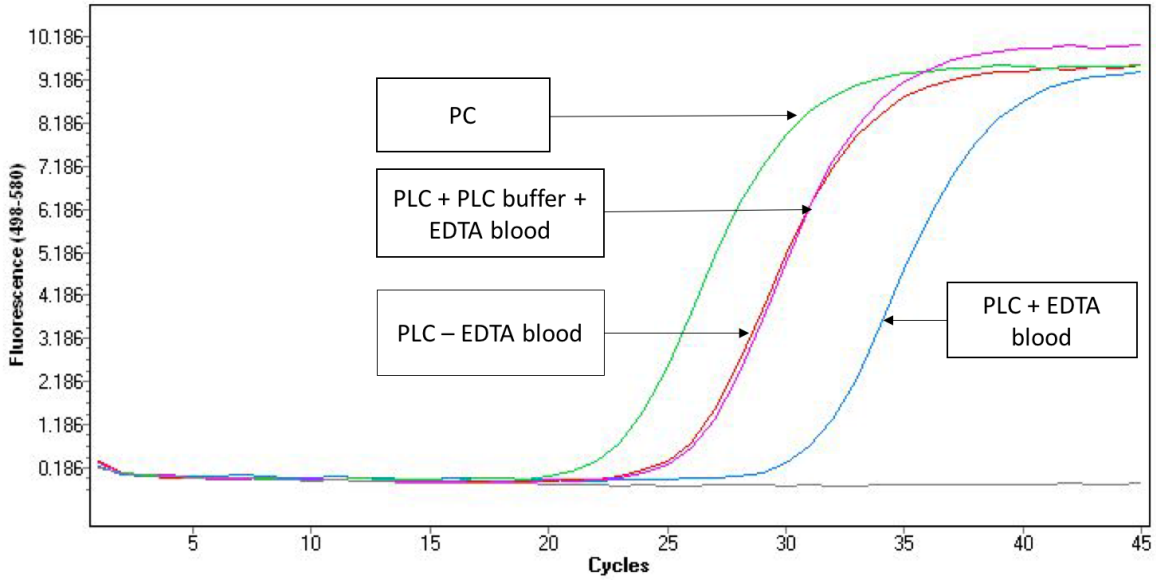


Figure 3- 20 Human qPCR results showing lysis activity of PLC on blood +/- EDTA and PLC buffer.

Further experiments were performed to test whether PLC had any activity on Gram-negative and Gram-positive bacteria and to increase the blood sample volume from 200 µL to 1 mL. When PLC was tested on 1 mL of spiked EDTA preserved blood, two problems were observed. First was poor human DNA depletion in increased blood sample volume and second was the loss of Gram-positive (*S. aureus*) but not Gram-negative (*E. coli*) signal during the depletion process. To improve the level of depletion, PLC incubation was performed in large sample tubes whilst mixing using a Hulamixer® (Life technologies). qPCR showed mixing during incubation improved

depletion from 32 Cq to 38 Cq (100-fold improvement) (**Table 3-17**). The introduction of slow mixing in large sample tubes assisted efficient PLC host cell lysis, resulting in almost complete depletion of human DNA (approximately 1 cell equivalent human DNA remaining; a depletion of ~2.6 x 10^5) (**Table 3-17** and **Figure 3-21**) for 1 ml samples and complete depletion of human DNA for 200 µl samples (a depletion of at least 10^6). However, there was still loss of *S. aureus* but not *E. coli* (**Table 3-17, Figures 3-21, 3-22** and **3-23**).

Table 3- 17 qPCR results showing human, *E. coli* and *S. aureus* DNA quantification after PLC and HL-SAN human DNA depletion on 1mL and 200µL EDTA blood samples.

Sample ID	**Human qPCR (Cq)**	***E. coli* qPCR (Cq)**	***S. aureus* qPCR (Cq)**
PLC_1mL	38.04*	19.83	28.46
PLC_200µL	-**	22.84	32.47
PC	22.38	23.03	26.34
NTC	-	-	-

* Cq value suggests approx.1 cell (<10 cells in total input)

*** No target DNA detected

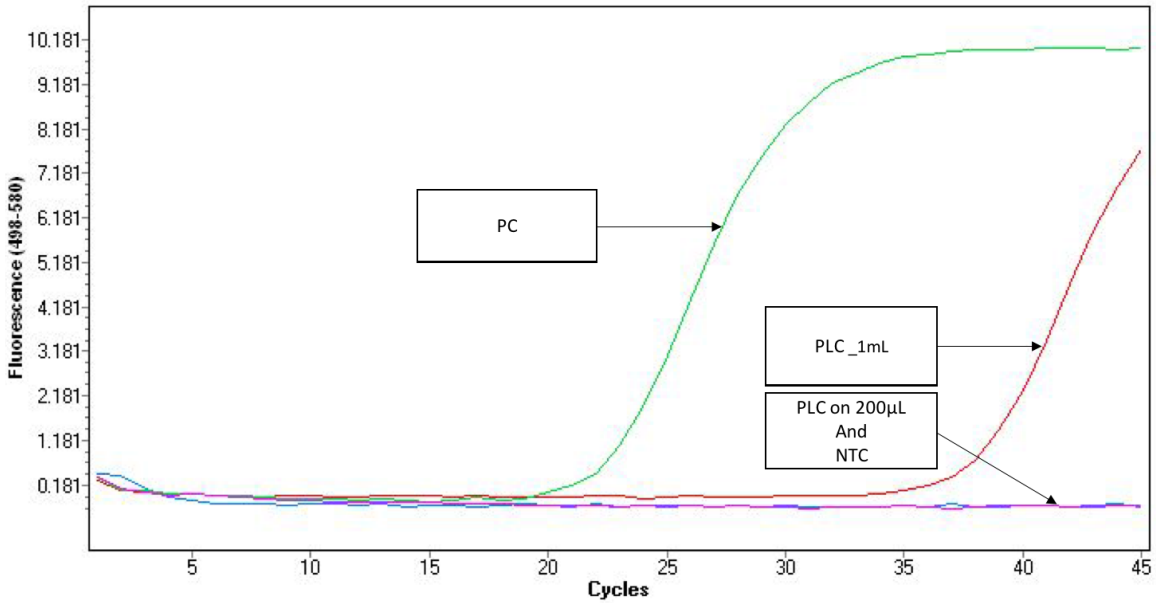


Figure 3- 21 qPCR results showing quantification of human DNA after depletion using PLC and HL-SAN on 1mL and 200µL EDTA blood

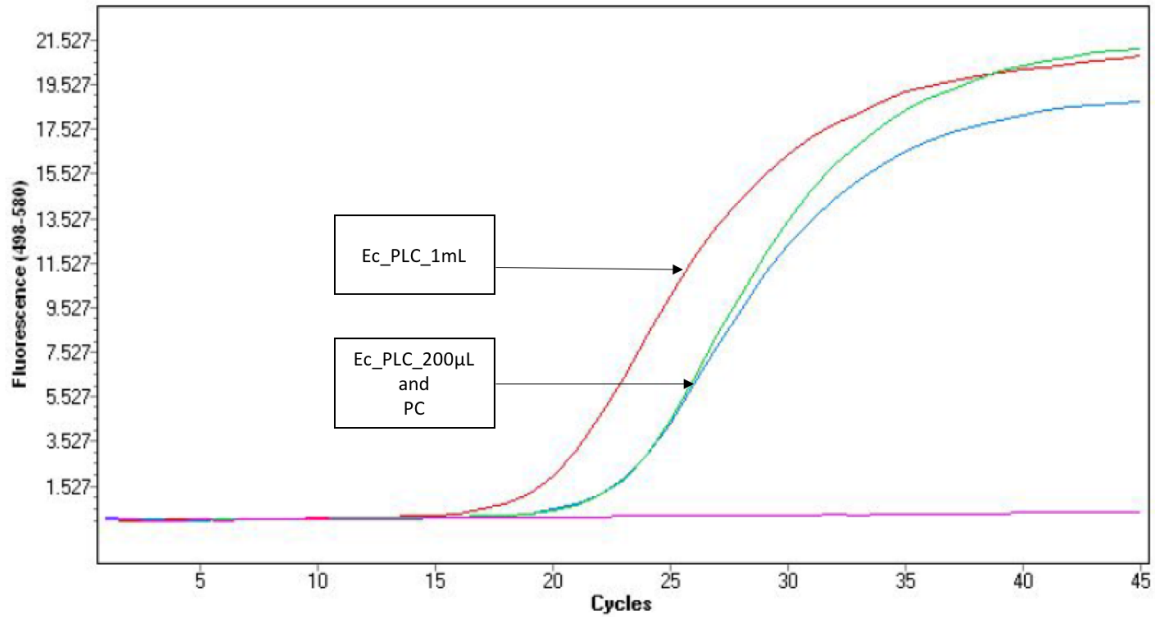


Figure 3- 22 qPCR amplification curve results showing *E. coli* quantification after depletion of host DNA.

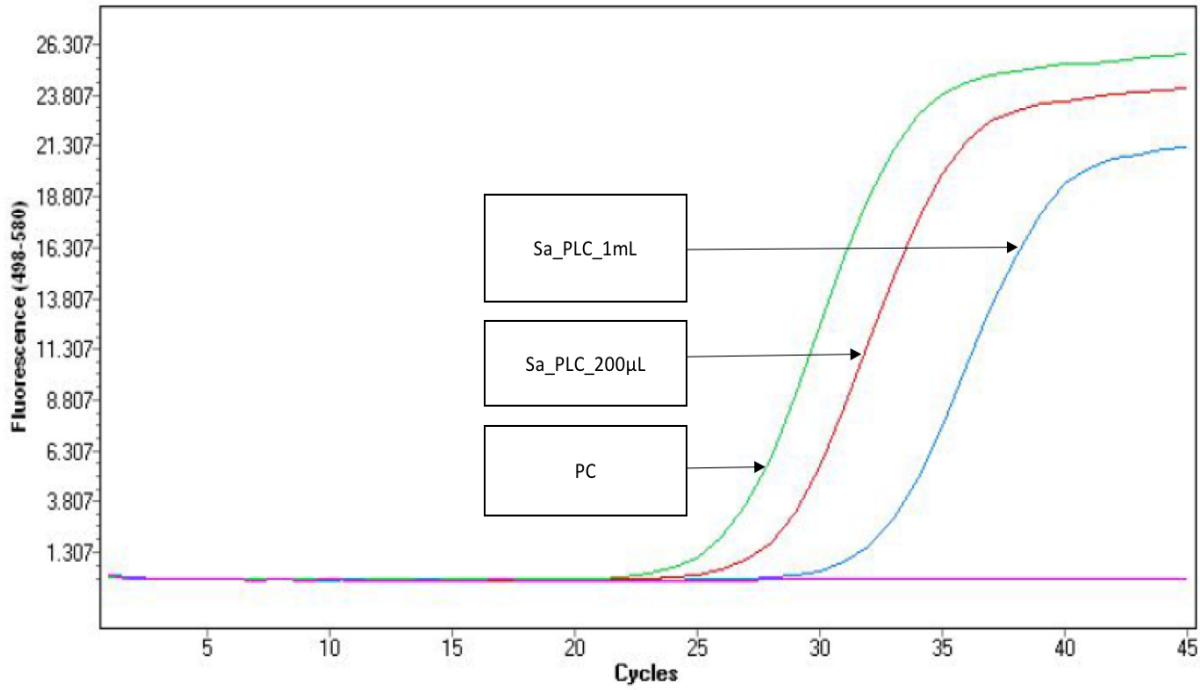


Figure 3- 23 qPCR results showing loss of *S. aureus* during depletion of host DNA

It was hypothesised that heat inactivation of HL-SAN DNase was affecting the cell wall of *S. aureus,* reducing the efficiency of cell lysis, thereby resulting in low recovery levels of DNA. An alternative method of inactivating HL-SAN enzyme by removing the high salt conditions required for its activity was investigated (section

2.3.4.3). After human DNA depletion, DNA was extracted from all samples, followed by qPCR for human, *E. coli* and *S. aureus* DNA.

HL-SAN enzyme inactivation by removing the HL-SAN buffer rather than heat inactivation resulted in a high level of human DNA depletion as seen before but with no loss of *E. coli* or *S. aureus* DNA. Higher amounts of bacterial DNA (approx. 5-fold/2.5 Cq) in a 1 mL blood sample (compared to 200µl control) meant there was no loss of bacterial DNA that has been initially observed (**Table 3-18** and **Figures 3-24**, **3-25** and **3- 26**).

Table 3- 18 qPCR quantitative cycles showing human, *E. coli* and *S. aureus* DNA quantification after PLC and HL SAN host DNA depletion by improved experimental conditions

Sample ID	**Human qPCR (Cq)**	***E. coli* qPCR (Cq)**	***S. aureus* qPCR (Cq)**
PLC_1mL	37.36	17.65	19.04
PC _200µL	21.90	20.17	21.47
NTC	-	-	-

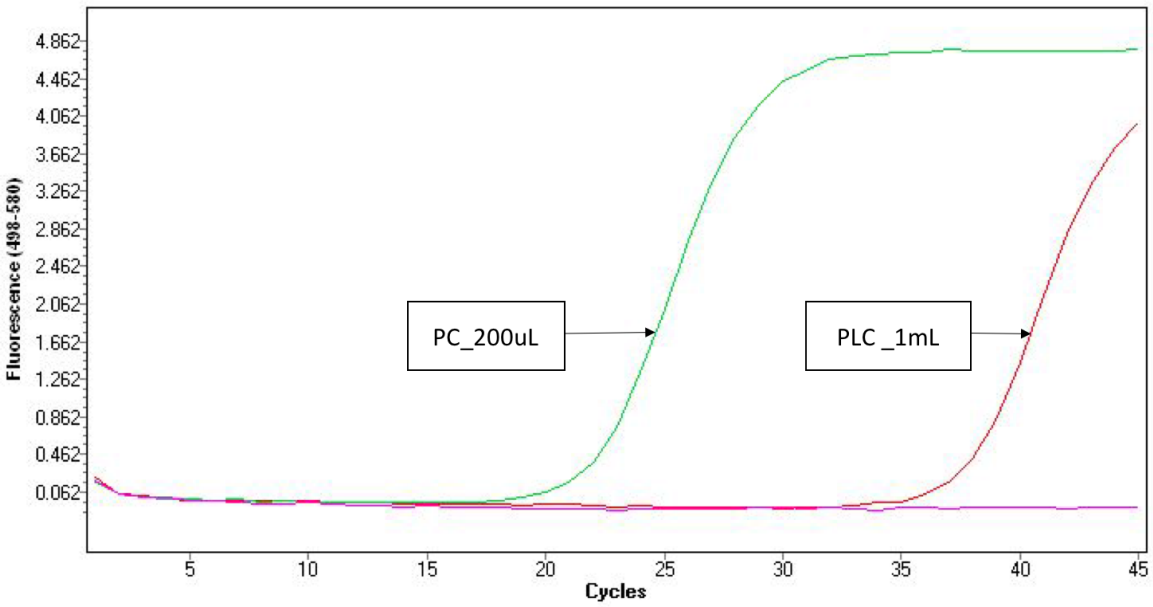


Figure 3- 24 qPCR curves showing quantification of human DNA after depletion using PLC and HL-SAN using optimised conditions

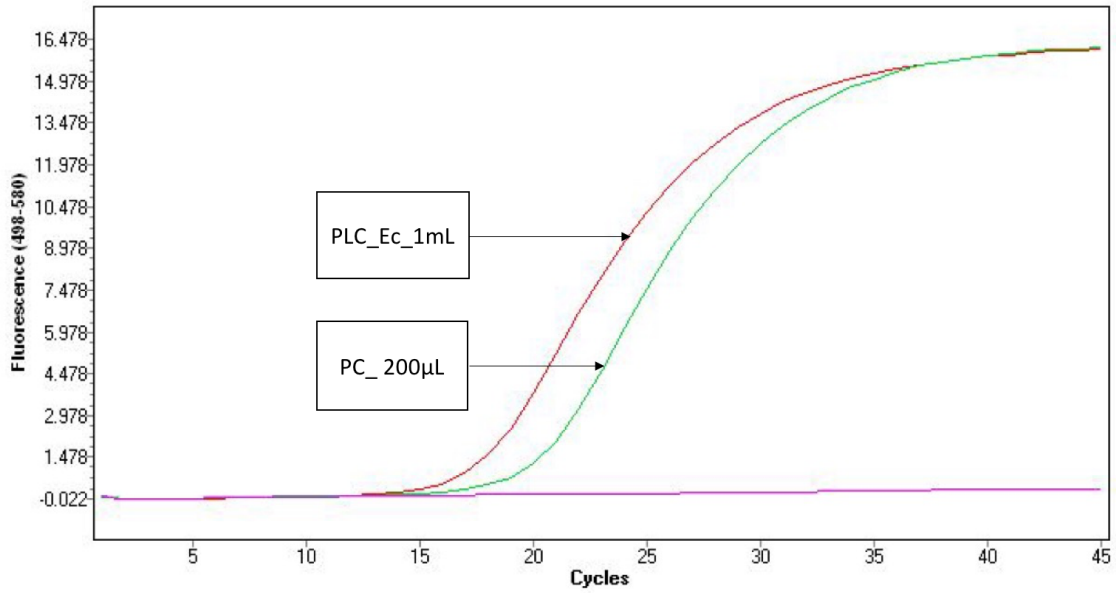


Figure 3- 25 qPCR amplification curves showing quantification of *E. coli* DNA after depletion

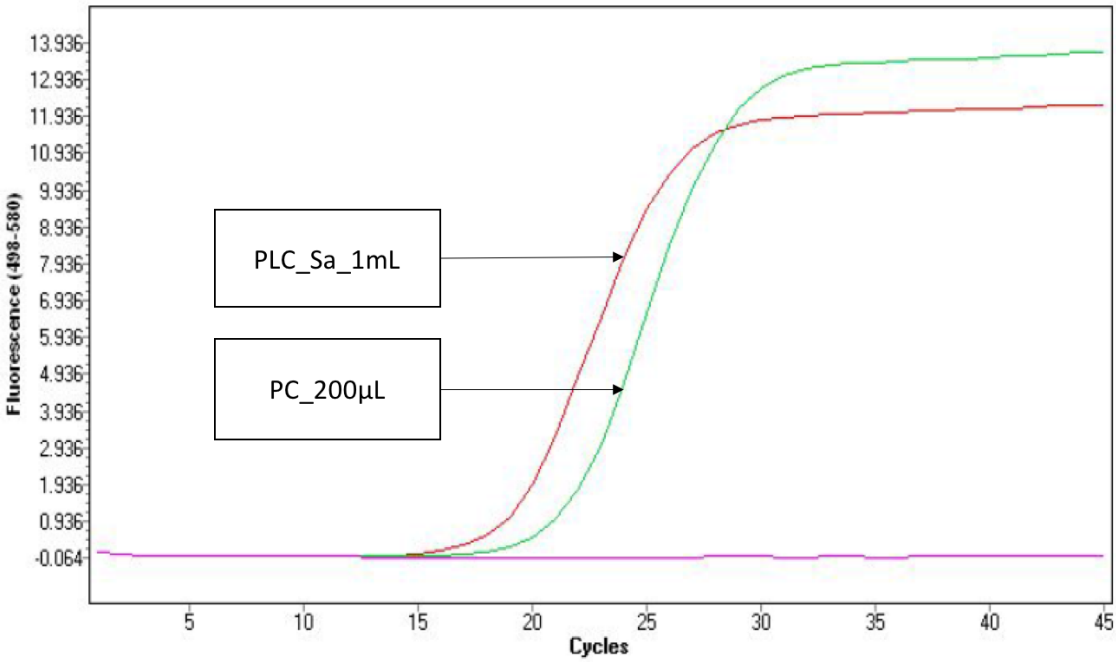


Figure 3- 26 qPCR amplification curves showing quantification of *S. aureus* DNA after depletion

Introducing a buffer exchange to inactive HL-SAN DNase instead of heat inactivation, improved the recovery of *S. aureus* cells. This method alteration enabled efficient *S. aureus* DNA recovery and there was no detrimental effect on *E. coli* DNA recovery or on human DNA depletion. Hence, an efficient cytolysin human DNA depletion procedure had been developed for 1ml blood that did not result in the loss of the microbial component of the blood sample.

After several methodology alterations, the finalised procedure is detailed below.

Optimized Human DNA Depletion Method:

PLC solution: 4mg in 100µl nuclease free water

HL-SAN buffer: 5M NaCl, 100mM $MgCl_2$, pH8.5 in nuclease free water

100µL PLC solution was added to 1ml blood

↓

Incubated at 37°C with gentle mixing for 15min

↓

500µl HL-SAN buffer, 15µl HL-SAN DNase was added and mixed by vortexing, then incubated at 37°C for 15min

↓

Bacterial cells were pelleted at 12,000xg for 10min

↓

Supernatant was discarded

↓

Bacterial cell pellet was resuspended in 1.5ml PBS

↓

Pellet bacterial cells at 12,000xg for 5mins and remove supernatant

↓

Proceeded to DNA extraction of choice

[Depletion time: 45min.]

DNA extraction:

Bacterial cell pellet resuspended in 350µL bacterial lysis buffer (Roche) and vortexed

↓

30µL enzyme cocktail (lysozyme, mutanolysin and lysostaphin – lyticase optional) was added and incubated at 37°C for 15min at 1000rpm

↓

20µL proteinase K (Roche) was added

↓

Mixed by vortexing

↓

Incubated at 65°C for 5min

↓

Proceed to MagnaPure (Roche) for DNA extraction

[DNA extraction time: 45 min; Total time: 90 min]

To confirm the robustness of this procedure, we compared it to the commercially available MolYsis™ and our in-house CD45 IMS combined with MolYsis procedure as described below.

3.2.2.3: Comparison of Cytolysin Human DNA Depletion against MolYsis and the MolYsis plus CD45 IMS method

To test the robustness of the newly developed human DNA depletion procedure, the method was compared to the commercially available MolYsis™ host DNA depletion protocol and the in-house CD45 IMS combined with MolYsis™ protocol (CD45 IMS + MolYsis). The cytolysin human DNA depletion procedure was carried out as described in the optimized protocol above. The CD45 IMS and MolYsis procedure was carried out according to Methods section **2.3.3** and the MolYsis method was performed as per the manufacturer's instructions, explained in Methods section **2.3.1.** DNA was extracted from all samples (including PC) and qPCR was used to quantify human, *E. coli* and *S. aureus* DNA respectively for all methods as detailed in the Methods (**2.6.1**).

When comparing the PLC based human DNA depletion method to the commercially available MolYsis, results showed an approximate 10^4-fold difference in the reduction of human DNA (Δ Cq 12) but comparable levels of bacterial DNA recovery. The CD45 IMS and MolYsis protocol showed the same level of human DNA depletion as the PLC method (**Table 3-19** and **Figures 3-27**, **3-28** and **3-29**).

Table 3- 19 qPCR Cq results showing human, *E. coli* and *S. aureus* DNA remaining before and after human DNA depletion using MolYsis, modified MolYsis and PLC methods.

Sample ID	**Human qPCR (Cq)**	***E. coli* qPCR (Cq)**	***S. aureus* qPCR(Cq)**
PLC + HL SAN	36.05	17.58	18.98
CD45 + MolYsis	36.13	18.74	18.89
Standard Molysis	24.54	17.25	21.33
PC_200µL	21.87	20.13	21.31
NTC	-	-	-

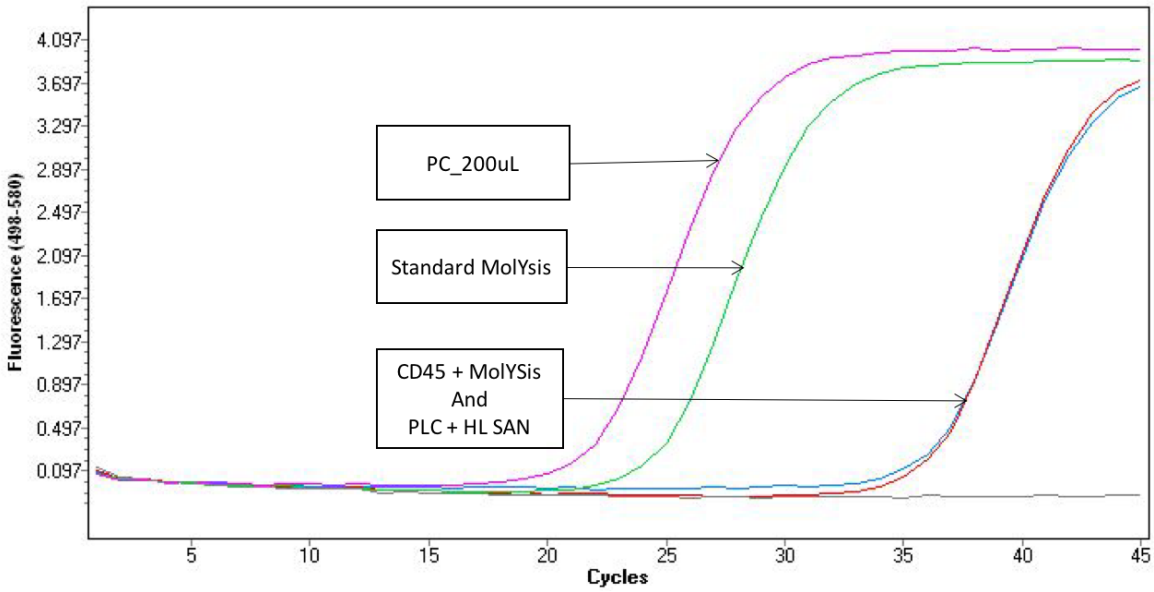


Figure 3- 27 qPCR amplification curves showing human DNA depletion

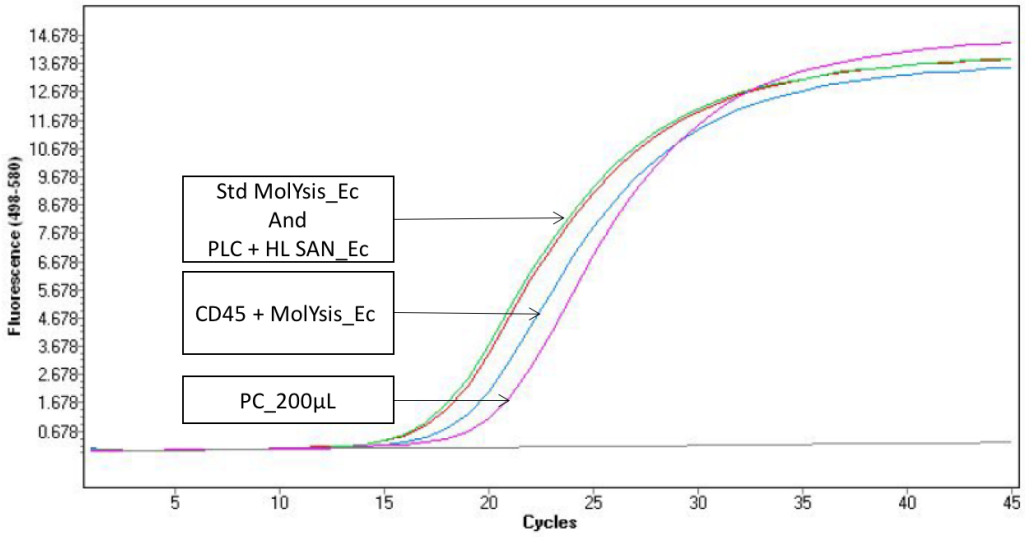


Figure 3- 28 qPCR amplification curves showing *E. coli* quantification after human DNA depletion

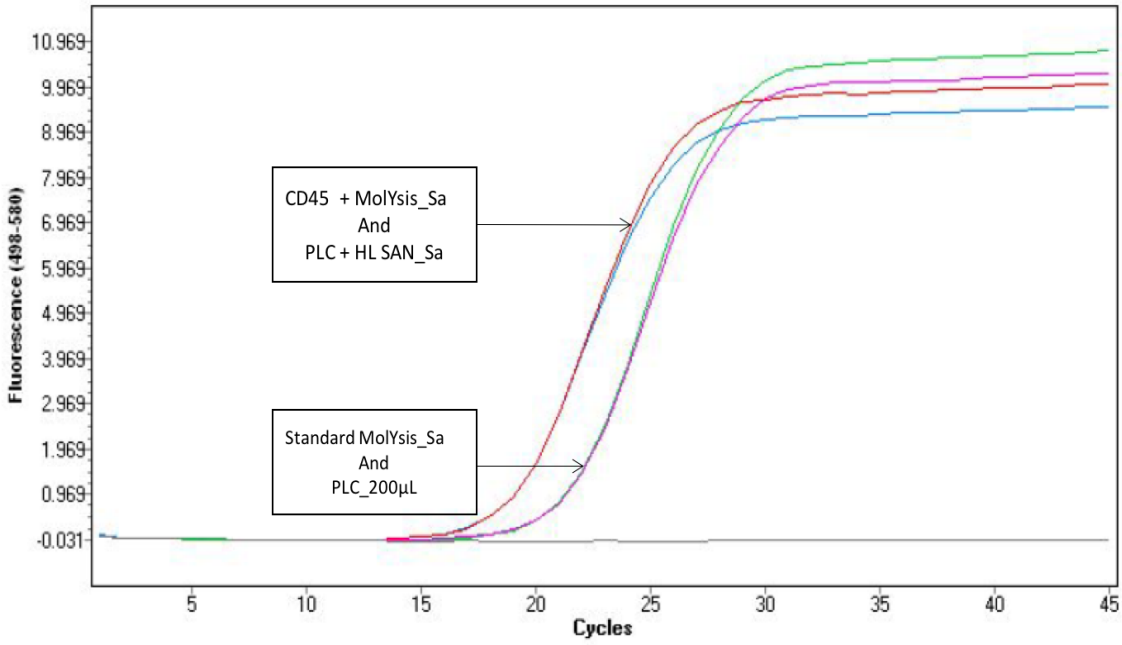


Figure 3- 29 qPCR amplification curves presenting *S. aureus* quantification post human DNA depletion

3.2.3: Method 3: Human DNA Depletion using Propidium monoazide (PMA)

In this method, we investigated the use of an alternative to DNase digestion for the depletion of human DNA. Leukocytes were differentially lysed using the MolYsis differential lysis buffer (Buffer CM) as explained in Methods section **2.3.1,** followed by PMA treatment to remove human DNA. Bacterial DNA was then extracted and qPCR of human and bacterial DNA was performed. qPCR results showed Buffer CM

combined with PMA depletion of human DNA was not sufficient to provide the level of depletion necessary for pathogen identification using metagenomics analysis. **Table 3-20** shows results of human and *S. pneumoniae* following Buffer CM and PMA treatment. The treatment resulted in approximately 90% depletion of human DNA (∆ Cq 5.07), and over 70% loss of bacterial DNA (∆ Cq 1.38) (**Table 3-20** and **Figures 3-30** and **3-31**).

Table 3- 20 showing qPCR results for human and *S. pneumoniae* after depleting human DNA using PMA.

	Sample ID	**Human qPCR (Cq)**	**S. pneumo qPCR (Cq)**
1	PMA_ treated_ 200µL	26.96	31.48
2	PC_ 200µL	21.89	30.10
3	NTC	_	_

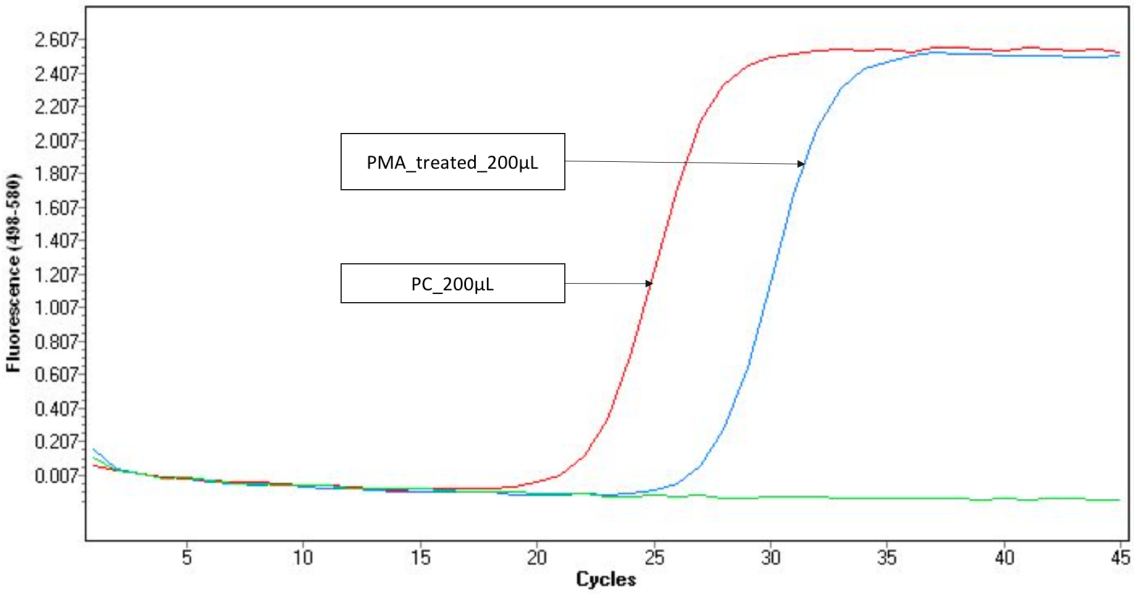


Figure 3- 30 amplification curves obtained using qPCR for human showing host DNA depletion using PMA.

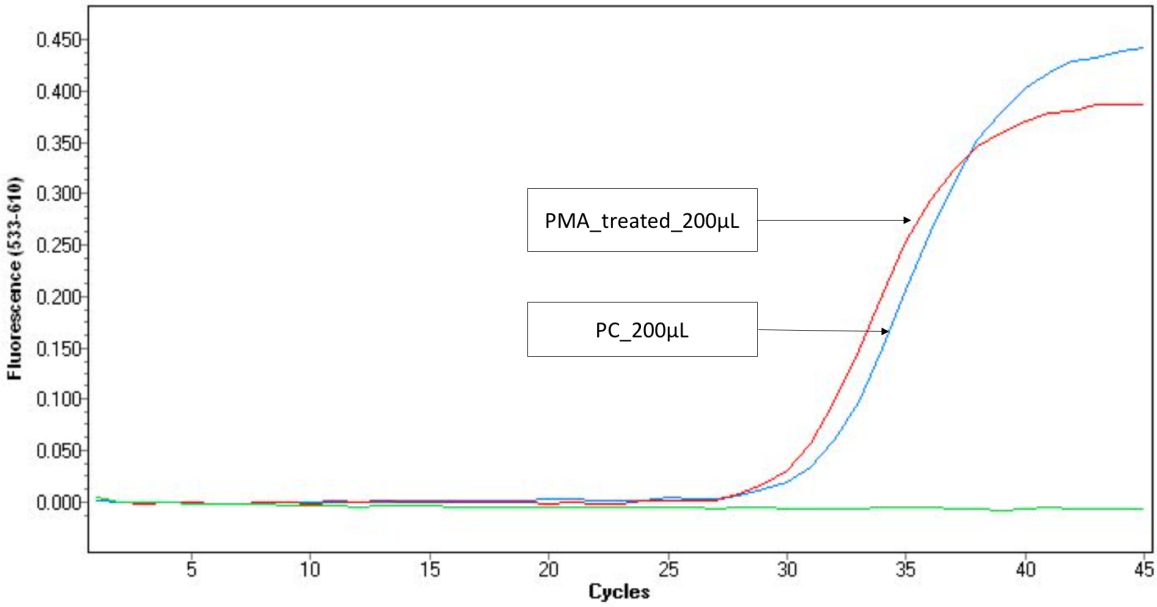


Figure 3- 31 amplification curves obtained using qPCR for *S. pneumoniae* showing host DNA depletion using PMA.

The method was combined with CD45 IMS to increase the level of host DNA depletion. The combined method resulted in 10^6 -fold reduction of human DNA (approx. Δ16 Cq and taking into account the control had 5-fold less sample volume) (**Table 3- 21** and **Figure 3-32**).

Table 3- 21 human qPCR results showing 15.88 quantitative cycles shift (106-fold depletion of host DNA) after depletion using CD45 IMS, differential lysis and PMA DNA degradation.

	Sample ID	**Human qPCR (Cq)**
1	CD45 + PMA_1mL	37.85
2	PC_ 200µL	21.97
3	NTC	_

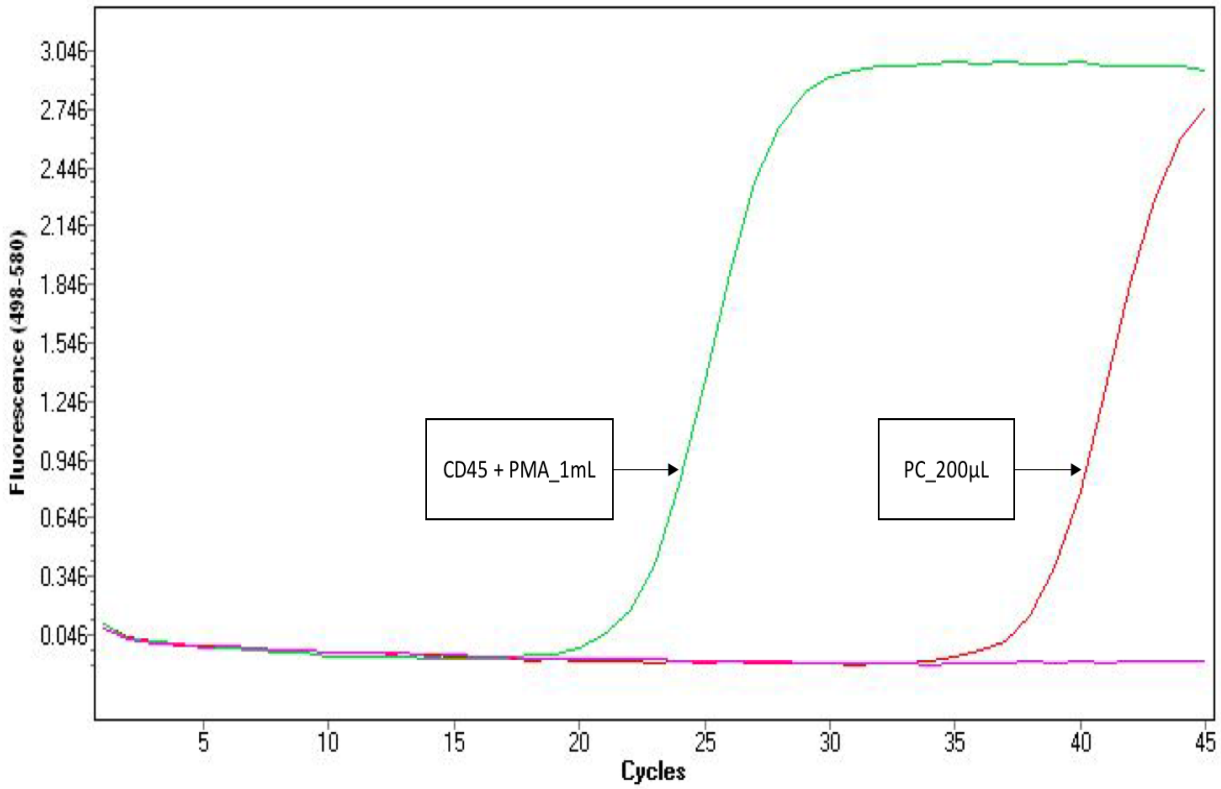


Figure 3- 32 Human qPCR results showing host DNA depletion by CD45 IMS, differential lysis and PMA DNA degradation.

However, the degradation of host DNA by PMA had an undesired effect on bacterial signal. The loss of spiked bacterial DNA was too high (Δ Cq 3.47), > 95% when the 5- fold difference in volume between the sample and the PC was taken into account (**Table 3-20** and **Figure 3-33**).

Table 3- 22 qPCR results showing *S. pneumoniae* DNA loss in the PMA DNA degradation process (ΔCq 3.47).

	Sample ID	***S. pneumo* qPCR (Cq)**
1	CD45 + PMA_1mL	31.65
2	PC_200µL	28.18
3	NTC	_

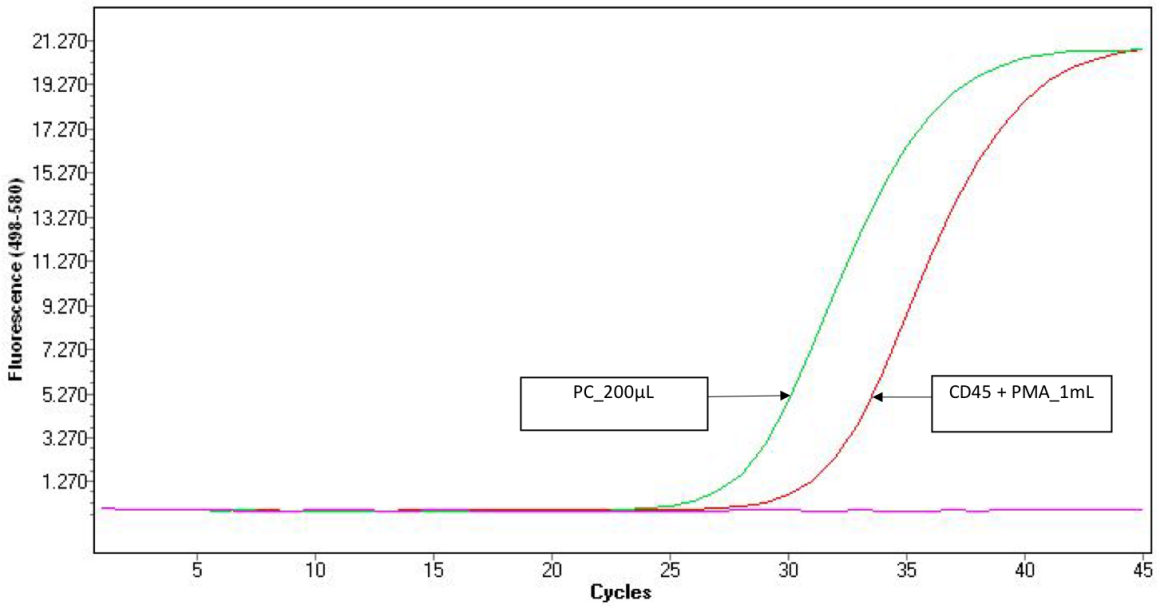


Figure 3- 33 qPCR results showing *S. pneumoniae* DNA loss when DNA was degraded by PMA (ΔCq 3.47).

A possible reason for the observed higher loss of bacterial signal was due to the penetration of PMA into bacterial cell walls which have been damaged by the selective host cell lysis buffer (chaotropic buffer) thus rendering bacterial DNA unamplifiable. This loss of bacteria during the depletion process might have been reduced by titrating the strength of lysis chaotropic buffer or the concentration of PMA and optimizing the photo active incubation time. Troubleshooting of the assay was not feasible within the timeframe of this PhD project so the method was not pursued further.

3.3: Human DNA Depletion Methods in Clinical Urine Samples

3.3.1. Method 1: Host DNA Depletion using NEBNext® Microbiome DNA Enrichment Kit

The NEBNext Microbiome DNA Enrichment Kit (New England Biolabs, MA, USA) facilitates the enrichment of microbial DNA from samples containing methylated vertebrate host DNA (including human), by using beads to selectively bind and remove CpG-methylated host DNA. Up to 90% of human DNA at CpG sites can be methylated but in contrast, methylation at CpG sites in microbial species are rare (Feehery et al., 2013). This method differs from others used in this study as it is not

based on the differential lysis of human cells, rather it is used post DNA extraction to separate human and bacterial DNA.

Human DNA depletion using NEBNext® Microbiome DNA Enrichment kit was performed according to the manufacturer's instruction as described in Methods section **2.4.1.** Total DNA was extracted from *E. coli* infected clinical urine sample (one 2 mL aliquot for depletion and the other as positive control) as explained in **2.6.2**. Host DNA was depleted by adding DNA to MBD2-Fc Magnetic beads, and the Supernatant Fraction containing enriched microbial DNA was retained and the bead fraction containing human DNA was discarded.

Human qPCR and 16S rRNA assays were used to assess the efficiency of host DNA depletion, as explained in Methods section **2.9**. The table below shows qPCR results of three clinical urine samples depleted of host DNA compared to their respective positive controls. In this experiment the best depletion (10^2 –fold depletion) was observed in sample 2, which had the highest initial amount of host DNA, the other two samples showed approximately 10-fold depletion of host DNA (**Table 3-23** and **Figure 3-34**). The bacterial 16s rRNA qPCR showed no loss of pathogen signal except for sample1 (Δ Cq 20.1), due to set-up error (DNA not added to the tube). 16s rRNA qPCR assay for the sample 1 was repeated with DNA template carefully added to sample 1 and the respective positive, results showed no loss of bacterial signal (Δ Cq 0.15) (**Figure 3-36**).

Table 3- 23 qPCR results of three clinical urine samples depleted of human DNA using Microbiome kit compared to their respective positive controls.

	Sample ID	Human qPCR (Cq)	Bacterial rRNA 16s (Cq)
1	Sample 1	34.86	11.33
	Sample 1 PC	31.50	31.43
2	Sample 2	31.29	18.92
	Sample 2 PC	24.44	18.22
3	Sample 3	28.36	12.87
	Sample 3 PC	26.35	13.81
4	NTC	-	30.26

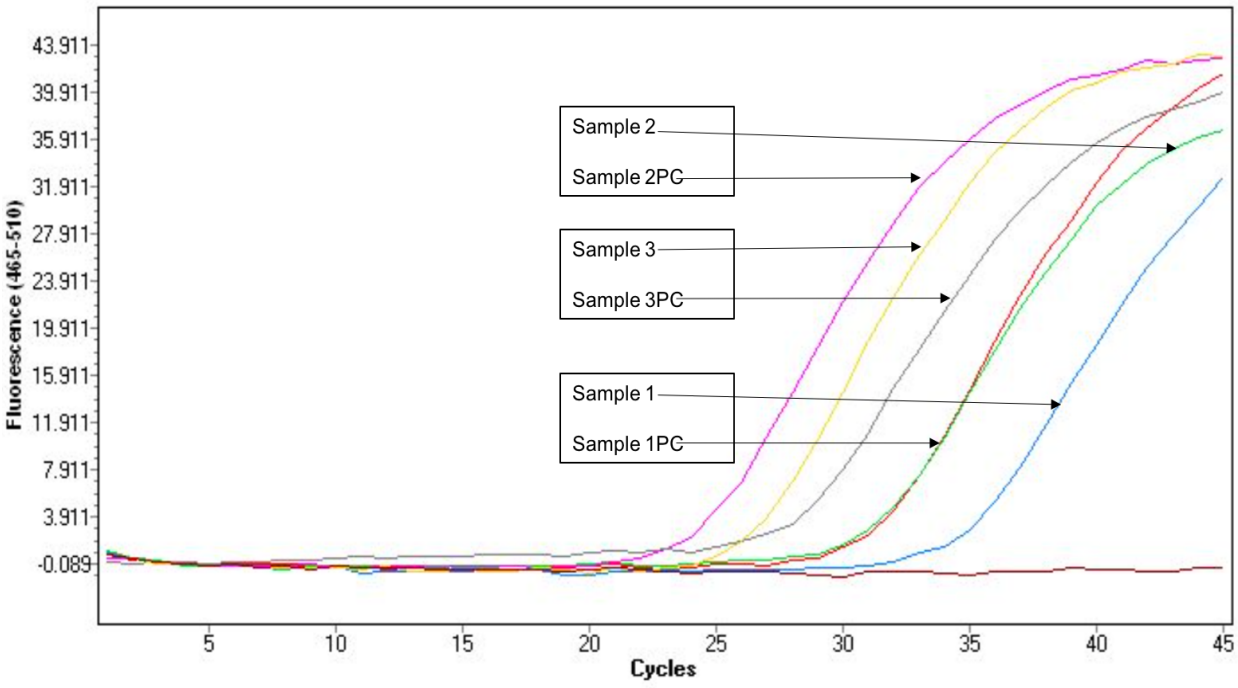


Figure 3- 34 qPCR amplification curves showing 3 urine samples depleted of human DNA using Microbiome Enrichment and their respective un-depleted samples (PC).

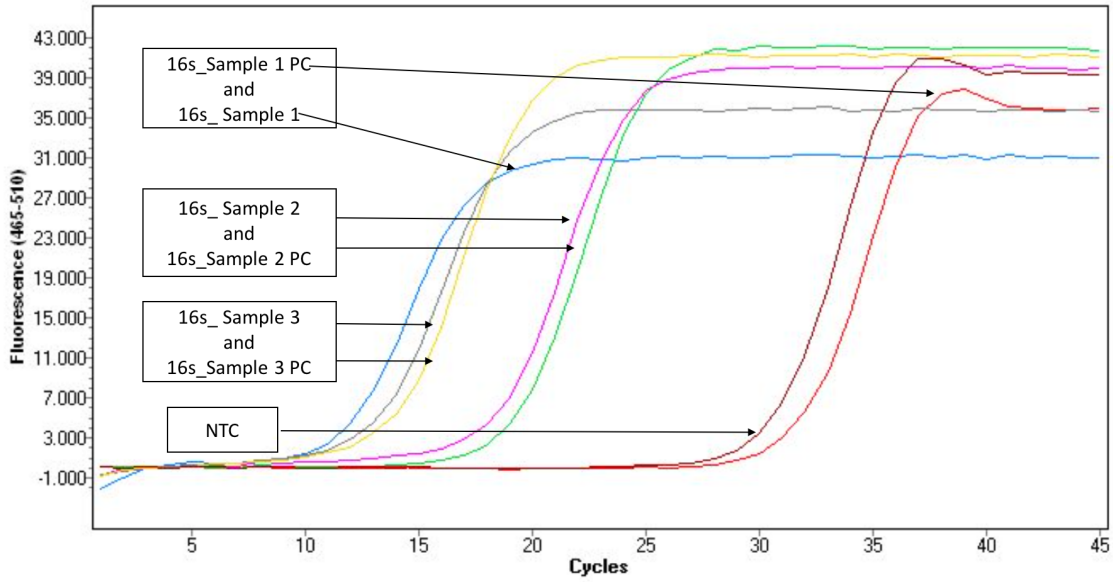


Figure 3- 35 16s rRNA qPCR amplification curves showing bacterial signal after human DNA depletion using Microbiome Enrichment kit in the 3 urine samples

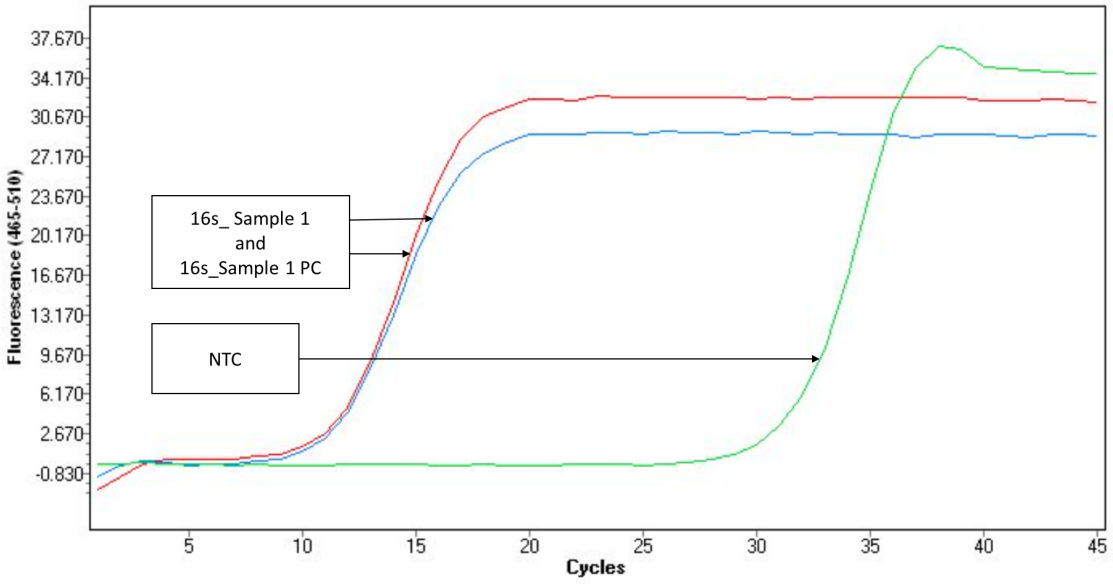


Figure 3- 36 repeated 16s rRNA qPCR assay showing no loss of bacterial signal in sample1 after human DNA depletion by Microbiome Enrichment.

3.3.2: Method 2: Human DNA Depletion by MolYsis™ Basic5 Kit

To test the MolYsis™ kit on urine, 2 mL of clinical sample (heavily infected with *E.coli*) was used for human DNA depletion following manufacturer's instructions as

detailed in Methods section **2.3.1,** and another 2 mL aliquot as positive control. Thereafter, DNA was extracted from depleted and un-depleted samples as explained in section **2.6.2** followed by human and 16S rRNA qPCR assays as described in section **2.9**. **Table 3-24** (sample 1) and **Figures 3-37** and **3-38** shows 10^2-fold depletion of host DNA using MolYsis and 10-fold loss of bacterial DNA during the process.

3.3.3: Method 3: Human DNA Depletion by Differential Centrifugation Combined with MolYsis™ Basic5 Kit

In this method, host cells were first separated from pathogens based on the difference in size and density by differential centrifugation. Differential centrifugation on its own did not provide sufficient depletion (removed approximately 90% of leukocytes). It was therefore combined with the MolYsis assay to deplete human DNA in 2 mL of clinical urine sample infected with *E.coli*. The method was performed as described in Methods section **2.4.3.** Human cells were pelleted by differential centrifugation at low relative centrifugal force (300 x g) for 2min leaving most of the bacteria in suspension. The pellet was discarded and the supernatant was depleted of the remaining leukocytes using MolYsis as described in section **2.3.1.** Thereafter, DNA was extracted as explained in section **2.6.2** followed by qPCR analysis. Results of the assay are shown in **Table 3-24** (sample 2**)** and **Figures 3-37** and **3-38,** in which approximately 10^3 –fold depletion of human DNA and 10-fold loss of bacterial DNA was observed.

3.3.4: Method 4: CD45 IMS in combination with MolYsis™ Basic5 kit

Host DNA depletion using CD45 IMS in combination with MolYsis was performed as described in section **2.3.2,** followed by DNA extraction from *E. coli* clinical urine samples (section **2.6.2**) and analyzed using qPCR as described in section **2.9. Table 3.22** below shows real time qPCR results of host DNA depletion using the MolYsis kit when combined with differential centrifugation (DC) (sample 3). The three methods were tested using the 2mL aliquots from the same clinical sample/patient and compared with the corresponding positive control. The minimum depletion of less than 100-fold (Δ Cq 5.22) was observed when using MolYsis alone. MolYsis combined with CD45 IMS or DC was more effective, depleting approximately 10^3 –

fold human DNA, (Δ Cq 9.21 and Δ Cq 9.7 for the MolYsis with DC and CD45 IMS respectively). The loss of bacterial signal across all three methods was similar, Δ Cq 3 for MolYsis by itself and Δ Cq 3.66 and Δ Cq 3.86 when combined with DC and CD45 IMS respectively (**Table 3-24** and **Figure 3-37** and **3-38**).

Table 3- 24 Human and 16S qPCR results showing host DNA depletion and some bacterial DNA loss using the three depletion methods compared to the positive control.

	Sample ID	**Human qPCR (Cq)**	**Bacterial rRNA 16s (Cq)**
1	MolYsis	30.10	18.11
2	DC + MolYsis	33.99	18.77
3	CD45 + MolYsis	34.48	18.97
4	PC	24.78	15.11
5	NTC	-	30.25

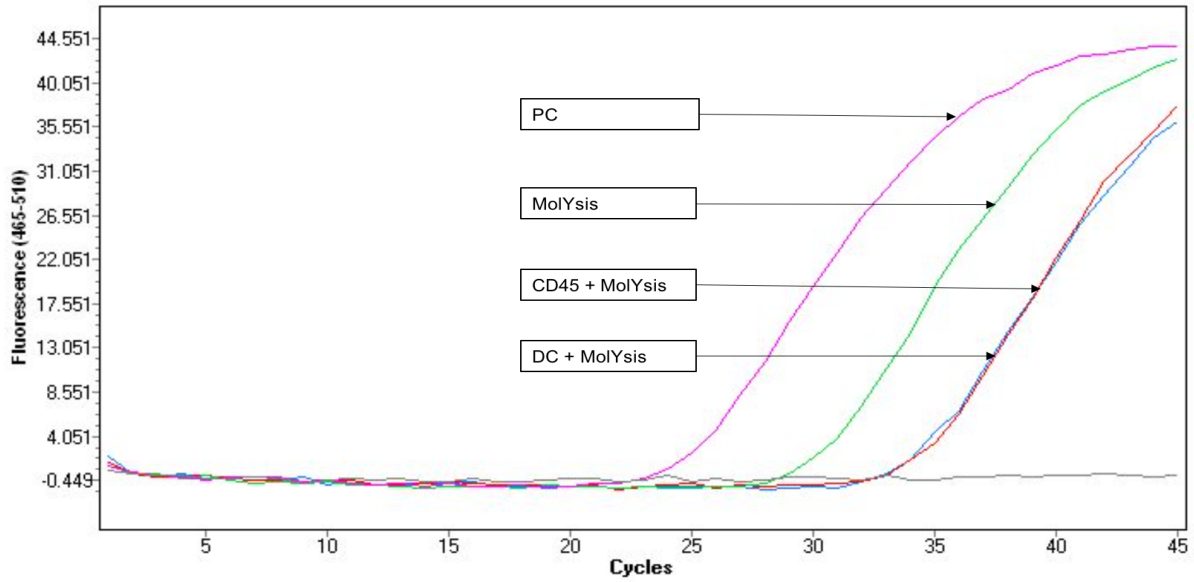


Figure 3- 37 qPCR amplification plots showing host DNA depletion in urine using the 3 depletion methods CD45 + MolYsis, DC+ MolYsis and MolYsis.

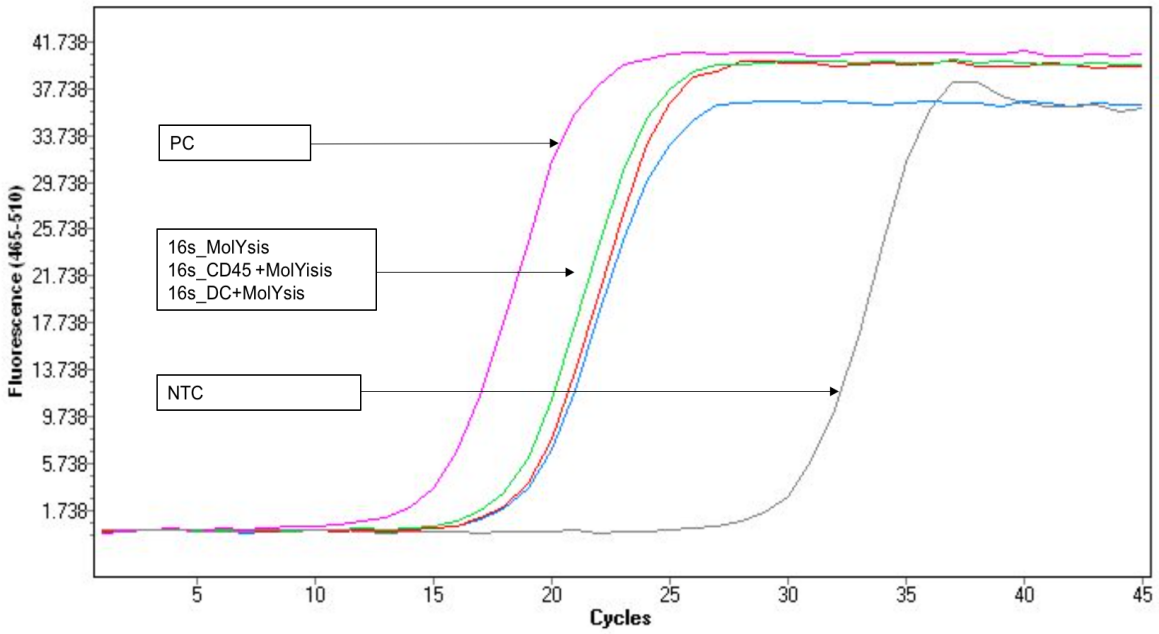


Figure 3- 38 16S qPCR amplification curves showing loss of bacterial DNA after host DNA depletion using the 3 methods

3.3.4.1: Comparison of four DNA depletion strategies in urine

After testing all four methods, three urine samples (biological replicates) from patients with a UTI infection with at least 10^5 CFU/mL were collected for comparing the four methods as described (section **2.4**). Each assay was repeated three times using 2 mL from each sample of clinical urine, all methods were performed as described above. Samples were assessed by qPCR after depletion. Cq values of the four qPCR assays were used to plot the graphs below. The higher the human Cq the better the depletion of host DNA. Results showed differential centrifugation combined with MolYsis (DC + Molysis) was the most efficient method by consistently depleting 10^3 to 10^4 – fold of human DNA. The least efficient methods were NEB microbiome enrichment kit and the MolYsis kit depleting up to 10 –fold human DNA. MolYsis combined with CD45 IMS depleted 10 to 10^3-fold human DNA (**Figure 3-39).**

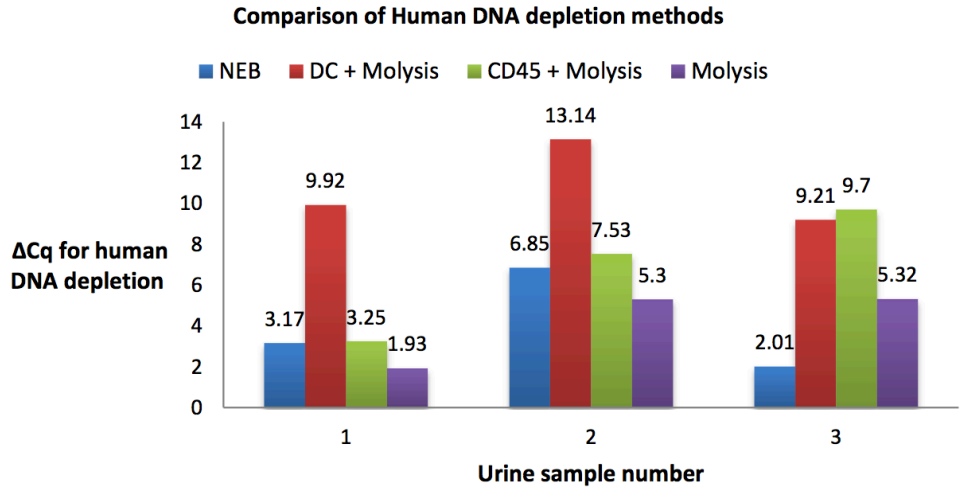


Figure 3- 39 Human DNA depletion using the 4 methods in replicate urine samples measured by ΔCq compared to an un-depleted control

Figure 3-40 below shows average human DNA depletion for the four methods with corresponding average loss of bacteria. The differential centrifugation and MolYsis (DC+M) method showed higher host DNA depletion (Approx 99.5%) and higher levels of bacterial loss (Approximately 10-fold). The least effective methods for human DNA depletion were MolYsis and NEB both with 10-fold depletion of host DNA, however, the methods recovered higher amounts of bacterial DNA (approx. 25% loss of bacterial signal) **Figure 3-40.**

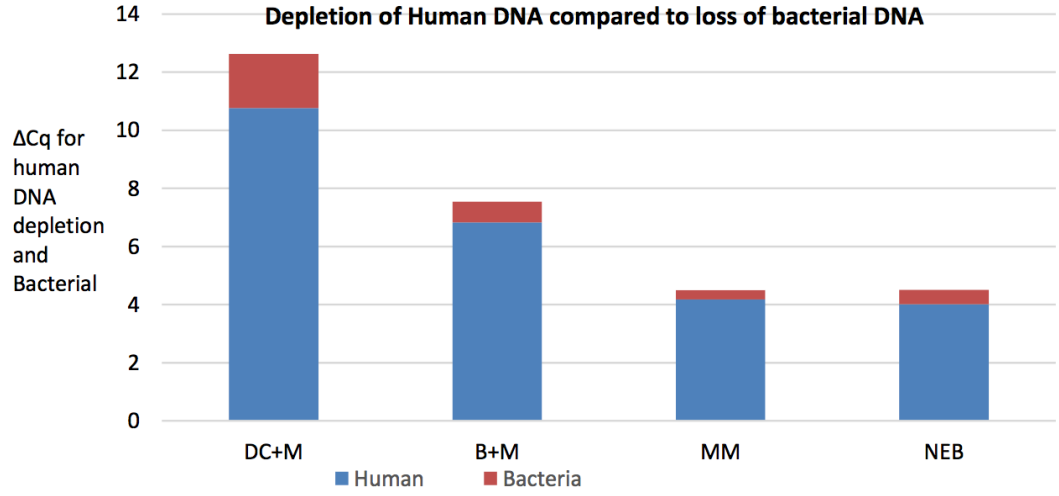


Figure 3- 40 Graph showing the average depletion of the four DNA depletion strategies and the corresponding average bacteria loss (DC+M = differential centrifugation plus MolYsis; B+M = CD45 IMS plus MolYsis; MM = MolYsis; NEB = New England Biolabs Microbiome enrichment kit)

Although the loss of bacterial signal was higher following processing by differential centrifugation combined with MolYsis, the level of human DNA depletion (more than 10^4 -fold) compensated for the loss of bacteria. Furthermore, DC + MolYsis was cheaper and quicker compared to CD45 IMS + MolYsis. Hands on time from receipt of the sample to the extracted DNA for the MolYsis combined with CD45 IMS was approximately two hours and 45 minutes whereas DC with MolYsis required less than two hours. Hence DC combined with MolYsis was the method of choice for depleting human DNA in urine samples for pathogen identification using metagenomics sequencing. Analysis of limit of the detection of the assay was not required as the method was performed on clinical urines infected with more than 10^6 bacterial cell per mL following initial CFU count analysis.

3.4. Methods for *C. difficile* enrichment from Clinical Stool Samples

3.4.1: Method 1: *C. difficile* Spore Separation using Nycodenz

The aim of this series of experiments was to develop novel methods to enable the diagnosis and genotyping of *C. difficile* directly from stool samples using metagenomics. Initial experiments in stool samples were focussed on enriching *C. difficile* spores form stool, separating them from human and commensal microbiota. *C. difficile* spores were separated using Nycodenz medium, a commonly used method to separate microbiota for metagenomic studies, followed by human DNA was depleted by MolYsis following the manufacturer's instruction.

Excess stool samples from patients with confirmed CDI were collected from NNUH clinical microbiology lab, commensal microbiota and human cells were separated as explained in section **2.5.1.** Collected cells were equally aliquoted into two different tubes. Total DNA from one cell aliquot (positive control) was extracted as described in section **2.6.2** (established method for DNA extraction from urine samples)**,** with the addition of RNase solution (Qiagen). The second aliquot was treated with MolYsis to remove human DNA as described in **2.3.1**. After MolYsis treatment, DNA from the host DNA depleted pellet was extracted as for un-depleted sample. qPCR analysis after MolYsis treatment showed approximately 99% (ΔCq 6.41) depletion of human DNA in sample 1141. Initial amount of human DNA in sample 1142 was low, of

which all was depleted (ΔCq 0.92). *C. difficile* qPCR probe based assay showed approximately 10-fold loss of *C. difficile* signal in sample 1141 (ΔCq 3.35) and 20-fold (95%) in sample 1142 (ΔCq 5.01). The 16S rRNA assay showed there was still a very high amount of other commensal bacteria (**Table 3-25** and **Figure 3-41**, **3-42** and **3-43**)

Table 3- 25 Human, *C. difficile* and 16S qPCR results of two stool samples and respective positive controls after depletion using MolYsis.

	Sample ID	**Human (Cq)**	**C. difficile (Cq)**	**16s rRNA (Cq)**
1	1141 MolYsis	_*	29.91	8.46
2	1141 PC	33.59*	24.90	9.47
3	1142 MolYis	_	27.12	_
4	1142 PC	39.08	23.77	_
5	NTC	_	_	33.49

*No target DNA was detected

*** Cq value suggests <10^2 cells

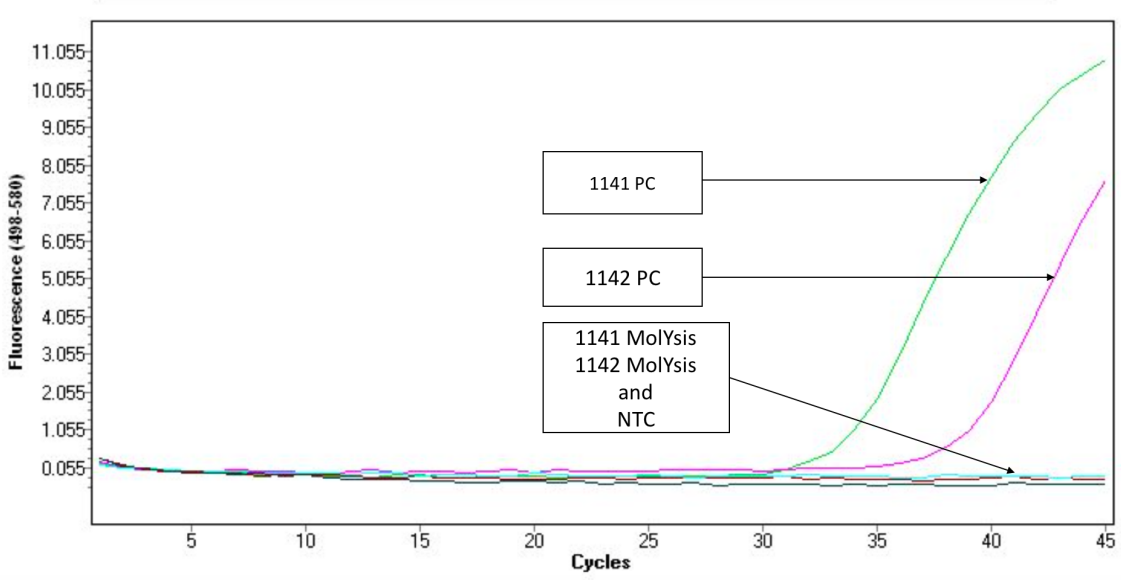


Figure 3- 41 Human qPCR amplification curves after depletion of human DNA in stool sample using MolYsis.

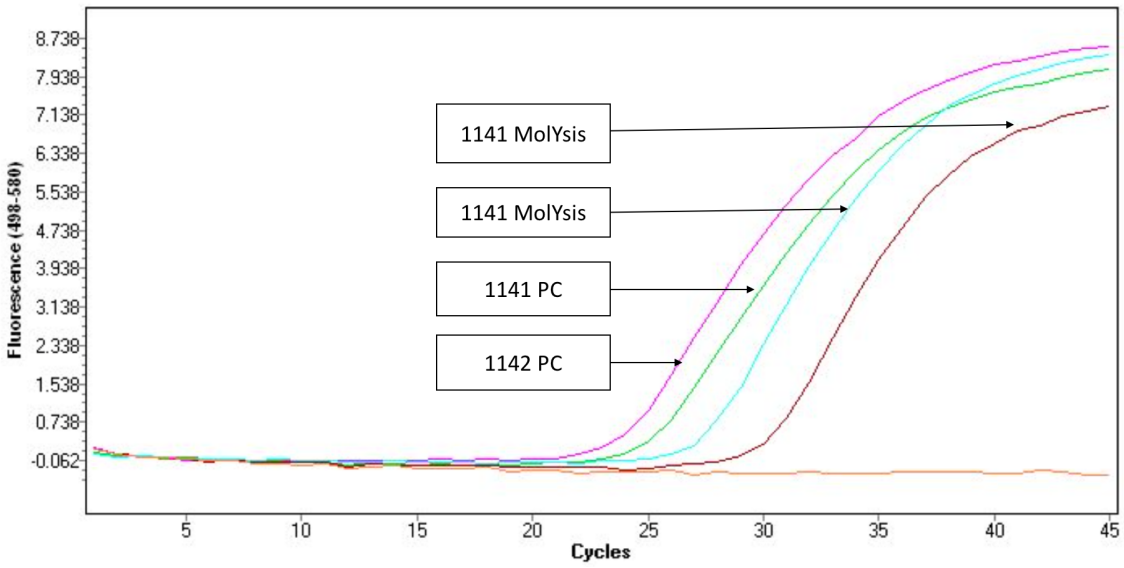


Figure 3- 42 *C. difficile* qPCR amplification curves after depletion of human DNA.

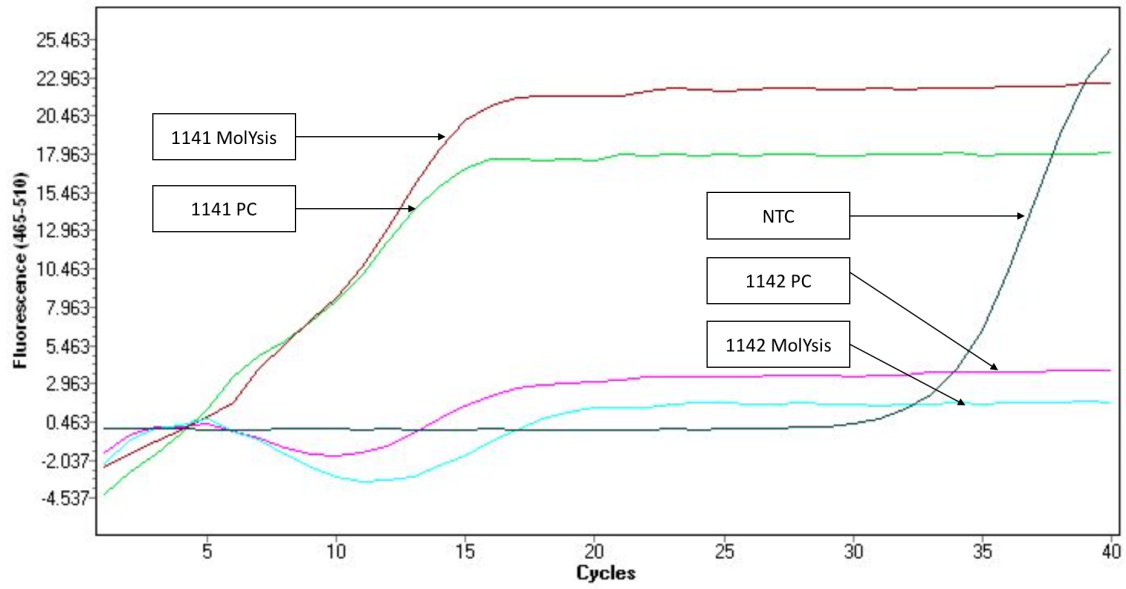


Figure 3- 43 qPCR results of 16s rRNA assay showing bacteria community before and depletion of human DNA in stool sample using MolYsis.

Depleted and un-depleted samples (1141 MolYsis, 1141 PC and 1142 PC) were sequenced on MiSeq (Illumina) for further analysis. Nextera XT Illumina library was prepared using the extracted DNA as described in section **2.11.2**. Sequences were identified and classified using kraken and presented in TSV files and Krona charts.

Illumina sequencing produced 1,302,146 total number of reads for the depleted sample 1141 (1141 MolYsis). Unclassified reads by Kraken were 63% and the rest were classified as bacteria (37%), only 0.02% were classified as *C. difficile* (**Table 3-27**). *Akkermansia muciniphila* was the most abundant species of classified reads with 15.78% of the total reads followed by *E. coli* 8.49% (**Table 3-27**).

Similar results were observed for the respective un-depleted sample (1141 PC), 972,597 total number of reads were produced of which 62% were unclassified, 38% were classified as bacteria. *C. difficile* accounted for only 0.01% of reads (**Table 3-27**). The most abundant species was *A. muciniphila* (11.59%) followed by *E. coli* (8.94%) and *Parabacteroides distasonis* (5.87%) (**Table 3-27**).

For the sample 1142 PC, total reads produced were 882,162 of which 29% were unclassified by Kraken and the rest were bacteria (71%), 18.73% of the total were *E. coli* and only 0.02% were classified as *C. difficile* (**Table 3-27**). The top hit was *E. coli* (18%) followed by *P. distasonis* (5.45%) and *Citrobacter koseri* (3.23%) (**Table 3-27**).

Table 3- 26. Illumina reads showing total reads, *E. coli* and *C. difficile* reads of the three stool samples as classified by Kraken.

Sample ID	**Total Reads**	**Unclassified** (%)	**Bacteria** (%)	***E. coli*** (%)	***Clostridioides difficile*** (%)
1141 MolYsis	1,302,146	63	37	8.86	0.02
1141 PC	972,597	62	38	9.33	0.01
1142 PC	882,162	29	71	18.73	0.02

Table 3- 27 presenting top five species (ranked by number of reads as classified by Kraken) of the three stool samples.

Sample ID	Top Specie (%)	2nd Specie (%)	3rd Specie (%)	4th Specie (%)	5th Specie (%)
1141 MolYsis	*A. muciniphila* (15.78)	*E. coli* (8.49)	*Parabacteroides distasonis* (1.59)	*Bacteroides vulgatus* (0.6)	*Methanobrevibacter smithii* (0.57)
1141 PC	*A. muciniphila* (11.59)	*E. coli* (8.94)	*P. distasonis* (5.87)	*B. vulgatus* (0.65)	*Odoribacter splanchnicus* (0.43)
1142 PC	*E. coli* (18.16)	*P. distasonis* (5.45)	*Citrobacter koseri* (3.23)	*Streptococcus pasteurianus* (2.26)	*Eggerthella lenta* (0.19)

It was interesting to see *A. muciniphila* and *E. coli* being the most abundant organisms in the patient sample that was positive for *C. difficile* by EIA Toxin assay and qPCR.

Due to the unexpected low number of *C. difficile* sequences in CDI cases, it was postulated that to obtain sufficient *C. difficile* DNA/reads for SNP typing from metagenomic data, three things had to be addressed.

1. Develop an efficient DNA extraction method for *C. difficile* spores
2. Develop an enrichment protocol for *C. difficile* spores from a high background of normal gut microbiota
3. Develop an efficient human and normal flora DNA depletion method for stool

It is well documented that *C. difficile* vegetative cells transform to spores once exposed to aerobic conditions. In anaerobic conditions in the human colon, *C. difficile* spores initiate infection by transforming to vegetative cells. After excretion and once the stool sample is exposed to air, the vegetative cells convert to spores again (Lawley et al., 2009). As *C. difficile* spores are hard to lyse, physical methods such as bead beating, nitrogen freezing and crushing have been used to successfully obtain DNA (Angelakis et al., 2016). However, these physical methods are known to damage and break genomic DNA, which makes it unsuitable for long-read rapid nanopore sequencing. Therefore, a non-physical DNA extraction method was required to obtain good quality genomic DNA for rapid MinION metagenomics sequencing.

The second challenge was the high background of normal flora (especially *A. muciniphila* and *E. coli*) compared to *C. difficile*. In both *C. difficile* infected patients tested, *E. coli* was one of the predominant bacteria. It was therefore necessary to develop a strategy that would enrich for *C. difficile* spores/DNA.

The third challenge was depleting human DNA from stool samples. Most human DNA in stool samples comes from epithelial cells from colon walls damaged by infection and differentiated colonic epithelial cells (Iyengar, Albaugh, Lohani, & Nair, 1991). Although the amount of host DNA is generally low in stool compared to other sample types, initial results showed the quantity varies significantly from patient to patient, possibly depending on infection damage to the colon. It was therefore necessary to deplete both vegetative cells (normal colon flora) and human cells to obtain sufficient *C. difficile* genome coverage to study transmission, virulence markers, resistance genes etc.

To tackle the challenges above, a non-physical DNA extraction method for genomic DNA extraction was developed followed by *C. difficile* enrichment methods as explained below.

3.4.2: Method 2: *C. difficile* DNA Extraction Method

To make enzymatic lysis of *C. difficile* spores easier, they were stimulated to transform into vegetative cells by inducing germination. During spore formation, the proteins required for germination are pre-packaged into the spore, priming the spore to germinate when conditions are appropriate (Chankhamhaengdecha et al., 2013)(Paredes-Sabja et al., 2014). It is known that bile salts (cholate) stimulate germination of *C. difficile* spores (Paredes-Sabja et al., 2014). Here, we used taurocholate, a derivative of cholate, and Thioglycollate medium (Sigma) to induce germination (Chankhamhaengdecha et al., 2013; Paredes-Sabja et al., 2014).

Appropriate growth conditions and growth stimulants lead to the activation of hydrolases embedded within the spore cortex to start cortex hydrolysis. Once the core is rehydrated and the cortex is degraded, vegetative cells begin to grow out from the germinated spore. This was achieved by incubating the spores for 1 hour at

37 ºC in a double strength germination matrix which was made of 59.6 g/L Thioglycollate medium (Sigma) and 7.4 g/L Sodium taurocholate (Sigma).

Several DNA extraction methods were tested before and after induced germination. Commercially available kits tested were Genomic-tip 20/G (QIAGEN) and purification using MagNA Pure Compact Nucleic Isolation Kit I (Roche). The following enzyme pretreatments were also tested under different conditions as recommended by manufacturers: Mutanolysin, Lysostaphin, Lyticase, Lysozyme and Proteinase K. Results of DNA extraction methods were compared using Nanodrop (methods section **2.8.1**), Qubit (section **2.8.3**) and qPCR (section **2.9**).

Table shows DNA quantification following treatment with enzymes (lysis enzymes) and bacterial lysis buffer combined with proteinase K (PK). All the methods produced around 0.5 ng/μL of DNA (eluted in 50 μL) from 100 μL of three days culture (**Table 2-28**)

Table 3- 28 showing DNA quantification following extraction from 100μl *C. difficile* spores solution.

DNA Extraction Method	**Normal Conditions (ng/μL)**	**Extended Incubation (ng/μL)**	**Induced Germination (ng/μL)**
Lysis Enzymes	0.554	0.536	0.780
Bacterial Lysis Buffer + Proteinase K	0.444	0.310	0.516

Obtained DNA using the two methods was too low, between 15ng and 39ng. At least 500 ng was expected from the 100 μL prepared solution of spores.

Other methods tested were the G-Tip kit (Qiagen), bead bating and pre-treatment using lysosome, bacterial lysis buffer (BLB) and proteinase K followed by DNA extraction using MagNa Pure. Extracted DNA was eluted in 50 μL of elution buffer. Initially, bead bating and pre-lysis treatment produced more DNA, but under extended incubations and induced germination pre lysis treatment was superior. The method produced 840 ng compared 223 ng using bead bating and 100 ng using G-Tip (**Table 3-29**).

Table 3- 29 showing DNA quantification after extraction from 100 μL C. difficile spores solution

DNA Extraction Method	Normal Conditions (ng/μL)	Extended Incubation (ng/μL)	Induced Germination (ng/μL)
G-Tip	2.00	2.16	1.08
Bead Beating	4.46	2.63	2.07
BLB + PK + Lysosome	5.22	7.40	16.8

The method that worked best was pre-treatment by Lysozyme, Proteinase K and Bacterial Lysis buffer. The enzymes released more than double the amount of DNA compared to others. The method has been detailed in section **2.6.3.** Briefly *C. difficile* spores were stimulated to grow by double strength germination matrix. Thereafter, bacterial lysis buffer (Roche) and 100 mg/mL lysozyme were added then incubate for 1 hr at 37 °C with shaking at 500 rpm in Thermomixer. Afterwards, Proteinase K was added and then incubated at 65 °C for 1 hour with shaking or alternatively overnight. Samples were then loaded onto MagNA Pure for DNA isolation and purification.

3.4.3: Method 3: *C. difficile* Spore Separation using Gastrografin™

This method was applied to clinical stool samples from patients with *C. difficile* infection to separate spores from other cells. The method was tested on four samples from patients suffering from diarrhoea. The samples were tested for toxin and Glutamate Dehydrogenase (GDH) using enzyme immune assays (EIA) and for GDH using PCR at the clinical laboratory as shown in **Table 3-30** below.

Table 3- 30 Microbiology laboratory diagnostic test results for stool samples with suspected *C. difficile* infection.

Sample ID	**EIA (Toxin)**	**EIA (GDH)**	**qPCR (GDH)**	**Microbiology diagnosis**
S17411	_	+	_	CDI +/-
S17664	_	+	+	CDI +/-
S17480	+	+	+	CDI +
S17687	+	+	+	CDI +

The four samples were aliquoted into approximately 200 mg, homogenized and processed as described in section **2.5.2**. After Gastrografin separation, the supernatant (thought to be mostly vegetative cells) was transferred to a new tube for DNA extraction and qPCR analysis. The middle layer (mostly Gastrografin medium) was discarded and the pellet, thought to contain mostly *C. difficile* spores, was retained for further analysis. DNA was extracted as described in section **2.6.3** and analyzed using *C. difficile*, *E. coli* and 16S rRNA qPCR assays. It was expected that most *C. difficile* spores would be recovered in the pellet (hence lower *C. difficile* Cq values) and less of *E. coli* and other bacteria 16S rRNA (higher Cq values). However, **Table 3-27** shows that although sample S17664 had more *E. coli* in the supernatant than pellet (ΔCq 3.35), the amount of *C. difficile* DNA was comparable in the supernatant and the pellet (ΔCq - 0.76). Sample S17480 had more *C. difficile* in the pellet compared to the supernatant (ΔCq 2.29) but the pellet also had more *E. coli* (ΔCq - 3.79) and other vegetative cells (ΔCq - 3.55). Sample S17687 showed good separation of *E. coli* (ΔCq 11.34) but the amount of *C. difficile* was similar in both parts (ΔCq 0.81). No *C. difficile* DNA was detected in sample S17411, this sample was not a confirmed positive for CDI according to the clinical microbiology results (**Table 3-31**).

Table 3- 31 qPCR analysis of the clinical samples after Gastrografin *C. difficile* spore purification

	Pellet			**Supernatant**		
Sample ID	**C. difficile**	**E. coli**	**16s rRNA**	**C. difficile**	**E. coli**	**16s rRNA**
S17411	_*	21.63	13.03	_	22.74	13.88
S17664	32.27	24.26	18.54	31.51	20.91	15.23
S17480	25.17	17.54	11.64	27.46	21.33	15.19
S17687	27.94	28.88	15.08	27.13	17.54	14.25

*No DNA was detected

From the table **3- 31** it can be concluded that spores separation by Gastrografin medium was not efficient when applied to the stool samples. The pellet, which was supposed to be predominately spores, was contained large proportional of *E. coli* and other microbial cells. Also, large amount of *C. difficile* spores remained in the supernatant. Therefore, a chemical lysis approach (section **3.4.4**) was adopted after failing to sufficiently enrich *C. difficile* spores with gradient centrifugation separation techniques using Nycodenz and Gastrografin media.

3.4.4: Method 4: *C. difficile* Spore Enrichment by Chemical Differential Lysis

This method takes advantage of diverse cell properties of the vegetative gut flora, *C. difficile* spores and human cells. Approximately 200 mg of stool sample was homogenized and thereafter host normal flora and human cells were differentially lysed using buffer FL, a bacterial cell lysis buffer from the GeneAll Exgene Stool gDNA purification kit (Seoul, South Korea) as described in section **2.5.3.** The released vegetative and human cell DNA was digested by endonuclease followed by DNA extraction of the *C. difficile* spores (as described in **2.6.3**) and analysis by qPCR.

Table 3-32 shows qPCR results after depletion. Human DNA was depleted by $>10^3$ -fold (ΔCq 11.72) compared to the unenriched sample (PC 6994) in sample S6994 (**Table 3-32** and **Figure 3-44**). In this sample, the initial amount of *E. coli* was low and consequently, all was depleted. Sample S7009 had a lower number of human cells to begin with, DNA of which was all depleted; the sample had a substantial amount of *E. coli* DNA depleted (ΔCq 7.11) and *C. difficile* DNA was not lost (ΔCq 0.09). Sample S7007 produced similar results (ΔCq 9.37 *E. coli* DNA), but the

sample had a negligible amount of human DNA and *C. difficile* was not detected by qPCR. The bacterial 16S rRNA assay results were variable, with sample S7007 showing ∆Cq 4 depletion of bacterial DNA (most of which was likely to be *E. coli*) but the other samples showed little difference between depleted and un-depleted aliquots (**Table 3-32**).

Table 3- 32 qPCR results for three samples with their respective positive controls after enrichment of spores by chemical differential lysis.

	Sample ID	**Human qPCR (Cq)**	***E. Coli* (Cq)**	***C. difficile* qPCR (Cq)**	**Bacterial rRNA 16s (Cq)**
1	S 7007	40.00	30.47	-*	15.59
	PC 7007	40.00	21.10	-	11.98
2	S 7009	-	30.68	32.77	9.89
	PC 7009	36.51	23.57	32.68	10.69
3	S 6994	34.05	-	24.74	16.03
	PC 6994	22.23	36.98	22.97	16.50
4	NTC	-	-	-	29.65

*No DNA was detected

The initial amount of host DNA in sample S6994 was the highest of the three samples (similar levels as seen in blood samples), of which approximately 10^3-fold was depleted (**Figure 3-44**).

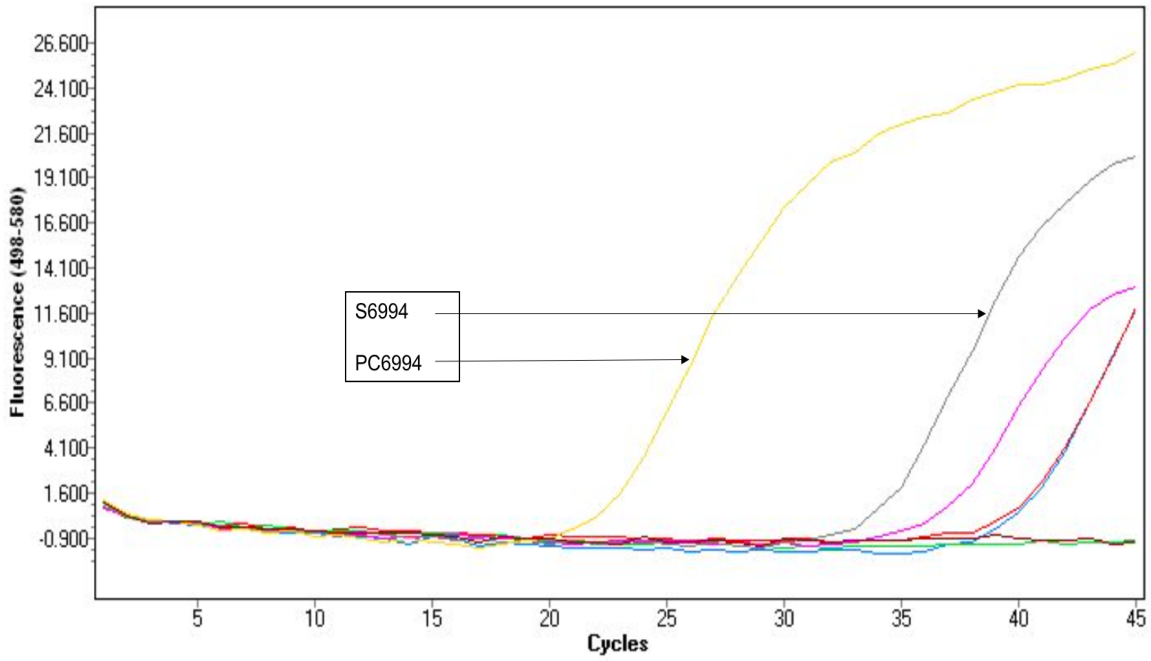


Figure 3- 44 Human qPCR amplification curves after depletion of host DNA in stool

There was depletion of *E. coli* DNA in all three samples, from 10 to 10^3-fold reduction. The initial amount of *E. coli* DNA in sample S7009 and S7007 was relatively high, of which most was depleted, ΔCq 9.37 and ΔCq 7.11 respectively (**Figure 3-45**).

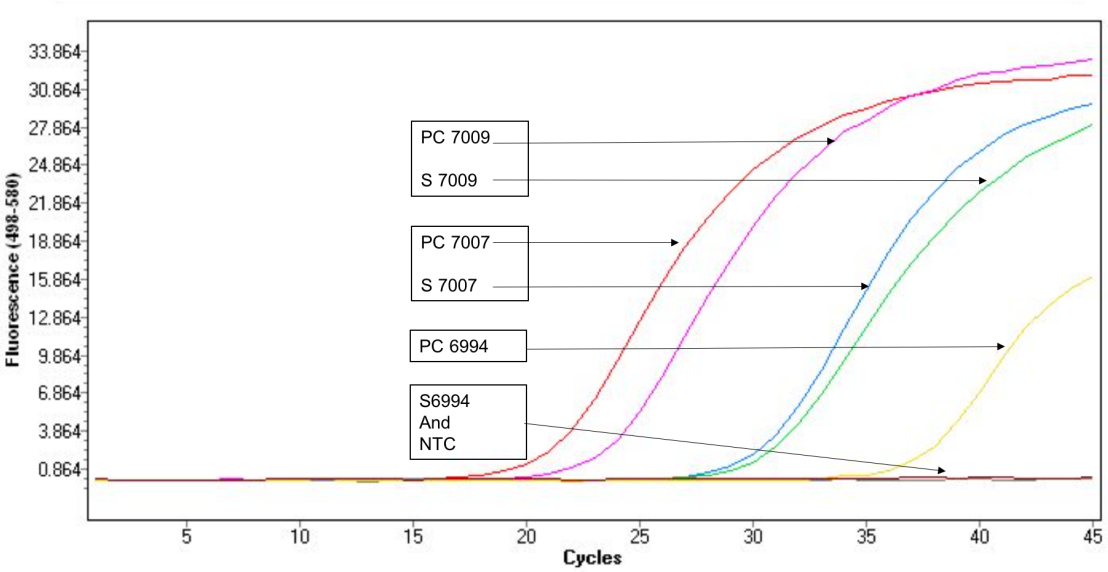


Figure 3- 45 *E. coli* qPCR amplification curves for the three stool samples post *C. difficile* enrichment

The qPCR results showed there was no significant loss of *C. difficile* DNA in the two samples positive for *C. difficile* using our assay (**Figure 3-46**).

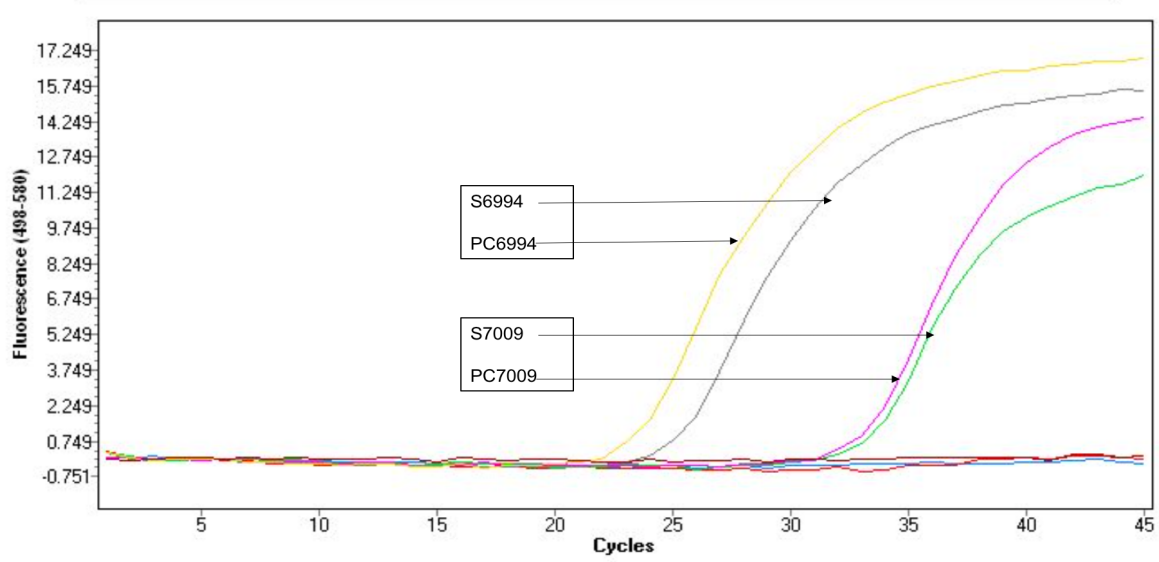


Figure 3- 46 *C. difficile* qPCR amplification curves for the three stool samples post enrichment

This technique, which could be completed in less than 8 hours, depleted host cells and normal gut flora to the satisfactory levels without losing *C. difficile* DNA. The samples were sequenced and analysed further as described in **Chapter 5.**

3.4: Chapter Discussion

For efficient, rapid, and cost effective metagenomic diagnosis of infectious diseases directly from clinical samples, depletion of human DNA or pathogen DNA enrichment is essential. In this chapter, I have developed rapid methods for DNA extraction and depletion of unwanted DNA (human and normal flora) for metagenomics based diagnosis of infectious diseases. These methods have been combined to form culture-independent pipelines designed for pathogen detection directly from clinical samples.

3.4.1: Blood

My findings and those of others (Hansen et al., 2009; Thoendel et al., 2016; Wiesinger-Mayr, Jordana-Lluch, Martró, Schoenthaler, & Noehammer, 2011) have shown current commercial tools for depletion of human DNA, e.g., MolYsis, Looxster and NEB Microbiome enrichment kit do not provide sufficient levels of host DNA depletion for metagenomics infection diagnosis. For example, the NEBNext

Microbiome DNA Enrichment Kit depletes 100-fold or less of human DNA in blood (Hewitt et al., 2017) and other sample types (Burke, McKenna, Cox, …, & 2016, n.d.; Feehery et al., 2013; Thoendel et al., 2016). A study by Loonen et al evaluated the efficiency of MolYsis™ Basic 5 kit and Polaris technology (BioCartis) in depleting host DNA in blood samples and the methods removed only approx. 10 to 100-fold of host DNA (Loonen et al., 2013). Most of these kits were designed to increase the sensitivity of qPCR based diagnostic tests which can sometimes fail in high human DNA background (Loonen et al., 2013). But despite host depletion, qPCR test sensitivity has not improved sufficiently to replace culture, and instead are used as supplementary tests to speed up clinical microbial diagnosis in cases where they are positive.

Although higher levels of host DNA depletion have been observed when MolYsis is applied to other sample types e.g.; 10^{4-} fold depletion in fluids from prosthetic joint infection (Thoendel et al., 2016) and oral samples (H.-P. Horz, Scheer, Huenger, Vianna, & Conrads, 2008; McCann & Jordan, 2014), and 10^2 –fold in urine in our experience, or more when combined with other methods (Schmidt et al., 2016). MolYsis™ Basic 5 kit is the most widely tested method for diagnosis of BSI despite poor performance (Gebert, Siegel, & Wellinghausen, 2008; Loonen et al., 2013; McCann & Jordan, 2014).

In this chapter I described the optimisation of MolYsis to improve the level of host DNA depletion. For human DNA depletion in blood, Dynabeads CD45 IMS was chosen to supplement the MolYsis procedure as it was relatively quick and easy to use and demonstrated no loss of bacterial DNA (**Figure 3-5**). CD45 IMS has never been used for depletion of host DNA for pathogen identification, however it has been used before for enrichment of circulating tumor cells where the same level of depletion (10^2 – folds) was observed (Yang et al., 2009). The combined method provided approximately 10^6-fold depletion of human DNA, resulting in a ratio of 1000:1 human: pathogen DNA as compared to approximately 10^9:1 before depletion. Compared to the commercially available and published assays (Benagli et al., 2013; Melnikov et al., 2011), this method provided higher levels of human DNA depletion for pathogen identification by metagenomic sequencing. None of the commercial methods tested alone provided the level of depletion necessary for metagenomics

based pathogen detection in blood. Hence, the developed strategy of combining CD45 IMS followed by MolYsis was a superior method for depleting human DNA in blood samples compared to commercially available kit methods used alone.

However, during depletion with MolYsis there was approximately 50 to 70% loss of bacterial DNA, this loss of bacterial signal has also been observed by (Hansen et al., 2009; H. P. Horz, Scheer, Vianna, & Conrads, 2010). Although loss of bacteria DNA is not ideal, the high levels of human DNA depletion make the level of bacterial loss acceptable. The method of depleting human DNA by Dynabeads CD45 IMS and MolYsis Basic 5 kit was incorporated into the workflow for detection and identification of BSI pathogens independent of blood culture **Figure 3-12.** The limit of detection of the workflow was later analysed and then tested in clinical blood samples as described in **Chapter 5**.

The lack of an effective commercially available kit for depleting host DNA from clinical samples when I started my PhD was the motivation for developing novel depletion methods. Several new approaches and methods were attempted without much success. Attempted methods/strategies included the use of PMA, development of leukocyte lysis buffers, differential centrifugation and some combinations of these strategies. The novel strategy that worked best was the utilization of properties of cytolysins (PLC) to specifically lyse human cells whilst leaving bacterial cells intact. DNA from the lysed host cells is then digested by endonucleases (HL-SAN) and finally pathogen DNA is extracted for further analysis. Cytolysins and the nucleases used in this method have never been used before for depletion of host DNA. Using this method, between 5×10^4 to 10^6-fold depletion of human DNA was achieved in blood within 45 minutes. This method has been patented and I am a named inventor on the patent application (**Appendix 1**).

Other researchers have developed several other methods for depleting host DNA based on different strategies e.g., utilizing a methylcytosine dependent restriction enzyme approach to enrich Plasmodium DNA in malarial blood samples (Oyola et al., 2013). This method uses restriction enzymes to recognise methylated sequences and selectively digest the host DNA in malarial samples. In human DNA, 4–6% of cytosines are methylated, and 60–90% of these methylated cytosines are at CpG sites (Lister et al., 2009), in contrast, methylation at CpG sites in lower species are

rare. However, the method depletes only approximately 80% of the host DNA, similar to NEBNext Microbiome enrichment kit which uses a similar strategy. Furthermore, the method has a long turnaround time including 16-hours incubation at 37°C (Oyola et al., 2013). Another published method uses a combination of selective lysis of leukocytes by Ox bile and enzymatic digestion of released DNA by Micrococcal nuclease. Ox bile can selectively lyse blood cells thereby releasing the human DNA for degradation. Due to its bile-resistance, *Salmonella* Typhi cells remain intact and cells are collected for the extraction of DNA (Zhou & Pollard, 2012). The method depletes approx. 0.5µg of human DNA in 200 µL blood sample thereby improving sensitivity of qPCR in detecting *S.* Typhi by 1000-fold (Zhou & Pollard, 2012). However, method has not been tested with other bacterial types and the level of depletion is not sufficient for metagenomics (there is approx. 30 µg human DNA in 1 mL blood).

When developing the cytolysin (PLC) based method, we used blood as a model clinical sample as blood represents one of the most complex samples to successfully apply metagenomics for infection diagnosis due to the very high ratio of human: pathogen DNA (as high as 10^9:1).

To the best of our knowledge, the PLC based method is novel and superior to all other commercially available kits and the published methods for depleting human DNA in biological samples. Use of PLC has increased the specificity of differential lysis of human cells that cannot be matched by chemical methods that use chaotropic buffers or saponin. There is no documented evidence suggesting bacteria, fungi or viruses can be lysed by PLC, however there is a possibility that bacteria without cell walls, e.g. Mycoplasma, may be lysed by PLC as the cell membrane is similar to the human cell membrane. Results using blood samples spiked with *E. coli* and *S. aureus* has shown no loss of bacterial DNA (**Table 3-19**). Enzymatic differential lysis followed by enzymatic DNA digestion at the same temperature (37°C) lends itself to a simple automatable procedure. The current procedure is also rapid, taking <45 minutes and this could be significantly reduced with automation. There is also the potential to combine the PLC and DNase treatment into one step, thereby reducing turnaround time even further.

The comparison of the new PLC based method and the commercially available MolYsis kit, showed the PLC based method was more efficient at human DNA depletion (demonstrating up to 10^6-fold depletion of human DNA in blood). Only the in-house MolYsis and CD45 IMS protocol showed the same level of efficiency compared to PLC.

This method can be considered a 'key enabling technology' for the efficient application of rapid metagenomics sequencing for infection diagnosis. As blood represents the most complex clinical sample type, with extremely high human to bacterial cell ratios, it was predicted that the method will work well in other clinical sample types providing similar levels of human DNA depletion and this has been demonstrated during the exemplification of the patent (data produced in Dr O'Grady's laboratory after I completed the laboratory work for my PhD).

3.4.2: Urine

Compared to other clinical samples, urine was not as challenging due to a less complex and abundant matrix and lower ratios of human to pathogen DNA. Hence depleting human DNA in clinical urine samples was not as challenging as blood. In fact, sequencing directly from urine samples without depletion of human DNA is possible (Hasman et al., 2013). However, depletion is important for rapid, cost effective and comprehensive metagenomics based diagnostics (Thoendel et al., 2016a).

Four methods were tested for human DNA depletion in clinical urine samples: MolYsis, CD45 IMS combined MolYsis, NEB Microbiome Enrichment kit, and MolYsis combined with differential centrifugation. These methods (except MolYsis combined with differential centrifugation) were designed and optimised to work in blood but we demonstrated that they also work in urine. The best of the four methods was differential centrifugation in combination with MolYsis. This method was relatively cheap, efficient, simple to perform and rapid compared to CD45 combined with MolYsis or the NEBnext Microbiome enrichment kit. Despite low ratios of human to pathogen DNA in some of the clinical urine samples from patients with UTI (ratio not defined but may be as low as 1:1); depletion of host DNA will still be advantageous, reducing sequencing time and cost. These benefits have been

observed in other samples with low contamination of human DNA such as skin swabs (Feehery et al., 2013; Grice et al., 2008; Paulino, Tseng, Strober, & Blaser, 2006), saliva (Feehery et al., 2013) and fluids from prosthetic joint infection (Thoendel et al., 2016). Therefore, development of cheap, rapid methods for depleting human DNA to suit other clinical sample types is important for metagenomics based diagnostics.

The loss of bacterial signal in urine samples was not critical because of the higher number of bacteria in urine samples and the efficiency of the method to deplete human DNA. Also, due to the abundance of bacterial cells in urine samples (as high as 10^8 CFU/ml) (Manoni et al., 2009; Schröder et al., 2015), even after depletion of host DNA, WGA was not required.

Differential centrifugation combined with MolYsis has shown an average of 10^3 fold depletion of human DNA, more than the other methods tested in urine. This method was later applied in combination with rapid nanopore sequencing for pathogen and antibiotic resistance marker identification in clinical urine samples (Schmidt et al., 2016).

3.4.3: Stool

In 1 gram of wet stool sample from a healthy 70kg male, there is estimated to be 10^{11} bacteria (Sender et al., 2016). This microbiota population consists of hundreds to thousands of species, dominated mostly by Bacteroidetes and Firmicutes phyla (Consortium, 2012), some of which exist nowhere else in nature (Donaldson, Lee, & Mazmanian, 2016). The complex ecosystem of the gut flora population makes individual characterization of pathogens by metagenomics sequencing a challenging task. As it was observed in this thesis, even in confirmed cases of CDI, there are still low amounts of pathogen cells compared to normal flora (mainly *E. coli* and *A. muciniphila*). Metagenomics data were sent to Prof. Mark Pallen at the University of Warwick for further genome analysis of *E. coli* in CDI cases. Further analysis (strain typing and virulence marker detection) by Prokka annotation, Nullarbor and BLAST search revealed that neither shiga toxin-producing *E. coli* nor locus of enterocyte effacement (LEE) genes were detected, hence, *C. difficile* was considered the likely causative agent of the diarrhoea in these cases.

To study specific bacteria/pathogens in this complex ecosystem, selective media are used to enrich for the organism of interest such as *Salmonella* (Gaillot, Di Camillo, Berche, Courcol, & Savage, 1999; Maddocks, Olma, & Chen, 2002; Sparbier, Weller, Boogen, & Kostrzewa, 2012), *Campylobacter jejuni* and *Campylobacter coli* (Rogol, Shpak, Rothman, & Sechter, 1985), *C. difficile* (Arroyo et al., 2005; Buchanan, 1984) or pathogens are targeted by qPCR assays (Dutta et al., 2001; Fukushima, Tsunomori, & Seki, 2003; Gerritzen, Wittke, & Wolff, 2011). To enrich for *C. difficile* spores, the alcohol shock method is recommended (Riley, Brazier, Hassan, Williams, & Phillips, 1987), followed by growing the bacteria on selective agar (Riley et al., 1987). In this PhD project, bacterial lysis by alcohol shock was briefly attempted but did not lyse enough cells to enable depletion for metagenomics sequencing application. Other methods tested were physical separation methods using gradient centrifugation in Nycodenz and Gastrografin. Although Nycodenz has been widely used for separating microbiota in samples including stool (Akiyoshi et al., 2003; Hevia et al., 2015; Manichanh et al., 2006), the method was not efficient in separating spores in our hands. The spores layer was contaminated by other microbial and host cells. Also, separating spores by gradient centrifugation in Gastrografin did not work. The pellet, which was supposed to be predominately spores, contained large proportional of host and vegetative cells. Gastrografin medium has been proven to work well in separating spores from bacterial cultures (**Figure 2-2**) (Barbosa, Serra, La Ragione, Woodward, & Henriques, 2005; Fichtel et al., 2007; Sixt et al., 2013), but not directly from stool samples. One possible reason for suboptimal and inconsistent separation of spores from other cells by Gastrografin, was that the large particles in stool samples disrupts the Gastrografin matrix thereby making channels in the medium that allow the large vegetative cells to pass through to the pellet.

To enrich for *C. difficile* spores, vegetative microbiota and human cells were lysed using buffer FL from the Exgene stool extraction kit. Released DNA from host and vegetative microbiota cells was thereafter degraded using DNase. Because spores exhibit resistant properties to extreme conditions such as heat, chemical decontamination (Leffler & Lamont, 2015) and low permeability to small molecules such as water (Paredes-Sabja et al., 2014), they remain intact in conditions where vegetative cells lyse.

The best methods for extracting DNA from *C. difficile* spores have been reported to involve bead beating and liquid nitrogen (Angelakis et al., 2016; Penders et al., 2005). Although effective, these methods produce shorter fragments of lower quality DNA which is not ideal for MinION sequencing. Longer reads are valuable for better genome assembly and linking resistance genes to specific pathogens from metagenomic data. I tested and compared several methods for DNA extraction in stool samples to identify the most efficient method. The best performing method tested involved germination of *C. difficile* spores using sodium taurocholate followed by extended incubation with lysozyme and Proteinase K in bacterial lysis buffer and extraction on the MagNA Pure (section **2.6.3)**.

Currently, PCR ribotyping is the most widely-used method for studying the epidemiology and transmission of *C. difficile* infection (S. H. Cohen et al., 2010). Typing using WGS has been demonstrated to provide higher resolution compared to PCR ribotyping and other restriction endonuclease based analysis such as AFLP, MLVA, PFGE and MLST (D W Eyre et al., 2013). However, application of SNP typing by WGS is hindered by dependence on culture methods, which requires up to 2 days to generate biomass and 3-10 days to sequence depending on the platform used. The developed method for *C. difficile* spore enrichment directly from stool samples could potentially enable rapid SNP typing from metagenomics data of CDI stool without depending on culture. Using this method, *C. difficile* enriched DNA can be obtained within five hours. Combined with rapid MinION metagenomics sequencing, CDI can potentially be diagnosed and typed within 24hrs of clinical diagnosis.

In summary, in this chapter, methods for depleting human DNA for culture independent metagenomics infection diagnosis in blood, urine and stool infections have been developed. In blood, the workflow has been designed to comprehensively diagnose BSI within 8 hours, which is an actionable turnaround time for the clinical management of sepsis patients, providing the necessary information to the clinician to refine antimicrobial therapy before second dose of empiric treatment. Furthermore, we have developed a completely novel method for depleting human DNA in blood samples. The simplicity, specificity and efficacy of this method make it an attractive new depletion tool for application in infection metagenomics. A patent

has been filed on the method and a number of companies are interested in licencing the technology. All the developed methods (except the patented cytolysins based method) were applied to clinical samples for the metagenomic diagnosis of infection without culture as described in **Chapter 5**.

Chapter 4. Results: Assessment and Capacity Development of MinION Nanopore Technology for Bacterial Genome Sequencing.

4.1: Introduction

Towards the beginning of my PhD (June 2014), MinION™ nanopore sequencing technology (Oxford Nanopore Technologies - ONT) was released. This technology had great potential for microbiology applications and as early adopters in the MinION Access Programme (MAP), we were at the cutting edge of demonstrating its potential. We applied MinION sequencing to various bacterial isolates from different environments (gut, environment, clinical samples) to better understand their genomes with long-read, accessible sequencing technology, as described in this chapter. This work resulted in our laboratory being one of the global leaders in nanopore sequencing application and enabled me to develop the skill to apply the technology for infection metagenomics.

The MinION system is a small device, about the size of a chocolate bar, operated and powered via laptop. The system uses flow cells, with integrated electronics, to sequence single molecules through protein based nanopores. The standout features of this technology are portability of the device, relatively quick library preparation, long sequence reads, and real-time sequencing and data analysis. These features make it possible to identify reads within 5 minutes of sequencing (Fox, 2014).

Libraries for ONT MinION sequencing are constructed by attaching tethering oligos, adapters and motors to double-stranded DNA (dsDNA). During the course of my PhD, no less than six genomic DNA sequencing kits (SQK), SQK-MAP002, SQK-MAP003, SQK-MAP004, SQK-MAP005 and SQK-MAP005.1 and SQK-MAP006, were developed by ONT, which we tested (highlighting the immature nature of the technology at the time). Although there were differences between sequencing kits, the general principles of nanopore sequencing remained the same. Each kit contained adapters, motors and tethers that were ligated to the DNA. The function of the adapters and motors is to mediate the movement of DNA through the pore, and to guide the DNA fragments to the vicinity of pores via binding to tethering oligos with

affinity for the polymer membrane (**Figure 4- 1B**). Nanopore sequencing begins at the single-stranded 5' end of the leader adapter (**Figure 4-1C**). Once the complementary (double-stranded) region of the leader adapter is reached, the motor protein loaded onto the leader adapter and bound enzyme that helps to unzip the dsDNA, allowing the first strand of the DNA fragment, the 'template', to be passed into the nanopore, one base at a time, while the sensor measures changes in the ionic current. After reaching the hairpin adapter, an additional protein, the 'hairpin protein', allows the complementary strand of DNA to pass through the nanopore in a similar manner. This was known as a 2D read and was the technology used throughout this PhD.

The current MinION flow cell has 512 channels, each containing four connected nanometer-scale biological wells/wells, known as nanopores, embedded in an electrically-resistant membrane bilayer. Each channel provides data from one of the four wells at a time allowing up to 512 independent DNA molecules to be sequenced simultaneously. During my PhD, I worked with three different flowcell chemistries - R7.0, R7.3, and R9.0 (and different protocols).

When a voltage is applied across the membrane, an ion current flows through the nanopore. The translocation of ssDNA through the nanopore causes a drop in the current that is characteristic of the bases in contact with the pore at that time (**Figure 4-1D**, (Ip et al., 2015). The raw current measurements are compressed into a sequence of 'events', each being a mean current value with an associated variance and duration (**Figure 4-1E**). A sensor in the nanopore measures the current in several thousand times per second, the data streams are passed to the ASIC (application-specific integrated circuit) and then to the MinKNOW™ software on the desktop where by 'event' in k-mer (possible sequence of 5 nucleotides) are predicted to have occupied the nanopore during the event. The MinKNOW™ software also controls the MinION and other sequencing parameters (Ip et al., 2015). The ONT base-calling software uses this information to assign nucleotides sequence of the DNA fragment. The raw current measurements or the corresponding events, plotted over time, are referred to and visualised as 'squiggle plot' (**Figure 4-1G**) (Ip et al., 2015). The base called sequencing reads are then retrieved as fast5 files using the Metrichor™ agent (ONT). Base-called fast5 files were initially placed in a single

folder but as the technology matured, reads were separated into pass and fail folders depending on quality score cut-offs.

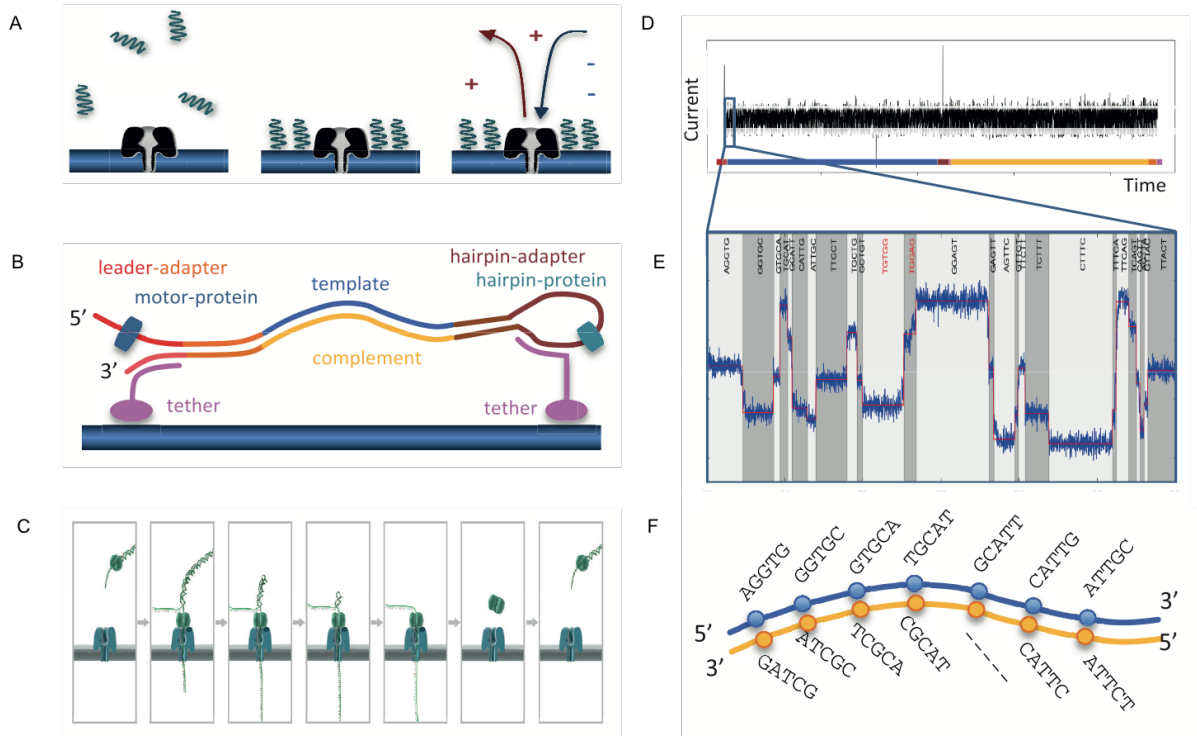


Figure 4- 1 The oxford Nanopore sequencing process. (A) Loaded library molecules are concentrated near nanopores embedded in the membrane. A voltage applied across the membrane induces a current through the nanopores. (B) Schematic of a library molecule, showing dsDNA ligated to a leader adapter pre-loaded with a motor protein and a hairpin adapter pre-loaded with a hairpin protein, and the tethering oligos (C) Sequencing starts from the 5' end of the leader adapter. The motor protein unwinds the dsDNA allowing single-stranded DNA to pass through the pore. (D) Perturbation in the current across the nanopore is measured 3,000 times per second as ssDNA passes through the nanopore. (E) The 'bulk data' are segmented into discrete 'events' of similar consecutive measurements. The 5-mer corresponding to each event is inferred using a statistical model (F) 1D base-calls for the strand and compliment (Ip et al., 2015).

In May 2014, we were one of the first few laboratories in the world to gain early access to the MinION through the MinION Eearly-Acess Program (MAP). The programme participants received free flow cells and library preparation kits from ONT for testing and validating MinION sequencing protocols and technology for various applications. At UEA, I developed laboratory capacity in MinION sequencing and assessed the potential of the MinION system for rapid pathogen and antibiotic resistance gene identification and whole genome sequencing. This was done by studying different bacterial genomes through collaboration with different partners from the Norwich Research Park, Public Health England (PHE), and other institutes

across the world. Below are descriptions of bacterial genomes sequenced and the purpose for sequencing. Also, features and performance of library preparation kits, SQK-MAP 002, SQK-MAP 003, SQK-MAP 004, SQK-MAP 005, SQK-MAP 005.1 and SQK-MAP 006 and flow cells from ONT, have been detailed.

4.2: Results

4.2.1 Characterisation of a previously undescribed antibiotic resistant island in *Salmonella enterica* subsp. *enterica* serovar Typhi haplotype H58 using MinION sequencing (SQK-MAP002, R7 flow cell)

Salmonella Typhi (*S*. Typhi) haplotype, H58 is a multidrug-resistant strain. In this study two strains of *S.* Typhi haplotype H58 were studied, strain H125160566 and 08-0446. Strain H125160566 was isolated in 2012 from a patient returning to the UK from Bangladesh and sent to the Salmonella Reference Service Laboratory at PHE in Colindale, London. *S*. Typhi strain 08-04776 was isolated in 2008 at the Robert Koch Institute, Wernigerode, Germany, from a patient returning from Iraq. *S*.Typhi strains were confirmed as being Haplotype 58 from the Illumina sequencing data by mapping to the *S.* Typhi CT18 reference genome (NC003198) and determining the presence of haplotype 58 specific SNPs. The Illumina sequence revealed that the *Salmonella* isolates harbored multiple resistance elements including, *strA*, *strB*, *sulI*, *sulII*, *dfrA7* and bla_{TEM-1}. The specific resistance plasmid (plasmid PST6 (incHI1) typical of H58 isolates (Holt et al., 2011) was, however, not present, raising the possibility that the antibiotic resistance island had integrated into the H58 chromosome. Short Illumina reads alone could not characterize the resistance island (due to the presence of repetitive insertion sequences causing the assemblies to break). In this study the MinION sequencing was performed at UEA and the Illumina sequencing and bioinformatic analysis was performed by collaborators at PHE.

De novo assembly of Illumina sequence for strain 08-0446 resulted in 143 contigs, an N50 of 124 Kbp and average genome coverage of 78 × (374 million bases of >Q30 data). *De novo* assembly of strain H125160566 resulted in 86 contigs, an N50 of 154 kbp and average genome coverage of 38 × (182 million base pairs of >Q30 data). At the UEA, MinION sequencing for the two strains was performed using 1µg

of *S.* Typhi DNA and SQK-MAP002 (June, 2014), and run on R7 flow cell as described in section **2.11.3** and **Table 2-2**.

Strain H125160566 was sequenced the on the MinION device for 18 h, the experiment resulted in a total of 16,401 sequencing reads with median length 5,412 bp, a maximum length of 66,748 bp, median accuracy of 68.4% and a total of 93.4 Mbp of sequence data (**Table 4-1**). MinION sequencing produced three types of reads, 8,209 template, 4,454 complement and 3,738 2D reads. The two-direction (2D) reads had the highest mean accuracy at 83.6%, followed by the complement reads (60.2%) and the template reads (49.9%).

Table 4- 1 Read statistics for a single MinION run of *S.* Typhi H125160566 broken down by read type

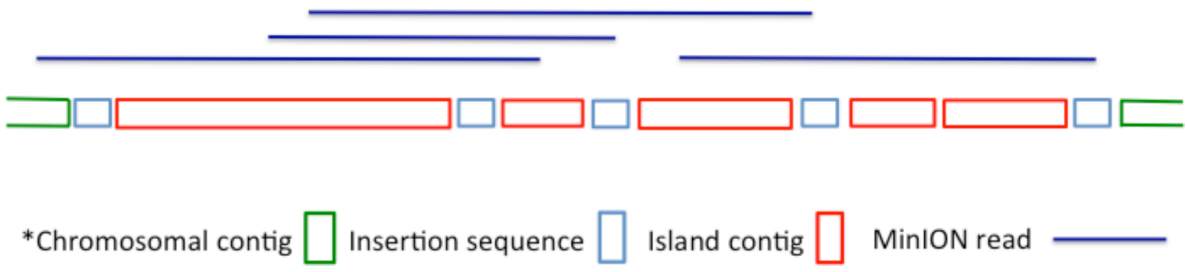
	Number of reads	**Median (bp)**	**Max length (bp)**	**Total length (Mbp)**	**Median accuracy (%)**
Total	16, 401	5,412	66,748	93.4	68.4
Template	8,209	5,614	58,810	49.6	49.9
Complement	4,454	4,728	66,748	21.1	60.2
2D	3,738	5,943	31,630	22.7	83.6

All *S. Typhi* H58 strain H125160566 MinION reads were mapped to the Illumina assembly of the same strain using the LAST sequence alignment tool (Frith et al., 2010). In total 68.7% (11,278/16,401) of the reads mapped at least once to the Illumina assembly, with 16,337 non duplicate alignments giving an average coverage of 14 ×. Mismatches between MinION reads and the Illumina assembly were identified, giving a mean percentage accuracy for all read types of 64.2%. Complement reads had the lowest similarity (61.6%), followed by template (64.3%). 2D reads had the highest accuracy as expected (71.5%) (**Table 4-2**).

Table 4- 2 Mapping Stats of nanopore sequencing reads for *S.* Typhi strain H125160566

	Number of reads aligned	**Number of alignment**	**Total length of alignment (Mbp)**	**Median accuracy (%)**	**Median gaps (%)**
Total	11, 278	16,337	70.8	64.2	17.8
Template	4,705	6,194	31.4	64.3	17.1
Complement	3,171	5,275	17.4	61.6	23.5
2D	3,402	4,868	22.0	71.5	14.8

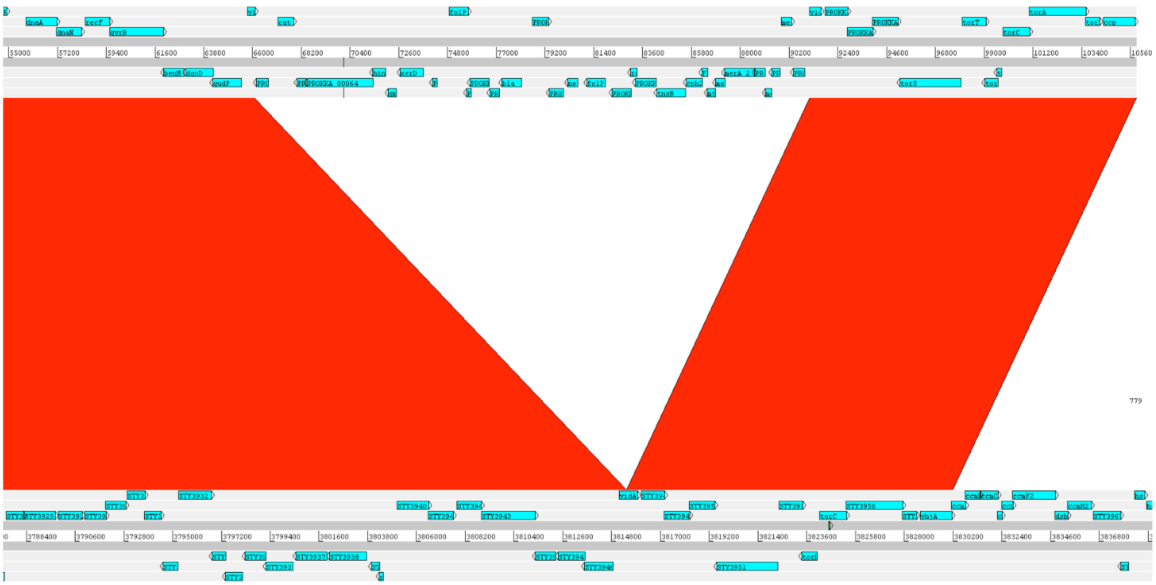
The long MinION reads were then used to scaffold the H125160566 Illumina contigs to determine the structure and chromosomal insertion site of the antibiotic resistance island. Fourty MinION reads that were informative as to the structure and insertion site of the island were identified and these reads were used to link the island contigs across the insertion sequences (**Figure 4-2**). This method resolved the genome into 34 contigs, with an N50 of 319kbp, whereas the Illumina-only assembly produced 86 contigs with an N50 of 154kbp.



*Chromosomal contigs flanking the island (15 kb and 65 kb) were identified based on similarity to IncHI1 plasmids on the 5' or 3' terminus

Figure 4- 2 Schematic of how the MinION reads were used to scaffold the Illumina contigs.

The MinION reads identified the insertion site as the *yidA* gene. Mapping from both ends of this disrupted gene into the island showed that it is flanked by IS1 elements (**Figure 4-3**).



Top: *S.* Typhi haplotype H58 - H125160566; bottom: *S.* Typhi CT18. Red indicates 99% similarity and the cut-off minimum was set at 800.

Figure 4- 3 Artemis Comparison Tool graphical representation of the context of the insertion point (yidA gene) of the S. Typhi chromosomal resistance island

The island contains several of the resistance genes and/or elements found on IncHI1 plasmids in MDR *S.* Typhi, including *strA*, str*B*, *sulI*, *sulII* and *blaTEM-1* (**Figure 4-4**). The structure and insertion site of the island were confirmed by PCR followed by Sanger sequencing. This PCR confirmation was not possible until the island structure was solved (using the hybrid assembly) because of the low-coverage and misassembled Illumina data.

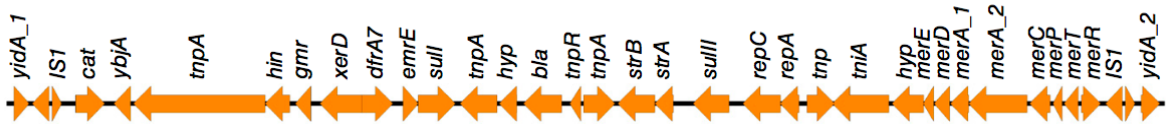


Figure 4- 4 Genetic organisation of the *S.* Typhi chromosomal resistance island. Gene names were assigned using BLAST analysis and manual annotation.

The second strain, *S.* Typhi H58 strain 08-04776 was sequenced the on the MinION device for 48 h, resulting in a total of 10,227 sequencing reads with median length 3,353 bp, a maximum length of 83,819 bp, median accuracy of 60.19% (derived from Phred score) and a total of 42.2 Mbp of sequence data. Number of template reads produced was 7227, complement was 1,840 and 1,160 reads were 2D. Similar to the

previous strain, 2D reads had the highest mean accuracy at 84.15 %, followed by the complement and template reads both at 49.88% (**Table 4-3**).

Table 4- 3 Read statistics for a single 48 hour run of strain 08-04776 broken down by read type

	Number of reads	**Median (bp)**	**Max length (bp)**	**Total length (Mbp)**	**Median accuracy (%)**
Total	10227	3353	83819	42.265801	60.19
Template	7227	2723	83819	28.075045	49.88
Complement	1840	3752	17257	7.419795	49.88
2D	1160	5937.5	19044	6.770961	84.15

In total 46.7% (4,773/10,227), of MinION reads for strain 08-04776*,* mapped at least once to the Illumina assembly of the same strain by the LAST. Mismatches between MinION reads and the Illumina assembly were identified, giving a mean percentage accuracy for all read types of 65.96%. Complement reads had the lowest similarity (62.70%), followed by template (65.61%) then 2D reads (70.18%) (**Table 4-4**).

Table 4- 4 Statistics on mapping of reads from for strain 08-04776 to the 08-04776 Illumina assembly

	Number of reads aligned	**Number of alignment**	**Total length of alignment (Mbp)**	**Median accuracy (%)**	**Median gaps (%)**
Total	4773	7262	25.456267	65.96	17.28
Template	2816	4014	14.78298	65.61	16.49
Complement	973	1639	4.924954	62.70	22.21
2D	984	1609	5.748333	70.18	14.24

The H58 *S.*Typhi strain 08-04776 contained the antibiotic resistance island and an IncN plasmid. The Illumina data from 08-04776 contained one 7 kb contig with a dihydopteroate synthase gene (*sull*) and there were no other *sull* genes in the assembly. When MinION reads were mapped to this contig, they did not map contiguously from this contig to other island contigs, but rather, mapped only to the

sull encoding section. Therefore, we hypothesised that 08-04776 encodes two *sull* genes; one on the IncN plasmid and one on the chromosomally inserted island. The assembly of the Illumina data collapsed these two genes into a single copy, which represented a misassembly that confounded analysis and was only resolved using the MinION reads (Ashton et al., 2015).

4.2.2: Comparison of long reads NGS technologies for *de novo* assembly of *Pseudomonas fluorescens*

In this study, we have collaborated with the Earlham Institute (EI) and the John Innes Center (JIC) to compare PacBio and MinION sequencing platforms for the assembly of antifungal producing *P. fluorescens* strains to gain a better understanding of how the natural products are synthesized within these strains. In this study, *P. fluorescens* strain was Illumina sequenced by collaborators at the JIC, and MinION data generated at UEA was combined with the Illumina for hybrid assemblies performed by collaborators at EI.

In October 2014, we received the new sequencing kit SQK-MAP003 (ONT) and new flow cells (R7.3). This upgrade was released by ONT in September 2014. The key feature change from SQK-MAP002 to SQK-MAP003 was the HP motor was prebound to the hairpin adapter, eliminating the extra incubation and washing steps. There were also minor changes in washing and mixing steps aiming at improving ligation reactions and the overall yield of 2D reads. ONT introduced a new nanopore, version R7.3, and the packaging and shipping of R.7.3 flow cell was improved to maintain stability and quality of the fragile flow cells/pores. There was also general improvement in manufacturing of the R.7.3 flow cells demonstrated by the number of the initial available pores (active pores) and the sequencing run time before the flow cell stopped producing data.

Library for SQK-MAP003 was made as described in section **2.11.3** and **Table 2-2**. The first attempt to obtain enough sequence for the *de novo* genome assembly of *P. fluorescens* using SQK-MAP003 and R7.3 flow cell was not successful. The sequence yield still wasn't sufficient for the *de novo* assembly of the relatively large *P. fluorescens* genome (expected size >6Mb), where 30-50 × coverage would mean

180-300Mb 2D data would be required. In this run, only a total of 9,523 reads were obtained; however, there was improvement in the proportion of high quality 2D reads from (up to 28.6% - 2,727 of 9,523 reads) using the new chemistry. The average read length was 4795bp with accuracy of 66.17%. All *P. fluorescens* MinION reads reads were aligned to Illumina scaffold by BWA-MEM (MinION data parameters) (Li, 2013) and BLASR programmes (default parameters) (Chaisson and Tesler, 2012). The average genome coverage of *P. fluorescens* was 2.69 × (**Table 4-5**).

Table 4- 5 Total reads produced, percentage of 2D reads, average read length accuracy of 2D reads of the first run of *P. fluorescens*.

Date	**November 2014 (MAP003, R7.3)**
Total Reads	9,523
Pass 2D	2,727 (28.6%)
Average 2D read length	4795 bp
Average alignment accuracy SBW25 (2D Pass)	66.17%
Average depth coverage SBW25 (2D)	2.69 ×

The Genomic sequencing kit SQK-MAP003 was not available for long – in December 2014 it was replaced with SQK-MAP004, which we then used to repeat the *P. fluorescens* sequencing. One of the new features of the SQK-MAP004 kit was the use of His-tag pull down beads (Life Technologies) rather than the Agencourt AMPure XP beads (Beckman-Coulter) for library purification. The His-tag pull down beads w**ere** used to isolate only the adapter ligated DNA fragments by attaching to the hairpin protein (histidine-tagged proteins) in the hairpin adapter (Figure 4-1B). This resulted in increased proportion of 2D reads.

Another change was the reduced amount of HP adapter used per ligation reaction, from 10µl to 2µl. The requirement for less HP adapter was related to the specific capture of adapter ligated DNA fragments using His-tag beads. The ligation reaction was also performed in Protein LoBind 1.5 ml tubes (Eppendorf) instead of normal

tubes to reduce loss of enzymes and the 'sticky' hairpin protein, which was thought to be reducing library preparation efficiency.

On the second attempt to obtain sufficient sequence data for the *de novo* assembly of *P. fluorescens,* SQK MAP004 was used to generate the library. The run performed in mid-December 2014 resulted in a total of 11,208 reads with 32.2% 2D reads (3,290 reads). The run was later repeated, producing 46,810 reads of which 30.4% (14,274) were 2D. This was one of the highest yielding MinION sequencing runs globally at the time. In trying to improve the proportion of 2D reads, in late December 2014, high quality *P. fluorescens* DNA was extracted from fresh cultured cells using the G-tip DNA extraction method (described in section **2.7.1**). Sequencing using high quality genomic DNA did not improve the yield of the run (total number of reads 44,356), however, there was increase in percentage of 2D reads, from 30.4 to 48.3% (21,467 reads); the highest we observed in the six months using the technology (**Table 4-6**).

Table 4- 6 Total reads produced, proportion of 2D reads, average read length and accuracy of 2D reads of the three runs of *P. fluorescens* performed in December 2014.

Date	**9th December 2014 (MAP004, R7.3)**	**17th December 2014 (MAP004, R7.3)**	**19th December 2014 (MAP004, R7.3)**
Total Reads	11,208	46,810	44,356
Pass 2D	3,290 (32.4 %)	14,274 (30.4 %)	21,467 (48.3 %)
Average 2D read length	6449	6437	7717
Average alignment accuracy SBW25 (2D Pass)	70.41 %	67.86 %	64.46 %
Average depth coverage SBW25 (2D)	5.95X	18.5X	42.01X

These data were sent to EI for detection of natural products biosynthesis gene clusters and comparison with PacBio and Illumina sequencing of the same strain (performed at EI).

NGS reads were assembled using SPAdes (Bankevich et al., 2012) for Illumina, HGAP (Rhoads and Au, 2015) for RSII PacBio, and using Miniasm (Li, 2016) for ONT reads. Performance of three platforms and tools for *de novo* genome assemblies of *P. fluorescens* using of three platforms is shown in **Table 4-7.** The Illumina platform produced the most data (and higher genome coverage) for the lowest cost. Data produced by long reads technologies per flow cell were similar, but faster assembly using MinION. SMRT cells were cheaper than ONT's flowcells. Mean read length for PacBio was 9kb and for ONT was 7kb (**Table 4-7**), however size selection was performed during the PacBio library preparation but not during the MinION library preparation.

Table 4- 7 Comparison of de novo assemblies of Illumina (Spades), PacBio(HGAP) and ONT (Miniasm*)*

	PacBio RS (HGAP)	**Illumina HiSeq (SPAdes)**	**ONT MinION (Canu)**
	2 SMRT cells	151 bp PE	3 flow cells
No. of reads	46,083 (after filter)	4,937,510 x 2	39,295 (2D pass)
Mean read size	~ 9 kb	151 bp	~ 7 kb
Input reads	539 Mb	1.4 Gb	288 Mb
No. of contigs	64	16	32
Coverage	82 x	223 x	43 x
Elapsed time	5h16m	2h32m	47m
Approx. cost	$1100	$170	$3000

4.2.3: Comparison of long reads NGS technologies for *de novo* assembly of the microbiota member *Bifidobacterium longum*

B. longum is an important member of the gut microbiota. Numerous animal and human studies have suggested that the presence of *Bifidobacterium* is associated with a range of health benefits, including immune system programming, reduction in chronic inflammatory diseases and pathogen protection (Hevia et al., 2015, Fanning et al., 2012). *Bifidobacterium* species have approximately a 60% GC content and genomes varying between 2.2-2.6 Mb. Previously it has been thought that a

significant selection of the microbiota was 'un-culturable', however a recent study (Browne et al., 2016), has highlighted that a sizeable proportion of the microbiota can be isolated and in turn genetically characterized via whole genome sequencing. Also, from studies like these it appears that a significant proportion of these microbes come from novel families, genera and species, thus appropriate reference genomes for suitable genome assembles and annotation is limited. This is particularly problematic when using a short-read sequencing approach. One way to overcome these issues is to utilise technology that generates sufficiently long reads that span repetitive elements of the genome and enable the user to assemble poorly characterised or indeed unknown species in an accurate manner. The aim of this study to determine the utility of nanopore long read sequencing technology for *de-novo* genome assembly of the microbiota member *B. longum* (strain 8809) and compare the results to the current gold standard long-read technology, PacBio. This study was performed in collaboration with Dr. Lindsay Hall from Institute of Food Research (IFR), Norwich.

Illumina and PacBio sequencing were performed at the Wellcome Trust Sanger Institute and MinION sequencing at the UEA.The first set of MinION sequening was performed in July 2014 using SQK-MAP002 with R7 flow cell. A total of 3352 reads were obtained of which only 188 were classified as 2D, with average read accuracy of 67.9%. BWA-MEM (using MinION data parameters) and BLASR programmes (default parameters) were used to align the MinION reads to the Illumina scaffold. The average depth coverage of 2D reads was 1× and the average read length was 3470 bp, **Table 4-8**.

Table 4- 8 Total reads produced, proportion of 2D reads, average read length and average accuracy of 2D reads of the first run of *B. longum* 8809.

Date	**July 2014 (MAP002, R7)**
Total reads	3352
Proportional of 2D reads	188 (5.6%)
Average 2D read length	3470 bp
Average alignment accuracy (2D Pass)	67.9%
Average depth coverage Illumina assembly (2D)	1.01×

Low number of high quality 2D reads generated in this run was thought be due in part to the poor quality of the input DNA used to generate the sequencing library. *B. longum* is difficult to lyse, therefore bead-beating (the standard method in microbiome studies) was used for DNA extraction. The physical nature of bead beating DNA extraction fragments and damages genomic DNA leading to short fragments and hence short reads, thereby failing to utilise the full potential of the technology. Nicks in the DNA backbone also result in few 2D reads, as the 2D construct falls apart when the DNA is denatured while passing through the nanopore during sequencing.

To obtain high quality DNA for long-read sequencing, bead-beating was not used for further experiments. Instead a new method based on combination of high concentration of lysis enzymes and extended incubation times was developed and used to obtain high quantity of genomic DNA as explained in section **2.7.2**. The concentration of enzymes and incubation time had to be optimised for the method to work in this difficult to lyse bacterium.

MinION sequencing of *B. longum* was repeated using Genomic Sequencing Kit MAP004 and R7.3 flow cell. Genomic DNA was extracted using the Genomic-tip 100/G (QIAGEN) as explained in section **2.7.2**. Although there was improvement in total number of reads produced from a single run, the proportion 2D reads remained at 8% (3,246 reads) (**Table 4-9**). Initially it was assumed to be poor quality of library or a flow cell, the experiment was repeated with similar outcome. This problem was subsequently observed by other MAP users and it was later confirmed by ONT that there was a fault in voltage and raw current calibration which affected sequencing and base calling system.

Table 4- 9 Total reads produced, percentage of 2D reads, average read length and accuracy of 2D reads of *B. Longum* using SQK MAP004 and R7.3 Flow cell

Date	**February 2015 (MAP004, R7.3)**
Total Reads	40,360
Pass 2D	3,246 (8%)
Average 2D read length	4451
Average alignment accuracy (2D Pass)	71.29%
Average depth coverage NCIMB assembly (2D)	6.61 ×

In March 2015, ONT released Genomic Sequencing Kit SQK–MAP005, an upgrade from SQK–MAP004. There was one major change and few other minor changes in SQK–MAP005 protocol. The major change was the priming of the flow cell before the run, the flow cell was primed with 1mL of 1 × running buffer and fuel mix instead of 300µL used in previous kit. The change aimed at boosting affinity of the polymer membrane and the nanopores to the library and hence the movement of the library to the nanopores. The minor changes in the SQK-MAP005 sequencing procedure included gentle rotation of beads on a Hula Mixer (Life Technologies) during incubation period. Also, 1.5 ml DNA LoBind Eppendorf tubes used previously in all steps before the ligation reaction were changed to 1.5 ml protein LoBind Eppendorf tubes to stop some of the library preparation enzymes binding to tube surfaces.

Further attempts were made to sequence *B. longum* using Genomic-tip 100/G extracted DNA (section **2.7.2**) and SQK-MAP005 and R7.3 flow cell. The first run using this kit was performed in May, the second in June and the third in July 2015. There was varying number of total reads produced: 36,178 in the first run, 27,383 in the second and 37,190 in the third. The proportional of reads classified as 2D 'pass' reads were 4,281 reads (11%), 7672 reads (28%) and 10350 (27%) for the three runs respectively, **Table 4-10**. The accuracy of 2D reads was similar, 75%, across all three runs (the accuracy value can vary slightly depending on the alignment method used). In this run alignments were carried out using BWA-MEM (MinION data

parameters), and BLASR programmes (default parameters) were used to align the MinION reads to the Illumina scaffold. Genome coverage of *B. longum* was 3.67 ×, 18.64 × and 26.48 × for the first, second and third runs, respectively (**Table 4-10**). However, there was still insufficient data generated from a single MinION run to perform *de novo* assembly of *B. longum*. A theoretical fold coverage of at least 60 was required to call 99.99% of the reference sites accurately from the majority consensus.

Table 4- 10 showing stats of reads produced from three runs of *B. Longum*, between May and July 2015.

Date	**May 2015 (MAP005, R7.3)**	**June 2015 (MAP005, R7.3)**	**July 2015 (MAP005, R7.3)**
Total Reads	36,178	27,383	37,190
Pass 2D	4,281	7672	10350
Average 2D read length	3145 bp	5340 bp	5760 bp
Average alignment accuracy (2D pass reads)	75.76%	75.75%	75.87%
Average depth coverage on Illumina assembly (2D pass reads)	3.67 ×	18.64 ×	26.48 ×

In August 2015, ONT released Genomic Sequencing Kit SQK-MAP006, there were three major changes in reagents required for the library preparation. The SQK-MAP006 protocol required the use of NEBNext Ultra II End-repair / dA-tailing Module (NEB) as a substitute to the separate End-Repair and dA-tailing modules steps. Also, the Dynabeads His-Tag pulldown beads for library purification were replaced with Dynabeads® MyOne™ Streptavidin C1 (Life technologies). Lastly, NEBNext FFPE DNA Repair Mix (NEB) was used for the optional DNA repair step rather than PreCR Repair Mix (NEB). In terms of consumables, Eppendorf DNA LoBind tubes were used throughout library preparation procedure and protein LoBind tubes were no longer required.

The use of NEBNext Ultra II End-repair / dA-tailing Module (NEB) in place of the separate End-Repair and dA-tailing modules steps eliminated extra incubation and

washing steps thus making library preparation simpler and shorter. Also, NEBNext FFPE DNA Repair Mix (NEB) required 15 minutes incubation at room temperature which was quicker and easier compared to 30 minutes at 37ºC required using PreCR Repair Mix (NEB). These changes made library preparation much easier and quicker.

The Genomic DNA Sequencing Kit SQK-MAP006 and flow cell R7.3 were used to obtain MinION reads for *de novo* assembly and genome analysis of the *B. longum.* A total of 62,655 reads were obtained of which 25798 (41%) were classified as high quality 2D pass reads. The estimated genome coverage was 64 × however the read accuracy went down to 71% from 75% observed with MAP005 (**Table 4-11**). Reduced accuracy of 2D reads was related to changes made by ONT to the default current used for driving DNA through the nanopores, which had an effect on the base calling process.

Table 4- 11 Showing the proportion of 2D 'pass' reads, average read length and accuracy of 2D reads of *B. Longum* using SQK MAP006 and R7.3 Flow cell.

Date	**November 2015 (MAP006, R7.3)**
Total Reads	62,655
Pass 2D	25798
Average 2D read length	4491 bp
Average alignment accuracy (2D Pass)	71.27%
Average depth coverage Illumina assembly (2D)	64.42 ×

This run produced sufficient reads for MinION-only assemblies and was used for subsequent comparison with dataset produced by Illumina HiSeq and PacBio RS platforms. Three separate nanopore-only assemblies were generated using alternative assembly programs and compared to a reference-based Illumina assembly and an RSII PacBio assembly. A Miniasm (H. Li, 2016) assembly was formed from the nanopore data with a processing time of under five minutes to produce a single contig assembly. A Canu (Koren et al., 2017) assembly was carried out using the nanopore data with a processing time of approximately 1 hour

using 4 cores and 8 threads to produce a single contig assembly. Nanopolish (N J Loman, Quick, & Simpson, 2015) was used to improve the accuracy of the Canu assembly using event alignments from raw fast5 files of Nanopore data.

B. longum 8809 was also sequenced using the PacBio RSII at and assembled using HGAP 3 at Wellcome Trust Sanger Institute for comparison with Illumina and MinION and technologies. Illumina data was assembled using SDAdes (Benagli et al., 2013). Assembiles were analysed using *B. longum* JDM301 (Wei et al., 2010) as the reference since it was the most closely related sequence to *B. longum* 8809 in the public databases.

The *de novo* assembly statistics of *B. longum* 8809 using the three platforms (PacBio, Illumina and MinION) and the different data analysis pipelines are shown in **Table 4-12**. Performance of long read assemblers for nanopore and PacBio data, showed similar total read length and similarity to *B. longum* JDM301.

The fastest pipeline for long read assembly was Miniasm, which does not include a base error correction nor a consensus step. Because of the high number of indels in the nanopore data, Miniasm failed to accurately determine similarity of *B. longum* 8809 to *B. longum* JDM301. Similarity of 8809 to JDM301 was comparable between PacBio and nanopore data (67%) and lower using illumina data (65%). GC content was similar between all the pipelines (59-60%) except for Miniasm (57%).

To assess the completeness of the newly generated assemblies, we checked for the known list of genes as predicted by GeneMark. Despite similar genome length, number of genes were higher for nanopore data using Canu pipeline (3,191) and Canu + Napolish (2,506). Similar number of gene were identified in Illumina (1,957), PacBio (1,959) and nanopore Miniasm (1,872) (**Tables 4-12**).

Table 4- 12 Comparison of de novo assemblies of Illumina, PacBio and ONT datasets. 2D-pass reads (ONT) were assembled using Minasm and Canu and polished by Nanopolish.

Assembly	De novo & reference-based Illumina	PacBio RSII	Nanopore miniasm	Nanopore Canu	Nanopore Canu + nanopolish
No. of contigs	1	1	1	1	1
No. of scaffolds	1	1	1	1	1
No. of genes predicted by GeneMark	1,957	1,959	1,872	3,191	2,506
Similarity to JDM301	65.5	67.84	- *	67.37	67.51
Total length (bp)	2,278,920	2,372,061	2,441,997	2,341,355	2,351,897
% GC Content	59.85	60.11	57.12	60.07	60.11
Mismatch to JDM 301 per 100 kb	2219.55	2260.39	2777.78	2214.10	2259.48

*Alignment percentage could not be accurately determined

4.2.4: The MinION Analysis and Reference Consortium

In the early stages of the MinION Access Programme, when MinION sequencing was unreliable and unproven, an international consortium was set up (by Dr Ewan Birney, Director, European Bioinformatics Institute), to evaluate the performance and reproducibility of MinION sequencing in laboratories across the world, known as MinION Analysis and Reference Consortium (MARC). The first phase of the consortium's work was to generate MinION genome sequence data for a well-studied and heavily sequenced reference genome, *E. coli* strain K-12 substrain MG1655, following a standardised laboratory protocol. Five laboratories on two continents performed the sequencing, including our laboratory (**Table 4-13**), and others were involved in the data analysis (Ip et al., 2015).

Table 4- 13 Five Laboratories that formed the MinION Analysis and Reference Consortium (MARC) from MAP participants.

SITE	SITECODE	SITE NAME
Lab1	CSHL	Cold Spring Harbor Laboratory, NY, USA
Lab2	UCSC	University of California Santa Cruz, CA, USA
Lab3	UEA	University of East Anglia, Norwich, UK
Lab4	WTCHG	Wellcome Trust Centre for Human Genetics, University of Oxford.
Lab5	ZF	ZF-SCREENS B.V., The Netherlands

We used SQK–MAP005 and SQK–MAP005.1 library preparation chemistries to assess yield, accuracy, and reproducibility of MinION data by undertaking replicate sequencing experiments across multiple sites using *E. coli* MG1655, with the intention of identifying technical factors important for consistently good MinION sequencing performance.

The five laboratories participated in generating the sequence data, in duplicate, using a single shared protocol for growing the bacteria (section **2.2**), genomic DNA extraction from fresh cultures of *E. coli* (section **2.7.1**), library preparation, and sequencing (section **2.11.3** and **Table 2-2)**. The rational for using *E. coli* MG1655 was that it has a single circular chromosome of 4.6 Mb with 50% G/C, that sufficient depth could be obtained from a single MinION run for *de novo* assembly. Also, the strain was sequenced multiple times using different technologies so high quality reference sequence was available (NCBI RefSeq NC_ 000913).

Data generated using SQK-MAP005 (March 2015) was known as MARC Phase 1a and using SQK–MAP005.1 (June 2015) was Phase 1b (Ip et al., 2015b). Libraries were prepared the day after DNA extraction from 1µg and 1.5µg for the Phase 1a and 1b experiments, respectively. In the two phases, a total of 20 individual flow cell experiments were performed in five laboratories according to the shared protocol. Each laboratory undertook two identical replicate experiments for each kit version. The 20 experiments were referred to as P1a-Lab1-R1 to P1b-Lab5-R2, following a ‘phase-lab-replicate’ format.

To compare sequencing data from the five laboratories, several parameters were

assessed by bioinformaticians in the MARC consortium. There was a high level of variability among the 20 experiments, partially attributable to protocol deviations. Below is a description of the most important analyses, the full analysis can be found in the published paper (Ip et al., 2015).

4.2.4.1: Total Event Yield

A median of 60,600 reads (inter-quartile range (IQR) of 38,000 to 74,000, max. 139,000) containing 650,000 bases (IQR 434,000 to 750,000, max. 1.9 million) were generated from the 20 flow cells. The median read lengths indicated most experiments had a broad distribution with a peak around 10,700 bases and a long tail containing a very small number of reads that reached the upper limit of 230,000 bases (**Figure 4-5**) (Ip et al., 2015).

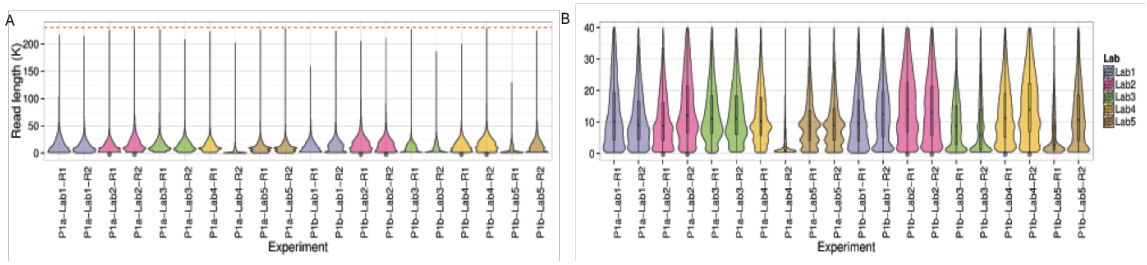


Figure 4- 5 Showing read length of the 20 Phase 1 experiments in kilo bases as (A) the entire distribution of callable read lengths and (B) a subset showing the lower part in more detail.

The highest data yield was from experiment P1a-Lab3-R1 (UEA run), **Figure 4-2**, which commenced sequencing with the highest number of active channels (506/512 = 98.8%) to produce over 138 thousand reads and almost 2 billion bases. This flow cell yield was the first to break the 1Gb benchmark and highest reported in the field (outside ONT laboratories) globally, a record which we held for several months. The two experiments with the highest event yield, P1a-Lab3-R1 and P1b-Lab4-R1, used 60 ng and 9.1 ng of input DNA (median for 20 experiments was 70 ng), respectively, which confirmed that DNA quantity was not the limiting factor relating to flow cell yield.

All experiments demonstrated event accumulation rates that decreased for the first 24h, experienced a sharp increase at 24h following a switch to a new group of pores and library reload, then steadily decreased again until the run was terminated. A

typical run such as P1b-Lab2-R2 generated 47% of the data (367 million events) in the first quarter (12h) of the experiment and 69% of the data (544 million events) in the first half (24h) of the experiment (Ip et al., 2015).

4.2.4.2: Proportion of 2D pass and fail reads, and error rate

Although there was substantial variability in the proportion of 2D pass reads produced during the experiments, there was a clear decrease in median percentage of 2D pass reads from 85% to 20% over the course of the first 21h of the experiment. The drop in 2D pass yield coincided with voltage adjustments (-5mV every 2 hours) suggesting the reads produced during these changes do not have correctly calibrated base qualities. To quantify the error rate of reads produced by the MinION, we referred to as 'total percent error' of a read, defined as the percentage of a read that is inaccurate due to miscalled bases, inserted bases in the read, and deleted bases that are missing from the read but present in the reference sequence. The 2D pass reads had a total error of 10.5%, the 2D fail reads a much higher value of 20.7% (Ip et al., 2015).

4.2.4.3: Read Length

One distinguishable feature of nanopore sequencing is the ability to generate long reads, a proportion of which are over 10,000 bp. Template reads were longest, with 7.6% of total reads over 10 kb. 2D pass reads had the lowest percentage of 10 kb reads at 3.6% (**Figure 4-6**). 50% of reads had a length of at least 5,500, 5,600, 6,000 and 6,300 bases for the template, complement, 2D, and 2D 'pass' base-calls respectively. 5% of the reads had a length of at least 14.5, 13.0, 13.5 and 13.6 Kb for the template, complement, 2D, and 2D 'pass' base-calls. The longest template, complement, 2D, and 2D 'pass' base-calls observed in this study were 291.6, 300.5, 59.7 and 59.7 Kb, respectively.

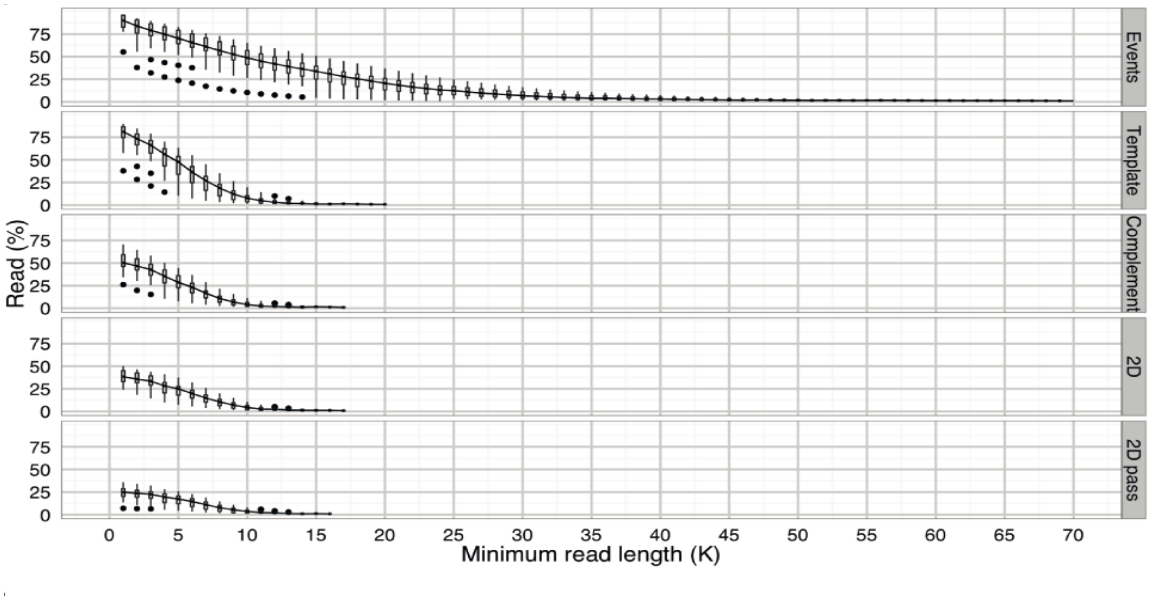


Figure 4- 6 Showing proportion of long reads and data in long reads. A boxplot of the percentage of reads was plotted for each read in multiples of 1000 until the read percentage dropped to 1%. 7.6%, 4.0%, 4.4% and 3.6% of the reads were >10 kb for template, complement, 2D, and 2D 'pass', respectively (Ip et al., 2015a).

4.2.4.4: Genome Coverage

The median theoretical fold coverage of the target *E. coli* genome achieved in the 20 experiments was 25× for 2D reads (min=5.2, Q1=16.3, median=24.9, mean=29.0, Q3=36.5, max=78.5) and 16× if restricted to 2D 'pass' reads (min=1.7, Q1=11.3, median=15.9, mean=20.3, Q3=27.0, max=47.9). A theoretical fold coverage of at least 60× was required to call 99.99% of the reference sites accurately from the majority consensus. Our record 2Gb run produced >60× 2D coverage, enough for *de novo* genome assemble and consensus base-calling (Ip et al., 2015).

4.2.4.5: Contamination

Between 63% and 99% (median 92%) of the reads were allocated to the target sample and most of the remainder to the control sample. Two Phase 1a experiments omitted to include the control sample (both UEA phase 1a runs - P1a-Lab3-R1 and P1a-Lab3-R2). Phase 1b experiments P1b-Lab3-R1 and P1b-Lab3-R2 (both UEA) contained a larger proportion of reads (3.7% and 15.3%, respectively) that did not map to either the target or the control reference, suggesting contamination. We later traced this back to a lab Tris-Cl elution buffer contaminated with *Pseudomonas*

putida. Taxonomic classification of all 2D reads using Kraken version 0.10.5-beta (Wood and Salzberg, 2014) found only two experiments with non-*E. coli* bacterial matches: P1b- Lab3-R1 had 2.3% of the reads classified as *Pseudomonadales* and P1b-Lab3-R2 had 10.7% of reads as *Pseudomonales* and 2.2% as Burkholderiales, at percentages comparable to those inferred from the BWA-MEM alignments (Ip et al., 2015).

4.2.5: Identification of resistant island in *Salmonella* enterica subsp. enterica serovar Blockley using MinION sequencing

Collaborators at the Salmonella Reference Laboratory, at PHE, isolated 667 *Salmonella* isolates from 2012-15 that were part of a WGS validation project. The isolates were screened for known acquired resistance genes, including those previously associated with resistance to azithromycin in *Enterobacteriaceae*. The presence of azithromycin resistance determinants *mphA*, *mphB* and *mefB*, amongst other resistance determinants conferring resistance to ß-lactams, aminoglycosides, quinolones, tetracycline and sulphonamides, was identified in nine genomes, all *Salmonella* Blockley. Phenotypic susceptibility testing confirmed the multidrug resistance phenotypes of the corresponding nine isolates, which had MICs of azithromycin from 6 to ≥16 mg/L. Azithromycin resistance in nine isolates was also confirmed by WGS and PCR methods in the nine isolates.

When the genomic context of azithromycin resistance genes was investigated, it was found that they were on contigs that showed homology to either chromosomes or plasmids. Bandage, Blast, Prokka and Artemis analysis of the 9 *Salmonella* Blockley isolates harbouring the mphA gene and the 10 *mphA*-negative *Salmonella* Blockley isolates showed mphA being inserted down-stream from a *livF* gene on a chromosomal contig. The chromosomally mediated macrolide inactivation gene cluster mphA-mrx-mphr(A), which is flanked by IS6100 and IS26 elements, is part of a larger composite transposon inserted within the coding sequence of the ribokinase gene (rbsK) in all nine *Salmonella* Blockley isolates. However, it was not possible resolve the full island structure using Illumina data, so long-read nanopore technology was used to sequence a representative isolate (H123780513).
At UEA we used SQK-MAP006 and flow cell R9 to generate a total of 10913 2D

MinION reads with a mean length of 3133 bp (both pass and fail 2D reads were used). When mapped using BWA-MEM, 9076 reads (83%) mapped back to the Illumina-only assembly of H123780513, giving an average depth of 5.8× (**Table 4-14**).

Table 4- 14 showing read statistics for a MinION run of *Salmonella* Blockley

Date	**April 2016 (SQK-MAP006, R9)**
Total 2D reads	10913
Pass 2D	53%
Average 2D read length	3133
Average alignment accuracy (2D Pass)	83%
Average depth coverage Illumina assembly (2D)	5.8×

The 5.8× depth of coverage was not sufficient for *de novo* assembly, so a hybrid assembly approach was used. The hybrid Illumina – MinION assembly resolved the complete structure of a previously undescribed 17kb *Salmonella* azithromycin resistance genomic island (Nair et al., 2016). The island harboured tetracycline and aminoglycoside resistance genes as well as phage and plasmid reminants (**Figure 4-7**). This is the first known chromosomally mediated *mphA* gene cluster described in salmonellae (Nair et al., 2016).

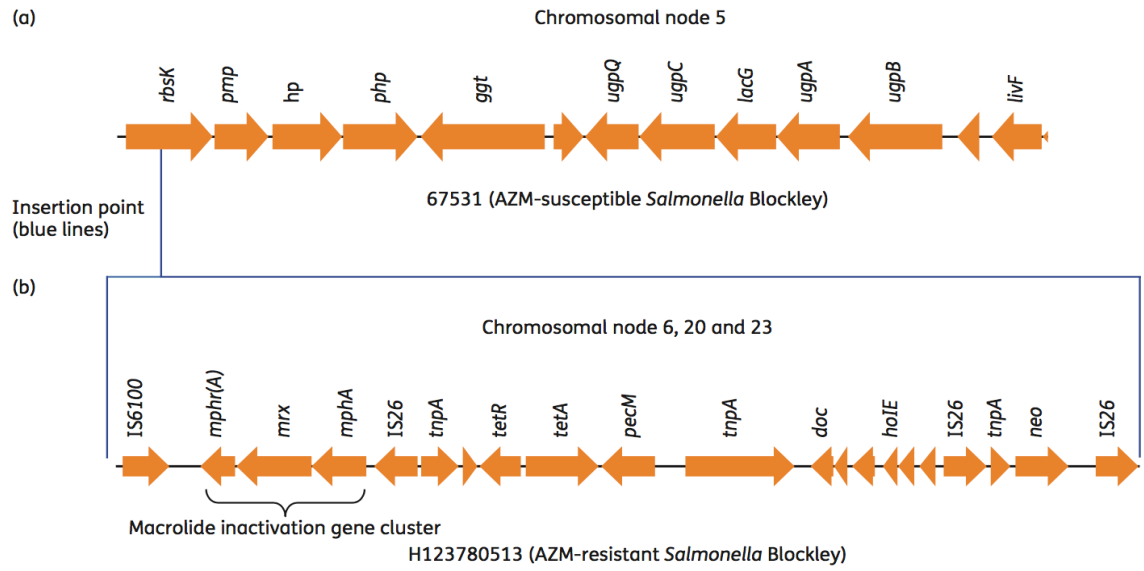


Figure 4- 7 Chromosomal insertion site of the azithromycin resistance gene (mphA) and possible structure of the Salmonella azithromycin resistance genomic island. Chromosomal nodes are based on Bandage assembly.

4.3: Chapter discussion

High-throughput next generation sequencing platforms have made a substantial impact in studying and understanding bacterial genomes. However, despite rapid technology transformation and introduction of novel platforms, NGS is yet to be introduced into clinical microbiology laboratories. Deployment of high-throughput NGS is faced with the challenges of complicated data analysis, complex library preparation procedure, and long turn-around time. The MinION has the potential to eliminate these limitations with simple library preparation, long-read sequences that uncomplicate (some) data analysis, small footprint, very low instrument cost and real-time data generation. The overall objective of this chapter was to assess MinION sequencing technology for various WGS applications in order to develop our expertise in this area and to help drive the technology forward for future implementation in metagenomic infectious diseases diagnosis.

In the last three years, we observed rapid progress of the technology in terms of total sequence yields, reads accuracy, quality of flow cells, increased proportion of 2D reads per run and more rapid protocols for library preparation. Whilst nanopore technology was in its infancy, we sequenced *B. longum* and *P. fluorescens* genomes and compared the nanopore results to established sequencing platforms (Illumina

and PacBio). MinION showed great promise to liberate the consumer is from reliance on sequencing centres, allowing scientist to work faster and on their own timetable whlile obtaining the same results as other established platforms.

We also assessed utility of nanopore long reads to describe novel genomic features, the structure of which could not be solved using short read sequencing technology alone, by studying novel genomic resistance islands in 2 separate Salmonella projects (*S.* Typhi H58 and *S.* Blockley). Using relatively few long reads generated by the early technology, SQK-MAP002 and R7 flow cell, we demonstrated the utility of the technology for hybrid genome assembly with Illumina reads to resolve complex genome structure of the *S.* Typhi H125160566 strain (Ashton et al., 2015). This was the first MinION application paper published globally, which put our group at the forefront of the field. We applied the technology to a similar problem more recently, resolving the structure of a *Salmonella* azithromycin resistance genomic island in *Salmonella* Blockley using SQK-MAP006 kit and R9 flow cells (Nair et al., 2016). These papers demonstrated the potential of the technology for applications in public health microbiology and antimicrobial resistance and MinION is now used by a number of groups at PHE Colindale for various applications.

High-throughput short-read sequencing technologies cannot unambiguously assemble repetitive elements that are longer than sequencing read-length into a single contig (Liu et al., 2012; Treangen & Salzberg, 2011). The use of MinION for public health and clinical diagnosis applications will enable microbiologists to resolve complex structures such as chromosomal resistance islands and resistance plasmids much faster than is currently possible. Although it is possible to determine whether an insertion site is occupied using traditional methods (PCR and Sanger sequencing), the process takes a long time and is not always possible when the inserted DNA fragment is too long. Using accessible rapid long-read technology, scientists will be able to describe structures of insertions, resistant islands, phage, plasmids and other genome elements of repetitive nature and to put them in context. This can now be done rapidly without extensive, post-sequencing, laboratory-based analysis which can take weeks. It has been shown that using MinION is possible to generate reads of >100Kb (Ip et al., 2015) in length with accuracies from 92% (Jain et al., 2017) to 98% (N J Loman et al., 2015) after correction. More recently, ultra-

long reads have been described (up to almost 1Mb) in a study sequencing the human genome on the MinION (our laboratory participated in this study) (Jain et al., 2017). These features will likely make MinION the technology of choice for assembly of complex genomes and for rapid pathogen detection in the near future (Karlsson, Larkeryd, Sjodin, Forsman, & Stenberg, 2015; Joshua Quick et al., 2016).

Despite demonstrated success, there were problems with variability of proportion of 2D reads, and reliability and reproducibility of the sequencing kits and flow cells over the course of my PhD. If MinION sequencing to be applied routinely in public health and clinical microbiology laboratories, the technology must be capable of providing *de novo* assembled bacterial genomes. This will enable bacterial identification, subtyping, resistance gene detection at the point of clinical need and outbreak characterisation. Although we have demonstrated the usefulness of long reads in to resolve complex antibiotic resistance islands in *S.* Typhi H58 and *S.* Blockley, assemblies had to be complemented with Illumina reads due to insufficient data for *de novo* assembly of these genomes. However, higher yields were achieved in some runs, e.g. *P. fluorescens* and *E. coli* MG1655 but there was no consistency of performance. Development is still occurring at a rapid pace and the introduction of R9.3 flow cells in late 2016 resulted in significantly increased and consistent yields, generally around 5Gb per flow cell, and yields > 5Gb per flow cell (up to 17Gb) have been seen in our lab and have been reported in the literature (Jain et al., 2017).

Reason for poor reproducibility, as highlighted in MARC study, are likely related to: DNA quality used for library preparation; the many steps of library preparation being sensitive to the quality of the materials and reagents used; user experience in making libraries; sheared length of DNA; volume/mass of library loaded on a flow cell; quality of library and flow cells, and faults with computer software or hardware. Also, at the beginning of MAP there were major issues with flow cell manufacturing, particularly bubble formation and impurities that damaged the nanopores. These issues have now been resolved and the majority of R9.3 flow cells received in our lab are of high quality (up to 500 available channels per flow cell).

Some of the MAP users, including myself, tried to optimise different steps of the library preparation from DNA extraction to washing, mixing, and loading of the library to improve total yield and proportional of long high-quality 2D reads. Several DNA

extraction methods were tested to improve quality of input DNA - the best results were observed using g-Tip DNA extraction kit (Qiagen). This method yields high molecular weight undamaged DNA through a gentle filter column purification, ideal for MInION sequencing. However, these methods are not practical for clinical microbiology laboratories, hence we routinely use an automated device commonly used in clinical labs (MagNA Pure, Roche) for extraction, with good results.

MinION has shown its capacity of producing long reads similar in length to other platforms such as PacBio. Portability, turn-around time and low cost of the device makes it practical for many scientists to have long-read sequencing technology available in their laboratory. However, the technology is yet to achieve its full potential due to remaining issues with the basecallers not accurately calling homopolymers and methylated bases, hence MinION consensus accuracy is lower that PacBio (Jain et al., 2017).

Over the last three years, there have been several changes to the existing sequencing kits, introduction of new sequencing kits for different applications (e.g. cDNA kits, rapid kits, low input kits, barcoding kits and the direct RNA kit) and changes to the flow cells and nanopores aimed at improving MinION sequencing. In my experience of using the technology, rapid improvement was seen from using SQK-MAP002 and R7 flow cells to SQK-MAP003 and R7.3 flow cells. Generally, the percentage 2D pass reads from a sequencing run was between 5 and 10%, with error rate averaging approx. 30%, and flow cells generally produced data for less than 18 hours. The change from SQK-MAP002 and R7 flow cells to SQK-MAP003 and R7.3 flow cells saw a jump of high quality 2D reads from 8% to 28.6%, similar results were observed by others (Karlsson et al., 2015). There was also improvement in read accuracy from 70% to an average of 79%; and total sequence yield improved from generating data enough for only one-fold genome coverage to *de novo* assembly of *E. coli* (J Quick et al., 2014).

The technology continues to develop rapidly and there have been several changes since I stopped working on MinION in 2016. 2D sequencing has been replaced with $1D^2$ and 1D sequencing has been introduced. The basecallers have improved over time and the latest combinations of chemistry, flow cell and basecalling technologies produce average identities of 90% for 1D and 95% for $1D^2$ reads. Yield and

consistency of performance have increased dramatically since the introduction of R9.4 flow cells and ‘active unblocking’ (reversing the current in a single pore when it becomes blocked to reactivate it). Data can be generated to Fastq files, rather than Fast5, which reduces the file size significantly and speeds up real-time data analysis. In Jan 2018, the rapid development of MinION sequencing was highlighted in a paper sequencing and *de novo* assembling the human genome (Jain et al., 2017). While several flow cells were required to produce sufficient sequence data, vertebrate *de novo* genome assembly is now possible on the MinION, in stark contrast to hybrid assemblies of small bacterial genomes that was possible in 2015. Overall, the technology is maturing rapidly and is likely to be a major competitor in the sequencing market in the near future.

Chapter 5. Results: Rapid Metagenomics Identification of Pathogens in Clinical Samples

5.1: Chapter Introduction

The aim of the work in this chapter was to identify pathogenic organisms directly from clinical samples using culture-independent rapid metagenomics sequencing. Pathogen DNA was sequenced after rapid depletion of human DNA, and normal flora DNA in stool, using in-house methods developed during this PhD (**Chapter 3**). These depletion (and DNA extraction) methods were applied to blood, urine and stool clinical specimens to enable rapid and comprehensive metagenomics pathogen identification.

Blood samples were collected from patients with suspected BSI/sepsis, a major cause of morbidity and mortality globally (Fleischmann et al., 2016). Current culture-based methods for diagnosing BSI have poor sensitivity and take 24-48 hours to identify pathogen in positive samples and a further 24 hours to determine antimicrobial sensitivity. For negative samples, it takes up to 5 days to confirm absence of a pathogen (Vincent et al., 2006). Applying metagenomics based tests to the blood samples would allow rapid pathogen identification, along with antibiotic resistance profiling (Mwaigwisya et al., 2015).

Urinary tract infections (UTIs) are among the most common bacterial infections acquired in the community and in hospitals (Foxman, 2010). Again, current culture based methods for UTI diagnosis take 24 hours to identify the pathogen and a further 24 hours to characterise the antimicrobial susceptibility profile. Combining in-house host DNA depletion methods with rapid MinION nanopore metagenomics sequencing we could profile pathogens and antibiotic resistance directly from clinical urines within 4 hours.

Toxigenic *C. difficile* strains are the major cause of diarrhoea and life-threatening colitis, and is associated with substantial morbidity and mortality mainly in elderly, hospitalised patients and patients receiving antibiotics (Leffler & Lamont, 2015). Currently, PCR ribotyping is the main method for studying CDI transmission in the

UK (PHE, 2010). However, the method is faced with several drawbacks including long turn-around time (currently 14 days) (PHE, 2010), poor resolution of the technology compared to SNP typing methods (Knetsch et al., 2013) and dependence on biomass produced by culture methods, which often fail because of anaerobic conditions required and sporulation (Edwards, Suárez, & McBride, 2013). In this thesis, we have developed a method of isolating and extracting DNA from *C. difficile* spores (the infectious and transmissible morphotype of CDI) directly from stool for SNP typing from metagenomics data.

We have used sample preparation methods developed in **Chapter 3** and two NGS platforms, MinION and MiSeq, for the rapid and comprehensive profiling of pathogens and resistance markers directly from clinical samples.

5.2: Identification of pathogens in blood samples

As previously discussed, the major challenge to metagenomics detection of pathogens in blood is the ratio of human: pathogen DNA (as high as 10^9:1). We developed a highly efficient human DNA depletion method for blood and combined it with pathogen DNA extraction, library preparation and metagenomics sequencing, to form a diagnostic pipeline for the culture-independent detection of pathogens in clinical blood samples (see Chapter 3, **Figure 3-12**). The first step in assessing the optimised pipeline was to determine its limit-of-detection.

5.2.1: The limit-of-detection of the blood pipeline

The “limit of detection” (LoD) or “analytical sensitivity” of the workflow was defined as the ability of the assay to consistently sequence the lowest concentrations of a given substance in a blood sample and distinguish from negative samples. The lower the detectable concentration of analyte, the greater the analytical sensitivity of the assay. LoD is very important in blood testing as sepsis can potentially be caused by very low numbers of pathogen per ml of blood (as low as 1-10 CFU/ml) (Kellogg et al., 2000; John Wain et al., 1998). The LoD was determined empirically by testing serial dilutions of known bacterial numbers (in colony forming units per one millilitre (CFU/mL)) then spiked into 1 ml of blood followed by processing through the

established pipeline. LoD experiments were performed in triplicate using blood samples spiked with *S. aureus* or *E. coli* (10, 50 and 100 CFU) as described in section **2.3.3.1**.

After sample processing, all three sets of replicates of each dilution were analysed by qPCR (before WGA) and one set of replicates by Illumina metagenomics sequencing (after WGA). The metagenomics data was profiled using Kraken (Illumina Basespace) to identify bacteria, fungi and viruses present in the sample. The top five organisms (ranked by number of reads as classified by Kraken) are presented in tables below. Further bioinformatics analysis was performed using Qualimap to determine the genome coverage of the top pathogen (mapped against the Illumina sequence of the spiked pathogen) as detailed in section **2.12.2**.

Table 5.1 shows qPCR results of the three replicates of each concentration of *E. coli* and *S. aureus* after depletion of human DNA. The average CFU count of spiked *E. coli* in one milliliter blood sample was 12, 47 and 113 for the samples '*E. coli* 10 CFU', '*E. coli* 50 CFU' and '*E. coli* 100 CFU' respectively. For *S. aureus*, the average colony count was 6, 58, and 91 for the samples '*S. aureus* 10 CFU', '*S. aureus* 50 CFU' and '*S. aureus* 100 CFU' respectively.

After host DNA depletion, the spiked bacteria in all samples were consistently detected by qPCR except for the samples spiked with an estimated 10 CFU/ mL. In these samples, *E. coli* and *S. aureus* were detected in one replicate at Cq 38.83 and 38.13 respectively. *E. coli* was detected at averaged Cq 37.3 and 36.28 in sample '*E. coli* 50 CFU' and '*E. coli* 100 CFU' respectively. *S. aureus* was detected at averaged Cq 37.17 and Cq 37.48 in samples '*S. aureus* 50 CFU' and '*S. aureus* 100 CFU' respectively (**Table 5.1**).

Table 5- 1 showing average colony count and qPCR results of three biological replicates of *E. coli* and *S. aureus* for LoD of the workflow for detection of BSI pathogens independent of culture.

Sample ID	**Average colony count (10-100 CFU/ mL)**	**1st replicate (Cq)**	**2nd replicate (Cq)**	**3rd replicate (Cq)**	**Average (Cq)**
E. coli 10 CFU	12	≥ 40	≥ 40	38.83	39.87
E. coli 50 CFU	47	38.22	36.55	37.13	37.3
E. coli 100 CFU	113	37.91	36.42	34.52	36.28
S. aureus 10 CFU	7	-	≥ 40	38.13	39.37
S. aureus 50 CFU	58	36.78	38.49	36.24	37.17
S. aureus 100 CFU	91	38.49	38.03	35.91	37.48

Replicate 3 of *E. coli* spiked samples was whole genome amplified and analyzed further by metagenomics sequencing. DNA quantity was measured using Qubit, Tapestation and qPCR as explained in section **2.8**. A Nextera library was made for sequencing on MiSeq (Illumina) (section **2.11.2**).

Total number of reads produced for the sample *E. coli* 10 CFU was 761,590. Most reads, 735,062 (96.52%) were not classified, and the remaining, 25,738 (3.48 %), were classified as non-human. The most abundant organism was *E. coli* with 1.17% of the total reads. The second most abundant organism was *Shigella sonnei* with 0.06% of the total reads and third was *Alteromonas macleodii* (0.03%) (**Table 5-2)**

Sample *E. coli* 50 CFU produced 1,011,256 total number of reads but only 1.94% were classified by Kraken analysis. *E coli* was most abundant species with 0.62% of the total reads followed by *P. fluorescens* (0.06%) *and A. macleodii* (0.03%) (**Table 5-2)**

A total of 850,580 reads were produced for the '*E. coli* 100 CFU' sample of which 799899 (93.85%) were unclassified and 50,645 (5.94%) were classified. *E. coli* was

the most abundant organism with 2.13% of the total reads, followed by *P. acnes* (0.08%) and *S. sonnei* (0.06%) (**Table 5-2**).

Table 5- 2 showing five most abundant species of Kraken analysis of three samples spiked with approximately 10, 50 and 100 CFU/mL of *E. coli.*

Sample ID	**Unclassified reads (%)**	**Top species (%)**	**2nd Species**	**3rd Species**	**4th Species**	**5th Species**
E. coli 10 CFU	96.38	*E. coli* (1.17)	*S. sonnei* (0.06)	*A. macleodii* (0.03)	Human endogenous retrovirus K (0.03)	*P. fluorescens* (0.01)
E. coli 50 CFU	97.98	*E. coli* (0.62)	*P. fluorescens* (0.06)	*A. macleodii* (0.03)	*S. aureus* (0.02)	Pandoravirus salinus (0.02)
E. coli 100 CFU	93.85	*E. coli* (2.13)	*P. acnes* (0.08)	*S. sonnei* (0.06)	Pandoravirus salinus (0.06)	Pandoravirus dulcis (0.04)

Further analysis was performed to determine genome coverage of the spiked *E. coli* by mapping to the reference genome using Qualimap. The '*E. coli* 10 CFU' sample had 1.1× genome coverage, the *'E. coli* 50 CFU' sample had 0.6 × and genome coverage of '*E. coli* 100 CFU' was 2.2 × (**Table 5-3**). Notably, genome coverage of the '*E. coli* 50 CFU' sample was lower than the '*E. coli* 10 CFU' sample. The difference may be due to the level of host DNA depletion achieved in that sample or unexpected loss of bacteria during the depletion process (this sample had the lowest proportion of *E. coli* reads) (**Table 5-3)**.

Table 5- 3 Genome coverage of *E. coli* in spiked blood when mapped to the reference genome

E. coli **(CFU/ml)**	**Total read count (thousands)**	**Reads mapping to hg19 (thousands)**	**Reads mapping to *E. coli* H141480453 (thousands)**	**Coverage (x)**	**Ratio (Human DNA: *E. coli* DNA)**
E. coli 10 CFU	762	69 (90.72%)	28 (3.74%)	1.1	24:1
E. coli 50 CFU	1,012	929 (91.87%)	18 (1.85%)	0.6	50:1
E. coli 100 CFU	852	818 (95.96%)	51 (6.06%)	2.2	16:1

The limit of detection of the workflow for *E. coli* was therefore concluded to be >10 but <50 CFU/mL.

For the *S. aureus*, the second replicate (Table 5-1) samples were analysed further by Illumina sequencing following WGA**.** Sample '*S. aureus* 10 CFU' (3 CFU/mL) produced a total of 403,446 reads of which 2,457 (0.6 %) were classified. The top species of classified reads was *E. coli* (0.15%) followed by Torque teno virus 16 (0.12%). *S. aureus* was third most abundant species with 0.05% (**Table 5-4**).

Sample '*S. aureus* 50 CFU' (57 CFU/mL) produced 652,870 reads of which 13,021 reads (1.9%) were classified. *S. aureus* was the top hit with 1.32% of classified reads (and 36% of the non-human reads) followed by *E. coli* (0.15%) (**Table 5-4**).

The '*S. aureus* 100 CFU' (138 CFU/mL) sample produced 477,143 reads of which 1.99% were classified. *S. aureus* was again the top hit with 1.42% of the total reads (and 45% of the non-human reads) followed by *Alteromonas macleodii* at 0.02% (**Table 5-4)**.

Table 5- 4 Top five species of Kraken analysis of three samples were spiked with dilutions of 10,50 and100 CFU/mL of *S. aureus*.

Sample ID	**Unclassif -ed reads (%)**	**Top Specie (%)**	**2nd Specie (%)**	**3rdSpecie (%)**	**4th Specie (%)**	**5th Specie (%)**
S. aureus 10 CFU	99.39	*E. coli* (0.15)	Torque teno virus 16 (0.12)	*S. aureus* (0.05)	*A. macleodii* str. (0.02)	-
S. aureus 50 CFU	98.01	*S. aureus* (1.32)	*E. coli* (0.16)	*A. macleodii* str. 'Ionian Sea U8' (0.05)	*P. aeruginosa* (0.02)	*Mycoplasma hyopneumoniae* (0.01)
S. aureus 100 CFU	98.16	*S. aureus* (1.42)	*A. macleodii* (0.02)	*E. coli* (0.01)	*Mycoplasma hyopneumoniae* (0.01)	Pandoravirus salinus (0.01)

Mapping the metagenomics reads to the *S. aureus* reference strain showed genome coverage of the approximately 0.1×, 1.3×, and1.4× for the samples *S. aureus* 10 CFU, *S. aureus* 50 CFU and *S. aureus* 100 CFU respectively (**Table 5-5)**.

Table 5- 5 Genome coverage of *S. aureus* in spiked blood when mapped to the reference strain

***S. aureus* (CFU/ml)**	**Total read count (thousands)**	**Reads mapping to hg19**	**Reads mapping to *S. aureus* NCTC 6571**	**Coverage (x)**	**Ratio (Human DNA: *S. aureus* DNA)**
3	806	767 (95.05%)	3,188	0.1	241:1
57	977	915 (93.66%)	25,424	1.3	36:1
138	954	918 (96.18%)	24,534	1.4	37:1

In general, metagenomics data analysis showed there was little difference between the sample spiked with 57 CFU and 138 CFU/mL in terms of percentage reads, background contamination, reads aligned to *S. aureus,* and the genome coverage. The limit of detection of the workflow for *S. aureus* was therefore concluded to be

>10 and <50 CFU/mL. The workflow was later tested on patient samples using MinION and Illumina NGS platforms.

5.2.2: Pathogen identification from prospectively collected blood samples

A total of seven (P1-7) clinical blood samples were collected prospectively from patients with suspected sepsis from NNUH Critical Care Complex (section **2-1**) and processed immediately through the established workflow (**Figure 3-12**). After depletion of human DNA, WGA and sequencing of seven samples, only one (P4) was identified as positive, containing *Streptococcus pyogenes*. *S. pyogenes* was identified by both MinION and Illumina platforms. Interestingly, the sample taken for metagenomics analysis in the ICU was negative by culture; however, a sample taken eight hours earlier from the same patient in the Emergency Department of NNUH was positive for *S. pyogenes* by blood culture. The cultured isolate from the Emergency Department was sent to Public Health England, Colindale, UK as it was an invasive *S. pyogenes.* PHE sequenced the isolate on an Illumina HiSeq platform and the sequences were made available for bioinformatics analysis. Results for all seven prospectively collected samples tested using the established metagenomics pipeline are presented below (**Table 5-6**).

Table 5- 6 Blood culture, human DNA depletion, and metagenomic results of the seven blood samples collected and analyzed prospectively. One sample (P4) was confirmed positive by both metagenomics based methods and blood culture, the rest were negative.

Sample ID	**Clinical blood culture results**	**Total reads (Illumina MiSeq) (millions)**	**Reads mapping to human (millions)**	**Reads unmapped to human (thousands)**	**Ratio human:non-human DNA**	**Kraken results (Top specie)**	**Metagenomics pipeline results**
P1	Negative	3.8	3.1 (82.9%)	657 (17.1%)	5:1	*A. macleodii* str. 'Ionian Sea U8'	N/A
P2	Negative	3.0	2.5 (85.6%)	431 (14.4%)	6:1	*A. macleodii* str. 'Ionian Sea U8'	N/A
P3	Negative	3.0	2.3 (78.9%)	638 (21.1%)	4:1	*Staphylococcus epidermidis*	N/A
P4	Positive* (*Streptococcus pyogenes*)	9.4	8.0 (85.9%)	1,329 (14.1%)	6:1	*S. pyogenes*	*S. pyogenes*
P5	Negative	2.8	2.3 (82.2%)	511 (17.8%)	5:1	*P. acnes*	N/A
P6	Negative	3.6	2.8 (77.2%)	837 (22.8%)	3:1	*S. epidermidis*	N/A
P7	Negative	2.9	2.5 (86.4%)	402 (13.6%)	6:1	*S. epidermidis*	N/A

*First sample was positive by culture but sample taken eight later hours after initiation of empirical treatment was negative.

Approximately 23,000 raw Illumina metagenomics reads from sample P4 aligned to *S. pyogenes* sequenced at PHE (mapping accuracy of 99.999%) (**Table 5-6** and **Figure 5-2).** Kraken analysis confirmed the presence of *S. pyogenes* as the most abundant organism in sample P4. The most abundant organisms in the remaining samples were classified as skin/workflow contaminants i.e. *A. macleodii* (P1 and P2), *S. epidermidis* (P3, P6 and P7) and *P. acnes* (P7) (**Table 5-6**). Therefore, using fresh clinical blood samples from patients with suspected sepsis the developed metagenomics pipeline correctly identified the only culture positive sepsis case.

5.2.3: Pathogen identification from retrospective blood samples

As positive sepsis samples were difficult to obtain prospectively, a retrospective study was performed with samples collected as part of a sepsis diagnostics evaluation study conducted by the University of Manchester and funded by the National Institute for Health Research (Warhurst et al., 2015) (section **2.1**). After sample collection, blood samples were aliquoted, biobank stored and later shipped to UEA for further analysis. From eight (R1-8) frozen clinical blood samples, four (R1 *K. pneumoniae*, R4 *E. faecalis*, R6 *P. aeruginosa* and R7 *S. aureus*) were positive for sepsis by blood culture (**Table 5-7**).

After processing through the metagenomics pipeline, samples R4, R6 and R7 correlated with blood culture results (**Table 5-7**). *Enterococcus faecalis* was isolated from sample R4 and diagnosed as the infecting pathogen, this correlated with the sequencing pipeline output. However, metagenomics analysis identified *Staphylococcus haemolyticus* as a potential second infecting pathogen which was not isolated via blood culture. Dual infection with *E. faecalis* and *S. haemolyticus* was also detected by SeptiFast (Roche), the PCR based method tested in the Manchester study. In sample R6, Kraken analysis detected *P. aeruginosa* but was not the top hit, the most abundant organism was *K. pneumoniae*. Sample R7 correlated with the blood culture result with the identification of *S. aureus* (approximately 14,000 reads). However, *K. pneumoniae* in R1 was not detected by metagenomics analysis and sample R5 contained sequences which mapped to *E. coli* despite negative blood culture results. For the other negative samples, metagenomics analysis detected *Edwardsiella tarda, Alteromonas macleodii,*

Yersinia enterocolitica in samples R2, R3 and R8 respectively, all considered workflow contaminants (**Table 5-7**).

Table 5- 7 Blood culture, metagenomics and depletion results of the eight retrospective blood samples. Four samples were positive for blood culture of which three correlated with metagenomic results (R4, R6 and R7).

Sample ID	**Clinical blood culture results**	**Total reads (Illumina MiSeq) (millions)**	**Reads mapping to hg19 (millions)**	**Reads unmapped to hg19 (millions)**	**Ratio human:non-human DNA**	**Kraken results (Top specie)**	**Metagenomics pipeline results**
R1	*K. pneumoniae*	13.0	3.9 (30.5%)	9.0 (69.5%)	1:2	*Achromobacter xylosoxidans*	N/A
R2	Negative	5.0	1.6 (32.5%)	3.3 (67.5%)	1:2	*Edwardsiella tarda*	N/A
R3	Negative	8.1	4.5 (55.7%)	3.6 (44.3%)	1.3:1	*A. macleodii* str. 'Ionian Sea U8'	N/A
R4	*Enterococcus faecalis*	7.0	3.9 (56.4%)	3.0 (43.6%)	1.3:1	*E. faecalis* *Staphylococcus haemolyticus*	*E. faecalis* *S. haemolyticus*
R5	Negative	4.8	3.1 (65.6%)	1.6 (34.4%)	2:1	*E. coli*	N/A
R6	*P. aeruginosa*	7.2	2.9 (40.4%)	4.3 (59.6%)	1:1.5	*K. pneumoniae*	*P. aeruginosa*
R7	*S. aureus*	10.5	5.2 (49.6%)	5.2 (50.4%)	1:1	*Achromobacter xylosoxidans*	*S. aureus*
R8	Negative	9.8	5.2 (53.3%)	4.6 (46.7%)	1.1:1	*Yersinia enterocolitica*	N/A

5.2.4: MinION metagenomic sequencing of blood samples

All samples identified as sepsis positive by Illumina sequencing, P4, R4 and R7, were sequenced using the MinION platform (with the exception of R6). Sample P4 was sequenced using Genomic Sequencing Kit SQK-MAP002 and R7 flow cell and SQK-MAP005 and R 7.3 flow cells was used for R4 and R7 (section **2.11.2**). MinION metagenomics data was analysed using kraken (**Table 5-6** and **5- 7**) and WIMP analysis (real-time analysis of MinION metagenomics data provided by ONT). In sample R4, both *E. faecalis* (488 reads) and *S. haemolyticus* (151 reads) in sample R4 were identified by WIMP (**Figure 5-1**). When discounting known contaminants (after genome coverage analysis, section **5.2.8**), *S. aureus* was identified as the pathogen in sample R7 (approximately 150 reads). *S. pyogenes* was detected in sample P4, however, the total number of reads from the MinION run was extremely low (817 reads in total of which 2 were *S. pyogenes*) as this was early MinION technology.

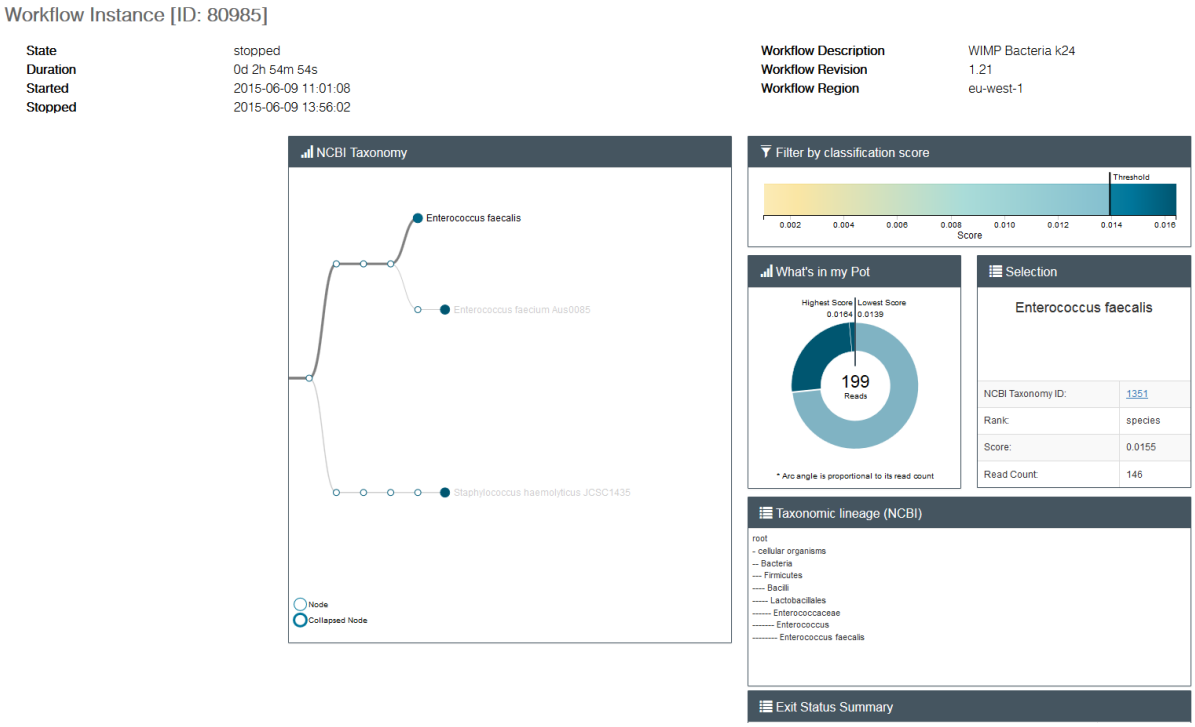


Figure 5- 1 WIMP analysis of R4 metagenomics data, *E. faecalis* and *S. haemolyticus* were correctly identified (first reads identified within minutes of sequencing).

5.2.5: Antibiotic resistance gene detection in metagenomics data

Positive samples confirmed for BSI by metagenomics (P4, R4, and R7) were analysed for the presence of antibiotic resistance genes using the CARD database (using Illumina data) as described in section **2.12.2**. Genome coverage was not sufficient to for analysis of resistance genes in *P. aeruginosa* (R6). In sample R4 (containing *E. faecalis* and *S. haemolyticus*), a number of antibiotic resistance genes were identified with ≥98% identity (strict criteria); *ermB*, *ermC*, *sat-4* and *fusB*. The *E. faecalis* isolated by blood culture was erythromycin and penicillin resistant, so detection of the *ermB* macrolide resistance gene correlated with culture. The penicillin resistance was likely to be inherent and the *fusB* gene was likely to be from the *S. haemolyticus*. R7 had no antibiotic resistance genes using perfect and strict criteria but two membrane proteins of an efflux complex were identified on loose criteria with ≥97% identity (*smeF* and *smeB* respectively),blood culture isolate was sensitive to all drugs tested (**Table 5-8**). Using all criteria there were no antibiotic resistance genes identified in P4 and the *S. pyogenes* isolated by blood culture was fully sensitive to antibiotics. However, genome for P4, R6 and R7 was low for reliable identification of resistance genes.

Table 5- 8 Resistance genes identified using the CARD database, percentage identity and expectation value of sample R4 and R7. The lower the E value, the more significant the score and the alignment.

Sample	Gene	E value	% identity
R4	ermB	2.42×10^{-180}	99
	ermC	8.97×10^{-178}	99
	fusB	2.04×10^{-91}	100
	sat-4	4.13×10^{-132}	98
R7	smeB	4.06×10^{-63}	99
	smeF	2.69×10^{-98}	97

5.2.6: Depletion of human DNA

The initial estimate of human to pathogen DNA in blood samples was approximated to be 10^9:1. The method for removing human DNA that combined CD45 IMS and

MolYsis was estimated to remove 10^6-fold (99.9999%) of human DNA. These were estimates based on qPCR results, however, NGS showed the level of depletion was higher than 10^6. The highest ratio of human to non-human reads was 6:1 (P2 and P7) and the lowest was 1:2 (R1 and R2) (**Table 5-9**). However, varying proportions of non-human reads were background reads, so the true level of host depletion was hard to estimate. Highest level of depletion of human DNA was observed in frozen blood samples (R1-R8), this was because most of the human cells were lysed due to freeze-thaw cycles, but consequently, freeze-thaw was thought to have also lysed pathogen cells, particularly Gram-negative bacteria. Sample number R2 had a lower initial amount of human DNA, which was because the sample was coagulated (was not collected in anticoagulant tubes) (**Table 5-9**).

Table 5- 9 qPCR determined human depletion levels and ratios of human: non-human reads after samples processed through the metagenomics pipeline using Illumina platform.

Patient number	**Human qPCR before depletion (Cq)**	**Human qPCR after sample (Human Cq)**	**Approximate human DNA fold depletion**	**Reads ratio of human: non-human**
P1	21.27	37.08	10^5	5:1
P2	20.89	33.45	10^4	6:1
P3	21.54	34.56	10^4	4:1
P4	21.23	30.63	10^3	6:1
P5	21.18	>45	10^6	5:1
P6	21.04	36.19	10^5	3:1
P7	20.11	34.02	10^4	6:1
R1	24.14	40	10^5	1:2
R2	35.33	>45	$>10^3$	1:2
R3	23.89	32.28	10^3	1.3:1
R4	25.74	>45	10^6	1.3:1
R5	23.15	>40	10^6	2:1
R6	23.66	39.67	10^5	1:1.5
R7	22.93	>40	10^6	1:1
R8	25.31	35.87	10^3	1.1:1

5.2.7: Analysis of Genome coverage

In sample P4, approximately 23,000 Illumina reads aligned to the *S. pyogenes* isolate sequenced at PHE with 0.2× genome coverage. The coverage plot, **Figure 5-2**, shows coverage across the genome but with some spikes, likely caused by uneven WGA amplification. In the retrospective study, analysis of sample R4 showed 1.8 million reads aligning to *E. faecalis* (51 × genome coverage), **Figure 5-3**, and 350,000 aligning to the second infecting pathogen *S. haemolyticus* (9× genome coverage) **Figure 5-4**. The genome coverage plots for these organisms were similar to *E. faecalis*. Sample R7 correlated with the blood culture result by the identification of *S. aureus* (approximately 14,000 reads) but genome coverage was very patchy (**Figure 5-5**). In sample R6, approximately 30,321 reads aligned to *P. aeruginosa* with estimated 0.13× genome coverage, **Figure 5-6.**

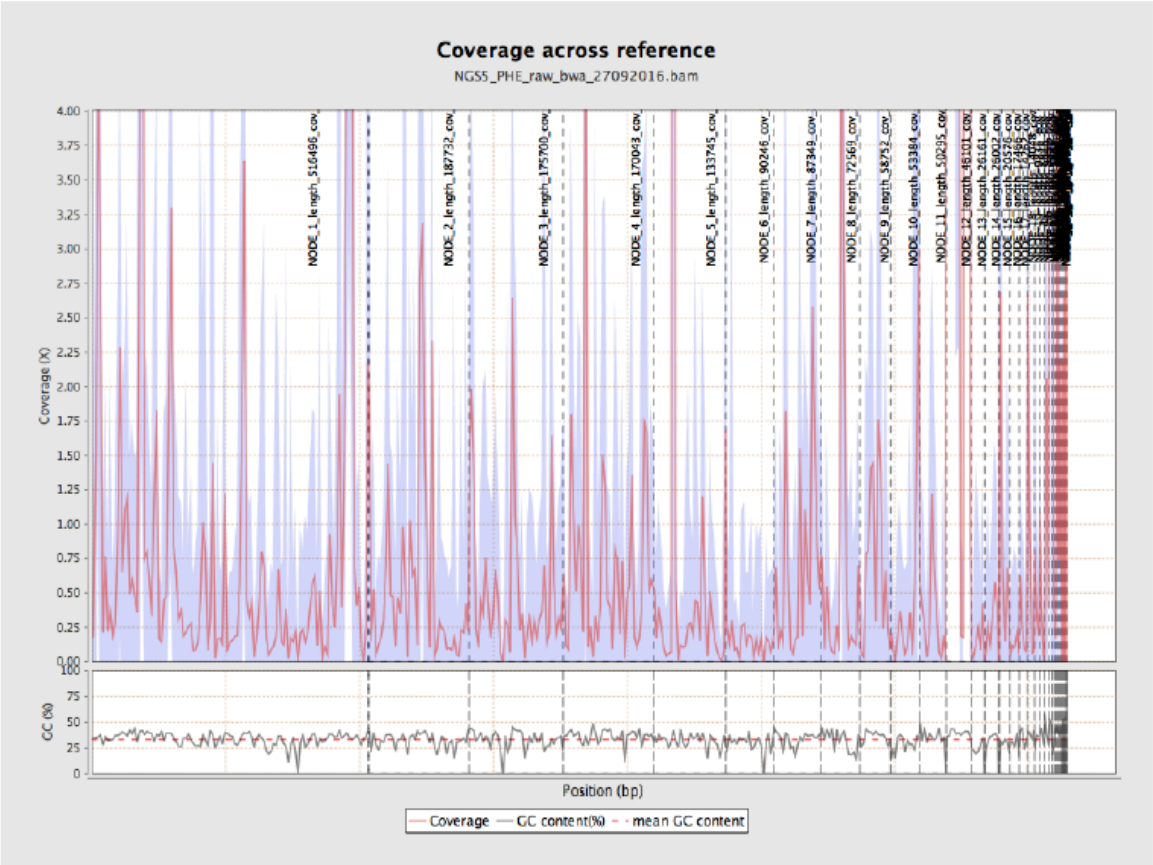


Figure 5- 2: Coverage plot of P4 against PHE *S. pyogenes* isolate from Illumina data

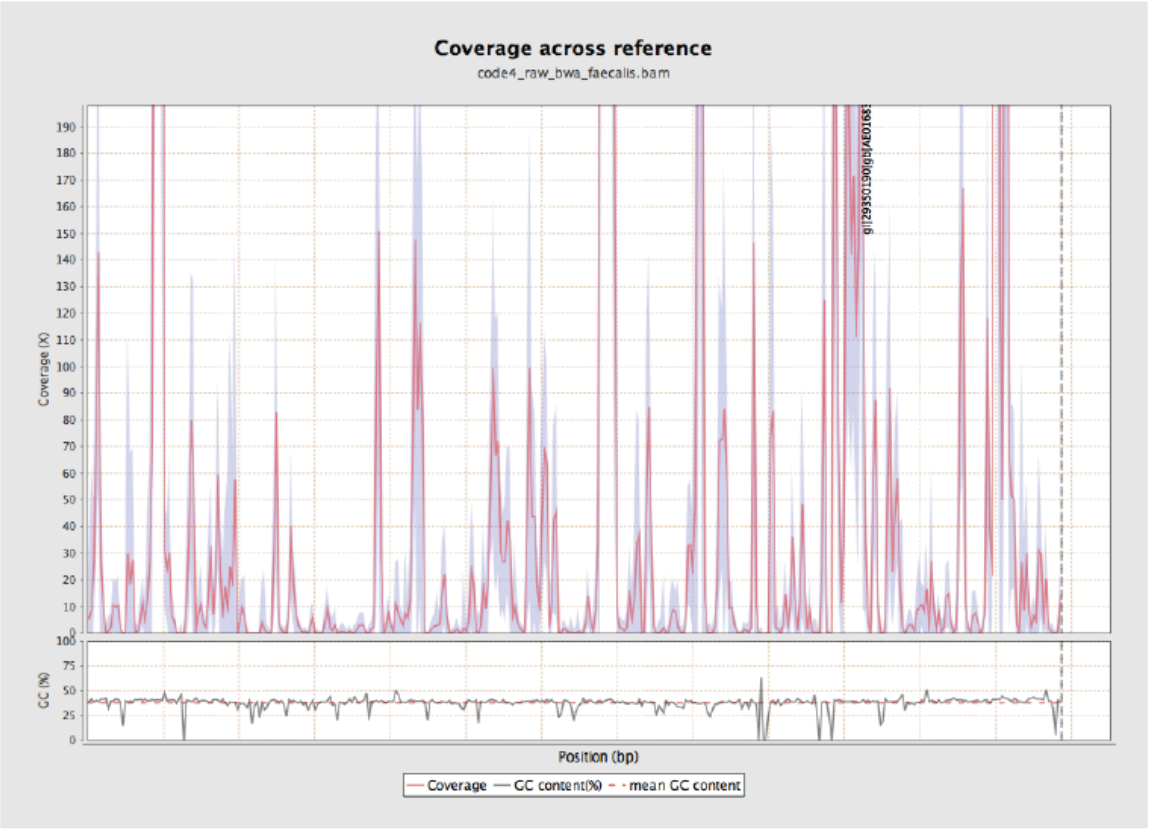


Figure 5- 3 showing coverage of *E. faecalis* in sample R4 using Illumina data

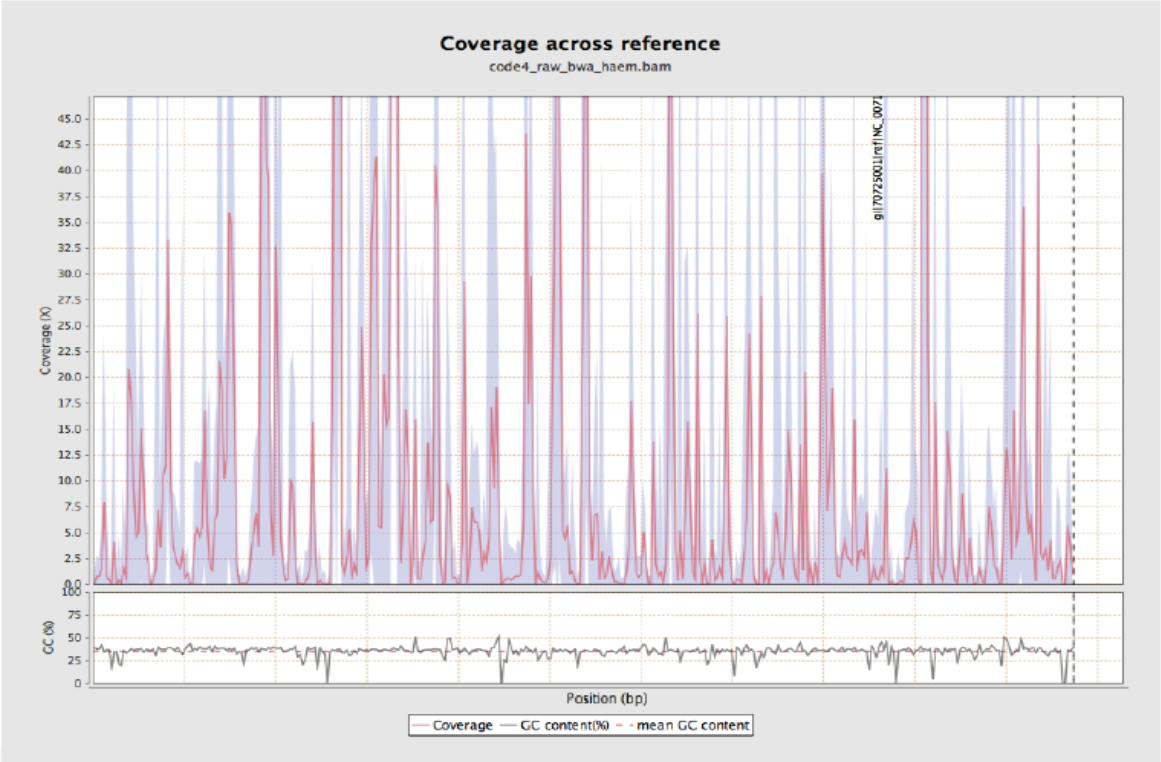


Figure 5- 4 Coverage plot of R4 of *S. haemolyticus* by Illumina data.

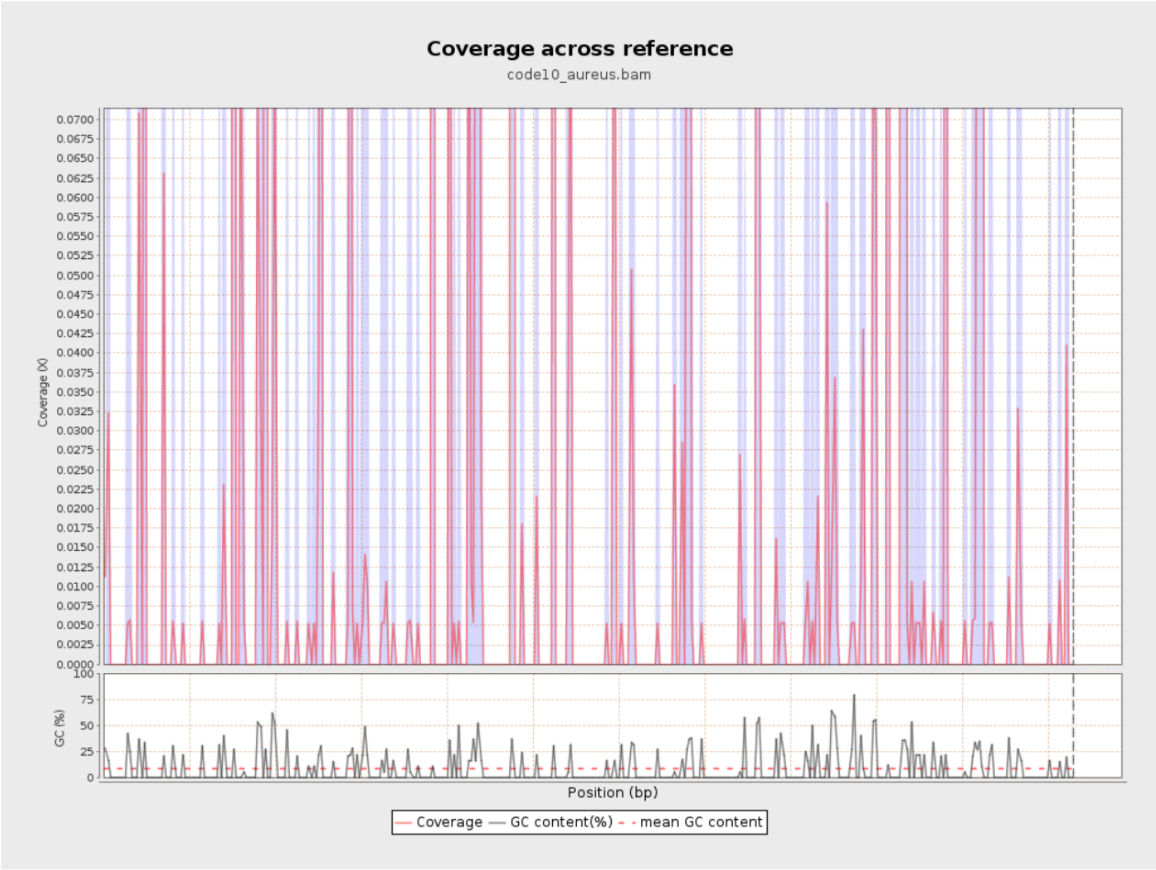


Figure 5- 5 showing coverage plot of *S. aureus* in sample R7 using Illumina data

Figure 5- 6 showing coverage plot of *P. aeruginosa* in sample R6 using Illumina data

As seen from the above coverage plots, although depth of coverage was not high, the breath of coverage is high and the chance of missing resistance genes is low (drop outs in coverage can be seen on the GC plot where the GC content drops to zero). In some samples, there was insufficient depth of coverage to detect mutational resistance.

5.2.8: Analysis of contaminants

Detection of *A. macleodii* in all six samples in LoD experiments (**Table 5-2** and **5-4**) raised a concern about contamination and background sequences. When molecular grade water (provided with Qiagen WGA kit) was amplified without adding template DNA, 57.5% of reads were classified as human and the rest were bacteria reads. The most abundant specie was *Staphylococcus lugdunensis* with 19.88% of the total reads followed by *Pseudomonas putida* (17.87%) and *Stenotrophomonas maltophilia* (10.51%) (Table **5-10**).

Table 5- 10 showing five most abundant species of Kraken analysis of the whole genome amplified sample without adding template DNA.

Sample ID	**Unclassified reads (%)**	**Top Species (%)**	**2nd Species (%)**	**3rd Species (%)**	**4th Species (%)**	**5th Species (%)**
WGA NTC	42.5	*Staphylococcus lugdunensis* (19.88)	*Pseudomonas putida* (17.87)	*Stenotrophomonas maltophilia* (10.51)	*Delftia ssp* (1.68)	*Delftia acidovorans* (1.18)

Other common detected bacteria from spiked and clinical samples were *E. coli, S. epidermidis, P. acnes* and *K. pneumoniae.* But detection was not consistent throughout the samples, different samples were affected by different contaminants . Most these bacteria were believed to have been contaminants present in kits, plastics and reagents, laboratory personnel and environment, however unexpected detection of *K. pneumoniae* and *E. coli,* major causes of BSI, was a major concern. Therefore, further analysis was required to identify if contaminants were introduced during the laboratory protocol. Additional bioinformatics analysis showed very patchy genome coverage (breath of coverage was <1% of the genome) of most of the contaminants suggesting the workflow was contaminated with DNA fragments rather than the whole organism/genome. This included *Alteromonas macleodii* (**Figure 5-7**)*, Bradyrhizobium diazoefficien, Edwardsiella tarda, Achromobacter xylosoxidans*, and *Y. enterocolitica.* Most of these organisms are known be environmental or kit contaminants, we hypothesise that these DNA fragments come from the kits, personnel, lab environment or plastics and get amplified repeatedly by multiple displacement amplification (WGA) and sequenced.

However, some contaminants had significant breath of coverage and it was not easy to rule them out as contamination. This included *E. coli* (**Figure 5- 8**), *S. epidermidis* and *K. pneumoniae.* It is likely the *E. coli* came from enzymes used in the pipeline that have been produced in recombinant *E. coli*, the *S. epidermidis* is likely from the blood sample when it was taken or from lab personnel and *K. pneumoniae* was probably cross-contamination in the laboratory.

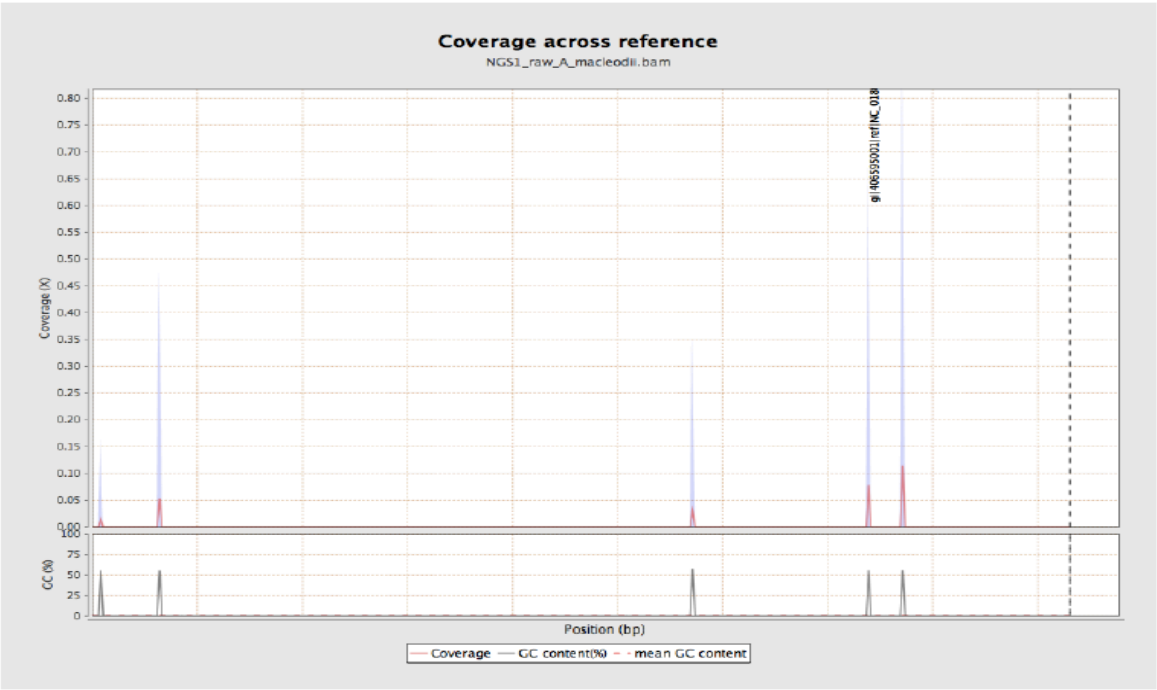


Figure 5- 7 Coverage plot showing very patchy coverage across the *Alteromonas macleodii* reference suggesting DNA fragment contamination.

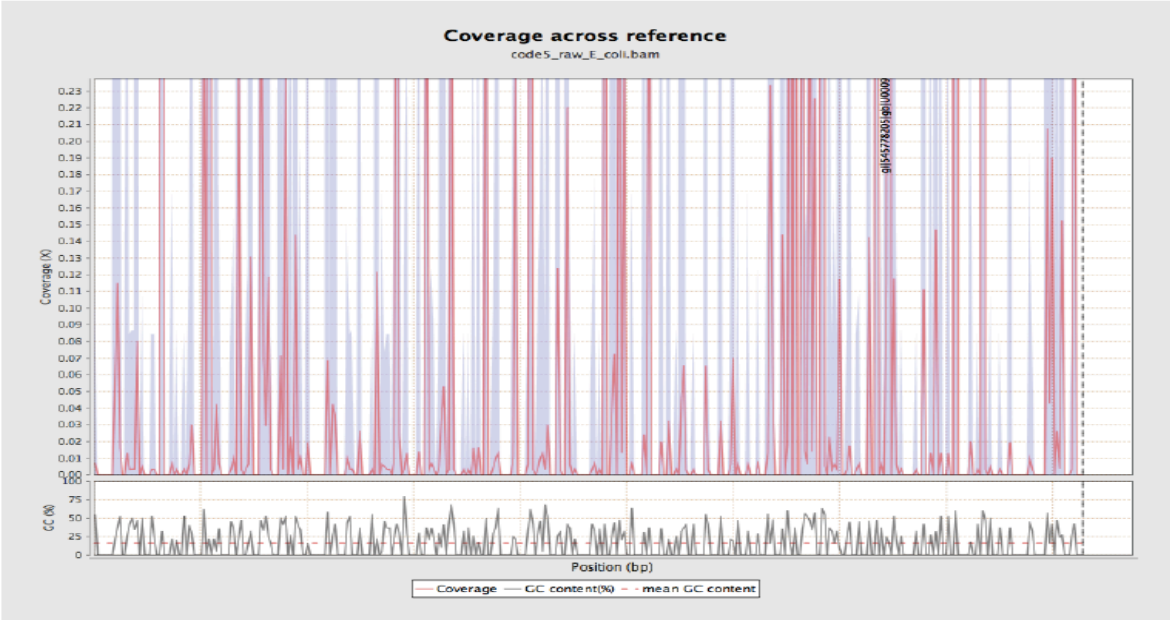


Figure 5- 8 Coverage plot showing relatively broad coverage across the *E. coli* reference, likely to be caused by whole genome rather than genome fragment contamination.

5.3: MinION Metagenomic Sequencing of Urine Samples

To deplete host DNA in urine samples, four methods were tested to find the most efficient and reliable method as described in section **3.3**. The best method was human DNA depletion by differential centrifugation (DC) followed by differential lysis of the residual human cells using the MolYsis kit. The method was rapid, cheap and efficient.

5.3.1: Analysis of Depletion of Human DNA

To test the efficiency of the DC + MolYsis depletion method for use with metagenomics sequencing, two infected clinical urines ($\geq 10^7$ pathogen CFU/mL) from patients at the NNUH were tested. Each sample was divided into two aliquots of 2 mL; one aliquot was depleted of human DNA using DC + MolYsis method as explained in section **2.4.3,** and the other aliquot was not depleted (positive control). For the positive control, cells were collected by pelleting at ≥ 12,000 × g for 5 minutes and thereafter total DNA was extracted as explained in section **2.6.2**. All samples were sequenced using MinION as explained is section **2.11.3,** using Genomic Sequencing kit MAP002 and R7 flow cells.

Table 5-11 below shows analysis of urine 1 MinION sequencing data comparing depleted and undepleted aliquots. Total 2D pass reads obtained in the undepleted sample was 1865 and there were 1831 in depleted sample. The table shows that the percentage of human reads decreased from 78% in undepleted sample to 6.5% in depleted sample. Depletion of human DNA also led to an increase of bacterial (Proteobacteria) reads from 316 to 1322 reads (4-fold increase). Results show most reads mapping to *K. pneumoniae*, suggestive of *K. pneumoniae* infection. Reads mapped to *K. pneumoniae* increased from 310 in undepleted to 887 in depleted, with the percentage of genome coverage doubling from 29.16% to 55.4% (**Table 5-11, Figure 5-9**).

Table 5- 11 Analysis of urine 1 MinION data comparing depleted and undepleted aliquots

	Undepleted Urine 1 (Total no of 2D reads 1,865)	**Depleted Urine1 (Total no of 2D reads 1,831)**
Percentage of Human reads	78%	6.5%
Proteobacteria reads	316	1322
Human reads	1125	99
K. pneumoniae reads	310	887
% coverage of *K. pneumoniae* CG43	29.16%	55.4%

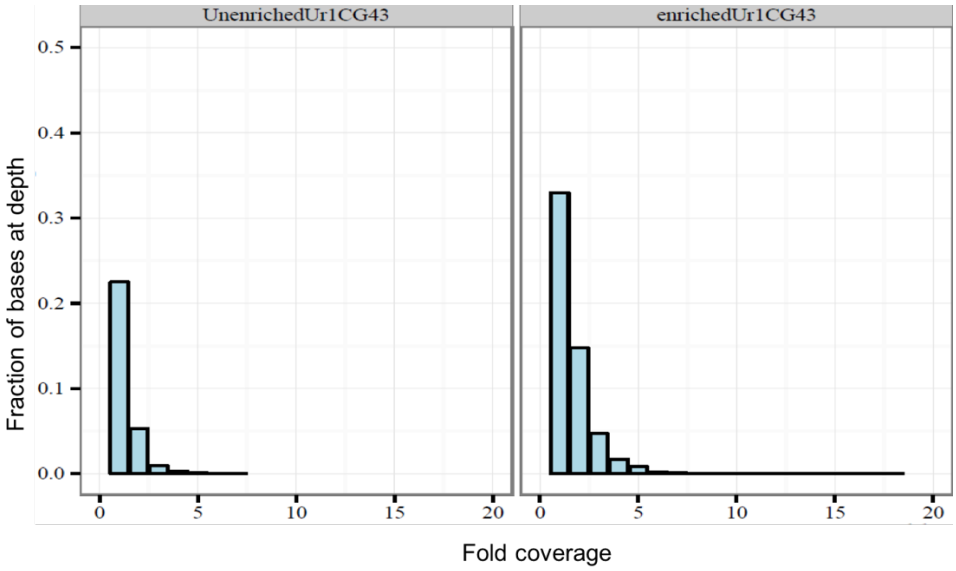


Figure 5- 9 Urine 1 *K. pneumoniae* GC43 coverage (left) before depletion and (right) after depletion of host DNA

Table 5-12 shows analysis of MinION sequencing data of Urine 2 before and after depletion of host DNA. Compared to urine 1, urine 2 had significantly less human contamination. Results show host DNA has been depleted to almost undetectable levels, from 14% human reads in undepleted sample to 0.2% in depleted sample (810 to 3 reads). Also, depletion of host DNA resulted in doubling of *E. coli* reads (From 746 to 1331) and increased percentage of genome coverage from 55% to 64% (**Table 5-12** and **Figure 5-10**).

Table 5- 12 Analysis of urine 2 MinION data comparing depleted and undepleted aliquots

	Unenriched Urine 2 (Total no of 2D reads 1146)	**Enriched Urine2 (Total no of 2D reads 1832)**
Percentage of Human reads	14.07%	0.2%
Hits on Proteobacteria	128	1533
Hits on Human	810	3
Escherichia coli	746	1331
E. coli Xuzhou 21	54.99%	64.5%

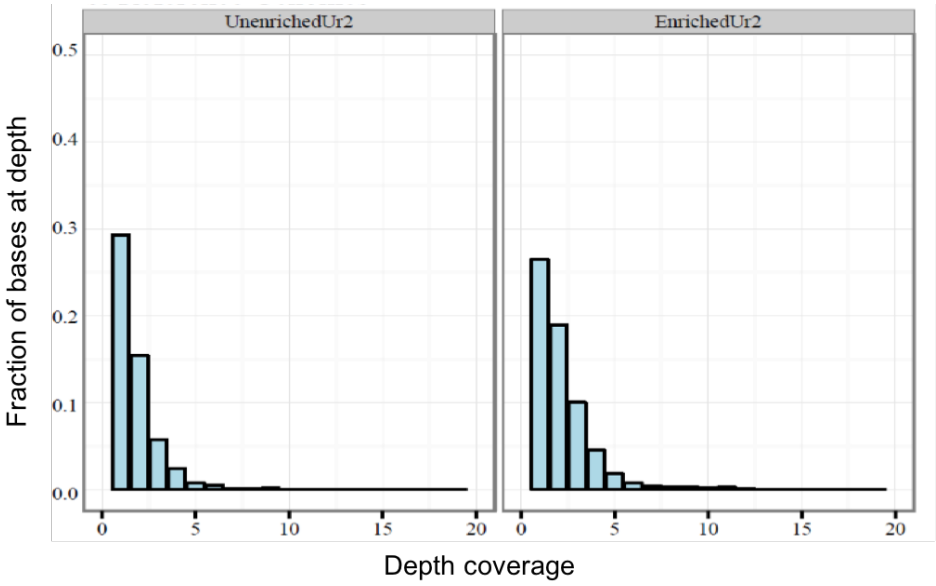


Figure 5- 10 Urine 2 *E. coli* genome coverage (left) before depletion and (right) after depletion of host DNA

Depletion of host DNA resulted in increased pathogen genome coverage which is important for identifying strain type and antibiotic resistance markers. Sequencing times (and cost) are also reduced by removing unwanted human DNA.

5.3.2: Analysis of Series of Clinical Urine Samples

We then applied the depletion method to a series of clinical urine samples collected at the NNUH between January and October, 2015. In this study, ten suspected UTI urine samples were collected and processed at the UEA Medical School. Additionally, urine from a healthy volunteer was spiked with 10^8 CFU/mL of the MDR

E. coli strain H141480453 and another *E. coli* strain (cultivated from clinical sample) was also analysed by MinION sequencing.

Host DNA was depleted from clinical and spiked urine samples as explained in section **2.4.3**. Human depleted DNA was then sequenced using MinION by genomic sequencing kits SQK MAP002, 003, and MAP SQK-005 and flow cells R7 and R7.3 (section **2.11.3**). Sequences were analysed using external databases and bioinformatics pipelines, also, using integrated real-time analysis applications ONT's Metrichor's WIMP and Antimicrobial Resistance Mapping Application (ARMA) for real-time detection of antibiotic resistance genes. Results were compared with Illumina data and resistance phenotypes (Schmidt et al., 2016).

Fifteen MinION runs were performed: Ten on clinical urines (CUs), four using urine spiked with *E. coli* H141480453 and one spiked with the *E. coli* from CU6 (**Table 5-13**). Early attempts failed because: (i) human DNA was insufficiently depleted using NEB Microbiome Enrichment kit (CU1) (ii) flow cells were poor quality (CU2 and 4) (iii) Poor quality/degraded DNA (CU3). Improved depletion method, flow cells (R7.3) and sequencing kits resolved these issues. From CU5 onwards, MinION produced 6536 to 34330 2D reads per run and a mean read-length of 3452 to 6076 bp. The longest single read was 46213 bp and single-read identity to reference sequences improved from 70% to 85% (**Table 5-13**).

Furthermore, improved methods increased sequence yield and depth. Breadth of coverage was from 82.6% to 100%; depth was least for CU5 (2.71×) and greatest (21.55– 22.84×) for spiked urine run 2 and CU8. In all cases, MinION correctly identified the pathogen, WIMP achieved this within 15 min. Human DNA accounted for only 1.6%–12.3% of reads, confirming that depletion was effective (**Table 5-14**). Results showed MinION sequencing identified the pathogens as culture results and detected 51 of the 55 acquired resistance genes that were identified by Illumina. However, MinION coverage was too low for SNP calling (Schmidt et al., 2016).

Sample and date	Flow cell	Sequencing time (hours)	Total number of reads	Mean read length	Number of 2-D pass reads	Mean read-length of 2-D pass (bp)
CU1 09-07-2014	R7.0	24	12 295	3647	1645	—
CU2 12-07-2014	R7.0	24	8299	2859	621	—
CU3 04-09-2014	No Results					
CU4 09-09-2014	R7.0	21	3829	1728	184	
Urine spiked with *E. coli* H141480453 run 1; 06-11-2014	R7.3	30	45 652	2827	15 216	4103
CU5 16-01-2015	R7.3	25.5	22 968	3292	8191	3980
CU6 24-01-2015	R7.3	23	57 289	4700	15 932	5510
CU7 05-02-2015	R7.3	17.5	76 499	4473	17 050	5414
Urine spiked with *E. coli* from CU6 09-03-2015	R7.3	14	56 394	5419	13 206	6076
CU8 02-03-2015	R7.3	33	86 294	4664	20 799	5324
CU9 30-03-2015	R7.3	26	28 767	4926	6536	5741
CU10 16-05-2015	R7.3	35	141 511	3107	34 330	3452
Urine spiked with *E. coli* H141480453 run 2; 04-05-2015	R7.3	48	138 720	4424	33 589	3452
Urine spiked with *E. coli* H141480453 run 3; 23-10-2015	R7.3	7.5	97 961	4308	28 787	4416
Urine spiked with E. coli H141480453 run 4; 26-01-2016	R7.3	29	21 441	2043	—	—

			spiked with *E. coli* from CU6					**with *E. coli* H141480453 run 1**	**with *E. coli* H141480453 run 2**	**with *E. coli* H141480453 run 3**
Reads used	2-D pass only	2-D pass only	2-D pass only	2-D pass only	2-D pass and fail	2-D pass and fail	2-D pass and fail	2-D pass only	2-D pass only	2-D pass only
Non-human reads matching bacteria	76%	84%	83%	84%	81%	95%	85%	98%	89%	—
Reads matching human	6.6%	8.5%	8.5%	8.1%	12.3%	1.7%	9.7%	1.6%	4.2%	—
Best match species	*K. pneumoniae* CG43	*E. coli* JJ1886	*E. coli* JJ1886	*E. coli* PMV-1	*E. coli* 536	*E. cloacae* NCTC 9394	*K. pneumoniae* CG43	*E. coli* APEC O78	*E. coli* K-12	*E. coli* APEC O78
Breadth coverage of matching bacteria	82.57%	99.59%	100%	92.19%	99.9%	86.25%	96.70%	95.13%	96.13%	—
Genome coverage	2.71×	15.65×	10.58×	10.77×	22.84×	9.16×	17.61×	7.25×	21.55×	21.51×
Run time	25.5	23	14	17.5	36	26	35	30	48	7.5

5.4: Culture independent metagenomics identification of *Clostridioides difficile* directly from stool samples.

To be able to diagnose and type *C. difficile* directly from stool samples using metagenomics, a novel method of depleting host DNA and high levels of commensal gut floral was developed. The designed method used buffer FL (GeneAll) to lyse background host and commensal flora cells and enrich for *C. difficile* spores, the infectious and transmissible morphotype of CDI (section **2.5.3**), followed by DNA extraction from spores (section **2.6.3**). Depletion of normal flora and host DNA was completed in less than 8 hours and thereafter the spore enriched DNA was available for sequencing on the Illumina MiSeq (section **2.11.2**).

5.4.1: Analysis of depletion of human/commensal DNA

Three stool samples, confirmed to be *C. difficile* toxin positive by either diagnostic PCR assay, EIA or both, were collected from NNUH and processed at UEA as described above. Initial qPCR analysis showed depletion of *E. coli* (used as a marker for commensal depletion as we found high levels of *E. coli* in patients with CDI) and host cells at range of 10 to 10^3 – fold. The *C. difficile* remained relatively unchanged and surprisingly the 16S rRNA results showed no commensal depletion in 2 of 3 samples tested **Table 5-15.**

Table 5- 15 qPCR results three stool samples, before depletion (PC) and after depletion.

Sample ID	**Human qPCR (Cq)**	***E. coli* (Cq)**	***C. difficile* qPCR (Cq)**	**Bacterial 16S rRNA (Cq)**
S7007	40	30.47	-	15.59
S7007 PC	40	21.10	-	11.98
S7009	-	30.68	32.77	9.89
S7009 PC	36.51	23.57	32.68	10.69
S6994	34.05	-	24.74	16.03
S6994 PC	22.23	36.98	22.97	16.50

Further analysis using NGS data (MiSeq 2 x 300bp) was performed by mapping reads to the Human (hg19), *C. difficile* 630 and E coli UMN026 reference genomes using Bwa-mem alignment (section **2.12.2**). The analysis showed decrease in human sequences by 20% in sample S6994, 25% in sample S7009, and 50% in sample S7007. The proportion of *E. coli* reads was reduced from 7% to 3.0% in sample S6994, 1.7% in sample S7007 to 0.01%, and remained the same in sample S7009. The percentage of *C. difficile* reads remained mostly the same except for the sample S6994 in which there was an increase from 3.6% to 6.3%, **Table 5-16.**

Table 5- 16 Proportion of Human, *E. coli* and *C. difficile* reads in the clinical stool metagenomes.

Sample ID	**Total reads**	***C. difficile* reads (%)**	**Human reads (%)**	***E. coli* reads (%)**
S6994	282,520	17,921 (6.3)	220,455 (78)	8,578 (3.0)
S6994 PC	155,098	495 (0.3)	139,588 (89)	10,856 (7)
S7007	275,768	8,055 (2.9)	69,278 (25)	30 (0.01)
S7007 PC	387,944	12,517 (3.2)	293,737 (75)	6,737 (1.7)
S7009	204,362	15,206 (7.4)	75,891 (37)	5,496 (2.6)
S7009 PC	238,770	16,175 (6.8)	197,144 (82)	6,286 (2.6)

Despite high reduction of human reads in sample S7007 and S7009 (approx. 50%), there was no substantial increase of *C. difficile* reads, the rest of the reads were thought to be for commensal flora.

By depleting unwanted DNA by differential lysis of cells, and using efficient DNA extraction from spores method, we were able to reduce human and *E. coli* reads without affecting *C. difficile* reads.

5.4.2: Analysis of Genome coverage

In sample S6994, approximately 17,921 Illumina reads aligned to the *C. difficile* 630, the mean genome coverage was 0.2×, **Figure 5-10.** The respective un-depleted

control, sample S6994 PC, had approximated 5,636 *C. difficile*, the mean genome coverage was ten times lower 0.02× **Figure 5-11.**

Figure 5- 11 Illumina data showing coverage plot of *C. difficile* in depleted sample S6994.

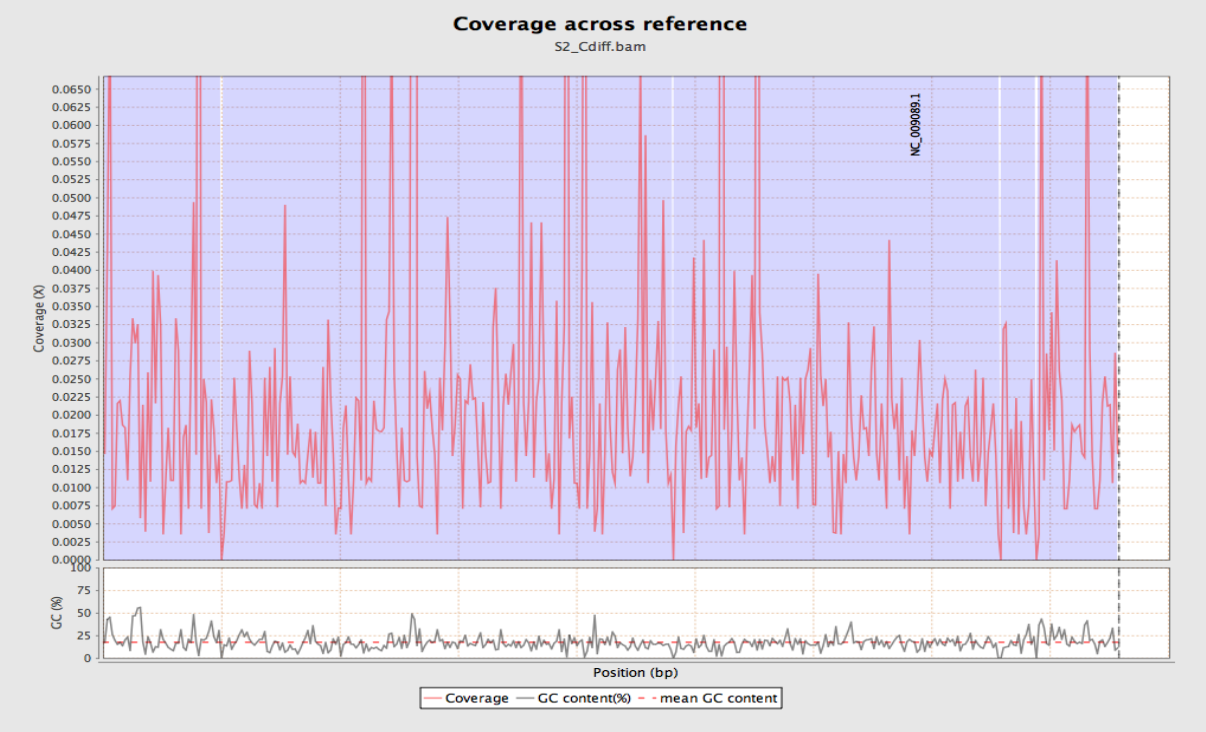


Figure 5- 12 *C. difficile* genome coverage plot of in un-depleted sample S6994 PC

Sample S7007 had approximately 8,055 *C. difficile* reads with genome coverage estimated to be 0.2 × **Figure 5-13**. The respective positive control, S7007 PC, had 12,517 *C. difficile* reads with approximately 0.1 × genome coverage **Figure 5-14**.

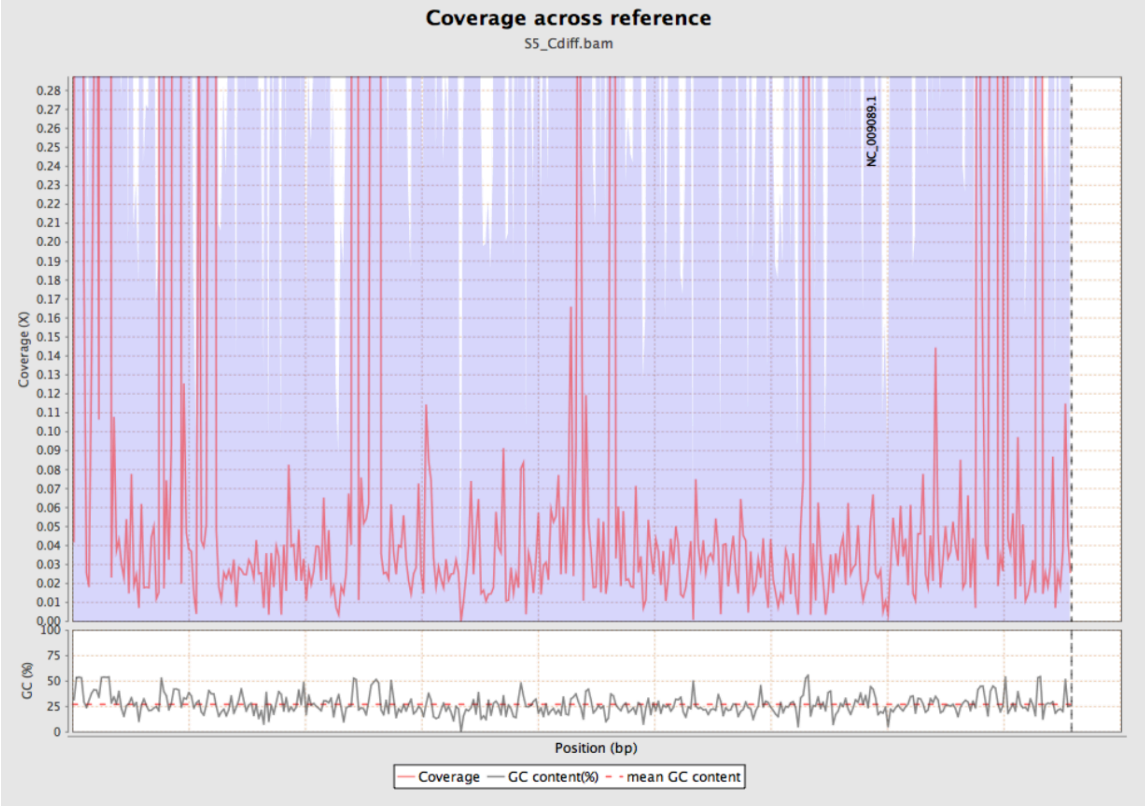


Figure 5- 13 showing coverage plot of *C. difficile* in depleted sample S7007

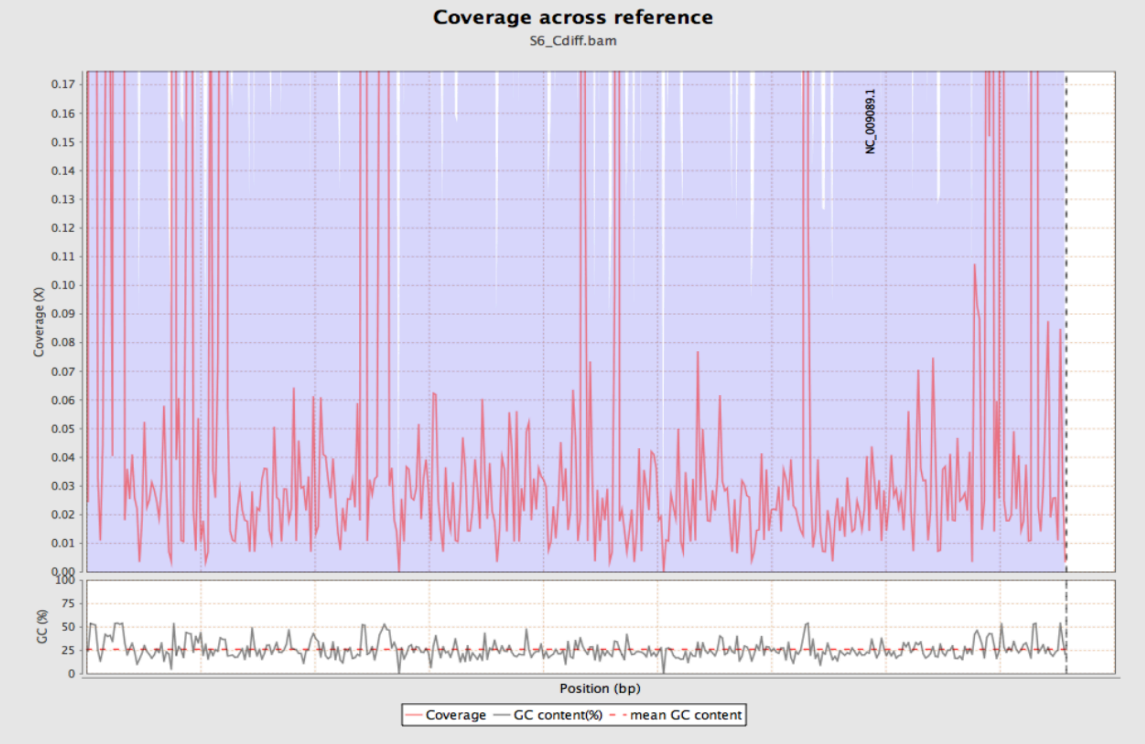


Figure 5- 14 showing coverage plot of *C. difficile* in un-depleted sample S7009 PC

There was no big difference between sample S7009 and the respective positive control S7009 PC. Number of *C. difficile* reads were 15,206 and 16,175 for depleted and un-depleted sample respectively, with approximately 0.3 × genome coverage for both **Figure 5-15** and **Figure 5-16**.

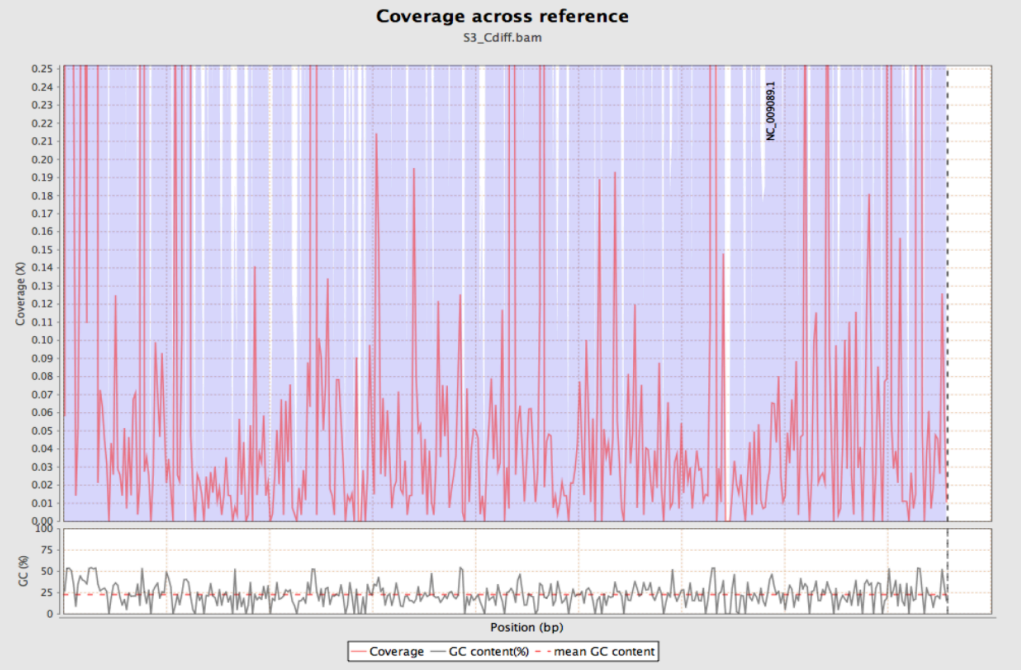


Figure 5- 15 showing coverage plot of *C. difficile* in depleted sample S7009

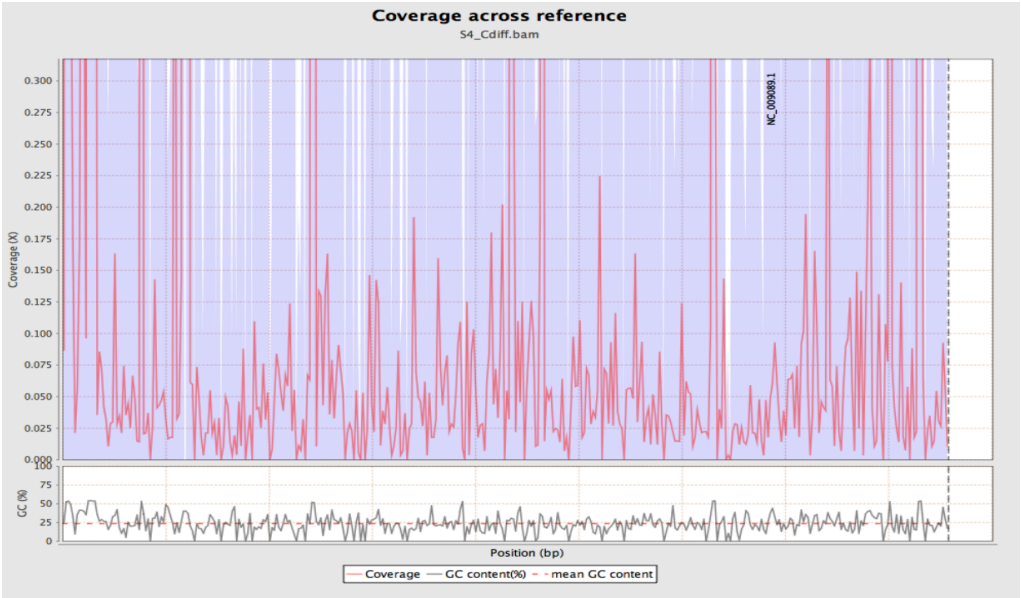


Figure 5- 16 showing coverage plot of *C. difficile* in un-depleted sample S7009 PC

Although all the above genome coverage plots showed a good uniform genome coverage in breadth and depth, the depth of coverage was not high enough for SNP typing, identification of virulence markers (Toxin A and B) or identification of resistance genes.

5.5: Chapter discussion

In this chapter, we have shown a proof-of-concept that pathogenic organisms in clinical samples can be comprehensively characterised by metagenomics sequencing. Unbiased metagenomic diagnosis of pathogens and antibiotic resistance has been achieved using culture-independent pipelines.

5.5.1: Blood

Metagenomics diagnosis of BSI was performed by developing a rapid workflow based on depletion of vast amount of human DNA present in blood, and WGA of the remaining pathogen DNA. Turnaround time of the workflow was estimated to be 7 to 8 hours when sequencing using MinION and 3 days by Illumina platform. The LoD of was estimated to be between 10 and 50 CFU/mL for both Gram positive and negative bacteria. The established LoD is similar to other widely used commercial molecular assays of detecting pathogens directly in whole blood including SeptiFast (Roche) 3 - 30 CFU/mL (Chang et al., 2013), SeptiTest (Molzym) 20 - 460 CFU/mL (Kuhn et al., 2011; Wellinghausen et al., 2009), and Magicplex Sepsis (SeeGene, Seoul, Korea) 30 CFU/mL (Ziegler, Fagerström, Strålin, & Mölling, 2016). Turnaround time of these molecular assays is between 6 to 8 hours (Mwaigwisya et al., 2015; Schmiemann et al., 2010). Although improvement is still required as the circulating pathogen load in BSI can be as low as one and ten CFU/mL (Afshari et al., 2012; Kellogg et al., 2000; John Wain et al., 1998). Our results suggest the established metagenomics pipeline can detect pathogen at equivalent sensitivity (and turnaround times) as other currently available commercially molecular tests - but is significantly more comprehensive in the pathogens and antibiotic resistances it can detect.

When tested in seven clinical blood samples collected prospectively, diagnostics results of metagenomics pipeline correlated with culture results by identifying one sample as positive for *S. pyogenes* and the rest as negative. Interestingly, the sample taken for metagenomics analysis was negative by culture methods, however the sample taken eight hours earlier from the same patient, before initiation of broad spectrum antibiotics, was positive. This correlates with the published reports that microbiological cultures report no pathogen in about 40-70% of patients with severe sepsis or septic shock (Bochud, Bonten, Marchetti, & Calandra, 2004; Vincent et al., 2006). The reason for failed detection by culture in this particular sample was likely due to initiation of broad spectrum antibiotic treatment in the emergency room killing the *S. pyogenes* before the second sample was taken in the ICU. Detection of dead pathogens is important in management of BSI as it identifies initial cause of infection/illness, and it gives a clinician a snapshot of what is going on in patient bloodstream (Opota et al., 2015). Such information can help to make important decisions such as de-escalating empirical therapy to a tailored treatment. Benefits of tailored treatment include: use of less toxic agents; use of antibiotics that achieve higher concentrations; providing the option to select more effective agents if inherent or acquired resistance is detected; fewer side effects; and more favorable clinical outcomes (Garnacho-Montero et al., 2014).

In eight retrospectively collected blood samples, the metagenomics pipeline identified three of the four culture positive samples. In sample R4 dual infection with *E. faecalis* and *S. haemolyticus* was detected by metagenomics and PCR (SeptiFast test performed in a previous study (Warhurst et al., 2015) but culture identified *E. faecalis* only. This may be because it did not grow or, more likely, because it wasn't reported by the Biomedical Scientist who examined the plate (written off as skin contamination). This highlights an important point about the subjectivity and variability of the 'gold standard' diagnostic method. The metagenomics workflow failed to identify *K. pneumoniae* in sample R1 and *E. coli* (important causative of sepsis) was detected in in negative sample, R5. In sample R6, although *P. aeruginosa* reads were detected by Kraken, the most abundant species was *K. pneumonia*e. However further analysis using reference alignments revealed that the *K. pneumoniae* was likely to be a contaminant and the true pathogen was *P. aeruginosa*. This highlights the weaknesses in the current tools for the rapid

identification of pathogens from metagenomics data (typically kmer based). Better tools are required that can quickly and accurately assign species, and ultimately strains, from metagenomics data.

Retrospective samples were frozen at -80ºC for more than 18 months before being shipped in dry ice then defrosted for processing and metagenomics analysis. Possible reason for failing to detect *K. pneumoniae* in R1 by metagenomics pipeline detection and low signal *P. aeruginosa* in R6 was thought to be extended freezing and thawing of bacteria cells. Freeze-thawing at low temperatures has long been documented to lyse bacterial cells, especially Gram negatives (Kohn, 1960) (Ray & Speck, 1973). This has been supported by the fact that signal for the bacteria with tough cell wall, *E. faecalis, S. haemolyticus* and *S. aureus,* was high. DNA from dead (lysed) pathogens could not be detected using this workflow because cell free DNA is degraded for human DNA depletion purposes before recovery of intact pathogen cells for metagenomics analysis. Nonetheless, freezing blood samples is not anticipated in real-clinical settings where results for BSI/Sepsis diagnostics are required as soon as possible.

Using real-time MiniION sequencing and automated WIMP analysis detected *E. faecalis* and *S. haemolyticus* reads within 15 minutes of sequencing the clinical sample (Figure **5-1**). Recently similar results have been reported where MinION and WIMP detected *Yersinia pestis* in a spiked blood sample (Hewitt, Guertin, Ternus, Schulte, & Kadavy, 2017). However, this study by Hewitt et al was limited by low level of human DNA depletion (using NEB Microbiome kit which was abandoned earlier in thesis) and the absence of WGA which had consequences for limit of detection and turnaround time. The depleted blood sample spiked with approximately 2.58×10^6 cells/mL of *Y. pestis* produced only 57 pathogen reads out of 8,649 total reads in16 hours of MinION sequencing (Hewitt, Guertin, Ternus, Schulte, & Kadavy, 2017). The MinION run of sample P4 was poor (817 reads in total produced of which 2 were *S. pyogenes*). This was partly due to low levels of human depletion in the sample and early MinION sequencing technology producing low yield. Also, during the course of this PhD project, it was observed that whole genome amplified DNA degraded rapidly in storage at -20 ºC despite washing with

Ampure beads and that the branched nature of WGA DNA seemed to block nanopores and result in low flow cell yields.

The high level of DNA depletion achieved using our novel methods enabled the detection of pathogen with genome coverage ranging from 0.13× to 51×. To the best of my knowledge 51 × is the highest recorded bacterial genome coverage from metagenomics sequencing of a clinical blood sample from a septic patient. Elsewhere, 0.3 × coverage of pathogen genome has been reported but from pathogen cell free circulating DNA in plasma (Grumaz et al., 2016). Higher genome coverage enable detection of resistance and virulence markers and typing if required.

While metagenomics data can be used to predict resistance (more accurately in some organisms than others), predicting susceptibility is more challenging and it is debatable whether it will ever be possible with high accuracy in all pathogenic species. Using Illumina data, in sample R4 (containing *E. faecalis* and *S. haemolyticus*), antibiotic resistance genes, *ermB*, *ermC*, *sat-4* and *fusB* were detected. The *ermB* macrolide resistance gene correlated with isolated culture which was resistant to erythromycin. Although metagenomics (and other molecular tests like PCR) can accurately and reliably predict resistance in *S. aureus* and *M. tuberculosis* (Bradley et al., 2015), it is more challenging for other pathogens, for example, *P. aeruginosa* because of complex efflux and permeability characteristics and *Enterococci,* which can be intrinsically resistant to vancomycin (Cetinkaya, Falk, & Mayhall, 2000; Eliopoulos & Gold, 2001). Also, resistance genes may be detected by metagenomics sequencing but fail to cause resistance because of poor expression, silencing or inactivation.

Some of the challenges we faced in the metagenomics diagnosis of sepsis were unreliable WGA methods and background contamination. Protocols for library preparation using the MiSeq platform required 1ng input DNA and at least 1µg was required for the MinION. Following human DNA depletion by CD45 IMS and MolYsis, the amount of DNA remaining was estimated to be in the femtogram to low picogram range, which made the WGA step unavoidable. WGA was performed by multiple displacement amplification (MDA) technology using Phi 29 DNA polymerase and

random primers to amplify the entire genome. The proofreading activity of Phi 29 polymerase ensured high-fidelity amplification with an error rate of only 3×10^{-6} (in mutations/nucleotide) in the amplified DNA (Nelson et al., 2002) which was low compared to Taq DNA polymerase (2.4×10^{-5}) used for PCR (Tindall and Kunkel, 1988). Factors that determine good quality WGA is absence of contaminants, the ability to yield >1ng DNA from starting material of <50fg and uniform genome coverage in breadth and depth (Picher et al., 2016). In this thesis, WGA was initially performed using Qiagen REPLI-g Single Cell kit (Qiagen). Initial results using this kit were very good however, when a new batch was purchased, we noticed inconsistent performance in amplification. It was therefore decided to change to the Illustra Single Cell kit (GE Healthcare Life Sciences) which also uses Phi29 enzyme for WGA, but similar problems were observed. Several troubleshooting experiments were performed such as control reactions using high quality commercially sourced human DNA (Promega) and *E.coli* DNA (NEB), testing pH of buffers, and aliquoting of Phi29 enzyme to avoiding freezing-thawing, all without much success. Recently, a new modified method for WGA by MDA has been published (Picher et al., 2016). The method uses two enzymes, first *Thermus thermophilus* (Tth) PrimPol to synthesize short DNA primers and then Phi29 DNA polymerase to amplify the whole genome from single cells. The method is claimed to provide superior breadth and evenness of genome coverage with high reproducibility (Picher et al., 2016). Alternative kits and methods should be tested/developed in future to identify one suitable for metagenomics based diagnosis of infection.

Another challenge was identifying pathogens from metagenomics data in high background contamination. Potential sources of contaminants includes laboratory kits, personnel, plastics, and buffers (Salter et al., 2014). The most commonly occurring contaminant was *Alteromonas macleodii*, a marine Gram negative bacteria which widely occurs in temperate tropical waters and waters around Europe (López-Pérez et al., 2012; Vandecandelaere et al., 2008). When molecular grade water (provided with Qiagen WGA kit) was amplified without adding template DNA, *Staphylococcus lugdunensis, P. putida, Stenotrophomonas maltophilia* and *Delftia ssp* were isolated. This issue was highlighted in a recent study on impact of contaminating DNA in WGA kits used for sequencing for infection diagnosis (Thoendel et al., 2017). In this study, some samples with no template DNA were

amplified and sequenced, some of the identified contaminants were *Sphingomonas*, *Staphylococcus*, *Propionibacterium*, *Peptoniphilus*, *Finegoldia*, *Elizabethkingia* *Achromobacter*, *Rothia*, *Micrococcus* and *Delftia* among others. The most common contaminants co-amplified with template DNA were *Toxoplasma* when Illustra Single Cell kit was used and *Propionibacterium* using Qiagen REPLI-g Single Cell kit (M. Thoendel et al., 2017). Most of these bacteria can easily be categorized as contaminants, however detection of important organisms such as *E. coli* (in R5) and *K. pneumoniae* (R6) can signal false positive results. Although further bioinformatics analysis can help to identify true positives from metagenomics data, it is crucial to source reagents that are as clean as possible and to develop working practices that limit the potential for laboratory contamination.

Despite the challenges, we have demonstrated proof-of-concept that our rapid nanopore sequencing based metagenomics diagnostics pipeline enables agnostic pathogen and antibiotic resistance gene identification directly from whole blood. The main advantage of using this pipeline is the rapid turnaround time from sample collection to pathogen identification and the comprehensiveness of the results produced. It has been reported that the level of host to pathogen DNA has made it difficult to detect the low levels of bacteria present in clinical blood samples (Hewitt et al., 2017; Thoendel et al., 2016). The developed pipeline overcomes this issue by efficiently depleting human DNA prior to pathogen DNA extraction, thus enabling the identification of bacterial pathogens, at concentrations as low as 10 CFU/ml, within eight hours.

5.5.2: Urine

Rapid pathogen profiling from clinical specimens, without culture, could facilitate better treatment and antibiotic stewardship. Urine was taken as an exemplar, with a heavy load of infecting bacteria, thereby: (i) yielding sufficient DNA without WGA for MinION sequencing; (ii) minimizing the confounding effects of commensal bacteria and laboratory/reagent contamination on results; and (iii) ensuring a high bacterial cell/human cell ratio. However, despite high bacterial load, we have demonstrated that depletion of host DNA has a positive impact on urine metagenomics sequencing for pathogen identification independent of culture. Depletion of human DNA enabled increased yields of bacterial sequences (more than 76% of total sequences) and

increased genome coverage thereby reducing library running time. Depletion of human DNA enabled bacterial yields of up to 22.84× depth and 99.9% breadth of *E. coli* genome coverage within 8 hours. Similar benefits were seen by (M. Thoendel et al., 2017) whereby they used MolYis Basic 5 kit to deplete approximately 10^4 folds human DNA in prosthetic joint infection samples to enable metagenomics based diagnostics and guide therapy.

Using urine specimens, we have comprehensively demonstrated that MinION can rapidly identify pathogens and antimicrobial resistance genes with high accuracy given sufficient genome coverage. The advantage of using MinION over other platform is real time sequencing and data analysis. Unlike other platforms where data analysis is performed post-run, MinION data can be analysed at real-time using WIMP and ARMA (for antimicrobial detection) for real-time analysis of metagenomics data. With this approach, adding together times for human DNA depletion and DNA extraction (2.5 h) library preparation (3 h), sequencing (1 h) and data analysis (1 h), suggests a total 7 – 8 h turnaround, equating to one dosage interval for a 'typical', every 8 h, antibiotic. This turnaround would enable clinicians to refine antibiotic therapy before second dose of empiric therapy. Further acceleration was demonstrated using the rapid (15 min) library preparation chemistry from ONT (Votintseva et al., 2017) reducing turnaround to 4 h. This is a substantial improvement to a study by Hasman *et al* in which urine sequencing was performed using Ion Torrent PGM™. Despite using an automated bioinformatics analysis pipeline for pathogen identification, turnaround time was estimated between 18 hours and 24 hours (Hasman et al., 2014).

Although we have shown great potential of rapid metagenomics diagnostics, there are still challenges to the approach. First, heavily infected urines (>10^7 CFU/mL) were used to obtain 1µg DNA required for MinION library preparation sequencing. Bacteria load in these samples was high considering significant bacteriuria is defined as ≥10^5 CFU/mL (Krcmery, Hromec, & Demesova, 2001) (Schröder et al., 2015). Low-input library preparation kits are now available from ONT for the MinION which require as little as 1ng input DNA, equivalent to about 10^5 bacteria (Goodwin, Wappel, & McCombie, 2017). Another limitation was, allelic variants were poorly distinguished MinION data due high error rate of nanopore data and low pathogen

genome coverage. The error rate has reduced from 30% to about 10% in recent chemistry and the yields per flowcell have risen approx. 10 fold. SNPs and sequence variants can be accurately called when MinION data is polished using bioinformatic analysis such as nanopolish to improve consensus accuracy, facilitating precise identification.

Advantages of MinION's long reads is the ability to inform on the context of resistance genes, potentially enabling differentiation between plasmid-borne and chromosomal genes. Also, assembly of complete plasmids can be done with just a few MinION reads, in some cases, single reads can cover the full length of a plasmid. MinION metagenomics sequencing has the potential to rapidly diagnose patients with urosepsis, thus allowing refinement of antibiotic regimens within the first dosage interval after clinical diagnosis.

5.5.3: Stool

To study transmission and epidemiology of CDI, a typing method with high discriminatory power and reproducibility is needed (Huber et al., 2013). PCR ribotyping is the main method for *C. difficile* typing in the UK and is widely used across the world. Key advantage of the method is a straightforward protocol (Bidet et al., 2000), however, it is limited by poor resolution (Kuijper, Coignard, & Tüll, 2006), long turnaround time (Tenover et al., 2010) and lack of standardized protocol (Knetsch et al., 2013). In the UK, MLVA typing is used in addition to ribotyping, if further resolution is needed (PHE, 2010). Whole-genome sequencing may soon become the method of choice for the investigation of transmission of *C. difficile* strains (Dingle et al., 2013; He et al., 2013). However, dependence on culture and long sequencing turnaround time are still potential limitations.

In this thesis, we have developed methods for enriching for *C. difficile* spores and DNA extraction that upon further optimisation will enable direct SNP typing of CDI from metagenomics data (without PCR or culture methods). *C. difficile* spores were enriched by depleting human and vegetative cells by differential lysis and digestion of released DNA. Depletion by differential lysis, percentage of *C. difficile* reads increased from 0.02% or less using widely applied Nycodenz medium to 2.9 - 7.4% (150 to 350-fold increase). Although there was huge increase of *C. difficile* reads, the

methods did not provide enough breath/depth of genome coverage to enable toxin gene identification or SNP typing. Sufficient sequencing data could have been obtained by increasing the depth of sequencing on the MiSeq run (48 instead of 96 samples per run) or by using more powerful sequencing platform such as NextSeq (Illumina) or PromethION (ONT). Further optimization of incubation times and strength of buffers for depletion is required (could not be completed in the timeframe of this PhD). Depth of coverage of 7 × provides 99.9% breath of coverage, sufficient to reliably identify resistance gene and virulence markers (Schmidt et al., 2016).

Here we have described rapid sample processing methods for enriching *C. difficile* DNA from stool without the need for culture. Upon further optimisation, these methods have potential to shorten *C. difficile* typing from the current 14 days to within 24 hours (using MinION), or three days by Illumina.

6: General Discussion and Conclusions, and Future Directions

6.1: General Discussion and Conclusions

In the UK, reports including the UK 5-year (2013-2018) Antimicrobial Resistance Strategy from PHE and Department of Health (Davies & Gibbens, 2013), and the UK Prime Minister commissioned O'Neill report on Tracking Drug Resistance Infections Globally (O'Neill, 2014; Resistance, 2016), have emphasized the need for rapid diagnostics in transforming the way we use antimicrobials by reducing unnecessary use and slowing antimicrobial resistance, thereby making existing drugs efficacious for longer. The report states that, by 2020, a rapid diagnostic should be used where possible before antibiotics are administered. Not only that, a report by the American Academy of Microbiology Colloquium (April 2015) has gone further, highlighting the potential of NGS based diagnostics to dramatically revolutionize the clinical microbiology laboratory by replacing current time-consuming and labour-intensive techniques with a single all-inclusive diagnostic test (Weinstock et al., 2016). The implementation of rapid diagnostics will reduce hospitalisation, reduce unnecessarily prolonged use of broad-spectrum antibiotics, reduce costs and improve care of individual patients.

This aim of my PhD was to evaluate metagenomics sequencing as a method for replacing current traditional culture based methods, Gram-staining, biochemical, and molecular assays (including MALDI-TOF) to identify clinical pathogens. The metagenomics approach proved capable of delivering comprehensive pathogen identification and resistance profiling within a timeframe which would allow clinicians to tailor antibiotic prescription for the infecting pathogen before second dose of empiric treatment (i.e. 8 h). This has been achieved by addressing one of the major barriers preventing the deployment of NGS to the clinical microbiology, namely, pathogen DNA enrichment and/or depletion of human DNA.

In blood samples, human DNA has been depleted by two methods. The first method was based on a novel combination of the MolYsis kit (a commercially available host DNA depletion method) with immunomagnetic separation of leukocytes using anti-

human CD45-antibody coupled beads. This assay has proven that depletion of human DNA is a key enabling technology for the implementation of metagenomics based diagnostics. The second method is completely novel and has been patented. It is based on selective lysis of human cells using PLC and digestion of the human DNA using a specialised nuclease. This method is capable of removing 10^6 (1,000,000) fold host DNA from blood without loss of pathogen DNA in approximately 45 mins. The simplicity, specificity and efficacy of this method are unprecedented. Several companies are interested in licencing the technology to be used in conjunction with metagenomic sequencing for the diagnosis of pathogens in clinical samples. One company has already funded a 4 month proof-of-concept study for blood screening applications. Our depletion methods are superior to the commercial kits, and published assays of depleting human DNA in blood samples. The methods have highlighted the lack of efficient host depletion metagenomics diagnostic. I hope these findings will stimulate diagnostic developers to innovate in the area of host depletion for effective and rapid metagenomics based pathogen identification.

One of challenges encountered in the diagnosis of sepsis was low numbers of pathogen cells found in many septic patient samples. Low pathogen numbers mean: (i) WGA was required to obtain enough input DNA for sequencing; (ii) the WGA step, capable of amplifying femtogram quantities of DNA, made the pipeline prone to contamination from the sample, the laboratory and the kits; and (iii) data analysis and interpretation was complicated by background contamination. Environmental contamination has a serious impact on the detection of low level pathogens by metagenomics, and so extra care needs to be taken in handling the sample, from collection and processing to data analysis. If metagenomics is to be used routinely, diagnostic laboratories will need cleaner reagents, facilities and consumables (including plastics) (Weiss et al., 2014), WGA kits (Hosokawa, Nishikawa, Kogawa, & Takeyama, 2017; Thoendel et al., 2017), and buffers and enzymes that are derived from micro-organisms (Salter et al., 2014; Weiss et al., 2014). Nucleic acid free water rather that nuclease free water is required – in fact, as many consumables and reagents used in the pipeline as possible need to be nucleic acid free. It is well known that manufacturers reagents are contaminated with bacterial nucleic acids, but it is not something companies want to highlight. We propose a different approach – manufacturers selling reagents and plastics for sensitive applications should

provide a 16S profile of the DNA present in each batch of their reagents. This would promote confidence in the supplier and ensure the highest quality reagents. Presence of contaminants in WGA kits and other reagents can monitored using negative amplification. Negative samples can be used to create contamination databases that allow mapping based methods used to rapidly remove known contaminants, thereby simplifying data analysis. Further research is needed to address these challenges by either streamlining our method or developing new methods, to increase sensitivity and reduce background contamination and turnaround time.

The two main shortcomings of the current market leading sequencing platforms for application in infection diagnosis are long turnaround time and short reads. These limitations have been addressed by the MinION sequencing platform. Other benefits of the MinION include rapid library preparation methods, real-time sequencing and data analysis, and low capital cost and small size of the platform. However, the technology remained under active development whilst this PhD was undertaken and was faced with several challenges: a requirement for high input DNA concentration (at least 1 µg); no multiplexing options were available; expensive flow cells ($500-900 USD depending on number purchased); inconsistent and poor performing flow cells (caused by manufacturing issues). But despite the limitations, the qualities of the technology together with our sample preparation methods allowed rapid turnaround and comprehensive clinical diagnosis of UTIs and sepsis, independent of culture. Recent ONT technology advances have addressed most of the shortcomings of nanopore sequencing: Flongle, a $100 single use flow cell, also, low-input, multiplex library preparation kits are now available that require 1ng input DNA and up to 96 samples can be multiplexed (6 samples per flow cell runs have been performed in our laboratory these days); flow cell manufacturing has been optimised resulting in a consistently quality and active unblocking has helped to significantly increase flow cell performance.

6.2: Future Directions

The rapid uptake and implementation of MALDI–TOF in clinical microbiology highlights the potential impact of new technologies in the routine diagnosis of

infectious diseases. Many scientists believe metagenomics will revolutionise clinical microbiology by replacing current culture based methods (Burnham, Leeds, Nordmann, O'Grady, & Patel, 2017; Didelot et al., 2012; Goldberg, Sichtig, Geyer, Ledeboer, & Weinstock, 2015; J Wain & Mavrogiorgou, 2013). However, uptake, implementation and acceptance of metagenomics will depend on appropriate sequencing technology, efficient sample processing workflows, appropriately trained staff, rapid and accurate bioinformatics and infrastructure for 'big data' storage and sharing.

Important qualities to consider when choosing a sequencing technology for clinical use include reliability, reproducibility, read length, consensus read accuracy/error profiles, number of samples per run (including controls) that can be loaded, turnaround time, sequencing depth, throughput, and data analysis complexity. Currently, there is no technology that perfectly combines all the qualities we desire, however, ONT technology comes close and their development pipeline shows great potential for the future. In terms of scale-up, there is the GridION X5 and the PromethION. The GridION is a compact benchtop system capable of running five MinION flow cells at a time and has the inbuilt capacity to basecall the reads in real-time (Loose, 2017); PromethION is a bench top instrument for a higher number of flow cells (48) with more pores (3000 channels and 6000 nanopores per flow cell) (Loose, 2017). These platforms have the potential to rapidly produce much more data, hence more samples can be multiplexed. However, it is debatable whether highly multiplexed metagenomics runs in central labs is the best way forward for clinical microbiology. Multiplexing means you have to wait for multiple samples before you perform the run, hence increasing turnaround time. This may be less important for some diseases/sample types, e.g. UTI/urine (constant supply of samples), but much more important for others, e.g. sepsis/blood (where turnaround time is crucial and there are much fewer samples). ONT are aware of this and are releasing a scale-down version of the MinION, the Flongle. It will be a single use/single sample flow cell that will clip into an adapter and fit in the MinION. It will have fewer pores and expected to produce approx. 1Gb data for approx. $100. Furthermore, new applications, such as direct RNA sequencing kit (Garalde et al., 2018; Smith, Jain, Mulroney, Garalde, & Akeson, 2017) are now available. Real-time RNA sequencing will allow direct detection of RNA viruses such as Influenza virus,

measles virus and hepatitis C virus by metagenomics sequencing. Also, new technology using instrument called VolTRAx for rapid and automated library preparation (10 minutes for 1D library) has been introduced. These are important developments towards suitable technologies for metagenomic diagnostics.

Illumina and Pacific Biosciences sequencing platforms have been extensively used to determine the genomic sequences of microorganisms and identify emerging pathogens in infectious disease outbreaks but have not been approved for infection diagnosis (Goldberg et al., 2015). We expect both companies and ONT to develop infection diagnosis tests in the future, most likely through collaboration, that will be FDA/CE-IVD approved. The iSeq™ 100 Sequencing System (Illumina, San Diego, California, USA) is designed for fast and low-throughput sequencing and would appear to be Illumina's first technology for the infection metagenomics market.

Similar to current clinical microbiology practices, implementing metagenomics sequencing may require laboratories to define the range of specimen type that can be processed by a specific pipeline. Specimen type and range of pathogen in the sample may determine the preferred method for pathogen enrichment or human nucleic acid depletion, nucleic acid extraction, and sequencing (RNA and/or DNA). These factors may also determine sequencing time and/or depth and data analysis methods/tools. Also, implementation will require simplifying and automating the current metagenomics pipelines to reduce turnaround time and risk of contamination. Metagenomics data, as with culture, requires interpretation by biomedical scientists, clinical microbiologists and other clinicians in order to e.g distinguish between pathogens and organisms that constitute normal flora or are likely contaminants. Removal of human DNA will simplify data analysis by reducing the volume of sequence data that needs to be analysed and stored, and will eliminate ethics concerns related to inadvertently sequencing the patient genome.

Before metagenomics tests (laboratory developed tests or commercially developed products) can be implemented clinically, the entire workflow, from specimen collection to report generation, will require validation and accreditation by the appropriate authorities e.g. FDA, CE-IVD etc. This validation is complex for an agnostic technology and may be difficult in the absence of a good 'gold standard'. The FDA have recently released guidance on the approval process for NGS based

diagnostic devices for infection detection (FDA, 2016). The guideline intents to regulate/evaluate all the necessary components of a metagenomics workflow including specimen collection, instruments for sample processing and library preparation, software (if applicable), the sequencing instruments, all the workflow associated reagents, raw data collection and the data analysis pipeline (FDA, 2016). Other areas such as contamination (and negative controls) and limit of detection will also be thoroughly validated to minimise false positive/negative results. The FDA has stressed on the importance of bioinformatics as the key area in pathogen identification by metagenomics (FDA, 2016).

While metagenomics holds great promise, one current drawback of this technology is its inability to accurately predict antibiotic susceptibility (and, in certain cases, resistance). Emergence of new mechanisms of resistance and resistance due to altered expression of intrinsic genes further complicates the issue. Some studies have shown optimism that antibiotic resistance data can be extracted reliably from WGS data of pathogens such as *E. coli* (Tyson et al., 2015) and *Campylobacter ssp* (Zhao et al., 2016) for clinical decisions. However, a recent report by EUCAST suggests that for most bacterial species there is currently insufficient evidence to support the use of WGS-inferred AST to guide clinical decision making (Ellington et al., 2017), owing to poor knowledge of pathogen genomes, cut-off points, lack of standardised mutation database and bioinformatics tools, and lack of useful clinical evidence (Ellington et al., 2017). As research in this area gains momentum, I believe increased knowledge on the genomes of all common pathogens (e.g. genomic/transcriptomic basis of efflux and permeability) and better bioinformatics tools (e.g. metabolic modelling based on genome sequence to predict what proteins are made by the cell) will enable accurate metagenomic resistance and susceptibility testing in the future. The FDA is calling for more studies to establish the analytical and clinical performance of metagenomics based diagnostics by NGS for microbial identification, detection of antimicrobial resistance, and virulence markers (FDA, 2016) (Schlaberg et al., 2017). Such studies will help diagnostics developers to address current limitations of metagenomic diagnostics.

The implementation of sequencing based diagnostic technology to clinical microbiology may require staggered introduction to gain acceptance. Probably the

best approach is to initially use NGS for routine outbreak and surveillance investigation on isolates. This is a viable route because the cost of bacterial WGS continues to decline (now equivalent to other typing methods such as MLST) and the methods are already established in public health laboratories e.g. PHE Colindale. Once the capacity for NGS is built in clinical microbiology labs, the potential of the technology will be realised and the diagnostics industry/research community will have the pipelines in place for implementation of metagenomics, first for specific use cases and eventually broadly. I believe my research is part of the beginning of a paradigm shift in clinical microbiology and that these and similar methods will be improved upon and implemented in the clinic within the next 10 years.

7: Appendices

7.1: Patent

Mwaigwisya S, Kay G, Wain J, O'Grady J. 2017. Method for nucleic acid depletion. PCT/GB2017/053715.

Method for nucleic acid depletion

Field of the Invention

The invention relates to methods of depleting host nucleic acid from a biological sample.

Background to the Invention

Rapid and comprehensive infectious disease diagnostics are crucial for improved patient management and in the fight against antimicrobial resistance. Rapid diagnosis of life-threatening infectious diseases such as sepsis and pneumonia is paramount. These clinical syndromes have complex aetiologies and require pathogen recognition in challenging sample matrixes e.g. blood, sputum etc. Currently, the "gold standard" method for clinical diagnostics is microbial culture, which is labour intensive, has long turnaround times and poor clinical sensitivity. Currently available rapid molecular methods (e.g. PCR) improve turnaround time to result and sensitivity, but are limited by range and therefore rare pathogens and resistance markers can be problematic. The most applicable technology for rapid detection of microbial pathogens is nucleic acid amplification tests (NAATs). NAATs are available for sepsis diagnostics (e.g. Septifast (RTM), Roche) but complexity of use and suboptimal performance have prevented their widespread adoption. Most of the NAATs for respiratory tract infections (RTIs) focus on the detection of respiratory viruses (e.g. Biofire Filmarray Respiratory Panel, Seegene RV15). An exception is the Curetis Unyvero (RTM) test which is designed for health care associated pneumonia. NAATs, however, are not comprehensive (e.g. the Curetis test only covers 90% of the top pathogens), seeking only a pre-set range of targets, meaning that less common pathogens will be missed. Consequently, NAAT diagnostics are an adjunct to standard bacteriology, not a replacement, and adoption is limited.

A paradigm shift in diagnostics technology is urgently required - a universal diagnostic method which can detect any pathogen (e.g. viral, bacterial, fungal) and antibiotic resistance. Agnostic/shotgun metagenomic sequencing has the potential to be the technology of choice to drive this shift. Shotgun metagenomic sequencing can detect and provide relative proportions of viruses, bacteria and fungi in a sample without any prior knowledge of the microbial community present, and is increasingly being used to investigate complex metagenomes in clinical samples.

So why is shotgun metagenomics not currently being widely applied to infection diagnosis? One reason is that next generation sequencing (NGS) has traditionally been expensive, complex to perform and difficult to analyse. The development of MinION (RTM) nanopore sequencing technology has changed the NGS landscape with cheap portable sequencers, rapid simple library preparation (15 mins) and automated real-time analysis tools. Another major barrier is the large amount of human DNA present in clinical samples, which is often several orders of magnitude greater than the pathogen DNA present. Blood is a particularly challenging matrix for NGS-based pathogen characterization due to the vast amount of human vs. pathogen nucleic acid (particularly DNA) present (ratio is typically 10^8:1 to 10^9:1, based upon 10^6 leukocytes/ml [with ~6.6pg DNA/cell] but as few as 1-10 colony forming units [CFU] of pathogen/ml [with ~10fg DNA/cell]). A host DNA depletion of at least about 10^5, potentially resulting in a human:pathogen DNA ratio of 10^3:1, is required to facilitate NGS-based pathogen characterization, a level of depletion (giving rise to pathogen nucleic acid enrichment) not achieved by methods disclosed in the art, such as commercially available pathogen DNA enrichment methods (Looxster (RTM) Enrichment kit (Analytic Jena); NEBNext (RTM) Microbiome DNA Enrichment kit (NEB); MolYsis (RTM) Basic 5 kit (Molzym)).

It is among the objects of this disclosure to address the aforementioned problems.

Summary of the Invention

Accordingly, provided is a method for depleting host nucleic acid in a biological sample, said sample having been previously obtained from an animal host, said method comprising the steps of:

(a) adding a cytolysin, or an active variant thereof, to said sample; and

(b) carrying-out a process to physically deplete nucleic acid released from host cells within said sample or otherwise render such nucleic acid unidentifiable.

Preferably, step (b) comprises adding a nuclease to said sample and/or the method further comprises the step of extracting remaining nucleic acid from the sample.

Preferably, the method further comprises the step of subjecting the extracted nucleic acid to a purification process and/or further comprises the step of amplifying the extracted nucleic acid.

Preferably, the method further comprises the step of conducting a nucleic acid amplification test on the extracted nucleic acid or, preferably, conducting a sequencing process on the extracted nucleic acid.

In preferred embodiments, the cytolysin is a phospholipase, preferably a phospholipase C (PLC), more preferably is a bacterial PLC, more preferably a Group 1 PLC, most preferably PLC from *Clostridium perfringens*.

In preferred embodiments the biological sample is a blood sample.

In preferred embodiments the method results in at least a 5×10^4 fold depletion, preferably at least a 10^5 fold depletion, of host DNA originally contained within the sample.

Also provided is a kit comprising i) a cytolysin, or an active variant thereof, and ii) means to physically deplete free nucleic acid within a biological sample or otherwise render such nucleic acid unidentifiable. Preferably, said cytolysin is as defined as above and/or wherein said means comprises a nuclease.

Brief description of the Figures

Figure 1 shows amplification curves of human qPCR results after various endonuclease treatments.

Figure 2 shows amplification curves of human qPCR results after endonuclease treatment with various buffer volumes.

Figure 3 shows amplification curves of human qPCR results after HL-SAN DNase and MolDNase treatment with respective buffers.

Figure 4 shows amplification curves of human qPCR results after cytolysin treatment.

Figure 5 shows amplification curves of human qPCR results showing PLC activity in different sample conditions.

Figure 6 shows amplification curves of qPCR results after PLC and HL-SAN DNase treatment on increased volumes of bacterial spiked blood; A: Human qPCR; B: *E. coli* qPCR; C: *S. aureus* qPCR.

Figure 7 shows amplification curves of human qPCR results of PLC activity after the addition of efficient mixing during host cell lysis.

Figure 8 shows amplification curves of qPCR results after altered HL-SAN DNase inactivation; A: Human qPCR; B: *E. coli* qPCR; C: *S. aureus* qPCR.

Figure 9 shows Amplification curves of qPCR results for method comparison; A: Human qPCR; B: *E. coli* qPCR; C: *S. aureus* qPCR.

Figure 10 shows *C. albicans* genome coverage plot after *C. albicans* single-plex MinION sequencing.

Detailed Description of the Invention

General

Provided herein is a method for depleting host nucleic acid (particularly RNA and/or, most preferably, DNA) in a biological sample, said sample having been previously obtained from an animal host, said method comprising the steps of:

adding a cytolysin, or an active variant thereof, to said sample; and

carrying-out a process to physically deplete nucleic acid released

from host cells within said sample or otherwise render such nucleic acid unidentifiable.

The animal host can be a vertebrate, e.g. a bird, a fish or, preferably, a mammal, most preferably a human. The host may, at the time of sample collection, be alive or dead.

The biological sample can be any sample that comprises animal cells (in tissue form or otherwise). Particular (e.g. clinical) samples of interest include bile, nail, nasal/bronchial lavage, bone marrow, stem cells derived from the body, bones, non-fetal products of conception, brain, breast milk, organs, pericardial fluid, buffy coat layer, platelets, cerebrospinal fluid, pleural fluid, cystic fluid, primary cell cultures, pus, saliva, skin, fetal tissue, fluid from cystic lesions, stomach contents, hair, teeth, tumour tissue, umbilical cord blood, mucus and stem cells. Particularly preferred samples include, though, joint aspirates, faeces, urine, sputum and, especially, blood (including plasma). Preferably, the sample is in liquid form. An initial sample might need to be converted to liquid form before conducting the present methodology. A liquid sample might have a volume of between 10µl and 100ml, preferably between 10µl and 50ml, such as between 10µl or 100µl and 20ml (e.g. 0.2ml or 1ml).

The cytolysin causes (selective) lysis of the host cells, releasing host nucleic acid such that it can be (partially or completely) depleted. Nucleic acid within a non host cell or particle (e.g. pathogen) is essentially left intact (i.e. has not been significantly removed from the sample or digested) and identifiable, such that it can be subsequently collected and analysed and, in particular, identified (by e.g. sequencing or targeted PCR). A nucleic acid is identifiable e.g. if its sequence and/or biological origin can be ascertained. Preferably, therefore, the cytolysin is added to the sample and allowed to act for a period of time such that sufficient host cell lysis can occur. Steps (a) and (b) (“cytolysin incubation” and “depletion step”) can occur simultaneously, or step (b) follows step (a).

The method of depleting host nucleic acid comprises both physical depletion and (in the context of the present technology) virtual depletion (of nucleic acid released from

host cells within the sample). Physical depletion can involve e.g. digesting the nucleic acid (i.e. breaking down nucleic acid polymers to e.g. base monomers) or removing nucleic acid from the sample (e.g. by any nucleic acid capture method known to the skilled person, such as deploying nucleic acid-binding magnetic beads in the sample to bind DNA and/or RNA, which can subsequently be removed or harvested from the sample).

Virtual depletion involves rendering (released) nucleic acid unidentifiable (via, in particular, targeted PCR or, most preferably, sequencing). For DNA, this means rendering the DNA non-amplifiable (e.g. by PCR) and/or (preferably) non-sequenceable. For RNA, this means rendering the RNA non-amplifiable, non-reverse-transcribable and/or (preferably) non-sequenceable. A preferred process for such rendering (particularly for DNA) involves adding a photoreactive nucleic acid-binding dye, such as propidium monoazide (PMA) or ethidium monoazide (EMA), to the sample and inducing photoreaction.

Most preferably, however, the method of depletion is via digestion of nucleic acid, most preferably via enzymatic digestion. It is therefore preferred that step (b) comprises adding a nuclease to the sample. Preferably, the nuclease is added to the sample and allowed to act for a period of time such that sufficient nucleic acid digestion can occur. Preferably, therefore, a deoxyribonuclease (DNase) and/or a

ribonuclease (RNase) is added to the sample (and preferably allowed to act for a period of time such that sufficient DNA/RNA digestion can occur). The nuclease can have both DNase and RNase activity (e.g. HL-SAN DNase). Depletion of host DNA is important if analysis of non host (e.g. pathogen) DNA is to be carried out. Depletion of host RNA is important if analysis of non host (e.g. pathogen) RNA is to be carried out, and indeed can facilitate the optimisation of DNA analysis (e.g. DNA sequencing).

In such embodiments, the method preferably further comprises the subsequent step of neutralising the (or each) nuclease (i.e. decreasing or substantially eliminating the activity of the nuclease). The skilled person will recognise a range of neutralisation options, to be selected for each depletion protocol. This might include heat inactivation or, preferably, buffer exchange (i.e. the removal of a buffer in which the

nuclease is active and/or replacement with or addition of a buffer in which the nuclease is substantially inactive). Preferably, the temperature of the sample (at any/all stage(s) at/before extraction of remaining nucleic acid from the sample) is maintained at 50°C or less, preferably 45°C or less, preferably 40°C or less, to optimise subsequent release of nucleic acid from the pathogen (particularly from bacterial cells).

Further steps

In preferred embodiments, the method further comprises the step of extracting remaining (preferably non host) nucleic acid from the sample (or aliquot thereof). Part or all of the remaining nucleic acid (particularly non host nucleic acid) will be intact and identifiable.

Typically, the extraction process will involve a centrifugation step to collect, in particular, non host cells/particles (e.g. pathogens) (virus particles and/or, in particular, bacterial and/or non-animal (e.g. non-mammalian) (e.g. unicellular) eukaryotic cells, such as fungi), from which the nucleic acid can be obtained. Centrifugation conditions can be selected such that bacterial and non-animal cells, but not virus particles, are pelleted, or such that virus particles are pelleted in addition to bacterial and non-animal cells. If the former, standard virus detection tests could be performed on the supernatant. (Indeed, prior to any addition of cytolysin, one might centrifuge a clinical sample, keep the cell-containing pellet (for the method of the current technology), and keep the supernatant for virus detection using standard procedures, with or without enrichment using the present technology.)

Nucleic acid can be obtained from the pathogen(s) using methods known in the art, and might involve the addition of a lysis buffer, a lytic enzyme(s) (degrading or abrogating cell membranes, cell walls and/or viral capsids), and/or a protease, e.g. proteinase K. Preferred lytic enzymes include lysozyme, mutanolysin, lysostaphin, chitinase and lyticase.

Optionally, the extracted nucleic acid (or aliquot thereof) is subject to a purification process, such as one known in the art. During purification of DNA, RNase is

optionally used to facilitate the optimisation of subsequent DNA sequencing. However, RNase is omitted from any purification step if non host (e.g. pathogen) RNA extraction is of interest (for e.g. subsequent RNA sequencing) (and a DNase might be used to assist with purification).

In preferred embodiments, extracted nucleic acid (or aliquot thereof) is subject to an amplification process, such as whole genome amplification, to increase the copy number of the nucleic acid, particularly where the biological sample is a blood sample. For RNA, this might involve direct amplification or conversion of RNA to cDNA, followed by amplification of cDNA.

In preferred embodiments, the method further comprises the step of conducting a nucleic acid amplification test (e.g. targeted PCR amplification process, isothermal amplification, nucleic acid sequence-based amplification (NASBA)) on the extracted nucleic acid (RNA, DNA or cDNA) (or aliquot thereof) or, preferably, conducting a sequencing process on the extracted nucleic acid (or aliquot thereof), such as (e.g. short or long read) DNA or RNA sequencing, using e.g. nanopore or Illumina (RTM) sequencing.

In the preceding embodiments, nucleic acid (particularly host nucleic acid) previously rendered unidentifiable will not be amplified by any amplification process and/or (in particular) sequenced by any sequencing process.

The new method, in comparison with methods of the prior art (e.g. the MolYsis (RTM) technique, which deploys chaotropic agents to lyse host cells prior to host nucleic acid digestion), facilitates highly improved depletion of host nucleic acid (particularly DNA), while leaving non host (e.g. pathogen, particularly bacterial) nucleic acid intact (and identifiable), leading to highly improved non host (e.g. pathogen) nucleic acid enrichment, sufficient for subsequent sequencing-based (e.g. next-generation sequencing [NGS] based) (e.g. pathogen) diagnostics. A key factor in this advance has been the ability to achieve e.g. a 5×10^4 or greater, such as 10^5 or greater (e.g. 10^6 or greater), fold depletion of host DNA from within biological sample from a mammalian host, and these are preferable outcome features of the present technology (as is a fold depletion of 10 or greater, 10^2 or greater, 10^3 or greater, 5×10^3 or greater, or 10^4 or greater). It is particularly preferred that host

nucleic acid (e.g. DNA) is undetectable (e.g. via qPCR) following deployment of the method of the invention. In more general terms, the selective depletion of host nucleic acid enables enrichment of non host nucleic acid, and hence improved identification of non host organisms. This technology is thus applicable to fields other than medical microbiology, such as biological research, veterinary medicine/diagnostic, and agriculture/food safety

The cytolysin

A cytolysin (also known as a cytolytic toxin) is a protein secreted by a microorganism, plant, fungus or animal which is specifically toxic to a heterologous cell type(s), particularly promoting lysis of target cells. Preferred cytolysins are those secreted by microorganisms, particularly by bacteria, and/or those that are toxic to an animal (e.g. mammalian) cell type(s).

The cytolysin can be a cytolysin that has a detergent effect on the target cell membrane (e.g. a 26 amino acid delta toxin produced by *Staphylococcus*) or forms pores in the target cell membrane (e.g. Alpha hemolysin from *S. aureus*, Streptolysin O from *S. pyogenes*, and Perfringiolysin O produced by *C. perfringens*). See e.g.:

Alpha hemolysin from *S. aureus* –

https://www.ncbi.nlm.nih.gov/protein/BBA23710.1 (SEQ ID No. 2):

1 mktrivssvt ttlllgcilm npvanaadsd iniktgttdi gsnttvktgd lvtydkengm
61 hkkvfysfid dknhnkkilv irtkgtiagq yrvyseegan ksglawpsaf kvqlqlpdne
121 vaqisdyypr nsidtkeyms tltygfngnv tgddsgkigg liganvsigh tlkyvqpdfk
181 tilesptdkk vgwkvifnnm vnqnwgpydr dswnpvygnq lfmktrngsm kaadnfldpn
241 kassllssgf spdfatvitm drkaskqqtn idviyervrd dyqlywtstn wkgtntkdkw
301 tdrsseryki dwekeemtn

Streptolysin O from *S. pyogenes* –

https://www.ncbi.nlm.nih.gov/protein/BAD77794.2 (SEQ ID No. 3):
1 msnkktfkky srvaglltaa liignlvtan aesnkqntas tettttseqp kpesseltie
61 kagqkmddml nsndmiklap kemplesaek eekksedkkk seedhteein dkiyslnyne
121 levlaknget ienfvpkegv kkadkfivie rkkkninttp vdisiidsvt drtypaalql
181 ankgftenkp davvtkrnpq kihidlpgmg dkatvevndp tyanvstaid nlvnqwhdny
241 sggntlpart qytesmvysk sqieaalnvn skildgtlgi dfksiskgek kvmiaaykqi
301 fytvsanlpn npadvfdksv tfkdlqrkgv sneapplfvs nvaygrtvfv kletssksnd
361 veaafsaalk gtdvktngky sdilenssft avvlggdaae hnkvvtkdfd virnvikdna

421 tfsrknpayp isytsvflkn nkiagvnnrt eyvettstey tsgkinlshq gayvaqyeil
481 wdeinyddkg kevitkrrwd nnwysktspf stviplgans rnirimarec tglawewwrk
541 viderdvkls keinvnisgs tlspygsity k

Preferably, the cytolysin is a cytolysin that digests a cell membrane component, (e.g. phospholipids, i.e. is a phospholipase). An example is Sphingomylinease (also know as beta-toxin) from *S. aureus*, see e.g. https://www.ncbi.nlm.nih.gov/protein/CAA43885.1 (SEQ ID No. 4):

1 mmvkktksns lkkvatlala nlllvgaltd nsakaeskkd dtdlklvshn vymlstvlyp
61 nwgqykradl igqssyiknn dvvifneafd ngasdkllsn vkkeypyqtp vlgrsqsgwd
121 ktegsysstv aedggvaivs kypikekiqh vfksgcgfdn dsnkgfvytk iekngknvhv
181 igthtqseds rcgaghdrki raeqmkeisd fvkkknipkd etvyiggdln vnkgtpefkd
241 mlknlnvndv lyaghnstwd pqsnsiakyn ypngkpehld yiftdkdhkq pkqlvnevvt
301 ekpkpwdvya fpyyyvyndf sdhypikays k

The phospholipase can be a phospholipase A, B, C or D, such as PLD from Streptomyces, see e.g. https://www.ncbi.nlm.nih.gov/protein/BAL15170.1 (*Streptomyces vinaceus*) (SEQ ID No. 5):

1 mhrhtpslrr psahlpsala vraavpaall alfaavpasa apaagsgadp aphldaveqt
61 lrqvspgleg qvwertagnv ldastpggad wllqtpgcwg ddkctarpgt eqllskmtqn
121 isqatrtvdi stlapfpnga fqdaivsglk tsaargnklk vrvlvgaapv yhlnvlpsky
181 rdelvaklga darnvdlnva smttsktafs wnhskllvvd gqsvitggin dwkddyleta
241 hpvadvdlal rgpaaasagr yldelwswtc qnksniasvw fassngaacm pamakdtapa
301 apapapgdvp avavgglgvg ikrndpsssf rpalpsapdt kcvvglhdnt nadrdydtvn
361 peesalrtli ssanrhieis qqdvnatcpp lprydirvyd alaarmaagv kvrivvsdpa
421 nrgavgsggy sqikslseis dtlrdrlalv tgdqgaakat mcsnlqlatf rssqsptwad
481 ghpyaqhhkv vsvddsafyi gsknlypawl qdfgyvvesp aaaaqlnarl lapqwqysra
541 tatidheral cqs

Preferably the phospholipase is a phospholipase C (PLC) (i.e. a phospholipase that cleaves before the phosphate, releasing diacylglycerol and a phosphate-containing head group). Preferably the PLC is a bacterial PLC, selected from any of the following groups:
Group 1 – Zinc metallophospholipases

Group 2 – Sphingomyelinases (e.g. sphingomyelinase C)
Group 3 – Phosphatidylinositol
Group 4 – Pseudomonad PLC

A Group 1 PLC is preferred, particularly PLC from *Clostridium perfringens*, see e.g. https://www.ncbi.nlm.nih.gov/protein/EDT77687.1 (SEQ ID No.1):

1 mkrkickali caalatslwa gastkvyawd gkidgtgtha mivtqgvsil endmsknepe
61 svrknleilk enmhelqlgs typdydknay dlyqdhfwdp dtdnnfskdn swylaysipd
121 tgesqirkfs alaryewqrg nykqatfylg eamhyfgdid tpyhpanvta vdsaghvkfe
181 tfaeerkeqy kintagcktn edfyadilkn kdfnawskey argfaktgks iyyshasmsh
241 swddwdyaak vtlansqkgt agyiyrflhd vsegndpsvg knvkelvayi stsgekdagt
301 ddymyfgikt kdgktqewem dnpgndfmtg skdtytfklk denlkiddiq nmwirkrkyt
361 afpdaykpen ikviangkvv vdkdinewis gnstynik

This cytolysin provides for highly effective lysis of animal host cells in the present technology, despite reports in the literature that purified *C. perfringens* PLC when used alone has no cytotoxic activity against leukocytes.

The cytolysin can be a wild-type cytolysin or an active variant (produced e.g. by recombinant DNA technology). An active variant of a cytolysin is a variant of a cytolysin that retains the ability to lyse a target cell, demonstrating e.g. at least 10%, preferably at least 25%, preferably at least 50%, preferably at least 60%, preferably at least 70%, preferably at least 80%, preferably at least 90%, preferably at least 95% of the activity of the wild-type protein in any assay where lytic activity against a target cell can be shown for the wild-type protein.

"An active variant thereof" includes within its scope a fragment of the wild-type protein. In preferred embodiments, a fragment of the wild-type protein is selected that is at least 10% of the length of the wild-type protein sequence, preferably at least 20%, preferably at least 30%, preferably at least 40%, preferably at least 50%, preferably at least 60%, preferably at least 70%, preferably at least 80%, preferably at least 90% and most preferably at least 95% of the length of the wild-type protein sequence.

"An active variant thereof" also includes within its scope a protein sequence that has homology with the wild-type protein sequence, such as at least 50% identity, preferably at least 60%, preferably at least 70%, preferably at least 80%, preferably at least 85%, preferably at least 90%, preferably at least 95%, preferably at least 97%, and most preferably at least 99% identity, for example over the full wild-type sequence or over a region of contiguous amino acid residues representing 10% of the length of the wild-type protein sequence, preferably at least 20%, preferably at least 30%, preferably at least 40%, preferably at least 50%, preferably at least 60%, preferably at least 70%, preferably at least 80%, preferably at least 90% and most preferably at least 95% of the length of the wild-type protein sequence. Methods of measuring protein homology are well known in the art and it will be understood by those of skill in the art that in the present context, homology is calculated on the basis of amino acid identity (sometimes referred to as "hard homology").

The homologous active cytolysin variant typically differs from the wild-type protein sequence by substitution, insertion or deletion, for example from 1, 2, 3, 4, 5 to 8 or more substitutions, deletions or insertions. The substitutions are preferably 'conservative', that is to say that an amino acid may be substituted with a similar amino acid, whereby similar amino acids share one of the following groups: aromatic residues (F/H/W/Y), non-polar aliphatic residues (G/A/P/I/L/V), polar-uncharged aliphatics (C/S/T/M/N/Q) and polar-charged aliphatics (D/E/K/R). Preferred sub-groups comprise: G/A/P; I/L/V; C/S/T/M; N/Q; D/E; and K/R.

The cytolysin or active variant (as described above) may have any number of amino acid residues added to the N-terminus and/or the C-terminus provided that the protein retains lytic activity. Preferably, no more than 300 amino acid residues are added to either or both ends, more preferably no more than 200 amino acid residues, preferably no more than 150 amino acid residues, preferably no more than 100 amino acid residues, preferably no more than 80, 60 or 40 amino acid residues, most preferably no more than 20 or 10 or 5 amino acid residues.

Preferably, the sample is subject to mixing after the cytolysin has been added.

Preferably, to promote cytolysin activity, particular buffering conditions and/or incubation temperature might be provided for any one selected cytolysin. Cytolysin

incubation can take place at e.g. between 5°C and 50°C, such as between 15°C and 45°C (e.g. 37°C), and for between 1min and 120min, preferably between 1min and 60min, more preferably between 1min and 30min (e.g. 15min or 20min). For part or all of the cytolysin incubation, the sample is preferably subject to mixing/shaking, at e.g. between 1 and 1500rpm, preferably between 1 and 1000rpm (e.g. at 500rpm or 1000rpm).

Preferably, the cytolysin is used in the sample at a concentration of at least 0.1mg/ml, such as between 0.1mg/ml and 100mg/ml, preferably between 0.1mg/ml and 100mg/ml, preferably between 1mg/ml and 100mg/ml (e.g. at 40mg/ml).

The DNase

If a DNase is used in the present methodology, the DNase can be an endonuclease or an exonuclease (or a combination thereof can be provided), preferably an endonuclease.

Preferred DNases (particularly where the biological sample is a blood sample) include HL-SAN DNase (heat labile salt activated nuclease, supplied by Arcticzymes) and MolDNase (endonuclease active in the presence of chaotropic agents and/or surfactants, supplied by Molzym), and active variants are also contemplated, essentially as discussed above in relation to the cytolysin.

Preferably, the sample is subject to mixing after the DNase has been added.

Preferably, to promote DNase activity, particular buffering conditions and/or incubation temperature might be provided for any one selected DNase. DNase incubation can take place at e.g. between 5°C and 50°C, such as between 15°C and 45°C (e.g. 37°C), and for between 1min and 120min, preferably between 1min and 60min, more preferably between 1min and 30min (e.g. 15min). In particularly preferred embodiments, the DNase buffer is added to the sample, containing the cytolysin, and incubated (e.g. as described above) before pelleting. The pellet is then resupsended in DNase buffer and the DNase itself is added (ahead of further incubation).

The biological sample

Preferably, the biological sample is a blood sample. Preferably, where the sample is blood, the cytolysin targets/lyses (e.g. human) leukocytes.

Preferably, especially where the sample is blood and/or the cytolysin is PLC from Clostridium perfringens, the sample comprises a chelating agent (e.g. EDTA).

Kits

Also provided is a kit comprising a cytolysin (according to e.g. any of the aspects described above) (preferably with a buffer for the cytolysin) and means to physically deplete free nucleic acid within a biological sample or otherwise render such nucleic acid unidentifiable. Free nucleic acid includes nucleic acid not contained within a cell or virus particle (e.g. has been released/liberated from animal cells within the sample as a result of lysis of those cells).

The means can be e.g means for nucleic acid capture (using e.g. magnetic bead technology), means for rendering nucleic acid unidentifiable (e.g. PMA or EMA) or, preferably, a nuclease (e.g. a DNase) (preferably with a suitable buffer and/or a composition for inactivating the nuclease), according e.g. to any of the aspects described above.

General

Please note that wherever the term ‘comprising’ is used herein we also contemplate options wherein the terms ‘consisting of’ or ‘consisting essentially of’ are used instead. In addition, please note that the term ‘protein’ used herein can be used interchangeably with the term ‘polypeptide’.

Examples

In the context of medical microbiology, metagenomics sequencing needs to achieve sufficient genome coverage to identify the pathogenic species present and preferably detect all resistance markers, whether mutational or acquired. To deliver this we estimate that a minimum of 10x genome coverage is required. We directly sequenced (HiSeq) blood, spiked with pathogen cells (Escherichia coli), which delivered human reads only, highlighting the need for pathogen DNA enrichment

(data not shown). Hence, host DNA depletion is required to reliably and cost effectively apply metagenomics to infectious disease diagnosis.

Here, we describe the process of developing a simple, rapid and highly efficient human DNA depletion method to enable downstream metagenomic sequencing (and other molecular applications e.g. PCR) for the detection and identification of pathogens and associated antibiotic resistance markers.

For efficient and cost effective metagenomic diagnosis of infection, human DNA depletion or pathogen DNA enrichment is essential. We took the human DNA depletion approach focussing on differential lysis of human cells, and removal of human DNA, leaving intact non-human pathogens for further analysis. We used blood as a model sample type, as blood represents one of the most complex clinical samples to successfully apply metagenomic infection diagnosis due to the very high ratio of human:pathogen DNA (as high as 10^9:1).

We applied cytolysins for differential lysis of human cells and endonucleases (DNases) for digestion of liberated DNA. We tested a number of DNases to determine the most efficient in blood. We then combined the most efficient DNases with various cytolysins to determine whether and how efficiently these toxins would lyse the DNA-containing leukocytes in blood.

A positive control (PC) was added in to every experiment, which was DNA extracted from 200µl of blood. For cytolysin experiments, the blood was spiked with the most common sepsis causing pathogens (*E. coli* and *S. aureus*) to ensure that pathogens were not lysed during the procedure. For all qPCR reactions, a no template control (NTC; molecular grade nuclease free dH_2O) was included. A MolDNase control sample (from the MolYsis (RTM) kit, Molzym, Germany) was also included where appropriate as it has been proven to work in blood.

Subsequently, DNA was extracted as follows (unless otherwise stated in the experimental procedure):

1. Bacterial lysis buffer (to a maximum volume of 380µl) and proteinase K (20µl) was added to the treated sample and mixed by vortexing. No bacterial lysis buffer

was added to blood samples that were not spiked with bacteria (volume made up to 400 µl with PBS where necessary).

2. All samples were incubated at 65°C for 5min
3. Followed by purification on the MagNAPure (RTM)

For all experiments, human and bacterial DNA was quantified using qPCR. Specific hydrolysis probe assays were designed or taken from the literature to detect human, *E. coli* and *S. aureus* DNA (all were single copy gene targets; RNA polymerase II, *cyaA* and *eap* respectively). In addition, fungal and viral targets included *C. albicans* 5.8S rRNA, *A. niger* ITS1-2, HBV X gene, and HIV 5' nuclease assay in LTR gene. All qPCR results are presented as amplification curves and/or quantification cycle (Cq) values (this represents the cycle at which the fluorescence signal increases above background which is directly related to the quantity of starting template concentration). The relative concentration of DNA in samples was calculated using the ΔCq (every 3.3 cycles represents a 10-fold difference in concentration; the higher the Cq value the less starting template DNA was present in the sample).

Example 1 – Efficacy of endonucleases for DNA digestion in blood

Initial focus was on identifying an endonuclease that would digest DNA released from leukocytes so that the efficacy of cytolysins could be easily assessed in blood. In this experiment, blood samples were freeze thawed three times to release human DNA and an endonuclease; either DS-DNase, HL-SAN DNase (heat labile, salt active nuclease) or micrococcal nuclease from *S.aureus* was added, incubated at 37°C and DNA was extracted. Controls included a positive control (PC – DNA from 200µl spiked blood without DNase treatment), a MolDNase control (known to work in blood) and a negative control (NTC – nuclease free water), as detailed above. Human specific qPCR was performed on all DNA extracts and Cq values were compared to determine whether the endonuclease treatment worked.

Detailed procedure:

To lyse blood cells, samples were frozen at -70°C and thawed at room temperature (RT) three times

Freeze-thawed blood was aliquoted into 5x 200µl samples

To sample 1, 5µl of HL-SAN DNase (28.4U/µl) was added
To sample 2, 5µl of DS-DNase (2U/µl) was added
To sample 3, 20µl of nuclease micrococcal (resuspended in 100µl of nuclease free water; 0.62U/µl) was added

All samples were mixed by vortexing

Samples 1-3 and PC were incubated at 37°C for 30min

To the MolDNase control sample, 50µl of DB1 buffer was added followed by 5µl of MolDNase then incubated at RT for 15min

All reactions were stopped by adding 5µl of DNase inactivation buffer (Ambion (RTM), life technologies (RTM))

DNA was extracted and quantified by human qPCR (as described above)

Results:

As shown in Table 1 and Figure 1, DS-DNase (sample 2) and nuclease micrococcal (sample 3) showed no endonuclease activity on human DNA in blood samples, with ΔCq <1 compared with PC. With a ΔCq of 2.2, HL-SAN DNase (sample 1) showed endonuclease activity resulting in an approximate 4-fold reduction in human DNA when compared to the PC. As previously stated MolDNase was known to work in blood samples and showed the greatest endonuclease activity with the highest Cq value.

Table 1: Human qPCR results after various endonuclease treatments

Sample ID	Endonuclease	Human qPCR (Cq)
1	HL-SAN DNase	24.77
2	DS-DNase	22.80
3	Nuclease micrococcal	22.35
MolDNase control	MolDNase	29.47
PC	-	22.58

Conclusion:
From all the endonucleases tested in this experiment, HL-SAN DNase was the only one to show the potential to work effectively in blood. HL-SAN DNase was the endonuclease of choice selected for further testing. As HL-SAN DNase is known to be most active in high salt concentrations, we aimed to test a high salt buffer to improve activity, and Example 2 details buffer optimization.

Example 2 – Optimization of HL-SAN buffer conditions

From Example 1, HL-SAN DNase was chosen as the most promising endonuclease to work in blood. As HL-SAN DNase is a salt active enzyme, we tested the addition of a high salt buffer to optimize HL-SAN DNase activity on human DNA in blood samples. A high-salt buffer was made and added in various volumes to freeze-thawed blood samples with HL-SAN DNase, incubated at the known working temperature (37°C), DNase inhibitor was added and samples further incubated. MolDNase control, PC and NTC were included; all samples were subjected to DNA extraction and human qPCR (as detailed above).

HL-SAN buffer components:
10mM Tris HCl, 100mM magnesium and 1M NaCl pH8.5

Detailed procedure:

To lyse blood cells, 2ml of blood was frozen at -70°C and thawed at RT three times
Freeze-thawed blood was spiked with human DNA and aliquoted into 5x 200µl samples:
To sample 1, 20µl of HL-SAN buffer and 3µl of HL-SAN DNase was added
To sample 2, 100µl of HL-SAN buffer and 3µl of HL-SAN DNase was added
To sample 3, 180µl of HL-SAN buffer and 3µl of HL-SAN DNase was added
The above reactions were incubated at 37°C for 15min

To the MolDNase control, 50µl of DB1 buffer and 5µl MolDNase was added and incubated at RT for 15min

All reactions were stopped by adding 5µl of DNase inactivation buffer (Ambion (RTM), life technologies (RTM))

DNA was extracted and quantified by human qPCR (as described above)

Results:

Table 2 and Figure 2 show that the addition of HL-SAN buffer increases the activity of HL-SAN DNase in correlation with an increase in volume. The most effective amount of HL-SAN buffer was 180µl, which resulted in a similar activity to MolDNase (<1 Cq difference between DNase treatments) and reduced the level of human DNA approximately 32-fold (∆Cq5) compared to no endonuclease treatment (PC). In the absence of buffer, HL-SAN DNase alone, resulted in a human qPCR Cq value of 24.77 (Table 1), with the addition of 180µl HL-SAN buffer this increased to 27.02Cq (Table 2), showing an increase in HL-SAN DNase activity to reduce human DNA approximately 4-fold (∆Cq 2).

Table 2: Human qPCR results after endonuclease treatment with various buffer volumes

Sample ID	**Conditions**	**Human qPCR (Cq)**
1	20µl HL-SAN buffer + HL-SAN DNase	22.27
2	100µl HL-SAN buffer + HL-SAN DNase	24.64
3	180µl HL-SAN buffer + HL-SAN DNase	27.02
MolDNase control	50µl DB1 buffer + MolDNase	27.63
PC	-	22.32
NTC	-	-

Conclusion:

The addition of a high salt buffer (HL-SAN buffer) increased the efficiency of HL-SAN DNase to digest human DNA present in the blood samples after cell lysis by freeze-thawing. Using this combination (HL-SAN buffer and HL-SAN DNase) enabled approximately the same level of human DNA depletion as the known control (MolDNase). Therefore, to test the robustness of the optimized HL-SAN DNase method, the experiment was repeated (with an adjusted volume of HL-SAN buffer

required due to limitations of input volume for DNA extraction) against MolDNase with respective DB1 buffer (Example 3).

Example 3 – Comparison of HL-SAN DNase and MolDNase activity

Here, we tested the robustness of the optimized method selected from Example 2 and compared the activity of HL-SAN DNase and MolDNase with their respective buffers. The volume of HL-SAN buffer which provided the same level of activity between HL-SAN DNase and MolDNase was 180µl, however, due to the volume input limitation of the MagNAPure (RTM) for DNA purification, the volume of HL-SAN buffer was reduced to 150µl. Blood cells were lysed by freeze-thawing, spiked with human DNA and HL-SAN DNase or MolDNase was added with their respective buffer, incubated and followed by enzyme heat inactivation. PC was also included, and DNA was extracted from all samples and human qPCR carried out.

Detailed procedure:

1. To lyse blood cells, 2ml of blood was frozen at -70°C and thawed at RT three times
2. Freeze-thawed blood was spiked with human DNA and aliquoted into 4x 250µl samples
3. To the HL-SAN DNase sample, 150µl of HL-SAN buffer (Example 2) and 4µl of HL-SAN DNase was added, mixed by vortexing and incubated at 37°C for 15min
4. To the MolDNase control sample, 50µl of buffer DB1 and 4µl of MolDNase was added, mixed by vortexing and incubated at RT for 15min
5. To the MolDNase control sample and PC PBS was added to increase the sample volume to 400µl (the required input volume for the MagNAPure (RTM))
6. DNase activity was stopped by heat killing the enzymes at 65°C for 10min
7. DNA was extracted and quantified by human qPCR (as described above)

Results:

Table 3 and Figure 3 show that the optimized HL-SAN DNase method out performs the MolDNase control. There is a difference of approximately ΔCq 2 which equates to an approximate 4-fold reduction in human DNA.

Table 3: Human qPCR results of HL-SAN DNase and MolDNase treatment with respective buffers

Sample ID	**Human qPCR (Cq)**
HL-SAN DNase	30.54
MolDNase control	28.23
PC	21.80
NTC	-

Conclusion:
Under optimized buffer conditions, HL-SAN DNase can work as, if not more, effectively as MolDNase in blood to deplete human DNA. At this point we continued to work with HL-SAN DNase as our endonuclease of choice and began the process of selecting a suitable cytolysin. Example 4 details the different cytolysins that we initially chose to evaluate for leukocyte cell lysis ability/efficacy.

Example 4 - Host DNA depletion using Streptolysin O and Alpha hemolysin

After identifying HL-SAN DNase as an effective endonuclease for the digestion of DNA, we investigated the potential of cytolysins to target and lyse specific cell types. Here, we evaluated the activity of two membrane pore forming cytolysins, namely streptolysin O (*Streptococcus pyogenes*) and alpha hemolysin (*Staphylococcus aureus*), on leukocyte lysis. Cytolysins were added (individually and in combination) to blood to lyse host cells. Samples were then incubated and released DNA from lysed cells was digested with MolDNase and a DNase inactivation reagent added after further incubation. PC and NTC samples were included and DNA was extracted from all samples and DNA quantified by human qPCR (as detailed above).

Cytolysin purchase information:

Streptolysin O

- Cat number no. S5265-25ku
- Lot number 025M4059V
- 25,000-50,000 u/vial
- 0.71mg Solid
- 229577 Units/mg solid

- 4794117 Unts/mg protein

Alpha-hemolysin

- Cat no H9395-5MG
- Lot no 095M4057V
- 28840 Units/mg Solid
- 49647 units/mg protein

Detailed procedure:

1. Streptolysin O and alpha-Hemolysin (0.71 mg (163,000 units) and 5 mg (144,200 units) respectively) was resuspended in 350µl of nuclease-free water
2. To sample 1, 50µl of Streptolysin O was added to 200µl of blood
3. To sample 2, 50µl of alpha-hemolysin was added to 200µl of blood
4. To sample 3, 50µl of Streptolysin O and 50µl of alpha-hemolysin was added to 200µl of blood
5. All samples were mixed by vortexing and incubated at 37°C with shaking at 400rpm for 30 min
6. After incubation, 150µl of HL-SAN buffer was added, followed by 3µl of HL-SAN DNase
7. Samples were further incubated at 37°C for 15 min
8. DNase activity was stopped by heat killing the enzymes at 65°C for 10min
9. To samples 1-3, 100µl of bacterial lysis buffer was added and to the PC sample 180µl of bacterial lysis buffer was added
10. DNA was extracted from all samples and human qPCR used to quantify human DNA (as detailed above)

Results:

When used alone streptolysin O and alpha-hemolysin showed approximately the same leukocyte lysis efficacy (Table 3), providing an approximate 10^3 fold depletion of DNA. Using both cytolysins in combination (alpha-hemolysin and streptolysin O in combination) on the same blood sample, resulted in improved leukocyte lysis efficiency and improved human DNA depletion with an approximate further 10-fold reduction (ΔCq 3.3) in human DNA.

Table 4: Human qPCR results after cytolysin treatment

Sample ID	Cytolysin	Human qPCR (Cq)
1	Streptolysin O	31.96
2	Alpha-hemolysin	31.32
3	Alpha-hemolysin & streptolysin O	35.31
PC	-	21.79
NTC	-	-

Conclusion:

Here we show that membrane pore forming cytolysins are able to target human cells and enable host DNA depletion. Interestingly, it was the combination of the two cytolysins that produced the greatest human DNA depletion. As we had shown that cytolysins could target human cells and demonstrated that host DNA depletion was possible with this approach, we switched our focus to another member of the cytolysins, namely phospholipase C (PLC) from *C. perfringens* (which is a cytolysin that breaks down phospholipids in bilayer membranes of eukaryotic cells) (Example 5).

Example 5 – Investigation of PLC activity on host cell lysis

As previously mentioned, PLC is a cytolysin produced by *C. perfringens* and acts by targeting and breaking down phospholipids in the bilayer membrane of eukaryotic cells. We therefore wanted to test PLC for specific host cell lysis and subsequent host DNA digestion using HL-SAN DNase. PLC is a known zinc metallophospholipase and requires the presence of zinc for activity; it was however unknown whether the concentrations of zinc in human blood would be sufficient for PLC to work. Also required for PLC activity are calcium and magnesium ions. With these experiments using blood collected with EDTA preservative, there was a concern that EDTA would chelate the required calcium and metal ions necessary for PLC activity. Therefore, we tested PLC on blood with no preservative, blood containing EDTA preservative and on blood in the presence of a metal ion containing buffer. PLC was added to the various blood sample types and incubated with shaking for host cell lysis. HL-SAN DNase (with HL-SAN buffer) was then added and incubated for host DNA digestion followed by heat inactivation of HL-SAN DNase.

PC and NTC samples were included, and DNA was extracted from all samples followed by human qPCR (as detailed in above).

PLC buffer components:
0.1M $ZnCl_2$ and 0.1M $MgCl_2$

Detailed procedure:

1. PLC (4mg) was reconstituted in 100µl of molecular grade water (40µg/µl)
2. Blood was aliquoted into 4x 250µl
3. To sample 1 (without EDTA preservative) and sample 2 (with EDTA preservative), 20µl of PLC was added and mixed well by vortexing, followed by incubation at 37°C with shaking at 500rpm for 15min
4. After incubation, 150µl of HL-SAN buffer and 4µl of HL-SAN DNase was added to samples 1 and 2, mixed well by vortexing and incubated at RT for 15min
5. To sample 3, 150µl of HL-SAN buffer and PLC buffer was added followed by 4µl of HL-SAN DNase and 20µl of PLC then mixed by vortexing and incubated for 15 min at 37°C without shaking
6. PC was topped up with 150µl of PBS (total 400µl)
7. HL-SAN DNase was inactivated by incubating all samples at 65°C for 10 min
8. DNA was extracted from all samples and human qPCR used to quantify human DNA (as detailed in Section 3)

Results:
There was no improvement in human DNA depletion when PLC was tested on blood with no preservative or with PLC buffer (Table 5 and Figure 5) in fact, the lack of EDTA or addition of PLC buffer reduced the efficacy of depletion. Sample 2 showed the highest level of host DNA depletion with an approximate 100-fold reduction in human DNA compared to the PC (ΔCq6).

Table 5: Human qPCR results of PLC activity in different sample conditions

Sample ID	Conditions	Human qPCR (Cq)
1	PLC on raw blood	26.36
2	PLC on EDTA blood	29.45
3	PLC combined with PLC buffer	21.87
PC	-	23.59
NTC	-	-

Conclusion:

Despite PLC being known to require calcium, magnesium and zinc ions for activity, the addition of buffer containing these ions appeared to decrease the efficiency of PLC to lyse host cells. After concerns that the preservative EDTA would chelate the metal ions required for PLC activity, we observed that PLC worked better in blood samples preserved with EDTA and was less effective in blood without any preservative. All previous experiments were performed in a volume of 200-250µl of blood to test the efficiency of PLC and HL-SAN DNase on human DNA depletion. We next wanted to increase the working volume of blood due to the low number of bacterial cells known to be present per millilitre of septic blood (potentially as few as 1 colony forming unit per millilitre) (Example 6).

Example 6 - Investigation of PLC activity on host DNA depletion and bacterial DNA recovery in an increased volume of blood

The pauci-microbial nature of sepsis means that testing larger volumes of blood increases diagnostic sensitivity. Therefore, we wanted to test the activity of PLC in a larger volume of blood (1ml) and also determine if PLC had any unwanted activity on bacterial cells. Blood was spiked with the most common sepsis causing pathogens (*E. coli* and *S. aureus*). Spiked blood was incubated with PLC to enable host cell lysis, followed by the addition of HL-SAN DNase (with HL-SAN buffer) for DNA digestion and the endonuclease was heat inactivated. A PC sample was included and DNA was extracted from all samples, followed by qPCR for human, *E. coli* and *S. aureus* DNA (as detailed above).

Detailed procedure:

1. PLC (4mg) was reconstituted in 100µl of molecular grade water (40µg/µl)
2. Blood spiked with *E. coli* and *S. aureus* cultures was aliquoted into 1x 1ml and 1x 200µl samples
3. To 1ml of spiked blood, 100µl of PLC was added and incubated at 37°C for 20 min with shaking at 500 rpm
4. To 200µl of spiked blood, 20µl of PLC was added and incubated at 37°C for 20 min with shaking at 500 rpm
5. After incubation, 500µl or 150 µl of HL-SAN buffer was added to 1ml or 200 µl samples respectively, followed by 10µl or 3 µl of HL-SAN DNase for 1 ml or 200µl respectively, mixed briefly by vortexing then incubated at 37°C for 15 min
6. Samples were centrifuged for 10 min at 12,000xg
7. The supernatant was carefully decanted and the pellet was re-suspended in 200µl of PBS
8. HL-SAN DNase was inactivated by heat killing at 68°C for 10 min
9. DNA was extracted from all samples and qPCR was used to quantify human, *E.coli* and *S. aureus* DNA respectively (as detailed above)

Results:

Increasing the volume of blood resulted in less efficient human DNA depletion (Table 6 and Figure 6A). There was approximately 4-fold more human DNA remaining in 1ml of blood compared with 200µl of blood (ΔCq2). There was no loss of *E. coli* between the two volumes, with the 1ml sample showing an approximate 5-fold increase in *E. coli* DNA (ΔCq~2.5) as expected (Table 6 and Figure 6B). There was, however, loss of *S. aureus* DNA in the 200µl and 1ml samples, equivalent to approx. 100 fold reduction (ΔCq~6 in the 200µl sample [lower in the 1ml sample due to the 5 fold increase in volume tested compared to the PC]) (Table 6 and Figure 6C).

Table 6: Human, *E. coli* and *S. aureus* qPCR results after PLC and HL-SAN DNase treatment on increased volumes of bacteria spiked blood

Sample ID	**Human qPCR (Cq)**	***E. coli* qPCR (Cq)**	***S. aureus* qPCR (Cq)**
PLC on 1ml blood	32.11	18.72	24.97
PLC on 200µl blood	34.74	21.47	28.26
PC	22.20	21.54	22.89
NTC	-	-	-

Conclusion:

Increasing the volume of blood resulted in less efficient human DNA depletion. Loss of *S. aureus* DNA was observed suggesting PLC activity on Gram-positive cell walls or a reduction in *S. aureus* lysis efficiency compared to the PC (possibly due to heat deactivation of DNase). There was no loss of *E. coli* DNA confirming the Gram-negative bacterial cells were not lysed by PLC. We proceeded to attempt to improve the efficiency of human DNA depletion in 1ml of blood by ensuring effective mixing during incubation with PLC (Example 7). The loss of *S. aureus* was also investigated using the hypothesis that heat inactivation of HL-SAN DNase was affecting the cell wall of *S. aureus*, reducing the efficiency of cell lysis (Example 8).

Example 7 – Investigation of efficient mixing during targeted cell lysis in increased volumes of blood

Firstly, to investigate the loss of PLC efficiency on host cell lysis in 1ml of blood, we investigated the effect of efficient mixing. After the addition of PLC to the bacterial spiked blood, samples were aliquoted in larger volume sample tubes (5ml) and continuously mixed during the incubation period to enhance contact of PLC with the host cells present in the sample and increase lysis efficiency. HL-SAN DNase (plus HL-SAN buffer) was added to enable host DNA depletion and incubated, followed by heat inactivation. A PC sample was included and DNA was extracted from all samples, followed by qPCR for human, *E. coli* and *S. aureus* DNA (as detailed above).

Detailed procedure:

1. PLC (4mg) was reconstituted in 100µl of molecular grade water (40µg/µl)
2. Blood spiked with *E. coli* and *S. aureus* cultures was aliquoted into 1x 1ml (in a 5ml tube) and 1x 200µl samples
3. To 1ml of spiked blood, 100µl of PLC was added and incubated at 37°C for 20 min with slow mixing using a Hulamixer (RTM)
4. To 200µl of spiked blood, 20µl of PLC was added and incubated at 37°C for 20 min with shaking at 500 rpm
5. After incubation, 500µl or 150 µl of HL-SAN buffer was added to 1ml or 200 µl samples respectively, followed by 10µl or 3 µl of HL-SAN DNase for 1 ml or 200µl respectively, mixed briefly by vortexing then incubated at 37°C for 15 min
6. Samples were centrifuged for 10 min at 12,000xg
7. The supernatant was carefully decanted and the pellet was re-suspended in 200µl of PBS
8. HL-SAN DNase was inactivated by heat killing at 68°C for 10 min
9. DNA was extracted from all samples (including PC) and qPCR was used to quantify human, *E.coli* and *S. aureus* DNA respectively (as detailed above)

Results:

The introduction of a larger sample tube and slow mixing after the addition of PLC resulted in almost complete removal of human DNA (approximately 1 cell human DNA remaining; a depletion of ~2.6 x 10^5 fold (Table 7 and Figure 7) for the 1ml sample and complete removal of human DNA for the 200µl sample (a depletion of at least 10^6 fold).

Table 7: Human qPCR results of PLC activity after the addition of efficient mixing during host cell lysis

Sample ID	**Human qPCR (Cq)**
PLC on 1ml blood	38.04
PLC on 200µl blood	-
PC	22.38
NTC	-

Conclusion:

By ensuring efficient mixing during host cell lysis the activity of PLC was improved and provided the level of depletion necessary for detecting pathogen sequences in blood by sequencing. However, as described in Example 6, the loss of *S. aureus* DNA still needed to be investigated (detailed in Example 8).

Example 8 – Altered inactivation of HL-SAN DNase to improve Gram-positive bacterial DNA recovery

We hypothesised that heat inactivation of HL-SAN DNase was affecting the cell wall of *S. aureus,* reducing the efficiency of cell lysis, resulting in low recovery levels of DNA. The aim of this experiment was to try a new method of inactivating HL-SAN DNase in order to improve recovery of *S. aureus* DNA. Rather than heat inactivation of HL-SAN, we inactivated the DNase by removing the high salt conditions required for its activity. PLC was added to bacterial spiked blood samples, incubated and mixed slowly. HL-SAN DNase (+HL-SAN buffer) was added to enable host DNA depletion and incubated. Samples were centrifuged to pellet the intact bacterial cells and the supernatant containing high salt buffer was removed. A PC sample was included and DNA was extracted from all samples, followed by qPCR for human, *E. coli* and *S.* aureus DNA (as detailed above).

Detailed procedure:

1. PLC (4mg) was reconstituted in 100µl of molecular grade water (40µg/µl)
2. Blood spiked with *E. coli* and *S. aureus* cultures was aliquoted into 1x 1ml (in a 5ml tube) and 1x 200µl samples
3. To 1ml of spiked blood, 100µl of PLC was added and incubated at 37°C for 20 min with slow mixing using a Hulamixer (RTM)

4. To 200µl of spiked blood, 20µl of PLC was added and incubated at 37°C for 20 min with shaking at 500 rpm
5. After incubation, 500µl or 150 µl of HL-SAN buffer was added to 1ml or 200 µl samples respectively, followed by 10µl or 3 µl of HL-SAN DNase for 1 ml or 200µl respectively, mixed briefly by vortexing then incubated at 37°C for 15 min
6. Samples were centrifuged for 10 min at 12,000xg
7. The supernatant was carefully decanted and the pellet was re-suspended in 1.5ml PBS
8. Prior to DNA extraction, bacterial cells were pelleted by centrifuging at 12000xg for 5min
9. DNA was extracted from all samples (including PC) and qPCR was used to quantify human, *E.coli* and *S. aureus* DNA respectively (as detailed above)

Results:

Using buffer exchange rather than heat inactivation on HL-SAN DNase resulted in efficient human DNA depletion with no loss of *E. coli* or *S. aureus* DNA (Table 8 and Figure 8). Human DNA depletion was effectively ~ 2.3 x 10^5 fold when using a 1ml sample and (data not shown) at least 10^6 fold when using a 200µl sample (no human DNA detected).

Table 8: Human, *E. coli* and *S. aureus* qPCR results after altered HL-SAN DNase inactivation

Sample ID	**Human qPCR (Cq)**	***E. coli* qPCR (Cq)**	***S. aureus* qPCR (Cq)**
PLC	37.36	17.65	19.04
PC	21.90	20.17	21.47
NTC	-	-	-

Conclusion:

Introducing a buffer exchange to inactivate HL-SAN DNase instead of heat inactivation, improved the lysis efficiency of *S. aureus* cells (it is likely that this could also have been achieved by using a more robust lysis method such as bead beating or using an enzyme cocktail). This method alteration enabled efficient *S. aureus*

DNA recovery with no negative effect on *E. coli* DNA recovery (previously reported in Example 6) or on human DNA depletion (previously reported in Example 7). Hence an efficient cytolysin human DNA depletion procedure had been developed that did not result in the loss of the microbial component of the sample. In order to confirm the robustness of this procedure we compared it to the commercially available MolYsis (RTM) method and our in-house modified MolYsis (RTM) procedure (Example 9).

Example 9 – Comparison of cytolysin human DNA depletion against MolYsis (RTM) Basic 5 kit and a modified MolYsis (RTM) method

To test the robustness of our newly developed human DNA depletion procedure we compared it to the commercially available MolYsis (RTM) pathogen DNA isolation protocol and an in-house modified MolYsis (RTM) protocol. Our cytolysin human DNA depletion procedure was carried out as per Example 8 using the buffer exchange method rather than heat inactivation of HL-SAN DNase. The MolYsis (RTM) pathogen DNA isolation protocol was performed as detailed in the manufacturer's instructions. A modified MolYsis (RTM) protocol (developed in house) was also tested which initially removed leukocytes by immunomagnetic separation, followed by MolYsis (RTM) as per the manufacturer's instructions.

Method 1 (Cytolysin human DNA depletion):
As described in Example 8.

Method 2 (MolYsis (RTM)):
MolYsis (RTM) was used as per the manufacturer's instructions.

Method 3 (Modified MolYsis (RTM)):

1. Anti-CD45 coated magnetic beads were re-suspended by gentle mixing then the desired volume of beads (250μl per 1ml sample) was aliquoted
2. Beads were washed by re-suspending in 1ml of isolation buffer (25ml Ca^{2+}, Mg^{2+} free PBS, 100μl 0.5M EDTA and 0.025g BSA)
3. Beads were separated on a magnetic rack and the supernatant was discarded
4. Beads were re-suspended in 250μl of isolation buffer
5. Leukocytes were depleted by adding 250μl of washed beads to 1ml of blood

and mixed gently at 2-8°C for 30min using a Hulamixer (RTM)

6. Beads were separated on a magnetic rack and the supernatant was transferred to a new sterile tube
7. Intact bacterial cells and any remaining blood cells were pelleted by centrifugation at 12,000xg for 10min then the supernatant was discarded
8. The pellet was re-suspended in 1ml PBS
9. Samples were further processed using the MolYsis (RTM) protocol according to the manufacturer's instructions

DNA was extracted from all samples (including PC) and qPCR was used to quantify human, *E.coli* and *S. aureus* DNA respectively for all methods (as detailed above)

Results:

When comparing our human DNA depletion method to commercially available MolYsis (RTM) we observed approximately 10^4-fold more human DNA depletion (ΔCq12) and comparable levels of bacterial DNA recovery (Table 9 and Figure 9). Our modified MolYsis (RTM) protocol also showed an approximate 10^4-fold reduction in human DNA (ΔCq12) compared to MolYsis (RTM).

Table 9: Human, *E. coli* and *S. aureus* qPCR results for method comparison

Sample ID	**Human qPCR (Cq)**	***E. coli* qPCR (Cq)**	***S. aureus* qPCR (Cq)**
PLC	36.05	17.58	18.98
Modified MolYsis (RTM)	36.13	18.74	18.89
MolYsis (RTM)	24.54	17.25	21.33
PC	21.87	20.13	21.31
NTC	-	-	-

Conclusion:

In comparison to the commercially available MolYsis (RTM) kit, our human DNA depletion method was more efficient at human DNA depletion (showing ~ 9.3 x 10^4 fold depletion of human DNA). Only our modified MolYsis (RTM) protocol showed the same level of efficiency compared to our cytolysin human DNA depletion method. This demonstrates that the leading commercially available host depletion kit

does not provide sufficient host cell/DNA depletion to enable efficient pathogen DNA detection by sequencing.

Overview:
In conclusion, we have developed a rapid pathogen identification procedure which utilizes the properties of cytolysins (PLC) and endonucleases (HL-SAN DNase) to specifically target and lyse host cells present in clinical samples (i.e. blood), followed by DNA digestion. This procedure is a pre-step to enable sufficient pathogen DNA extraction for NGS. As blood represents the most complex clinical sample matrix type with extremely high human to bacterial cell ratios, we predict that the clinical sample type will be easily interchangeable without affecting the levels of human DNA depletion.

After a number of methodology alterations, the finalised procedure is detailed below.

<u>Initially optimised human DNA depletion method</u>

PLC solution: 4mg in 100µl nuclease free water
HL-SAN buffer: 10mM Tris HCL, 100mM Magnesium and 1M NaCl pH8.5 in nuclease free water

100µl PLC solution was added to 1ml blood
↓
Incubated at 37°C with gentle mixing for 20min
↓
500µl HL-SAN buffer, 10µl HL-SAN DNase was added and mixed by vortexing, then incubated at 37°C for 15min
↓
Bacterial cells were pelleted at 12,000xg for 10min
↓
Supernatant was discarded
↓
Bacterial cell pellet was resuspended in 1.5ml PBS
↓

Pellet bacterial cells at 12,000xg for 5mins and remove supernatant

↓

Proceeded to DNA extraction of choice

[Total time: 50min.]

DNA extraction

Bacterial cell pellet was resuspended in 350µl bacterial lysis buffer and vortexed

↓

30µl enzyme cocktail (lysozyme, mutanolysis and lysostaphin – lyticase optional) was added and incubated at 37°C for 15min at 1000rpm

↓

20µl proteinase K was added

↓

Mixed by vortexing

↓

Incubated at 65°C for 5min

↓

Proceed to MagNAPure (RTM) (Roche) for DNA extraction

[Total time: 45min.]

[Therefore current protocol turnaround time approximately 90min.]

Example 10 – verification of methodology for fungal enrichment

10.1: The protocol above was altered slightly to focus on fungal enrichment and the final protocol was carried out to verify bacterial enrichment. The protocol was tested using ~200 *E. coli* cells. Blood was spiked with ~200 *E. coli* cells and was processed as detailed in section 10.2.

10.2: Amended protocol ("Enrichment" procedure):

1. PLC was added (0.8 mg/20 µl) to the blood sample (200 µl), vortexed and

incubated at 37 °C for 15 min at 1000 RPM in a heatblock.

2. HL-SAN buffer (5M NaCl and 100mM $MgCl_2$) was added at a 1:1 volume ratio (200 µl) with 10 µl HL-SAN DNase, vortexed and incubated at 37 °C for 15 min at 1000 RPM in a heat block.
3. PBS was added to a total volume of 2 ml (1.5 ml).
4. Cells were pelleted by centrifugation at 12,000 xg for 10 min and the supernatant was discarded.
5. The cell pellet was resuspended in 1.5 ml PBS.
6. Cells were pelleted again by centrifugation at 12,000 xg for 10 min and the supernatant was discarded.
7. To any test samples; 350 µl bacterial lysis buffer, 20 µl enzyme cocktail (6 µl mutanolysin 25 ku/ml, 5 µl lysozyme 10 mg/ml, 4 µl lyticase 10 ku/ml, 3 µl lysostaphin 4 ku/ml, 2 µl chitinase 50 u/ml) and 5 µl RNase A was added.
8. All samples were incubated at 37 °C for 15 min at 1000 RPM in a heat block.
9. To all samples, 20 µl proteinase K was added and incubated at 65 °C for 10 min in a heat block.
10. Total nucleic acid was extracted using the MagnaPure (RTM) Compact automated machine using the DNA_bacteria_V3_2 protocol.
11. Host DNA/RNA depletion and fungal DNA enrichment was determined via qPCR or RT-qPCR.

Results:

After plate counts it was identified that 200 µl of blood was spiked with ~110 *E. coli* cells. This resulted in ~10^5 fold depletion of human DNA and no loss of *E. coli* DNA (Tables 10.1a/b).

Table 10.1a Human DNA qPCR results for ~110 *E. coli* cells spiked blood with and without fungal/bacterial enrichment.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood spiked	22.72	17.3
Blood spiked Enriched (Sample 1)	Undetectable	

Table 10.1b *E. coli* DNA qPCR results for ~110 *E. coli* cells spiked blood with and without fungal/bacterial enrichment.

Sample ID	***E. coli* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood spiked	36.88	0.9
Blood spiked Enriched Sample 1	37.78	

Whole blood was spiked with ~1000 *C. albicans* cells and two samples were processed as detailed in section 10.2. After the enrichment protocol there was between ~10^4 and ~10^5 fold depletion of human DNA and no loss of *C. albicans* DNA (Tables 10.2a/b).

Table 10.2a Human DNA qPCR results in duplicate for ≤1000 *C. albicans* cells spiked blood with and without bacterial/fungal enrichment.

Sample ID	**Human qPCR assay (Cq)**	**Average Human (Cq)**	**Average ΔCq against PC**
PC blood spiked (PC 1)	24.37	24.3	14.9
PC blood spiked (PC 2)	24.32		
Blood spiked Enriched (Sample 1)	Undetectable (>40)	39.2	
Blood spiked Enriched (Sample 2)	38.33		

Table 10.2b *C. albicans* DNA qPCR results in duplicate for ≤1000 *C. albicans* cells spiked blood with and without bacterial/fungal enrichment.

Sample ID	***C. albicans* qPCR assay (Cq)**	**Average *C. albicans* (Cq)**	**Average ΔCq against PC**
PC blood spiked (PC 1)	33.91	33.6	2.3
PC blood spiked (PC 2)	33.28		
Blood spiked Enriched (Sample 1)	30.81	31.3	
Blood spiked Enriched (Sample 2)	31.81		

Whole blood was then spiked with ~200 *C. albicans* cells and was processed as detailed in section 10.2. After plate counts of *C. albicans* on sabouraud agar, it was identified that 200 µl of blood was spiked with ~60 *C. albicans* cells. After the enrichment protocol this resulted in ~10^5 fold depletion of human DNA and no loss of *C. albicans* DNA (Tables 10.3a/b).

Table 10.3a Human DNA qPCR results in for ~60 *C. albicans* cells spiked blood with and without bacterial/fungal enrichment.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood spiked	24.8	15.2
Blood spiked Enriched (Sample 1)	40	

Table 10.3b *C. albicans* DNA qPCR results in for ~60 *C. albicans* cells spiked blood with and without bacterial/fungal enrichment.

Sample ID	***C. albicans* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood spiked	35.55	3.3
Blood spiked Enriched (Sample 1)	32.22	

Using the *A. niger* bioball known to be ~10^8 cfu/ml, serial dilutions were made to ~10^4 and ~10^3. Both samples were processed as described in section 10.2. After the enrichment protocol this resulted in ~10^5 fold depletion of human DNA and no loss of *A. niger* DNA (Tables 10.4a-b/10.5a-b).

Table 10.4a Human DNA qPCR results for ~200 *A. niger* cells (10^3 dilution) spiked blood with and without bacterial/fungal enrichment.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood spiked (PC 1)	22.91	14.71
Blood spiked Enriched (Sample 1)	37.62	

Table 10.4b *A. niger* DNA qPCR results for ~200 *A. niger* cells (10^3 dilution) spiked blood with and without bacterial/fungal enrichment.

Sample ID	***A. niger* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood spiked (PC 1)	39.21	0.79
Blood spiked Enriched (Sample 1)	40	

Table 10.5a Human DNA qPCR results for ~2,000 *A. niger* cells (10^4 dilution) spiked blood with and without bacterial/fungal enrichment.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood spiked (PC 2)	22.54	13.39
Blood spiked Enriched (Sample 2)	35.93	

Table 10.5b *A. niger* DNA qPCR results for ~2,000 *A. niger* cells (10^4 dilution) spiked blood with and without bacterial/fungal enrichment.

Sample ID	***A. niger* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood spiked (PC 2)	34.62	1.95
Blood spiked Enriched (Sample 2)	36.57*	

* Cq value suggests <10 cell (<100 cells in total input)

Conclusion: Using the protocol detailed in section 10.2, there is ~10^5 fold human DNA depletion with no loss of bacterial or fungal DNA.

Example 11 – verification of methodology for virus and phage enrichment

11.1: Protocol for viral enrichment in plasma

1. Whole blood was spiked with viral particles (max 200 µl per sample).
2. Samples were centrifuged at 20,000 xg for 5 min.
3. Supernatant was retained and used for the protocol (effectively working in plasma) after being aliquoted into equal volumes (max 200 µl).
4. 20 µl of PLC (0.8 mg) was added to each test sample and incubated at 37 °C for 15 min with shaking at 1000 RPM in a heat-block.
5. 200 µl of HL-SAN buffer (5 M NaCl and 100 mM $MgCl_2$) and 10 µl HL-SAN was added, incubated at 37 °C for 15 min with shaking at 1000 RPM in a heat-block.
6. 20 µl proteinase K was added to all samples and incubated at 65 °C for 10min.
7. Total nucleic acid was extracted using the MagnaPure (RTM) Compact automated machine using the DNA_bacteria_V3_2 protocol.
8. Host DNA/RNA depletion and viral DNA/RNA enrichment was determined via qPCR or RT-qPCR.

11.2: Protocol for viral enrichment in blood

1. Whole blood was spiked with viral particles (max 200 µl per sample).
2. 20 µl of PLC (0.8 mg) was added to each test sample and incubated at 37 °C for 15 min with shaking at 1000 RPM in a heat-block.
3. 200 µl of HL-SAN buffer (5 M NaCl and 100 mM $MgCl_2$) and 10 µl HL-SAN was added, incubated at 37 °C for 15 min with shaking at 1000 RPM in a heat-block.
4. Test samples were centrifuged at 20,000 xg for 5 min and the supernatant retained.
5. 20 µl proteinase K was added to all samples and incubated at 65 °C for 10 min.

6. Total nucleic acid was extracted using the MagnaPure (RTM) Compact automated machine using the DNA_Bacteria_V3_2 protocol.
7. Host DNA/RNA depletion and viral DNA/RNA enrichment was determined via qPCR or RT-qPCR.

Once the protocols described in sections 11.1 and 11.2 were established, samples were run in triplicate to access the reproducibility of the protocols (a second blood protocol was also tested at this stage which was the same as section 11.2 with an additional centrifugation step after step 4).

Results:
In total, each 200 µl blood sample was spiked with 10,000 IU HIV and 350 IU HBV. For this experiment, all three enrichment protocols were tested in triplicate (as previously described), After the viral enrichment protocols in blood there was consistently ~10^4 fold depletion in human DNA and human DNA was undetectable after enrichment when working in plasma (Tables 11.1a/b).

There was no loss of HBV viral DNA target in blood and plasma, although it should be noted that the number of HBV cells in the PCR reactions was ~35 and so Cq values were close to the limit of detection for the qPCR assay used (Tables 11.2a/b). With regards RNA viral targets, there was no loss of HIV in blood and plasma (Tables 11.3a/b).

Table 11.1a Human DNA qPCR results in triplicate for spiked blood with and without viral enrichment.

Sample ID	**Human qPCR assay (Cq)**	**Average Human (Cq)**	**Average ΔCq against PC**
PC blood spiked 1 (PC #1)	24.34	24.81	
PC blood spiked 2 (PC #2)	25.04		
PC blood spiked 3 (PC #3)	25.06		
Blood spiked Enriched 1 1 (T_1 #1)	37.32	37.54	12.73 (10^4)
Blood spiked Enriched 1 2 (T_1 #2)	37.68		
Blood spiked Enriched 1 3 (T_1 #3)	37.91		
Blood spiked Enriched 2 1 (T_2 #1)	37.94	38.16	13.35 (10^4)
Blood spiked Enriched 2 2 (T_2 #2)	38.64		
Blood spiked Enriched 2 3 (T_2 #3)	37.91		

Table 11.1b Human DNA qPCR results in triplicate for spiked plasma with and without viral enrichment.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC plasma spiked 1 (PC_SN #1)	34.64	Undetectable
PC plasma spiked 2 (PC_SN #2)	33.45	
PC plasma spiked 3 (PC_SN #3)	33.81	
Plasma spiked Enriched 1 (T_SN #1)	Undetectable	
Plasma spiked Enriched 2 (T_SN #2)	Undetectable	
Plasma spiked Enriched 3 (T_SN #3)	Undetectable	

Table 11.2a HBV DNA qPCR results in triplicate for spiked blood with and without viral enrichment.

Sample ID	**HBV qPCR assay (Cq)**	**Average HBV (Cq)**	**Average ΔCq against PC**
PC blood spiked 1 (PC #1)	38.02	37.9	
PC blood spiked 2 (PC #2)	36.95		
PC blood spiked 3 (PC #3)	38.76		
Blood spiked Enriched 1 1 (T_1 #1)	39.12	38	0.1
Blood spiked Enriched 1 2 (T_1 #2)	37.99		
Blood spiked Enriched 1 3 (T_1 #3)	36.81		
Blood spiked Enriched 2 1 (T_2 #1)	37.37	37.4	0.5
Blood spiked Enriched 2 2 (T_2 #2)	37.47		
Blood spiked Enriched 2 3 (T_2 #3)	Undetectable		

Table 11.2b HBV DNA qPCR results in triplicate for spiked plasma with and without viral enrichment.

Sample ID	**HBV qPCR assay (Cq)**	**Average ΔCq against PC**
PC plasma spiked 1 (PC_SN #1)	37.62	0.02
PC plasma spiked 2 (PC_SN #2)	36.92	
PC plasma spiked 3 (PC_SN #3)	36.95	
Plasma spiked Enriched 1 (T_SN #1)	37.22	
Plasma spiked Enriched 2 (T_SN #2)	Undetectable	
Plasma spiked Enriched 3 (T_SN #3)	Undetectable	

Table 11.3a HIV RNA RT-qPCR results in triplicate for spiked blood with and without viral enrichment.

Sample ID	**HIV qPCR assay (Cq)**	**Average HIV (Cq)**	**Average ΔCq against PC**
PC blood spiked 1 (PC #1)	32.76	33.2	
PC blood spiked 2 (PC #2)	33.60		
PC blood spiked 3 (PC #3)	33.14		
Blood spiked Enriched 1 1 (T_1 #1)	33.33	33.5	0.3
Blood spiked Enriched 1 2 (T_1 #2)	34.02		
Blood spiked Enriched 1 3 (T_1 #3)	33.08		
Blood spiked Enriched 2 1 (T_2 #1)	33.63	33.7	0.5
Blood spiked Enriched 2 2 (T_2 #2)	33.75		
Blood spiked Enriched 2 3 (T_2 #3)	33.65		

Table 11.3b HIV RNA RT-qPCR results in triplicate for spiked plasma with and without viral enrichment.

Sample ID	**HIV qPCR assay (Cq)**	**Average HIV (Cq)**	**Average ΔCq against PC**
PC plasma spiked 1 (PC_SN #1)	34.44	34.6	0.4
PC plasma spiked 2 (PC_SN #2)	33.75		
PC plasma spiked 3 (PC_SN #3)	35.64		
Plasma spiked Enriched 1 (T_SN #1)	35.66	34.9	
Plasma spiked Enriched 2 (T_SN #2)	35.00		
Plasma spiked Enriched 3 (T_SN #3)	34.03		

Next, for phage testing; in total, each 200 µl blood sample was spiked with either 10^4, 10^5, 10^6 or 10^7 phage. After the viral enrichment protocol in plasma (section 11.1) there was consistently ~10^3 fold depletion in human DNA with no loss of phage target (Tables 11.4a/b).

Table 11.4a Human DNA qPCR results for spiked blood with and without viral enrichment.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood spiked 10^4	28.01	11.99
Blood spiked Enriched 10^4	40	
PC blood spiked 10^5	28.68	11.32
Blood spiked Enriched 10^5	40	
PC blood spiked 10^6	28.72	11.28
Blood spiked Enriched 10^6	40	
PC blood spiked 10^7	28.43	9.52
Blood spiked Enriched 10^7	37.95	

Table 11.4b Phage DNA qPCR results for spiked blood with and without viral enrichment.

Sample ID	**Phage qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood spiked 10^4	31.10	2.36
Blood spiked Enriched 10^4	33.46	
PC blood spiked 10^5	28.14	0.65
Blood spiked Enriched 10^5	28.79	
PC blood spiked 10^6	23.96	0.34
Blood spiked Enriched 10^6	24.30	
PC blood spiked 10^7	20.66	0.07
Blood spiked Enriched 10^7	20.73	

Conclusion:

Here we described a complete protocol for the depletion of host DNA and enrichment of viral (both DNA and RNA) and phage (DNA). Two methods have been developed (one working in plasma; section 11.1, and one working in blood; section 11.2), and both provide human DNA depletion (~10^4 fold depletion in blood to undetectable in plasma). There is no loss of viral and phage DNA targets or viral HIV RNA target.

Example 12 – altering the cytolysin (blood samples)

For all testing with other cytolysins, 200 µl of blood was used following the protocol set out in section 10.2. The only alteration was the addition of different volumes/concentrations in place of PLC, i.e. no optimization was carried out.

Phospholipase D (PLD) from Streptomyces

PLD was purchased from Sigma-Aldrich (RTM) (P0065-25KU) with a stock made to 50KU/ml; varying volumes of PLD were used (2, 5 and 8 µl). Human DNA was depleted $<10^2$ fold (Table 12.1a) with no loss of bacterial or fungal targets (Tables 12.1b,c,d).

Table 12.1a Human DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using PLD.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	23.07	3.68
Blood 1 Enriched 2µl	26.75	
PC blood 2 Unenriched	23.05	5.76
Blood 2 Enriched 5µl	28.81	
PC blood 3 Unenriched	23.28	3.76
Blood 3 Enriched 8µl	27.04	

Table 12.1b *E. coli* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using PLD.

Sample ID	***E. coli* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	25.98	1.02
Blood 1 Enriched 2μl	24.96	
PC blood 2 Unenriched	25.29	0.2
Blood 2 Enriched 5μl	25.09	
PC blood 3 Unenriched	27.32	0.55
Blood 3 Enriched 8μl	26.77	

Table 12.1c *S. aureus* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using PLD.

Sample ID	***S aureus* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	24.01	0.69
Blood 1 Enriched 2μl	23.32	
PC blood 2 Unenriched	23.20	0.78
Blood 2 Enriched 5μl	23.98	
PC blood 3 Unenriched	23.06	0.33
Blood 3 Enriched 8μl	22.73	

Table 12.1d *C. albicans* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using PLD.

Sample ID	***C albicans* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	29.55	
Blood 1 Enriched 2μl	29.02	0.53
PC blood 2 Unenriched	29.58	
Blood 2 Enriched 5μl	29.93	0.35
PC blood 3 Unenriched	29.91	
Blood 3 Enriched 8μl	29.76	0.15

Sphingomyelinase from S. aureus

Sphingomyelinase was purchased from Sigma-Aldrich (RTM) (S8633-25UN) in solution and varying volumes were used (2, 5 and 8 μl). Human DNA was depleted $<10^2$ fold (Table 12.2a) with no loss of bacterial or fungal targets (Tables 12.2b,c,d).

Table 12.2a Human DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using sphingomyelinase.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	23.07	
Blood 1 Enriched 2μl	27.64	4.57
PC blood 2 Unenriched	23.05	
Blood 2 Enriched 5μl	30.58	7.53
PC blood 3 Unenriched	23.28	
Blood 3 Enriched 8μl	28.74	5.46

Table 12.2b *E. coli* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using sphingomyelinase.

Sample ID	***E. coli* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	25.98	
Blood 1 Enriched 2μl	24.67	1.61
PC blood 2 Unenriched	25.29	
Blood 2 Enriched 5μl	25.26	0.03
PC blood 3 Unenriched	27.32	
Blood 3 Enriched 8μl	26.67	0.65

Table 12.2c *S.aureus* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using sphingomyelinase.

Sample ID	***S aureus* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	24.01	
Blood 1 Enriched 2μl	22.65	1.36
PC blood 2 Unenriched	23.20	
Blood 2 Enriched 5μl	24.12	0.92
PC blood 3 Unenriched	23.06	
Blood 3 Enriched 8μl	22.66	0.73

Table 12.2d *C. albicans* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using sphingomyelinase.

Sample ID	***C albicans* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	29.55	
Blood 1 Enriched 2μl	27.82	1.73
PC blood 2 Unenriched	29.58	
Blood 2 Enriched 5μl	29.69	0.11
PC blood 3 Unenriched	29.91	
Blood 3 Enriched 8μl	28.99	0.92

Alpha hemolysin from S. aureus

Alpha hemolysin was purchased from Sigma-Aldrich (RTM) (H9395-5MG) and added at 0.01, 0.08 or 0.8mg in 20μl water. Human DNA was depleted $<10^2$ fold (Table 12.3a) with no loss of bacterial or fungal targets (Tables 12.3b,c,d).

Table 12.3a Human DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using alpha hemolysin.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	23.06	1.45
Blood 1 Enriched 0.01mg	24.48	
PC blood 2 Unenriched	23.28	4.38
Blood 2 Enriched 0.08mg	27.63	
PC blood 3 Unenriched	23.28	3.94
Blood 3 Enriched 0.8mg	27.22	

Table 12.3b *E. coli* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using alpha hemolysin.

Sample ID	***E. coli* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	26.96	
Blood 1 Enriched 0.01mg	26.37	0.59
PC blood 2 Unenriched	27.32	
Blood 2 Enriched 0.08mg	27.21	0.11
PC blood 3 Unenriched	27.32	
Blood 3 Enriched 0.8mg	27.34	0.02

Table 12.3c *S. aureus* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using alpha hemolysin.

Sample ID	***S aureus* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	22.71	
Blood 1 Enriched 0.01mg	23.12	0.41
PC blood 2 Unenriched	23.06	
Blood 2 Enriched 0.08mg	22.73	0.3
PC blood 3 Unenriched	23.06	
Blood 3 Enriched 0.8mg	23.12	0.06

Table 12.3d *C. albicans* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using alpha hemolysin.

Sample ID	***C albicans* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	28.57	0.41
Blood 1 Enriched 0.01mg	28.16	
PC blood 2 Unenriched	29.91	1.75
Blood 2 Enriched 0.08mg	28.16	
PC blood 3 Unenriched	29.91	0.1
Blood 3 Enriched 0.8mg	29.81	

Streptolysin O from S. pyogenes

Streptolysin O was purchased from Sigma-Aldrich (RTM) (S5265-25KU) and added at 0.08 or 0.8mg in 20μl water. Human DNA was depleted 10 fold (Table 12.4a) with no loss of bacterial or fungal targets (Tables 12.4b,c,d).

Table 12.4a Human DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using streptolysin O.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	23.28	2.87
Blood 1 Enriched 0.08mg	26.15	
PC blood 2 Unenriched	23.28	2.9
Blood 2 Enriched 0.8mg	26.18	

Table 12.4b *E. coli* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using streptolysin O.

Sample ID	***E. coli* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	27.32	0.08
Blood 1 Enriched 0.08mg	27.24	
PC blood 2 Unenriched	27.32	0.32
Blood 2 Enriched 0.8mg	27.00	

Table 12.4c *S. aureus* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using streptolysin O.

Sample ID	***S aureus* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	23.06	0.38
Blood 1 Enriched 0.08mg	22.68	
PC blood 2 Unenriched	23.06	0.28
Blood 2 Enriched 0.8mg	22.78	

Table 12.4d *C. albicans* DNA qPCR results for spiked blood with and without bacterial/fungal enrichment using streptolysin O.

Sample ID	***C albicans* qPCR assay (Cq)**	**Average ΔCq against PC**
PC blood 1 Unenriched	29.91	0.11
Blood 1 Enriched 0.08mg	29.80	
PC blood 2 Unenriched	29.91	1.66
Blood 2 Enriched 0.8mg	28.25	

Conclusion:

All cytolysins tested showed effective human DNA depletion and no bacterial or fungal DNA loss.

Example 13 – verification of methodology for other clinical sample types

Using the established protocol detailed in section 10.2, the initial 200µl of blood was replaced with 200µl of sputum, sonicated tissue or urine to verify the depletion method works effectively in other clinical sample types.

Clinical sputum samples

Human DNA was depleted up to 10^4 fold (Table 13.1a) with no loss of bacteria (Tables 13.1b/c) in clinical sputum samples.

Table 13.1a Human DNA qPCR results for clinical sputum with and without fungal/bacterial enrichment.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC sputum 1 Unenriched	19.81	8.08
Sputum 1 Enriched	27.89	
PC sputum 2 Unenriched	22.10	12.31
Sputum 2 Enriched	34.41	

Table 13.1b 16S rRNA gene fragment (V3-V4) qPCR results for clinical sputum with and without fungal/bacterial enrichment.

Sample ID	**16S rRNA qPCR assay (Cq)**	**Average ΔCq against PC**
PC sputum 1 Unenriched	17.96	3.97
Sputum 1 Enriched	13.93	
PC sputum 2 Unenriched	15.89	0.23
Sputum 2 Enriched	15.66	

Table 13.1c *S. aureus* DNA qPCR results for clinical sputum with and without fungal/bacterial enrichment.

Sample ID	***S aureus* qPCR assay (Cq)**	**Average ΔCq against PC**
PC sputum 2 Unenriched (suspected *S aureus*)	22.29	0.87
Sputum 2 Enriched (suspected *S aureus*)	22.96	

Peri-prosthetic tissue samples

Peri-prosthetic tissue sample biopsies spiked with *Staphylococcus epidermidis* cells (15TB0821), with $<10^5$ fold human DNA depletion (Table 13.2a) and no loss of bacterial target (Table 13.2b).

Table 13.2a Human DNA qPCR results for per-prosthetic spiked tissue samples with and without fungal/bacterial enrichment.

Sample ID	**Human qPCR assay (Cq)**	**ΔCq against PC**
PC tissue Unenriched	23.09	
Tissue 100 cells Enriched	37.87	14.78
Tissue 1000 cells Enriched	37.90	14.81
Tissue 10,000 cells Enriched	38.37	15.28

Table 13.2b *S. epidermidis* DNA qPCR results for peri-prosthetic spiked tissue samples with and without fungal/bacterial enrichment.

Sample ID	***S epidermidis* qPCR assay (Cq)**	**ΔCq against PC**
PC tissue 100 cells Unenriched	37.25	1.99
Tissue 100 cells Enriched	35.26	

Clinical urine samples

Human DNA was depleted $<10^4$ fold (Table 13.3a) with no loss of bacteria (Tables 13.3b/c) in clinical sputum samples.

Table 13.3a Human DNA qPCR results for clinical urine with and without fungal/bacterial enrichment.

Sample ID	**Human qPCR assay (Cq)**	**Average ΔCq against PC**
PC urine 1 Unenriched	24.01	10.99
Urine 1 Enriched	35	
PC urine 2 Unenriched	31.26	3.74
Urine 2 Enriched	35	
PC urine 3 Unenriched	24.98	10.32
Urine 3 Enriched	35	

Table 13.3b 16S rRNA gene fragment (V3-V4) qPCR results for clinical urine with and without fungal/bacterial enrichment.

Sample ID	**16S rRNA qPCR assay (Cq)**	**Average ΔCq against PC**
PC urine 1 Unenriched	13.60	0.32
Urine 1 Enriched	13.92	
PC urine 2 Unenriched	14.16	1.34
Urine 2 Enriched	15.50	
PC urine 3 Unenriched	10.90	0.36
Urine 3 Enriched	10.54	

Table 13.3c *E. coli* DNA qPCR results for clinical urine with and without fungal/bacterial enrichment.

Sample ID	***E coli* qPCR assay (Cq)**	**Average ΔCq against PC**
PC urine 2 Unenriched (suspected *E. coli*)	19.46	1.27
Urine 2 Enriched (suspected *E. coli*)	20.73	

Conclusion: All clinical sample types tested showed host DNA depletion with no loss of bacterial DNA.

Example 14 – host RNA depletion (HL-SAN RNase activity)

There was >10^2 fold host RNA depletion using the viral blood protocol (section 11.2 and Table 14.1a). Using the viral plasma protocol detailed in section 11.1, showed >10^2 fold depletion of host RNA (Table 14.1b and 14.2a) with no loss of HIV target (Table 14.2b).

Table 14.1a Human RNA RT-qPCR results in duplicate for non-spiked blood with and without viral enrichment (host RNA depletion).

Sample ID	**Human RNA qPCR assay (Cq)**	**Average ΔCq against PC**
Unenriched blood non-spiked 1	24.72	8.53
Enriched blood non-spiked 1	33.25	
Unenriched blood non-spiked 2	32.49	5.9
Enriched blood non-spiked 2	38.39	

Table 14.1b Human RNA RT-qPCR results in duplicate for non-spiked plasma with and without viral enrichment (host RNA depletion).

Sample ID	**Human RNA qPCR assay (Cq)**	**Average ΔCq against PC**
Unenriched plasma unspiked 1	36.26	8.74
Enriched plasma Unspiked 1	Undetectable	
Unenriched plasma unspiked 2	34.44	10.56
Enriched plasma Unspiked 2	Undetectable	

Table 14.2a Human RNA RT-qPCR results in duplicate for spiked plasma with and without viral enrichment (host RNA depletion).

Sample ID	**Human RNA qPCR assay (Cq)**	**Average ΔCq against PC**
Unenriched plasma spiked 1	36.35	8.65
Enriched plasma spiked 1	Undetectable	
Unenriched plasma spiked 2	30.87	3.28
Enriched plasma spiked 2	34.15	

Table 14.2b HIV RNA RT-qPCR results in duplicate for spiked plasma with and without viral enrichment (host RNA depletion).

Sample ID	**HIV RNA qPCR assay (Cq)**	**Average ΔCq against PC**
Unenriched plasma spiked 1	35.67	0.36
Enriched plasma spiked 1	36.03	
Unenriched plasma spiked 2	31.95	0.78
Enriched plasma spiked 2	32.73	

Conclusion:

Due to the variability of starting host RNA, it was established that HL-SAN RNase activity provided the greatest host RNA depletion with no loss of viral RNA target and therefore no alterations to the enrichment protocol (detailed in section 11.1) was necessary. Human RNA was typically not detectable in plasma post depletion using this method.

Example 15 – removal of human DNA without nuclease

Propidium monoazide (PMA) to remove human DNA

An altered method from that described in section 10.2 was needed to enable the activation of PMA by light. After PLC treatment, the sample was centrifuged at 12,000xg for 5min and resuspended in 1.5ml of PBS. PMA was added at a final concentration of 50μM and incubated in the dark with occasional shaking for 5min. The sample was then placed in a photolysis device for 15min exposure to blue light, the protocol in section 10.2 was then followed from step 6. Human DNA was depleted $<10^2$ fold (Table 15.1) with no loss of bacterial target DNA (Table 15.2).

Table 15.1 Human DNA qPCR results for spiked blood samples with and without fungal/bacterial enrichment using PMA to remove human DNA.

Sample ID	**Human qPCR assay (Cq)**	**ΔCq against PC**
PC blood Unenriched	22.90	
Blood PMA #1 Enriched	27.62	4.72
Blood PMA #2 Enriched	28.47	5.57

Table 15.2 *E. coli* DNA qPCR results for spiked blood samples with and without fungal/bacterial enrichment using PMA to remove human DNA.

Sample ID	***E. coli* qPCR assay (Cq)**	**ΔCq against PC**
PC blood Unenriched	20.65	
Blood PMA #1 Enriched	21.13	0.48
Blood PMA #2 Enriched	21.30	0.65

Conclusion:
Using PMA to remove human DNA after PLC treatment showed human DNA depletion and no loss of bacterial target DNA

Example 16 – revised protocol for 1ml blood sample

1. PLC was added (4 mg/100 µl) to the blood sample (1 ml in a 5 ml bijou tube), vortexed and incubated at 37 °C for 3 min in a water bath followed by 38 °C for 20 min with slow mixing at 15rpm in a hulamixer (RTM).
2. Sample was transferred to a 2 ml tube and 500 µl of HL-SAN buffer (5M NaCl and 100mM $MgCl_2$) was added and incubated 37 °C for 15 min in a heatblock at 1000 RPM.
3. Cells were pelleted by centrifugation at 8,000 xg for 5 min.
4. The cell pellet was resuspended in 200 µl PBS
5. HL-SAN buffer was added at a 1:1 volume ratio (200 µl) with 10 µl HL-SAN DNase, vortexed and incubated at 37 °C for 15 min at 1000 RPM in a heat block.
6. PBS was added to a total volume of 2 ml (1.5 ml).
7. Cells were pelleted by centrifugation at 12,000 xg for 10 min and the supernatant was discarded.
8. The cell pellet was resuspended in 1.5 ml PBS.
9. Cells were pelleted again by centrifugation at 12,000 xg for 10 min and the supernatant was discarded.
10. To any test samples; 350 µl bacterial lysis buffer, 20 µl enzyme cocktail (6 µl mutanolysin 25 ku/ml, 5 µl lysozyme 10 mg/ml, 4 µl lyticase 10 ku/ml, 3 µl lysostaphin 4 ku/ml, 2 µl chitinase 50 u/ml) and 5 µl RNase A was added.
11. All samples were incubated at 37 °C for 15 min at 1000 RPM in a heat block.
12. To all samples, 20 µl proteinase K was added and incubated at 65 °C for 10 min in a heat block.
13. Total nucleic acid was extracted using the MagnaPure (RTM) Compact automated machine using the DNA_bacteria_V3_2 protocol.

Changes to the 200μl protocol in section 10.2 to increase the starting volume to 1ml are described above. This gave >10^6 fold depletion of human DNA (Table 16.1a) with no loss of bacterial or fungal target DNA (Tables 16.1b,c,d).

Table 16.1a Human DNA qPCR results for 1ml spiked blood with and without fungal/bacterial enrichment.

Sample ID	**Human qPCR assay (Cq)**	**ΔCq against PC**
PC blood Unenriched	23.21	21.79
Blood 1ml Enriched	-	

Table 16.1b *E. coli* DNA qPCR results for 1ml spiked blood with and without fungal/bacterial enrichment.

Sample ID	***E coli* qPCR assay (Cq)**	**ΔCq against PC**
PC blood Unenriched	31.77	2.1
Blood 1ml Enriched	29.67	

Table 16.1c *S. aureus* DNA qPCR results for 1ml spiked blood with and without fungal/bacterial enrichment.

Sample ID	***S aureus* qPCR assay (Cq)**	**ΔCq against PC**
PC blood Unenriched	37.63	3.72
Blood 1ml Enriched	33.91	

Table 16.1d *C. albicans* DNA qPCR results for 1ml spiked blood with and without fungal/bacterial enrichment.

Sample ID	*C albicans* qPCR assay (Cq)	ΔCq against PC
PC blood Unenriched	32.84	2.6
Blood 1ml Enriched	30.24	

Conclusion:
A slightly altered method was developed to enable fungal enrichment when using 1ml blood and this resulted in ~10^6 fold depletion of human DNA with no loss of bacteria or fungi target DNA. Greater sample volumes (>1ml) could also be used.

This method can seemingly be used on any sample type where the host cells have a phospholipid membrane e.g. clinical samples (infectious disease diagnosis) or animal samples (food safety and veterinary medicine/diagnosis).

Example 17 – NGS after depletion method

Additional methodology

After the depletion protocol detailed in section 10.2, 4µl DNA was processed using REPLI-g single cell kit (Qiagen 150343) for whole genome amplification (WGA). The manufacturer's instructions were followed with the amplification time reduced to 1hr 30min. WGA sample (17µl) was debranched using T7 endonuclease I (NEB M0302S) according to the manufacturer's instructions. MinION library preparation used the rapid low input by PCR barcoding kit (ONT SQK-RLB001) as per the manufacturer's guideline with the following alterations:

- 2.5µl FRM with 7.5µl template DNA (~140ng)
- 40µl nuclease-free water, 50µl LongAmp Taq 2x, 2µl RLB
- PCR: [95°C 3min]x1, [95°C 15s, 56°C 15s, 65°C 4min]x20, [65°C 4min]x20, [65°C 6min]x1

The SpotON R9.4 MinION flowcell was prepared and loaded according to the manufacturer's instructions.

Bioinformatics data analysis: reads were aligned to the *C. albicans* reference genome (SC5314 NC_003977.2) using minimap2. Genome coverage and number of

aligned reads were identified using samtools and qualimap. Percentage reads are given as those which aligned to the reference genome out of the total number of reads.

Results

~300cfu/ml *Candida albicans* at ~15Mb genome = 4.5pg of DNA
Average concentration of human DNA in 1ml blood = 33µg of DNA
Therefore before enrichment the ratio of human:Candida DNA is ~10^7:1

From the sequencing data presented below, *C. albicans* reads are 1% of the total (1.3x genome coverage) therefore assuming all other reads are human = 100:1 (human:*Candida*)

Ratio of human:*Candida* DNA before depletion = 10^7:1
Ratio of human:*Candida* DNA after depletion = 100:1
This is the equivalent of 10^5 fold depletion.

Table 17 *C. albicans* genome alignment from single-plex MinION run (input ~300cfu/ml).

Sequencing time	**Total number of reads**	**Aligned reads to known pathogen**	**Pathogen genome coverage**	**Percentage of known pathogen reads (%)**
14hrs	1.2 million	12,422	1.3	1

C. albicans genome coverage plot after *C. albicans* single-plex MinION sequencing is shown in Figure 10.

CLAIMS

1. A method for depleting host nucleic acid in a biological sample, said sample having been previously obtained from an animal host, said method comprising the steps of:
    i. adding a cytolysin, or an active variant thereof, to said sample; and
    ii. carrying-out a process to physically deplete nucleic acid released from host cells within said sample or otherwise render such nucleic acid unidentifiable.

2. A method according to claim 1 wherein step (b) comprises adding a nuclease to said sample.

3. A method according to claim 1 or claim 2, further comprising the step of extracting remaining nucleic acid from the sample.

4. A method according to claim 3, further comprising the step of subjecting the extracted nucleic acid to a purification process.

5. A method according to claim 3 or claim 4, further comprising the step of amplifying the extracted nucleic acid.

6. A method according to any one of claims 3 to 5, further comprising the step of conducting a nucleic acid amplification test on the extracted nucleic acid or, preferably, conducting a sequencing process on the extracted nucleic acid.

7. A method according to any one of the preceding claims, wherein the cytolysin is a phospholipase.

8. A method according to claim 7 wherein the phospholipase is a phospholipase C (PLC).

9. A method according to claim 8 wherein the PLC is a bacterial PLC.

10. A method according to claim 9 wherein the bacterial PLC is a Group 1 PLC.

11. A method according to claim 10 wherein the Group 1 PLC is PLC from *Clostridium perfringens*.

12. A method according to any one of the preceding claims wherein the biological sample is a blood sample.

13. A method according to any one of the preceding claims that results in at least a 10 fold, preferably at least a 10^2 fold, preferably at least a 10^3 fold, preferably at least a 10^4 fold, most preferably at least a 10^5 fold depletion of host DNA originally contained within the sample.

14. A kit comprising i) a cytolysin, or an active variant thereof, and ii) means to physically deplete free nucleic acid within a biological sample or otherwise render such nucleic acid unidentifiable.

15. A kit according to claim 14, wherein said cytolysin is as defined within any one of claims 7 to 11.

16. A kit according to claim 14 or claim 15, wherein said means comprises a nuclease.

ABSTRACT

Provided is a method for depleting host nucleic acid in a biological sample, said sample having been previously obtained from an animal host, said method comprising the steps of (a) adding a cytolysin, or an active variant thereof, to said sample; and (b) carrying-out a process to physically deplete nucleic acid released from host cells within said sample or otherwise render such nucleic acid unidentifiable.

7.2: Papers

Paper I

Schmidt, K., S. Mwaigwisya, L. C. Crossman, M. Doumith, D. Munroe, C. Pires, A. M. Khan, N. Woodford, N. J. Saunders, J. Wain, J. O'Grady and D. M. Livermore (2017). "Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing." J Antimicrob Chemother 72(1): 104-114.

Journal of Antimicrobial Chemotherapy Advance Access published September 25, 2016

J Antimicrob Chemother
doi:10.1093/jac/dkw397

Journal of Antimicrobial Chemotherapy

Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing

K. Schmidt[1], S. Mwaigwisya[1], L. C. Crossman[2], M. Doumith[3], D. Munroe[4], C. Pires[5], A. M. Khan[5], N. Woodford[3], N. J. Saunders[5], J. Wain[1], J. O'Grady[1]* and D. M. Livermore[1,3]

[1]Norwich Medical School, University of East Anglia, Norwich, UK; [2]SequenceAnalysis.co.uk, Norwich Research Park, Norwich, UK; [3]AMRHAI Reference Unit, National Infection Service, Public Health England, London, UK; [4]Microbiology Department, Norfolk and Norwich University Hospital, Norwich, UK; [5]Brunel University London, Uxbridge, UK

*Corresponding author. Tel: +44-(0)1603-59-7005; E-mail: justin.ogrady@uea.ac.uk

Received 9 May 2016; returned 5 July 2016; revised 9 August 2016; accepted 21 August 2016

Objectives: The introduction of metagenomic sequencing to diagnostic microbiology has been hampered by slowness, cost and complexity. We explored whether MinION nanopore sequencing could accelerate diagnosis and resistance profiling, using complicated urinary tract infections as an exemplar.

Methods: Bacterial DNA was enriched from clinical urines ($n=10$) and from healthy urines 'spiked' with multi-resistant *Escherichia coli* ($n=5$), then sequenced by MinION. Sequences were analysed using external databases and bioinformatic pipelines or, ultimately, using integrated real-time analysis applications. Results were compared with Illumina data and resistance phenotypes.

Results: MinION correctly identified pathogens without culture and, among 55 acquired resistance genes detected in the cultivated bacteria by Illumina sequencing, 51 were found by MinION sequencing directly from the urines; with three of the four failures in an early run with low genome coverage. Resistance-conferring mutations and allelic variants were not reliably identified.

Conclusions: MinION sequencing comprehensively identified pathogens and acquired resistance genes from urine in a timeframe similar to PCR (4 h from sample to result). Bioinformatic pipeline optimization is needed to better detect resistances conferred by point mutations. Metagenomic-sequencing-based diagnosis will enable clinicians to adjust antimicrobial therapy before the second dose of a typical (i.e. every 8 h) antibiotic.

Introduction

The UK Government's O'Neill Commission, reviewing the threat of antibiotic resistance,[1] stresses the potential of rapid diagnostics to improve both treatment and antibiotic stewardship. Reducing the time needed to obtain a microbiological diagnosis shortens the duration of broad empirical therapy and its selective pressures.

PCR can detect pathogens and resistance genes in specimens without culture, but cannot cover the diversity of organisms and resistance determinants potentially present. Metagenomic sequencing could deliver this comprehensiveness,[2–7] but slow turnaround, cost and complexity have impeded introduction into clinical microbiology.

Oxford Nanopore Technologies' (ONT) MinION[8] is the first technology potentially able to deliver sequencing data from clinical samples in a timeframe allowing early de-escalation and refinement of antimicrobial treatment. We examined its applicability to investigation of urinary tract infections (UTIs). These account for over 8 million physician visits per annum in the USA.[9] Most are trivial, but, in severe cases, infection may ascend to the kidneys, with overspill to the bloodstream precipitating bacteraemia and urosepsis. Complicated UTIs are a growing cause of hospitalization, mostly of elderly patients,[10] and 35 676 *Escherichia coli* bloodstream infections were recorded in England in 2014–15,[11] over 60% with a urinary origin. There is growing resistance, particularly in severe and bacteraemic infections, to fluoroquinolones, cephalosporins and lactamase-inhibitor combinations, driving use of previously reserved carbapenems, even as 'empirical' therapy. With carbapenemases now proliferating, and few alternative therapies in reserve, escalating empiricism becomes increasingly untenable, underscoring the desirability of moving to early targeted therapy, guided by diagnostics.

Materials and methods

Urines

Ten heavily infected (>10^7 cfu/mL) clinical urines (CU1–10) from patients at the Norfolk and Norwich University Hospital (NNUH) were tested.

Paper II

Solomon Mwaigwisya, Rasha Assad M Assiri and Justin O'Grady (2015) "Emerging commercial molecular tests for the diagnosis of bloodstream infection" Expert Rev Mol Diagn. May 2015, Vol. 15, No. 5, Pages 681-692 (doi:10.1586/14737159.2015.1029459)

EXPERT REVIEWS

Emerging commercial molecular tests for the diagnosis of bloodstream infection

Expert Rev. Mol. Diagn. Early online, 1–12 (2015)

Solomon Mwaigwisya[1], Rasha Assad M Assiri[1,2] and Justin O'Grady*[1]

[1]*Norwich Medical School, University of East Anglia, Norwich, UK*
[2]*Prince Abdullah Bin Khalid Celiac Disease Research Chair, Division of Microbiology, King Khalid University Hospital, King Saud University, Riyadh, Kingdom of Saudi Arabia*
**Author for correspondence: Tel.: +44 016 0359 7005 justin.ogrady@uea.ac.uk*

Bloodstream infection (BSI) by microorganisms can lead to sepsis. This condition has a high mortality rate, which rises significantly with delays in initiation of appropriate antimicrobial treatment. Current culture methods for diagnosing BSI have long turnaround times and poor clinical sensitivity. While clinicians wait for culture diagnosis, patients are treated empirically, which can result in inappropriate treatment, undesirable side effects and contribute to drug resistance development. Molecular diagnostics assays that target pathogen DNA can identify pathogens and resistance markers within hours. Early diagnosis improves antibiotic stewardship and is associated with favorable clinical outcomes. Nonetheless, limitations of current molecular diagnostic methods are substantial. This article reviews recent commercially available molecular methods that use pathogen DNA to diagnose BSI, either by testing positive blood cultures or directly testing patient blood. We critically assess these tests and their application in clinical microbiology. A view of future directions in BSI diagnosis is also provided.

KEYWORDS: antibiotic resistance • blood culture • blood stream infection • MALDI-TOF • molecular diagnostics • next-generation sequencing • NGS • PCR • PCR/ESI-MS • sepsis

The presence of infectious organism in the blood, including bacteria (bacteremia), viruses (viremia), or fungi (fungemia) is generalized as blood stream infection (BSI). The prevalence of BSI is on the increase owing to factors, such as immunosuppression treatment, increased frequency of invasive procedures, aging population and rise of multidrug-resistant organisms in hospitals and care homes [1]. BSIs may be associated with infections, such as endocarditis, and procedures, such as central venous catheter, colonoscopy or dental extraction. BSIs may also be caused by secondary focal infections, such as pneumonia, urinary tract infections or abscesses [2]. BSI often develops to sepsis (or severe sepsis), which can be defined as a systemic response to infection with the presence of some degree of organ dysfunction. It can be categorized as septic shock when complicated by either hypotension that is refractory to fluid resuscitation or by hyperlactatemia [3–6]. In Europe and USA, combined sepsis causes more than 400,000 deaths every year, costing an estimated US$17 billion in the USA [7–10].

The high mortality rate associated with sepsis is influenced by a combination of factors, including the characteristics of the infecting pathogen and the host response to infection. Studies suggest that every hour of delay in administration of effective antimicrobial therapy in patients with septic shock is associated with a decrease in survival rate [11,12]. The Surviving Sepsis Campaign recommends confirming sepsis by objective evidence of infection, which is achieved by identifying live pathogens by blood culture methods. Culture analysis, including drug susceptibility testing, takes at least 48 h for positive samples and 5 days to identify specimens as negative [3,11].

Because of the temporal separation between initial clinical suspicion of sepsis and laboratory confirmation of BSI, a 'safety first' strategy of using potent broad-spectrum antibiotics within 1 h of onset of sepsis is employed, aimed at covering the most probable

  ISSN 1473-7159

Paper III

Ashton PM, Nair S, Dallman, T, Rubino, S, Rabsch, W, Mwaigwisya S, Wain J. O'Grady, J. (2014) MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. Nat Biotechnol. ISSN 1087-0156 doi: 10.1038/nbt.3103

nature biotechnology

MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island

Philip M Ashton[1,6], Satheesh Nair[1,6], Tim Dallman[1], Salvatore Rubino[2,3], Wolfgang Rabsch[4], Solomon Mwaigwisya[5], John Wain[5] & Justin O'Grady[5]

Short-read, high-throughput sequencing technology cannot identify the chromosomal position of repetitive insertion sequences that typically flank horizontally acquired genes such as bacterial virulence genes and antibiotic resistance genes. The MinION nanopore sequencer can produce long sequencing reads on a device similar in size to a USB memory stick. Here we apply a MinION sequencer to resolve the structure and chromosomal insertion site of a composite antibiotic resistance island in *Salmonella* Typhi Haplotype 58. Nanopore sequencing data from a single 18-h run was used to create a scaffold for an assembly generated from short-read Illumina data. Our results demonstrate the potential of the MinION device in clinical laboratories to fully characterize the epidemic spread of bacterial pathogens.

Short read, high-throughput, next-generation sequencing (NGS) technology has transformed our understanding of microbiology and is poised to become an integral tool in epidemiology[1]. Although the utility of whole genome sequencing (WGS) for public health infection control is clear, adoption in clinical microbiology laboratories has been limited[2]. This is partly because short-read technologies cannot unambiguously assemble repetitive elements that are longer than sequencing read-length into a single contig. This assembly problem generates multiple contigs and leaves gaps in whole genome assemblies. It is particularly difficult to correctly assemble regions in which genes have been acquired by horizontal gene transfer, such as resistance and pathogenicity islands, and prophage[3], owing to their inherent repetitive nature or the flanking of these elements by repetitive insertion sequences. Analyzing these regions is essential for determining key characteristics such as antibiotic resistance profiles and for identifying highly pathogenic variants of many bacterial species[4]. Currently gap closure requires extensive, post-sequencing, laboratory-based analysis, which can take several months and makes the results irrelevant for clinical diagnostics and for guiding public health interventions.

Sequencing technology that generates long reads, capable of spanning repetitive sequences and closing gaps in short read data, is commercially available (Pacific Biosystems PacBio RS II) but has significant capital cost outlay, a very large laboratory footprint and is technically demanding. DNA sequencing using nanopore technology is an alternative method for producing long-read sequence data but has been a specialized research tool until very recently[5] and is not, as of December 2014, available commercially. The recent distribution of the MinION by Oxford Nanopore Technologies Ltd. in an early-access program (named the MinION Access Programme) has made it possible to evaluate the utility of long-read sequencing using a device that resembles a large USB memory stick.

There were an estimated 26.9 million cases of typhoid fever in 2010 (ref. 6) with a very high proportion of those cases in urban slums[7]. A recent emergence of a globally distributed multidrug-resistant (MDR) *Salmonella enterica* serovar Typhi (*S.* Typhi) haplotype, H58, has been observed contributing to a reduction in genetic diversity of extant *S.* Typhi[8–12]. At Public Health England, Salmonella Reference Service, in Colindale, UK, we have observed a similar increase in isolates of MDR *S.* Typhi phage type E9 variant from patients with a travel history to the Indian subcontinent. The routine adoption of WGS technologies to identify and type *Salmonella* isolates here allowed these to be characterized as H58 harboring multiple resistance elements including, *strA*, *strB*, *sulI*, *sulII*, *dfrA7* and bla_{TEM-1} (ref. 13) encoded on Tn10 and Tn9. The specific resistance plasmid (plasmid PST6 (incHI1)) typical of H58 isolates was, however, not present, raising the possibility that an antibiotic resistance island has integrated into the H58 chromosome.

Here, we report a hybrid assembly of combined MinION and Illumina HiSeq data to identify the structure and insertion site of a chromosomal antibiotic resistance island in *S.* Typhi H58, which, despite many "whole genome" sequencing projects[14], has not been previously characterized.

Two *S.* Typhi H58 strains (H125160566 and 08-0446) were sequenced using the Illumina HiSeq, and SNP typing was used to confirm haplotype[12]. *De novo* assembly of Illumina sequence for strain 08-0446 (ENA accession number ERR668456) resulted in 143 contigs, an N50 (a statistical measure of average length of a set of sequences) of 124 kbp and average genome coverage of 78× (374 million bases of >Q30 data). *De novo* assembly of strain H125160566 (ENA accession number ERR668457) resulted in 86 contigs, an N50 of 154 kbp and

[1]Gastrointestinal Bacteria Reference Unit, Public Health England, Colindale, London, UK. [2]Department of Biomedical Sciences, University of Sassari, Sassari, Italy. [3]Department of Infection and Immunity, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia. [4]National Reference Centre for Salmonellae and other Enterics, Robert Koch Institute, Wernigerode, Germany. [5]Norwich Medical School, University of East Anglia, Norwich, UK. [6]These authors contributed equally to this work. Correspondence should be addressed to J.O'G. (justin.ogrady@uea.ac.uk) or T.D. (tim.dallman@phe.gov.uk).

Received 7 September; accepted 21 November; published online 8 December 2014; doi:10.1038/nbt.3103

Paper IV

Ip, C. L., Loose, M., Tyson, J. R., de Cesare, M., Brown, B. L., Jain, M., Leggett, R. M., Eccles, D. A., Zalunin, V., Urban, J. M., Piazza, P., Bowden, R. J., Paten, B., Mwaigwisya, S., Batty, E. M., Simpson, J. T., Snutch, T. P., Birney, E., Buck, D., Goodwin, S., Jansen, H. J., O'Grady, J., Olsen, H. E. (2015) MinION Analysis and Reference Consortium: Phase 1 data release and analysis in F1000Research 4.

F1000Research F1000Research 2015, 4:1075 Last updated: 19 JAN 2016

RESEARCH ARTICLE

MinION Analysis and Reference Consortium: Phase 1 data release and analysis [version 1; referees: 2 approved]

Camilla L.C. Ip[1*], Matthew Loose[2*], John R. Tyson[3*], Mariateresa de Cesare[1*], Bonnie L. Brown[4*], Miten Jain[5*], Richard M. Leggett[6*], David A. Eccles[7], Vadim Zalunin[8], John M. Urban[9], Paolo Piazza[1], Rory J. Bowden[1], Benedict Paten[5], Solomon Mwaigwisya[10], Elizabeth M. Batty[1], Jared T. Simpson[11], Terrance P. Snutch[3], Ewan Birney[8*], David Buck[1*], Sara Goodwin[12*], Hans J. Jansen[13*], Justin O'Grady[10*], Hugh E. Olsen[5*], MinION Analysis and Reference Consortium

[1]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
[2]School of Life Sciences, Queens Medical Centre, University of Nottingham, Nottingham, UK
[3]Michael Smith Laboratories and Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, Canada
[4]Virginia Commonwealth University, Richmond, VA, USA
[5]University of California, Santa Cruz, Santa Cruz, CA, USA
[6]The Genome Analysis Centre, Norwich Research Park, Norwich, UK
[7]Malaghan Institute of Medical Research, Wellington, New Zealand
[8]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, UK
[9]Division of Biology and Medicine, Brown University, Providence, RI, USA
[10]Norwich Medical School, University of East Anglia, Norwich, UK
[11]Informatics and Biocomputing, Ontario Institute for Cancer Research, ON, Canada
[12]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA
[13]ZF-screens B.V., Leiden, Netherlands

[*] Equal contributors

v1 **First published:** 15 Oct 2015, **4**:1075 (doi: 10.12688/f1000research.7201.1)
Latest published: 15 Oct 2015, **4**:1075 (doi: 10.12688/f1000research.7201.1)

Abstract

The advent of a miniaturized DNA sequencing device with a high-throughput contextual sequencing capability embodies the next generation of large scale sequencing tools. The MinION™ Access Programme (MAP) was initiated by Oxford Nanopore Technologies™ in April 2014, giving public access to their USB-attached miniature sequencing device. The MinION Analysis and Reference Consortium (MARC) was formed by a subset of MAP participants, with the aim of evaluating and providing standard protocols and reference data to the community. Envisaged as a multi-phased project, this study provides the global community with the Phase 1 data from MARC, where the reproducibility of the performance of the MinION was evaluated at multiple sites. Five laboratories on two continents generated data using a control strain of *Escherichia coli* K-12, preparing and sequencing samples according to a revised ONT protocol. Here, we provide the details of the protocol used, along with a preliminary analysis of the characteristics of typical runs including the

Open Peer Review

Referee Status: ✓ ✓

	Invited Referees 1	2
version 1 published 15 Oct 2015	✓ report	✓ report

1 **Michael Quail**, Wellcome Trust Sanger Institute UK, **Louise Aigrain**, Wellcome Trust Sanger Institute UK

Paper IV

Nair, S., Ashton, P., Doumith, M., Connell, S., Painset, A., Mwaigwisya, S., Langridge, G., de Pinna, E., Godbole, G., Day, M. (2016) WGS for surveillance of antimicrobial resistance: A pilot study to detect the prevalence and mechanism of resistance to azithromycin in a UK population of non-typhoidal Salmonella in Journal of Antimicrobial Chemotherapy 71.pp. 3400-3408

Journal of Antimicrobial Chemotherapy Advance Access published September 1, 2016

J Antimicrob Chemother
doi:10.1093/jac/dkw318

Journal of Antimicrobial Chemotherapy

WGS for surveillance of antimicrobial resistance: a pilot study to detect the prevalence and mechanism of resistance to azithromycin in a UK population of non-typhoidal *Salmonella*

Satheesh Nair[1]*, Philip Ashton[1], Michel Doumith[1], Steve Connell[1], Anais Painset[1], Solomon Mwaigwisya[2], Gemma Langridge[2], Elizabeth de Pinna[1], Gauri Godbole[1] and Martin Day[1]

[1]Gastrointestinal Bacteria Reference Unit, Public Health England, London NW9 5EQ, UK; [2]Norwich Medical School, University of East Anglia, Norwich NR4 7UQ, UK

*Corresponding author. Gastrointestinal Bacteria Reference Unit, Public Health England, 61 Colindale Ave, London NW9 5EQ, UK. E-mail: satheesh.nair@phe.gov.uk

Received 29 January 2016; returned 11 March 2016; revised 27 May 2016; accepted 8 July 2016

Objectives: WGS and phenotypic methods were used to determine the prevalence of azithromycin resistance in *Salmonella enterica* isolates from the UK and to identify the underlying mechanisms of resistance.

Methods: WGS by Illumina HiSeq was carried out on 683 *Salmonella* spp. isolates. Known genes associated with azithromycin resistance were detected by WGS using a mapping-based approach. Macrolide resistance determinants were identified and the genomic context of these elements was assessed by various bioinformatics tools. Susceptibility testing was in accordance with EUCAST methodology (MIC ≤16 mg/L).

Results: Fifteen isolates of non-typhoidal *Salmonella enterica* belonging to serovars *Salmonella* Blockley, *Salmonella* Typhimurium, *Salmonella* Thompson, *Salmonella* Ridge and *Salmonella* Kentucky showed resistance or decreased susceptibility to azithromycin (from 6 to >16 mg/L) due to the presence of macrolide resistance genes *mphA*, *mphB* or *mefB*. These genes were either plasmid or chromosomally mediated. Azithromycin-resistant *Salmonella* Blockley isolates harboured a macrolide inactivation gene cluster, *mphA-mrx-mphr(A)*, within a novel *Salmonella* azithromycin resistance genomic island (SARGI) determined by MinION sequencing. This is the first known chromosomally mediated *mph*A gene cluster described in salmonellae. Phylogenetic analysis and epidemiological information showed that *mphA Salmonella* Blockley isolates were not derived from a single epidemiologically related event. The azithromycin MICs of the 15 *Salmonella* spp. isolates showed that the presence of the *mphA* gene was associated with MIC ≥16 mg/L, while the presence of *mefB* or *mphB* was not.

Conclusions: Azithromycin resistance due to acquisition of known macrolide resistance genes was seen in four different *Salmonella* serovars and can be either plasmid-encoded or chromosomally encoded.

Introduction

The increased resistance to a broad range of antibiotics in both *Salmonella* strains that cause enteric fever and non-typhoidal *Salmonella* (NTS) is an emerging threat.[1–8] Widespread resistance to amoxicillin, chloramphenicol, trimethoprim/sulfamethoxazole and fluoroquinolones has led to azithromycin being used as the preferred antimicrobial agent to treat cases of uncomplicated enteric fever reporting travel to the Indian subcontinent and South-East Asia.[4] It is also used to treat infections with MDR NTS in vulnerable patients who have prolonged or invasive infections.[7] Azithromycin is an azalide that has excellent tissue penetration, concentrates in the reticuloendothelial cells and has the advantage of oral administration and a long half-life. Clinical trials have shown it to be equivalent or superior to chloramphenicol, fluoroquinolones and third-generation cephalosporins for the management of uncomplicated typhoid fever.[9–11] However, reports are emerging of azithromycin resistance in cases of enteric fever as well as invasive NTS infection.[1,10–12]

Acquired resistance to macrolides/azalides may be caused by several different mechanisms.[13] They include: (i) target site modification by methylases encoded by *erm* genes;[14,15] (ii) modifying enzymes such as esterases encoded by *ereA* and *B* genes or phosphotransferases encoded by *mphA*, *B* and *D* genes;[16,17] (iii) efflux pumps, e.g. *mefA* and *msrA* found mainly in Gram-positive bacteria, with *mefA* also identified in Gram-negative strains;[15] (iv) mutations in the *rrl* and *rpl* genes encoding ribosomal proteins L22, L4 and 23S rRNA, which also confer resistance in Gram-positive bacteria.[18] The presence of more than one of the

8: Figures and Tables

8.1: Figures

8.2: Tables

9: References

Ackermann, R. J., & Monroe, P. W. (1996). Bacteremic urinary tract infection in older people. *Journal of the American Geriatrics Society*, *44*(8), 927–933. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med4&NEWS=N&AN=8708302

Afshari, A., Schrenzel, J., Ieven, M., & Harbarth, S. (2012). Bench-to-bedside review: Rapid molecular diagnostics for bloodstream infection--a new frontier? *Crit Care*, *16*(3), 222. https://doi.org/10.1186/cc11202

Akiyoshi, D. E., Dilo, J., Pearson, C., Chapman, S., Tumwine, J., & Tzipori, S. (2003). Characterization of Cryptosporidium meleagridis of human origin passaged through different host species. *Infection and Immunity*, *71*(4), 1828–1832. https://doi.org/10.1128/IAI.71.4.1828-1832.2003

Angelakis, E., Bachar, D., Henrissat, B., Armougom, F., Audoly, G., Lagier, J. C., … Raoult, D. (2016). Glycans affect DNA extraction and induce substantial differences in gut metagenomic studies. *Sci Rep*, *6*, 26276. https://doi.org/10.1038/srep26276

Arroyo, L. G., Rousseau, J., Willey, B. M., Low, D. E., Staempfli, H., McGeer, A., & Weese, J. S. (2005). Use of a selective enrichment broth to recover Clostridium difficile from stool swabs stored under different conditions. *J Clin Microbiol*, *43*(10), 5341–5343.

Ashton, P. M., Nair, S., Dallman, T., Rubino, S., Rabsch, W., Mwaigwisya, S., …

O'grady, J. (2015). MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nat Biotechnol*, *33*(3), 296–300.

Bader, O., Weig, M., Taverne-Ghadwal, L., Lugert, R., Gross, U., & Kuhns, M. (2011). Improved clinical laboratory identification of human pathogenic yeasts by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clinical Microbiology and Infection*, *17*(9), 1359–1365.

Baltekin, Ö., Boucharin, A., Tano, E., Andersson, D. I., & Elf, J. (2017). Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proceedings of the National Academy of Sciences*, 201708558. https://doi.org/10.1073/pnas.1708558114

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Prjibelski, A. D. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477.

Baquero, F. (2017). Emerging Concepts and Strategies in Clinical Microbiology and Infectious Diseases. *Clinical Infectious Diseases*, *65*, 1.

Baraban, L., Bertholle, F., Salverda, M. L. M., Bremond, N., Panizza, P., Baudry, J., … Bibette, J. (2011). Millifluidic droplet analyser for microbiology. *Lab on a Chip*, *11*(23), 4057. https://doi.org/10.1039/c1lc20545e

Barbosa, T. M., Serra, C. R., La Ragione, R. M., Woodward, M. J., & Henriques, A. O. (2005). Screening for Bacillus isolates in the broiler gastrointestinal tract. *Applied and Environmental Microbiology*, *71*(2), 968–978.

Benagli, C., Rossi, V., Dolina, M., Tonolla, M., Petrini, O., Mellmann, A., … Moehring, R. W. (2013). Evaluation of new preanalysis sample treatment tools and DNA isolation protocols to improve bacterial pathogen detection in whole blood. *J Clin Microbiol*, *6*(1), 2629–2631. https://doi.org/10.1001/jama.2013.3231

Bernardo, K., Krut, O., Wiegmann, K., Kreder, D., Micheli, M., Schäfer, R., … Meyer, H. E. (2000). Purification and characterization of a magnesium-dependent neutral sphingomyelinase from bovine brain. *Journal of Biological Chemistry*, *275*(11), 7641–7647.

Bidet, P., Lalande, V., Salauze, B., Burghoffer, B., Avesani, V., Delmée, M., … Petit, J.-C. (2000). Comparison of PCR-ribotyping, arbitrarily primed PCR, and pulsed-

field gel electrophoresis for typing Clostridium difficile. *J Clin Microbiol*, *38*(7), 2484–2487.

Bilecen, K., Yaman, G., Ciftci, U., & Laleli, Y. R. (2015). Performances and reliability of Bruker Microflex LT and Vitek MS MALDI-TOF mass spectrometry systems for the identification of clinical microorganisms. *BioMed Research International*, *2015*.

Bizzini, A., & Greub, G. (2010). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification. *Clinical Microbiology and Infection*, *16*(11), 1614–1619.

Bochud, P.-Y., Bonten, M., Marchetti, O., & Calandra, T. (2004). Antimicrobial therapy for patients with severe sepsis and septic shock: an evidence-based review. *Critical Care Medicine*, *32*(11), S495–S512.

Boedicker, J. Q., Li, L., Kline, T. R., & Ismagilov, R. F. (2008). Detecting bacteria and determining their susceptibility to antibiotics by stochastic confinement in nanoliter droplets using plug-based microfluidics. *Lab on a Chip*, *8*(8), 1265. https://doi.org/10.1039/b804911d

Bradley, P., Gordon, N. C., Walker, T. M., Dunn, L., Heys, S., Huang, B., … De Cesare, M. (2015). Rapid antibiotic-resistance predictions from genome sequence data for Staphylococcus aureus and Mycobacterium tuberculosis. *Nat Commun*, *6*, 10063.

Britton, R. A., & Young, V. B. (2014). Role of the intestinal microbiota in resistance to colonization by Clostridium difficile. *Gastroenterology*, *146*(6), 1547–1553.

Broeren, M. A. C., Bahçeci, S., Vader, H. L., & Arents, N. L. A. (2011). Screening for urinary tract infection with the Sysmex UF-1000i urine flow cytometer. *J Clin Microbiol*, *49*(3), 1025–1029.

Buchanan, A. G. (1984). Selective enrichment broth culture for detection of Clostridium difficile and associated cytotoxin. *J Clin Microbiol*, *20*(1), 74–76.

Burillo, A., Rodríguez-Sánchez, B., Ramiro, A., Cercenado, E., Rodríguez-Créixems, M., & Bouza, E. (2014). Gram-Stain Plus MALDI-TOF MS (Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry) for a Rapid Diagnosis of Urinary Tract Infection. *PLoS ONE*, *9*(1), e86915. https://doi.org/10.1371/journal.pone.0086915

Burke, R., McKenna, J., Cox, C., … P. C.-J. of, & 2016, undefined. (n.d.). A comparison of different pre-lysis methods and extraction kits for recovery of

Streptococcus agalacticae (Lancefield group B Streptococcus) DNA from whole blood. *Elsevier*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0167701216302275

Burnham, C.-A. D., Leeds, J., Nordmann, P., O'Grady, J., & Patel, J. (2017). Diagnosing antimicrobial resistance. *Nature Reviews Microbiology*, *15*(11), 697.

Burns, J. L., Saiman, L., Whittier, S., Krzewinski, J., Liu, Z., Larone, D., … Jones, R. N. (2001). Comparison of two commercial systems (Vitek and MicroScan-WalkAway) for antimicrobial susceptibility testing of Pseudomonas aeruginosa isolates from cystic fibrosis patients. *Diagnostic Microbiology and Infectious Disease*, *39*(4), 257–260. https://doi.org/10.1016/S0732-8893(01)00234-6

Carrara, L., Navarro, F., Turbau, M., Seres, M., Moran, I., Quintana, I., … Cordon, O. (2013). Molecular diagnosis of bloodstream infections with a new dual-priming oligonucleotide-based multiplex PCR assay. *Journal of Medical Microbiology*, *62*(11), 1673–1679.

Cetinkaya, Y., Falk, P., & Mayhall, C. G. (2000). Vancomycin-resistant enterococci. *Clinical Microbiology Reviews*, *13*(4), 686–707.

Chang, S.-S., Hsieh, W.-H., Liu, T.-S., Lee, S.-H., Wang, C.-H., Chou, H.-C., … Lee, C.-C. (2013). Multiplex PCR system for rapid detection of pathogens in patients with presumed sepsis–a systemic review and meta-analysis. *PLoS One*, *8*(5), e62323.

Chankhamhaengdecha, S., Hadpanus, P., Aroonnual, A., Ngamwongsatit, P., Chotiprasitsakul, D., Chongtrakool, P., & Janvilisri, T. (2013). Evaluation of multiplex PCR with enhanced spore germination for detection of clostridium difficile from stool samples of the hospitalized patients. *BioMed Research International*, *2013*. https://doi.org/10.1155/2013/875437

Chen, C. H., Lu, Y., Sin, M. L. Y., Mach, K. E., Zhang, D. D., Gau, V., … Wong, P. K. (2010). Antimicrobial susceptibility testing using high surface-to-volume ratio microchannels. *Analytical Chemistry*, *82*(xx), 1012–1019. https://doi.org/10.1021/ac9022764

Cho, I., & Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. *Nature Reviews. Genetics*, *13*(4), 260.

Choi, J., Yoo, J., Lee, M., Kim, E.-G., Lee, J. S., Lee, S., … Kwon, S. (2014). A rapid antimicrobial susceptibility test based on single-cell morphological analysis. *Science Translational Medicine*, *6*(267), 267ra174.

https://doi.org/10.1126/scitranslmed.3009650

Clements, A. C. A., Magalhães, R. J. S., Tatem, A. J., Paterson, D. L., & Riley, T. V. (2010). Clostridium difficile PCR ribotype 027: assessing the risks of further worldwide spread. *The Lancet Infectious Diseases*, *10*(6), 395–404.

Cohen, J. (2002). The immunopathogenesis of sepsis. *Nature*, *420*(6917), 885–891. https://doi.org/10.1038/nature01326

Cohen, J., Vincent, J.-L., Adhikari, N. K. J., Machado, F. R., Angus, D. C., Calandra, T., … Opal, S. (2015). Sepsis: a roadmap for future research. *The Lancet Infectious Diseases*, *15*(5), 581–614.

Cohen, S. H., Gerding, D. N., Johnson, S., Kelly, C. P., Loo, V. G., McDonald, L. C., … Wilcox, M. H. (2010). Clinical practice guidelines for Clostridium difficile infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). *Infect Control Hosp Epidemiol*, *31*(5), 431–455. https://doi.org/10.1086/651706

Consortium, H. M. P. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, *486*(7402), 207–214.

Cooper, M. A., & Shlaes, D. (2011). Fix the antibiotics pipeline. *Nature*, *472*(7341), 32.

Daniels, R. (2011). Surviving the first hours in sepsis: getting the basics right (an intensivist's perspective). *J Antimicrob Chemother*, *66 Suppl 2*, ii11-23. https://doi.org/10.1093/jac/dkq515

Dark, P., Blackwood, B., Gates, S., McAuley, D., Perkins, G. D., McMullan, R., … Warhurst, G. (2015). Accuracy of LightCycler((R)) SeptiFast for the detection and identification of pathogens in the blood of patients with suspected sepsis: a systematic review and meta-analysis. *Intensive Care Med*, *41*(1), 21–33. https://doi.org/10.1007/s00134-014-3553-8

Dark, P., Dunn, G., Chadwick, P., Young, D., Bentley, A., Carlson, G., & Warhurst, G. (2011). The clinical diagnostic accuracy of rapid detection of healthcare-associated bloodstream infection in intensive care using multipathogen real-time PCR technology. *BMJ Open*, *1*(1), e000181. https://doi.org/10.1136/bmjopen-2011-000181

Davies, S. C., & Gibbens, N. (2013). UK five year antimicrobial resistance strategy 2013 to 2018. *London: Department of Health*.

De Angelis, G., Posteraro, B., De Carolis, E., Menchinelli, G., Franceschi, F.,

Tumbarello, M., … Sanguinetti, M. (2018). T2Bacteria magnetic resonance assay for the rapid detection of ESKAPEc pathogens directly in whole blood. *Journal of Antimicrobial Chemotherapy*, *73*, iv20-iv26. https://doi.org/10.1093/jac/dky049

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., … Lawley, T. D. (2012). The Clostridium difficile spo0A gene is a persistence and transmission factor. *Infect Immun*, *80*(8), 2704–2711.

Dellinger, R. P., Levy, M. M., Rhodes, A., Annane, D., Gerlach, H., Opal, S. M., … Campaign, S. S. (2013). Surviving Sepsis Campaign: International Guidelines for Management of Severe Sepsis and Septic Shock, 2012. *Intensive Care Medicine*, *39*(2), 165–228. https://doi.org/DOI 10.1007/s00134-012-2769-8

Delmont, T. O., Robe, P., Clark, I., Simonet, P., & Vogel, T. M. (2011). Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J Microbiol Methods*, *86*(3), 397–400.

Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E. A., & Crook, D. W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nature Reviews. Genetics*, *13*(9), 601.

Dingle, K. E., Didelot, X., Ansari, M. A., Eyre, D. W., Vaughan, A., Griffiths, D., … Crook, D. W. (2013). Recombinational Switching of the Clostridium difficile S-Layer and a Novel Glycosylation Gene Cluster Revealed by Large-Scale Whole-Genome Sequencing. *The Journal of Infectious Diseases*, *207*(4), 675–686. https://doi.org/10.1093/infdis/jis734

Dodémont, M., De Mendonça, R., Nonhoff, C., Roisin, S., & Denis, O. (2014). Performance of the Verigene Gram-negative blood culture assay for rapid detection of bacteria and resistance determinants. *J Clin Microbiol*, *52*(8), 3085–3087.

Dolinger, D. L., & Jacobs, A. A. (2015). Molecular diagnostics and active screening for health care-associated infections: stepping-up the game. *Laboratory Medicine*, *42*(5), 267–272.

Donaldson, G. P., Lee, S. M., & Mazmanian, S. K. (2016). Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology*, *14*(1), 20–32.

Driks, A. (1999). Bacillus subtilis spore coat. *Microbiology and Molecular Biology Reviews*, *63*(1), 1–20.

Dunne Jr., W. M., Westblade, L. F., & Ford, B. (2012). Next-generation and whole-

genome sequencing in the diagnostic clinical microbiology laboratory. *Eur J Clin Microbiol Infect Dis*, *31*(8), 1719–1726. https://doi.org/10.1007/s10096-012-1641-7

Dutta, S., Chatterjee, A., Dutta, P., Rajendran, K., Roy, S., Pramanik, K. C., & Bhattacharya, S. K. (2001). Sensitivity and performance characteristics of a direct PCR with stool samples in comparison to conventional techniques for diagnosis of Shigella and enteroinvasive Escherichia coli infection in children with acute diarrhoea in Calcutta, India. *Journal of Medical Microbiology*, *50*(8), 667–674.

Edwards, A. N., Suárez, J. M., & McBride, S. M. (2013). Culturing and maintaining Clostridium difficile in an anaerobic environment. *Journal of Visualized Experiments: JoVE*, (79).

Eliopoulos, G. M., & Gold, H. S. (2001). Vancomycin-resistant enterococci: mechanisms and clinical observations. *Clinical Infectious Diseases*, *33*(2), 210–219.

Ellepola, A. N., & Morrison, C. J. (2005). Laboratory diagnosis of invasive candidiasis. *J Microbiol*, *43 Spec No*, 65–84.

Ellington, M. J., Ekelund, O., Aarestrup, F. M., Canton, R., Doumith, M., Giske, C., … Hopkins, K. L. (2017). The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. *Clinical Microbiology and Infection*, *23*(1), 2–22.

Emmadi, R., Boonyaratanakornkit, J. B., Selvarangan, R., Shyamala, V., Zimmer, B. L., Williams, L., … Wilson, J. A. A. (2011). Molecular methods and platforms for infectious diseases testing: a review of FDA-approved and cleared assays. *The Journal of Molecular Diagnostics*, *13*(6), 583–604.

Eyre, D. W., Cule, M. L., Wilson, D. J., Griffiths, D., Vaughan, A., O'connor, L., … Finney, J. M. (2013). Diverse sources of C. difficile infection identified on whole-genome sequencing. *New England Journal of Medicine*, *369*(13), 1195–1205.

Eyre, D. W., Fawley, W. N., Best, E. L., Griffiths, D., Stoesser, N. E., Crook, D. W., … Wilcox, M. H. (2013). Comparison of multilocus variable-number tandem-repeat analysis and whole-genome sequencing for investigation of Clostridium difficile transmission. *J Clin Microbiol*, *51*(12), 4141–4149. https://doi.org/10.1128/jcm.01095-13

Fair, R. J., & Tor, Y. (2014). Antibiotics and bacterial resistance in the 21st century.

Perspectives in Medicinal Chemistry, (6), 25–64. https://doi.org/10.4137/PMC.S14459

Falbo, R., Sala, M. R., Signorelli, S., Venturi, N., Signorini, S., & Brambilla, P. (2012). Bacteriuria screening by automated whole-field-image-based microscopy reduces the number of necessary urine cultures. *J Clin Microbiol*, *50*(4), 1427–1429.

Feehery, G. R., Yigit, E., Oyola, S. O., Langhorst, B. W., Schmidt, V. T., Stewart, F. J., … Pradhan, S. (2013). A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. *PLoS One*, *8*(10), e76096. https://doi.org/10.1371/journal.pone.0076096

Felt, J. R., Yurkovich, C., Garshott, D. M., Kamat, D., Farooqi, A., Fribley, A. M., … Hebert, K. (2017). The Utility of Real-Time Quantitative Polymerase Chain Reaction Genotype Detection in the Diagnosis of Urinary Tract Infections in Children. *Clin Pediatr (Phila)*, *56*(10), 912–919. https://doi.org/10.1177/0009922817706144

Fenollar, F., & Raoult, D. (2007). Molecular diagnosis of bloodstream infections caused by non-cultivable bacteria. *International Journal of Antimicrobial Agents*, *30*, 7–15.

Ferreira, L., Sánchez-Juanes, F., González-Ávila, M., Cembrero-Fuciños, D., Herrero-Hernández, A., González-Buitrago, J. M., & Muñoz-Bellido, J. L. (2010). Direct identification of urinary tract pathogens from urine samples by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol*, *48*(6), 2110–2115.

Ferrer, R., Martin-Loeches, I., Phillips, G., Osborn, T. M., Townsend, S., Dellinger, R. P., … Levy, M. M. (2014). Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. *Critical Care Medicine*, *42*(8), 1749–1755.

Fichtel, J., Köster, J., Rullkötter, J., & Sass, H. (2007). Spore dipicolinic acid contents used for estimating the number of endospores in sediments. *FEMS Microbiology Ecology*, *61*(3), 522–532.

Fleischmann, C., Scherag, A., Adhikari, N. K. J., Hartog, C. S., Tsaganos, T., Schlattmann, P., … Reinhart, K. (2016). Assessment of global incidence and mortality of hospital-treated sepsis. Current estimates and limitations. *American*

Journal of Respiratory and Critical Care Medicine, *193*(3), 259–272.

Fournier, P.-E., Drancourt, M., Colson, P., Rolain, J.-M., La Scola, B., & Raoult, D. (2013). Modern clinical microbiology: new challenges and solutions. *Nature Reviews Microbiology*, *11*(8), 574–585.

Fournier, P., Drancourt, M., … P. C.-N. R., & 2013, undefined. (n.d.). Modern clinical microbiology: new challenges and solutions. *Nature.Com*. Retrieved from https://www.nature.com/articles/nrmicro3068

Fox, J. L. (2014). Technology comes to typing. *Nat Biotechnol*, *32*(11), 1081–1084. https://doi.org/10.1038/nbt.3067

Foxman, B. (2010). The epidemiology of urinary tract infection. *Nature Reviews Urology*, *7*(12), 653.

Frith, M. C., Hamada, M., & Horton, P. (2010). Parameters for accurate genome alignment. *BMC Bioinformatics*, *11*(1), 80.

Fukushima, H., Tsunomori, Y., & Seki, R. (2003). Duplex real-time SYBR green PCR assays for detection of 17 species of food-or waterborne pathogens in stools. *J Clin Microbiol*, *41*(11), 5134–5146.

Gadsby, N. J., Russell, C. D., McHugh, M. P., Mark, H., Conway Morris, A., Laurenson, I. F., … Templeton, K. E. (2016). Comprehensive molecular testing for respiratory pathogens in community-acquired pneumonia. *Clinical Infectious Diseases*, *62*(7), 817–823.

Gaillot, O., Di Camillo, P., Berche, P., Courcol, R., & Savage, C. (1999). Comparison of CHROMagar Salmonella medium and Hektoen enteric agar for isolation of salmonellae from stool samples. *J Clin Microbiol*, *37*(3), 762–765.

Gangaram, R., Ojwang, P. J., Moodley, J., & Maharaj, D. (2005). The accuracy of urine dipsticks as a screening test for proteinuria in hypertensive disorders of pregnancy. *Hypertension in Pregnancy*, *24*(2), 117–123. https://doi.org/10.1081/PRG-200059849

Garalde, D. R., Snell, E. A., Jachimowicz, D., Sipos, B., Lloyd, J. H., Bruce, M., … Warland, A. (2018). Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods*.

García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L. M., Götz, S., Tarazona, S., … Conesa, A. (2012). Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics*, *28*(20), 2678–2679.

Garnacho-Montero, J., Gutierrez-Pizarraya, A., Escoresca-Ortega, A., Corcia-

Palomo, Y., Fernandez-Delgado, E., Herrera-Melero, I., … Marquez-Vacaro, J. A. (2014). De-escalation of empirical therapy is associated with lower mortality in patients with severe sepsis and septic shock. *Intensive Care Med*, *40*(1), 32–40. https://doi.org/10.1007/s00134-013-3077-7

Gebert, S., Siegel, D., & Wellinghausen, N. (2008). Rapid detection of pathogens in blood culture bottles by real-time PCR in conjunction with the pre-analytic tool MolYsis. *Journal of Infection*, *57*(4), 307–316.

Gerner-Smidt, P., Hise, K., Kincaid, J., Hunter, S., Rolando, S., Hyytiä-Trees, E., … Taskforce, P. (2006). PulseNet USA: a five-year update. *Foodbourne Pathogens & Disease*, *3*(1), 9–19.

Gerritzen, A., Wittke, J.-W., & Wolff, D. (2011). Rapid and Sensitive Detection of Shiga Toxin-Producing Escherichia Coli Directly From Stool Samples by Real-Time PCR in Comparison to Culture, Enzyme Immunoassay, and Vero Cell Cytotoxicity Assay. *Clinical Laboratory*, *57*(11), 993.

Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B., Jorden, J. R., & Chiodini, R. J. (2015). Changes in 16S RNA gene microbial community profiling by concentration of prokaryotic DNA. *J Microbiol Methods*, *119*, 239–242.

Goldberg, B., Sichtig, H., Geyer, C., Ledeboer, N., & Weinstock, G. M. (2015). Making the leap from research laboratory to clinic: challenges and opportunities for next-generation sequencing in infectious disease diagnostics. *MBio*, *6*(6), e01888-15.

Goldenberg, S. D., & French, G. L. (2011). Diagnostic testing for Clostridium difficile: a comprehensive survey of laboratories in England. *Journal of Hospital Infection*, *79*(1), 4–7.

Goldenberg, S. D., Price, N. M., Tucker, D., Wade, P., & French, G. L. (2011). Mandatory reporting and improvements in diagnosing Clostridium difficile infection: an incompatible dichotomy? *Journal of Infection*, *62*(5), 363–370.

Goodwin, S., Wappel, R., & McCombie, W. R. (2017). 1D Genome Sequencing on the Oxford Nanopore MinION. *Current Protocols in Human Genetics*, 18.11. 1-18.11. 14.

Goswitz, J. J., Willard, K. E., Eastep, S. J., Shanholtzer, C. J., Olson, M. L., Pinnell, M., … Peterson, L. R. (1993). Utility of slide centrifuge Gram's stain versus quantitative culture for diagnosis of urinary tract infection. *American Journal of Clinical Pathology*, *99*(2), 132–136.

Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., … Chiu, C. Y. (2015). Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med*, *7*, 99. https://doi.org/10.1186/s13073-015-0220-9

Grice, E. A., Kong, H. H., Renaud, G., Young, A. C., Bouffard, G. G., Blakesley, R. W., … Segre, J. A. (2008). A diversity profile of the human skin microbiota. *Genome Res*, *18*(7), 1043–1050. https://doi.org/10.1101/gr.075549.107

Grumaz, S., Stevens, P., Grumaz, C., Decker, S. O., Weigand, M. A., Hofer, S., … Sohn, K. (2016). Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Med*, *8*(1), 73. https://doi.org/10.1186/s13073-016-0326-8

Guido, M., Tumolo, M. R., De Donno, A., Verri, T., Serio, F., Bagordo, F., & Zizza, A. (2016). In vitro diagnosis of sepsis: a review. *PATHOLOGY AND LABORATORY MEDICINE INTERNATIONAL*, *8*, 1–14.

Gurtler, V. (1993). Typing of Clostridium difficile strains by PCR-amplification of variable length 16S-23S rDNA spacer regions. *Journal of General Microbiology*, *139*(12), 3089–3097. https://doi.org/10.1099/00221287-139-12-3089

Gyarmati, P., Kjellander, C., Aust, C., Song, Y., Öhrmalm, L., & Giske, C. G. (2016). Metagenomic analysis of bloodstream infections in patients with acute leukemia and therapy-induced neutropenia. *Sci Rep*, *6*.

Haag, H., Locher, F., & Nolte, O. (2013). Molecular diagnosis of microbial aetiologies using SepsiTest™ in the daily routine of a diagnostic laboratory. *Diagnostic Microbiology and Infectious Disease*, *76*(4), 413–418.

Hansen, W. L. J., Bruggeman, C. A., & Wolffs, P. F. G. (2009). Evaluation of new preanalysis sample treatment tools and DNA isolation protocols to improve bacterial pathogen detection in whole blood. *J Clin Microbiol*, *47*(8), 2629–2631.

Harris, D. M., & Hata, D. J. (2013). Rapid identification of bacteria and Candida using PNA-FISH from blood and peritoneal fluid cultures: a retrospective clinical study. *Annals of Clinical Microbiology and Antimicrobials*, *12*(1), 2.

Hasman, H., Saputra, D., Sicheritz-Ponten, T., Lund, O., Svendsen, C. A., Frimodt-Moller, N., & Aarestrup, F. M. (2014). Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. *J Clin Microbiol*, *52*(1), 139–146. https://doi.org/10.1128/JCM.02452-13

Hay, A., Birnie, K., Busby, J., Delaney, B., Downing, H., Dudley, J., … Hollingworth,

W. (2016). The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute i.

Hayden, R. T., Clinton, L. K., Hewitt, C., Koyamatsu, T., Sun, Y., Jamison, G., … Bankowski, M. J. (2016). Rapid antimicrobial susceptibility testing using forward laser light scatter technology. *J Clin Microbiol*, *54*(11), 2701–2706.

He, M., Miyajima, F., Roberts, P., Ellison, L., Pickard, D. J., Martin, M. J., … Lawley, T. D. (2013). Emergence and global spread of epidemic healthcare-associated Clostridium difficile. *Nat Genet*, *45*(1), 109–113. https://doi.org/10.1038/ng.2478

Hevia, A., Delgado, S., Margolles, A., & Sanchez, B. (2015). Application of density gradient for the isolation of the fecal microbial stool component and the potential use thereof. *Sci Rep*, *5*, 16807. https://doi.org/10.1038/srep16807

Hewitt, F. C., Guertin, S. L., Ternus, K. L., Schulte, K., & Kadavy, D. R. (2017). Toward Rapid Sequenced-Based Detection and Characterization of Causative Agents of Bacteremia. *BioRxiv*, 162735.

Hill, J. T., Tran, K.-D. T., Barton, K. L., Labreche, M. J., & Sharp, S. E. (2014). Evaluation of the nanosphere Verigene BC-GN assay for direct identification of gram-negative bacilli and antibiotic resistance markers from positive blood cultures and potential impact for more-rapid antibiotic interventions. *J Clin Microbiol*, *52*(10), 3805–3807.

Holt, K. E., Phan, M. D., Baker, S., Duy, P. T., Nga, T. V, Nair, S., … Wain, J. (2011). Emergence of a globally dominant IncHI1 plasmid type associated with multiple drug resistant typhoid. *PLoS Negl Trop Dis*, *5*(7), e1245. https://doi.org/10.1371/journal.pntd.0001245

Horz, H.-P., Scheer, S., Huenger, F., Vianna, M. E., & Conrads, G. (2008). Selective isolation of bacterial DNA from human clinical specimens. *J Microbiol Methods*, *72*(1), 98–102.

Horz, H. P., Scheer, S., Vianna, M. E., & Conrads, G. (2010). New methods for selective isolation of bacterial DNA from human clinical specimens. *Anaerobe*, *16*(1), 47–53. https://doi.org/10.1016/j.anaerobe.2009.04.009

Hosokawa, M., Nishikawa, Y., Kogawa, M., & Takeyama, H. (2017). Massively parallel whole genome amplification for single-cell sequencing using droplet microfluidics. *Sci Rep*, *7*(1), 5199.

Huang, T. D., Bogaerts, P., Ghilani, E., Heinrichs, A., Gavage, P., Roisin, S., … Glupczynski, Y. (2015). Multicentre evaluation of the Check-Direct CPE(R) assay for direct screening of carbapenemase-producing Enterobacteriaceae from rectal swabs. *J Antimicrob Chemother*, *70*(6), 1669–1673. https://doi.org/10.1093/jac/dkv009

Huber, C. A., Foster, N. F., Riley, T. V, & Paterson, D. L. (2013). Challenges for standardization of Clostridium difficile typing methods. *J Clin Microbiol*, *51*(9), 2810–2814.

Ip, C. L. C., Loose, M., Tyson, J. R., de Cesare, M., Brown, B. L., Jain, M., … Urban, J. M. (2015). MinION Analysis and Reference Consortium: Phase 1 data release and analysis. *F1000Res*, *4*.

Ipe, D. S., Sundac, L., Benjamin Jr, W. H., Moore, K. H., & Ulett, G. C. (2013). Asymptomatic bacteriuria: prevalence rates of causal microorganisms, etiology of infection in different patient populations, and recent advances in molecular detection. *FEMS Microbiol Lett*, *346*(1), 1–10.

Iyengar, V., Albaugh, G. P., Lohani, A., & Nair, P. P. (1991). Human stools as a source of viable colonic epithelial cells. *Faseb J*, *5*(13), 2856–2859.

Jain, M., Fiddes, I. T., Miga, K. H., Olsen, H. E., Paten, B., & Akeson, M. (2015). Improved data analysis for the MinION nanopore sequencer. *Nat Methods*, *12*(4), 351–356.

Jain, M., Koren, S., Quick, J., Rand, A. C., Sasani, T. A., Tyson, J. R., … Malla, S. (2017). Nanopore sequencing and assembly of a human genome with ultra-long reads. *BioRxiv*, 128835.

Jorgensen, J. H., & Ferraro, M. J. (2009). Antimicrobial susceptibility testing: a review of general principles and contemporary practices. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *49*(11), 1749–1755. https://doi.org/10.1086/647952

Juul, S., Izquierdo, F., Hurst, A., Dai, X., Wright, A., Kulesha, E., … Turner, D. J. (2015). What9s in my pot? Real-time species identification on the MinION. *BioRxiv*, 30742.

Kachrimanidou, M., & Malisiovas, N. (2011). Clostridium difficile infection: a comprehensive review. *Critical Reviews in Microbiology*, *37*(3), 178–187.

Karlsson, E., Larkeryd, A., Sjodin, A., Forsman, M., & Stenberg, P. (2015). Scaffolding of a bacterial genome using MinION nanopore sequencing. *Sci Rep*,

5, 11996. https://doi.org/10.1038/srep11996

Kellogg, J. A., Manzella, J. P., & Bankert, D. A. (2000). Frequency of low-level bacteremia in children from birth to fifteen years of age. *J Clin Microbiol*, *38*(6), 2181–2185.

Kelly, MD, C. P., & LaMont, MD, J. T. (1998). *CLOSTRIDIUM DIFFICILE* INFECTION. *Annual Review of Medicine*, *49*(1), 375–390. https://doi.org/10.1146/annurev.med.49.1.375

Kiełbasa, S. M., Wan, R., Sato, K., Horton, P., & Frith, M. C. (2011). Adaptive seeds tame genomic sequence comparison. *Genome Res*, *21*(3), 487–493.

Killgore, G., Thompson, A., Johnson, S., Brazier, J., Kuijper, E., Pepin, J., … van den Berg, R. J. (2008). Comparison of seven techniques for typing international epidemic strains of Clostridium difficile: restriction endonuclease analysis, pulsed-field gel electrophoresis, PCR-ribotyping, multilocus sequence typing, multilocus variable-number tandem-repeat an. *J Clin Microbiol*, *46*(2), 431–437.

Knetsch, C. W., Lawley, T. D., Hensgens, M. P., Corver, J., Wilcox, M. W., & Kuijper, E. J. (2013). Current application and future perspectives of molecular typing methods to study Clostridium difficile infections. *Euro Surveill*, *18*(4), 20381.

Kohn, A. (1960). Lysis of frozen and thawed cells of Escherichia coli by lysozyme, and their conversion into spheroplasts. *J Bacteriol*, *79*(5), 697.

Kollef, M. H. (2000). Inadequate antimicrobial treatment: an important determinant of outcome for hospitalized patients. *Clin Infect Dis*, *31 Suppl 4*, S131-8. https://doi.org/10.1086/314079

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/gr.215087.116

Köser, C. U., Ellington, M. J., Cartwright, E. J. P., Gillespie, S. H., Brown, N. M., Farrington, M., … Parkhill, J. (2012). Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathogens*, *8*(8), e1002824.

Krcmery, S., Hromec, J., & Demesova, D. (2001). Treatment of lower urinary tract infection in pregnancy. *International Journal of Antimicrobial Agents*, *17*(4), 279–282.

Kuczynski, J., Lauber, C. L., Walters, W. A., Parfrey, L. W., Clemente, J. C., Gevers,

D., & Knight, R. (2012). Experimental and analytical tools for studying the human microbiome. *Nature Reviews Genetics*, *13*(1), 47–58.

Kuhn, C., Disque, C., Muhl, H., Orszag, P., Stiesch, M., & Haverich, A. (2011). Evaluation of commercial universal rRNA gene PCR plus sequencing tests for identification of bacteria and fungi associated with infectious endocarditis. *J Clin Microbiol*, *49*(8), 2919–2923. https://doi.org/10.1128/jcm.00830-11

Kuijper, E. J., Coignard, B., & Tüll, P. (2006). Emergence of Clostridium difficile-associated disease in North America and Europe. *Clinical Microbiology and Infection*, *12*(s6), 2–18.

Kumar, A., Ellis, P., Arabi, Y., Roberts, D., Light, B., Parrillo, J. E., … Chateau, D. (2009). Initiation of inappropriate antimicrobial therapy results in a fivefold reduction of survival in human septic shock. *Chest*, *136*(5), 1237–1248. https://doi.org/10.1378/chest.09-0087

Kumar, A., Roberts, D., Wood, K. E., Light, B., Parrillo, J. E., Sharma, S., … Cheang, M. (2006). Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit Care Med*, *34*(6), 1589–1596. https://doi.org/10.1097/01.CCM.0000217961.75225.E9

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.

Lawley, T. D., Clare, S., Walker, A. W., Goulding, D., Stabler, R. A., Croucher, N., … Dougan, G. (2009). Antibiotic treatment of clostridium difficile carrier mice triggers a supershedder state, spore-mediated transmission, and severe disease in immunocompromised hosts. *Infect Immun*, *77*(9), 3661–3669. https://doi.org/10.1128/iai.00558-09

Leffler, D. A., & Lamont, J. T. (2015). Clostridium difficile infection. *New England Journal of Medicine*, *372*(16), 1539–1548.

Lehmann, L. E., Hauser, S., Malinka, T., Klaschik, S., Weber, S. U., Schewe, J.-C., … Book, M. (2011). Rapid qualitative urinary tract infection pathogen identification by SeptiFast® real-time PCR. *PLoS One*, *6*(2), e17146.

Lehmann, L. E., Hunfeld, K.-P., Steinbrucker, M., Brade, V., Book, M., Seifert, H., … Stüber, F. (2010). Improved detection of blood stream pathogens by real-time PCR in severe sepsis. *Intensive Care Medicine*, *36*(1), 49–56.

Lessa, F. C., Mu, Y., Bamberg, W. M., Beldavs, Z. G., Dumyati, G. K., Dunn, J. R.,

… Phipps, E. C. (2015). Burden of Clostridium difficile infection in the United States. *New England Journal of Medicine*, *372*(9), 825–834.

Li, B., Qiu, Y., Glidle, A., McIlvenna, D., Luo, Q., Cooper, J., … Yin, H. (2014). Gradient microfluidics enables rapid bacterial growth inhibition testing. *Analytical Chemistry*, *86*(6), 3131–3137. https://doi.org/10.1021/ac5001306

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv Preprint ArXiv:1303.3997*.

Li, H. (2016). Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics*, *32*(14), 2103–2110.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079.

Liesenfeld, O., Lehman, L., Hunfeld, K. P., & Kost, G. (2014). Molecular diagnosis of sepsis: New aspects and recent developments. *Eur J Microbiol Immunol (Bp)*, *4*(1), 1–25. https://doi.org/10.1556/EuJMI.4.2014.1.1

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., … Ecker, J. R. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, *462*(7271), 315–322. https://doi.org/10.1038/nature08514

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., … Law, M. (2012). Comparison of next-generation sequencing systems. *BioMed Research International*, *2012*.

Livermore, D. M., & Wain, J. (2013). Revolutionising bacteriology to improve treatment outcomes and antibiotic stewardship. *Infect Chemother*, *45*(1), 1–10. https://doi.org/10.3947/ic.2013.45.1.1

Lofgren, E. T., Cole, S. R., Weber, D. J., Anderson, D. J., & Moehring, R. W. (2014). Hospital-Acquired Clostridium difficile Infections Estimating All-Cause Mortality and Length of Stay. *Epidemiology (Cambridge, Mass.)*, *25*(4), 570.

Loman, N. J., Quick, J., & Simpson, J. T. (2015). A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat Methods*, *12*(8), 733–735. https://doi.org/10.1038/nmeth.3444

Loman, N. J., & Quinlan, A. R. (2014). Poretools: a toolkit for analyzing nanopore sequence data. *Bioinformatics*, *30*(23), 3399–3401.

Longtin, Y., Trottier, S., Brochu, G., Paquet-Bolduc, B., Garenc, C., Loungnarath, V., … Longtin, J. (2012). Impact of the type of diagnostic assay on Clostridium

difficile infection and complication rates in a mandatory reporting program. *Clinical Infectious Diseases*, *56*(1), 67–73.

Loonen, A. J., Bos, M. P., van Meerbergen, B., Neerken, S., Catsburg, A., Dobbelaer, I., … van den Brule, A. J. (2013). Comparison of pathogen DNA isolation methods from large volumes of whole blood to improve molecular diagnosis of bloodstream infections. *PLoS One*, *8*(8), e72349. https://doi.org/10.1371/journal.pone.0072349

Loose, M. W. (2017). The potential impact of nanopore sequencing on human genetics. *Human Molecular Genetics*, *26*(R2), R202–R207.

López-Pérez, M., Gonzaga, A., Martin-Cuadrado, A. B., Onyshchenko, O., Ghavidel, A., Ghai, R., & Rodriguez-Valera, F. (2012). Genomes of surface isolates of Alteromonas macleodii: The life of a widespread marine opportunistic copiotroph. *Scientific Reports*, *2*. https://doi.org/10.1038/srep00696

Lucke, K., Hombach, M., Hug, M., & Pfyffer, G. E. (2010). Rapid detection of methicillin-resistant Staphylococcus aureus (MRSA) in diverse clinical specimens by the BD GeneOhm MRSA assay and comparison with culture. *J Clin Microbiol*, *48*(3), 981–984.

Maddocks, S., Olma, T., & Chen, S. (2002). Comparison of CHROMagar Salmonella medium and xylose-lysine-desoxycholate and Salmonella-Shigella agars for isolation of Salmonella strains from stool samples. *J Clin Microbiol*, *40*(8), 2999–3003.

Mancini, N., Carletti, S., Ghidoli, N., Cichero, P., Burioni, R., & Clementi, M. (2010). The era of molecular and other non-culture-based methods in diagnosis of sepsis. *Clin Microbiol Rev*, *23*(1), 235–251. https://doi.org/10.1128/CMR.00043-09

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., … Marteau, P. (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut*, *55*(2), 205–211.

Manoni, F., Fornasiero, L., Ercolin, M., Tinello, A., Ferrian, M., Hoffer, P., … Gessoni, G. (2009). Cutoff values for bacteria and leukocytes for urine flow cytometer Sysmex UF-1000i in urinary tract infections. *Diagnostic Microbiology and Infectious Disease*, *65*(2), 103–107.

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, *24*(3), 133–141.

Marler, L. M., Siders, J. A., Wolters, L. C., Pettigrew, Y., Skitt, B. L., & Allen, S. D. (1992). Comparison of five cultural procedures for isolation of clostridium difficile from stools. *Journal of Clinical Microbiology*.

McArthur, A. G., Waglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A. J., … Ejim, L. (2013). The comprehensive antibiotic resistance database. *Antimicrobial Agents and Chemotherapy*, *57*(7), 3348–3357.

McCann, C. D., & Jordan, J. A. (2014). Evaluation of MolYsis Complete5 DNA extraction method for detecting Staphylococcus aureus DNA from whole blood in a sepsis model using PCR/pyrosequencing. *J Microbiol Methods*, *99*, 1–7. https://doi.org/10.1016/j.mimet.2014.01.013

McPherson, D., Griffiths, C., Williams, M., Baker, A., Klodawski, E., Jacobson, B., & Donaldson, L. (2013). Sepsis-associated mortality in England: an analysis of multiple cause of death data from 2001 to 2010. *BMJ Open*, *3*(8). https://doi.org/10.1136/bmjopen-2013-002586

Mellmann, A., Cloud, J., Maier, T., Keckevoet, U., Ramminger, I., Iwen, P., … Harmsen, D. (2008). Evaluation of matrix-assisted laser desorption ionization-time-of-flight mass spectrometry in comparison to 16S rRNA gene sequencing for species identification of nonfermenting bacteria. *Journal of Clinical Microbiology*, *46*(6), 1946–1954. https://doi.org/10.1128/JCM.00157-08

Melnikov, A., Galinsky, K., Rogov, P., Fennell, T., Van Tyne, D., Russ, C., … Neafsey, D. E. (2011). Hybrid selection for sequencing pathogen genomes from clinical samples. *Genome Biol*, *12*(8), R73. https://doi.org/10.1186/gb-2011-12-8-r73

Melzer, M., & Welch, C. (2013). Outcomes in UK patients with hospital-acquired bacteraemia and the risk of catheter-associated urinary tract infections. *Postgraduate Medical Journal*, *89*(1052), 329–334.

Menozzi, M. G., Eigner, U., Covan, S., Rossi, S., Somenzi, P., Dettori, G., … Fahr, A. M. (2006). Two-center collaborative evaluation of performance of the BD phoenix automated microbiology system for identification and antimicrobial susceptibility testing of gram-negative bacteria. *Journal of Clinical Microbiology*, *44*(11), 4085–4094. https://doi.org/10.1128/JCM.00614-06

Moir, A., Corfe, B. M., & Behravan, J. (2002). Spore germination. *Cellular and Molecular Life Sciences*, *59*(3), 403–409.

Morris, K., Wilson, C., & Wilcox, M. H. (2012). Evaluation of chromogenic meticillin-

resistant Staphylococcus aureus media: sensitivity versus turnaround time. *Journal of Hospital Infection*, *81*(1), 20–24.

Muhl, H., Kochem, A. J., Disque, C., & Sakka, S. G. (2010). Activity and DNA contamination of commercial polymerase chain reaction reagents for the universal 16S rDNA real-time polymerase chain reaction detection of bacterial pathogens in blood. *Diagn Microbiol Infect Dis*, *66*(1), 41–49. https://doi.org/10.1016/j.diagmicrobio.2008.07.011

Murray, C., Adeyiga, O., Owsley, K., & Di Carlo, D. (2015). Research highlights: microfluidic analysis of antimicrobial susceptibility. *Lab on a Chip*, *15*(5), 1226–1229. https://doi.org/10.1039/c5lc90017d

Mwaigwisya, S., Assiri, R. A., & O'Grady, J. (2015). Emerging commercial molecular tests for the diagnosis of bloodstream infection. *Expert Rev Mol Diagn*, *15*(5), 681–692. https://doi.org/10.1586/14737159.2015.1029459

Naccache, S. N., Federman, S., Veeraraghavan, N., Zaharia, M., Lee, D., Samayoa, E., … Enge, B. (2014). A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. *Genome Res*, *24*(7), 1180–1192.

Nair, S., Ashton, P., Doumith, M., Connell, S., Painset, A., Mwaigwisya, S., … Day, M. (2016). WGS for surveillance of antimicrobial resistance: a pilot study to detect the prevalence and mechanism of resistance to azithromycin in a UK population of non-typhoidal Salmonella. *Journal of Antimicrobial Chemotherapy*, *71*(12), 3400–3408.

Neely, L. A., Audeh, M., Phung, N. A., Min, M., Suchocki, A., Plourde, D., … Lowery, T. J. (2013). T2 magnetic resonance enables nanoparticle-mediated rapid detection of candidemia in whole blood. *Sci Transl Med*, *5*(182), 182ra54. https://doi.org/10.1126/scitranslmed.3005377

Norling, M., Karlsson-Lindsjö, O. E., Gourlé, H., Bongcam-Rudloff, E., & Hayer, J. (2016). MetLab: An in silico experimental design, simulation and analysis tool for viral metagenomics studies. *PLoS ONE*, *11*(8). https://doi.org/10.1371/journal.pone.0160334

O'Neill, J. (2014). Antimicrobial Resistance : Tackling a crisis for the health and wealth of nations. *Review on Antimicrobial Resistance*, (December), 1–16. https://doi.org/10.1038/510015a

Opota, O., Jaton, K., & Greub, G. (2015). Microbial diagnosis of bloodstream

infection: Towards molecular diagnosis directly from blood. *Clinical Microbiology and Infection*. https://doi.org/10.1016/j.cmi.2015.02.005

Oyola, S. O., Gu, Y., Manske, M., Otto, T. D., O'Brien, J., Alcock, D., … Quail, M. A. (2013). Efficient depletion of host DNA contamination in malaria clinical sequencing. *J Clin Microbiol*, *51*(3), 745–751. https://doi.org/10.1128/jcm.02507-12

Pancholi, P., Kelly, C., Raczkowski, M., & Balada-Llasat, J. M. (2012). Detection of toxigenic Clostridium difficile: comparison of the cell culture neutralization, Xpert C. difficile, Xpert C. difficile/Epi, and Illumigene C. difficile assays. *J Clin Microbiol*, *50*(4), 1331–1335.

Paredes-Sabja, D., Shen, A., & Sorg, J. A. (2014). Clostridium difficile spore biology: sporulation, germination, and spore structural proteins. *Trends in Microbiology*, *22*(7), 406–416.

Patel, H. D., Livsey, S. A., Swann, R. A., & Bukhari, S. S. (2005). Can urine dipstick testing for urinary tract infection at point of care reduce laboratory workload? *Journal of Clinical Pathology*, *58*(9), 951–954.

Patel, J. B. (2001). 16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. *Molecular Diagnosis*, *6*(4), 313–321.

Paulino, L. C., Tseng, C. H., Strober, B. E., & Blaser, M. J. (2006). Molecular analysis of fungal microbiota in samples from healthy human skin and psoriatic lesions. *J Clin Microbiol*, *44*(8), 2933–2941. https://doi.org/10.1128/jcm.00785-06

Penders, J., Vink, C., Driessen, C., London, N., Thijs, C., & Stobberingh, E. E. (2005). Quantification of Bifidobacterium spp., Escherichia coli and Clostridium difficile in faecal samples of breast-fed and formula-fed infants by real-time PCR. *FEMS Microbiol Lett*, *243*(1), 141–147. https://doi.org/10.1016/j.femsle.2004.11.052

Perry-O'Keefe, H., Rigby, S., Oliveira, K., Sørensen, D., Stender, H., Coull, J., & Hyldig-Nielsen, J. J. (2001). Identification of indicator microorganisms using a standardized PNA FISH method. *Journal of Microbiological Methods*, *47*(3), 281–292. https://doi.org/10.1016/S0167-7012(01)00303-7

Pezzlo, M. (1988). Detection of urinary tract infections by rapid methods. *Clinical Microbiology Reviews*, *1*(3), 268–280.

Poretsky, R., Rodriguez-R, L. M., Luo, C., Tsementzi, D., & Konstantinidis, K. T.

(2014). Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. *PLoS One*, *9*(4), e93827.

Putnam, L. R., Howard, W. J., Pfaller, M. A., Koontz, F. P., & Jones, R. N. (1997). Accuracy of the vitek system for antimicrobial susceptibility testing enterobacteriaceae bloodstream infection isolates: Use of “direct” inoculation from bactec 9240 blood culture bottles. *Diagnostic Microbiology and Infectious Disease*, *28*(2), 101–104. https://doi.org/10.1016/S0732-8893(97)00014-X

Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., … Mikhail, A. (2016). Real-time, portable genome sequencing for Ebola surveillance. *Nature*, *530*(7589), 228–232.

Quick, J., Quinlan, A. R., & Loman, N. J. (2014). A reference bacterial genome dataset generated on the MinION portable single-molecule nanopore sequencer. *Gigascience*, *3*, 22. https://doi.org/10.1186/2047-217x-3-22

Raoult, D., Fournier, P. E., & Drancourt, M. (2004). What does the future hold for clinical microbiology? *Nature Reviews Microbiology*, *2*(2), 151–159. https://doi.org/Doi 10.1038/Nrmicro820

Ray, B., & Speck, M. L. (1973). Freeze-injury in bacteria. *CRC Critical Reviews in Clinical Laboratory Sciences*, *4*(2), 161–213.

Reece, A., Xia, B., Jiang, Z., Noren, B., McBride, R., & Oakey, J. (2016). Microfluidic techniques for high throughput single cell analysis. *Current Opinion in Biotechnology*. https://doi.org/10.1016/j.copbio.2016.02.015

Resistance, R. on A. (2016). *Tackling drug-resistant infections globally: final report and recommendations*. Review on Antimicrobial Resistance.

Rhoads, D. D., Sintchenko, V., Rauch, C. A., & Pantanowitz, L. (2014). Clinical microbiology informatics. *Clin Microbiol Rev*, *27*(4), 1025–1047. https://doi.org/10.1128/cmr.00049-14

Rhodes, A., Evans, L. E., Alhazzani, W., Levy, M. M., Antonelli, M., Ferrer, R., … Dellinger, R. P. (2017). Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. *Intensive Care Medicine*, *43*(3), 304–377. https://doi.org/10.1007/s00134-017-4683-6

Richards, M., Knox, J., Elliott, B., Mackin, K., Lyras, D., Waring, L. J., & Riley, T. V. (2011). Severe infection with Clostridium difficile PCR ribotype 027 acquired in Melbourne, Australia. *Medical Journal of Australia*, *194*(7), 369–371.

Riley, T. V, Brazier, J. S., Hassan, H., Williams, K., & Phillips, K. D. (1987).

Comparison of alcohol shock enrichment and selective enrichment for the isolation of Clostridium difficile. *Epidemiology & Infection*, *99*(2), 355–359.

Rivers, E., Nguyen, B., Havstad, S., Ressler, J., Muzzin, A., Knoblich, B., … Early Goal-Directed Therapy Collaborative, G. (2001). Early goal-directed therapy in the treatment of severe sepsis and septic shock. *N Engl J Med*, *345*(19), 1368–1377. https://doi.org/10.1056/NEJMoa010307

Roberts, K. B. (2011). Urinary tract infection: clinical practice guideline for the diagnosis and management of the initial UTI in febrile infants and children 2 to 24 months. Am Acad Pediatrics.

Rogol, M., Shpak, B., Rothman, D., & Sechter, I. (1985). Enrichment medium for isolation of Campylobacter jejuni-Campylobacter coli. *Applied and Environmental Microbiology*, *50*(1), 125–126.

Rose, G., Wooldridge, D. J., Anscombe, C., Mee, E. T., Misra, R. V., & Gharbia, S. (2015). Challenges of the Unknown: Clinical Application of Microbial Metagenomics. *International Journal of Genomics*, *2015*. https://doi.org/10.1155/2015/292950

Sabat, A. J., Budimir, A., Nashev, D., Sá-Leão, R., Van Dijl, J. M., Laurent, F., … Markers, E. S. G. of E. (2013). Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill*, *18*(4), 20380.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., … Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*. https://doi.org/10.1126/science.2448875

Salipante, S. J., Sengupta, D. J., Hoogestraat, D. R., Cummings, L. A., Bryant, B. H., Natividad, C., … Hoffman, N. G. (2013). Molecular diagnosis of Actinomadura madurae infection by 16S rRNA deep sequencing. *J Clin Microbiol*, *51*(12), 4262–4265. https://doi.org/10.1128/jcm.02227-13

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., … Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol*, *12*(1), 87.

Schappert, S. M., & Rechtsteiner, E. A. (2011). Ambulatory medical care utilization estimates for 2007. *Vital and Health Statistics. Series 13, Data from the National Health Survey*, (169), 1–38.

Schlaberg, R., Chiu, C. Y., Miller, S., Procop, G. W., Weinstock, G., Committee, P. P., … Pathologists, M. R. C. of the C. of A. (2017). Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Archives of Pathology and Laboratory Medicine*, *141*(6), 776–786.

Schmidt, K., Mwaigwisya, S., Crossman, L. C., Doumith, M., Munroe, D., Pires, C., … Wain, J. (2016). Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *Journal of Antimicrobial Chemotherapy*, *72*(1), 104–114.

Schmieder, R., & Edwards, R. (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One*, *6*(3), e17288.

Schmiemann, G., Kniehl, E., Gebhardt, K., Matejczyk, M. M., & Hummers-Pradier, E. (2010). The diagnosis of urinary tract infection: a systematic review. *Deutsches Ärzteblatt International*, *107*(21), 361.

Schoepp, N. G., Khorosheva, E. M., Schlappi, T. S., Curtis, M. S., Humphries, R. M., Hindler, J. A., & Ismagilov, R. F. (2016). Digital Quantification of DNA Replication and Chromosome Segregation Enables Determination of Antimicrobial Susceptibility after only 15 Minutes of Antibiotic Exposure. *Angewandte Chemie - International Edition*, *55*(33), 9557–9561. https://doi.org/10.1002/anie.201602763

Schröder, U.-C., Bokeloh, F., O'Sullivan, M., Glaser, U., Wolf, K., Pfister, W., … Neugebauer, U. (2015). Rapid, culture-independent, optical diagnostics of centrifugally captured bacteria from urine samples. *Biomicrofluidics*, *9*(4), 44118.

Schulte, W., Bernhagen, J., & Bucala, R. (2013). Cytokines in sepsis: potent immunoregulators and potential therapeutic targets—an updated view. *Mediators of Inflammation*, *2013*.

Semeniuk, H., & Church, D. (1999). Evaluation of the leukocyte esterase and nitrite urine dipstick screening tests for detection of bacteriuria in women with suspected uncomplicated urinary tract infections. *Journal of Clinical Microbiology*, *37*(9), 3051–3052.

Sender, R., Fuchs, S., & Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol*, *14*(8), e1002533. https://doi.org/10.1371/journal.pbio.1002533

Seng, P., Abat, C., Rolain, J. M., Colson, P., Lagier, J.-C., Gouriet, F., … Raoult, D.

(2013). Identification of rare pathogenic bacteria in a clinical microbiology laboratory: impact of MALDI-TOF mass spectrometry. *J Clin Microbiol*, JCM. 00492-13.

Shigemura, K., Shirakawa, T., Okada, H., Tanaka, K., Kamidono, S., Arakawa, S., & Gotoh, A. (2005). Rapid detection and differentiation of Gram-negative and Gram-positive pathogenic bacteria in urine using TaqMan probe. *Clinical and Experimental Medicine*, *4*(4), 196–201.

Shrestha, N. K., Tuohy, M. J., Hall, G. S., Isada, C. M., & Procop, G. W. (2002). Rapid identification of Staphylococcus aureus and the mecA gene from BacT/ALERT blood culture bottles by using the LightCycler system. *J Clin Microbiol*, *40*(7), 2659–2661.

Simner, P. J., Uhl, J. R., Hall, L., Weber, M. M., Walchak, R. C., Buckwalter, S., & Wengenack, N. L. (2013). Broad-range direct detection and identification of fungi by use of the PLEX-ID PCR-electrospray ionization mass spectrometry (ESIMS) system. *Journal of Clinical Microbiology*, *51*(6), 1699–1706. https://doi.org/10.1128/JCM.03282-12

Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., … Coopersmith, C. M. (2016). The third international consensus definitions for sepsis and septic shock (sepsis-3). *Jama*, *315*(8), 801–810.

Sixt, B. S., Siegl, A., Müller, C., Watzka, M., Wultsch, A., Tziotis, D., … Horn, M. (2013). Metabolic features of Protochlamydia amoebophila elementary bodies–a link between activity and infectivity in Chlamydiae. *PLoS Pathogens*, *9*(8), e1003553.

Smith, A. M., Jain, M., Mulroney, L., Garalde, D. R., & Akeson, M. (2017). Reading canonical and modified nucleotides in 16S ribosomal RNA using nanopore direct RNA sequencing. *BioRxiv*, 132274.

Sparbier, K., Weller, U., Boogen, C., & Kostrzewa, M. (2012). Rapid detection of Salmonella sp. by means of a combination of selective enrichment broth and MALDI-TOF MS. *European Journal of Clinical Microbiology & Infectious Diseases*, *31*(5), 767–773.

Springer, J., Loeffler, J., Heinz, W., Schlossnagel, H., Lehmann, M., Morton, O., … Kurzai, O. (2011). Pathogen-specific DNA enrichment does not increase sensitivity of PCR for diagnosis of invasive aspergillosis in neutropenic patients. *J Clin Microbiol*, *49*(4), 1267–1273. https://doi.org/10.1128/jcm.01679-10

Stamm, W. E., & Norrby, S. R. (2001). Urinary Tract Infections: Disease Panorama and Challenges. *The Journal of Infectious Diseases*, *183*(s1), S1–S4. https://doi.org/10.1086/318850

Stamper, P. D., Babiker, W., Alcabasa, R., Aird, D., Wehrlin, J., Ikpeama, I., … Carroll, K. C. (2009). Evaluation of a new commercial TaqMan PCR assay for direct detection of the clostridium difficile toxin B gene in clinical stool specimens. *J Clin Microbiol*, *47*(12), 3846–3850. https://doi.org/10.1128/jcm.01490-09

Sterry-Blunt, R. E., Randall, K. S., Doughton, M. J., Aliyu, S. H., & Enoch, D. A. (2015). Screening urine samples for the absence of urinary tract infection using the sediMAX automated microscopy analyser. *Journal of Medical Microbiology*, *64*(6), 605–609.

Stevenson, M., Pandor, A., James, M.-S., Rafia, R., Uttley, L., Stevens, J., … McMullan, R. (2016). Sepsis: the LightCycler SeptiFast Test MGRADE®, SepsiTest™ and IRIDICA BAC BSI assay for rapidly identifying bloodstream bacteria and fungi–a systematic review and economic evaluation. *Health Technology Assessment*, *20*(46), 1–246.

Stewart, E. J. (2012). Growing unculturable bacteria. *J Bacteriol*, *194*(16), 4151–4160.

Stonehouse, M. J., Cota-Gomez, A., Parker, S. K., Martin, W. E., Hankin, J. A., Murphy, R. C., … Vasil, A. I. (2002). A novel class of microbial phosphocholine-specific phospholipases C. *Molecular Microbiology*, *46*(3), 661–676.

Stoneking, L. R., Patanwala, A. E., Winkler, J. P., Fiorello, A. B., Lee, E. S., Olson, D. P., & Wolk, D. M. (2013). Would earlier microbe identification alter antibiotic therapy in bacteremic emergency department patients? *J Emerg Med*, *44*(1), 1–8. https://doi.org/10.1016/j.jemermed.2012.02.036

Surawicz, C. M., Brandt, L. J., Binion, D. G., Ananthakrishnan, A. N., Curry, S. R., Gilligan, P. H., … Zuckerbraun, B. S. (2013). Guidelines for diagnosis, treatment, and prevention of Clostridium difficile infections. *The American Journal of Gastroenterology*, *108*(4), 478–498.

Telenti, A., Honoré, N., Bernasconi, C., March, J., Ortega, A., Heym, B., … Cole, S. T. (1997). Genotypic assessment of isoniazid and rifampin resistance in Mycobacterium tuberculosis: A blind study at reference laboratory level. *Journal of Clinical Microbiology*, *35*(3), 719–723.

Tenover, F. C., Novak-Weekley, S., Woods, C. W., Peterson, L. R., Davis, T., Schreckenberger, P., … Nomura, J. H. (2010). Impact of strain type on detection of toxigenic Clostridium difficile: comparison of molecular diagnostic and enzyme immunoassay approaches. *J Clin Microbiol*, *48*(10), 3719–3724.

Tessari, A., Osti, N., & Scarin, M. (2015). Screening of presumptive urinary tract infections by the automated urine sediment analyser sediMAX. *Clinical Chemistry and Laboratory Medicine (CCLM)*, *53*(s2), s1503–s1508.

Thoendel, M., Jeraldo, P., Greenwood-Quaintance, K. E., Yao, J., Chia, N., Hanssen, A. D., … Patel, R. (2017). Impact of Contaminating DNA in Whole-Genome Amplification Kits Used for Metagenomic Shotgun Sequencing for Infection Diagnosis. *J Clin Microbiol*, *55*(6), 1789–1801.

Thoendel, M., Jeraldo, P. R., Greenwood-Quaintance, K. E., Yao, J. Z., Chia, N., Hanssen, A. D., … Patel, R. (2016). Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. *J Microbiol Methods*, *127*, 141–145.

Tille, P. (2013). *Bailey & Scott's Diagnostic Microbiology-E-Book*. Elsevier Health Sciences.

Tissari, P., Zumla, A., Tarkka, E., Mero, S., Savolainen, L., Vaara, M., … Gant, V. (2010). Accurate and rapid identification of bacterial species from positive blood cultures with a DNA-based microarray platform: an observational study. *The Lancet*, *375*(9710), 224–230. https://doi.org/10.1016/S0140-6736(09)61569-5

Titball, R. W., Leslie, D. L., Harvey, S., & Kelly, D. (1991). Hemolytic and sphingomyelinase activities of Clostridium perfringens alpha-toxin are dependent on a domain homologous to that of an enzyme from the human arachidonic acid pathway. *Infect Immun*, *59*(5), 1872–1874.

Torio, C. M. (2013). *National inpatient hospital costs: the most expensive conditions by payer, 2011. Healthcare Cost and Utilization Project Statistical Brief# 160*.

Treangen, T. J., & Salzberg, S. L. (2011). Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nature Reviews Genetics*, *13*(1), nrg3117.

Tyson, G. H., McDermott, P. F., Li, C., Chen, Y., Tadesse, D. A., Mukherjee, S., … Loneragan, G. H. (2015). WGS accurately predicts antimicrobial resistance in Escherichia coli. *Journal of Antimicrobial Chemotherapy*, *70*(10), 2763–2769.

van der Zwet, W. C., Hessels, J., Canbolat, F., & Deckers, M. M. L. (2010).

Evaluation of the Sysmex UF-1000i® urine flow cytometer in the diagnostic work-up of suspected urinary tract infection in a Dutch general hospital. *Clinical Chemistry and Laboratory Medicine*, *48*(12), 1765–1771.

Vandecandelaere, I., Nercessian, O., Segaert, E., Achouak, W., Mollica, A., Faimali, M., … Vandamme, P. (2008). Alteromonas genovensis sp. nov., isolated from a marine electroactive biofilm and emended description of Alteromonas macleodii Baumann et al. 1972 (Approved Lists 1980). *International Journal of Systematic and Evolutionary Microbiology*, *58*(11), 2589–2596. https://doi.org/10.1099/ijs.0.65691-0

Viala, C., Le Monnier, A., Maataoui, N., Rousseau, C., Collignon, A., & Poilane, I. (2012). Comparison of commercial molecular assays for toxigenic Clostridium difficile detection in stools: BD GeneOhm Cdiff, XPert C. difficile and illumigene C. difficile. *J Microbiol Methods*, *90*(2), 83–85.

Vince, A., Lepej, S. Ž., Baršić, B., Dušek, D., Mitrović, Z., Serventi-Seiwerth, R., & Labar, B. (2008). LightCycler SeptiFast assay as a tool for the rapid diagnosis of sepsis in patients during antimicrobial therapy. *Journal of Medical Microbiology*, *57*(10), 1306–1307.

Vincent, J.-L., Sakr, Y., Sprung, C. L., Ranieri, V. M., Reinhart, K., Gerlach, H., … Payen, D. (2006). Sepsis in European intensive care units: results of the SOAP study. *Critical Care Medicine*, *34*(2), 344–353.

Votintseva, A. A., Bradley, P., Pankhurst, L., del Ojo Elias, C., Loose, M., Nilgiriwala, K., … Walker, T. M. (2017). Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. *J Clin Microbiol*, *55*(5), 1285–1298.

Wagner, M., Horn, M., & Daims, H. (2003). Fluorescence in situ hybridisation for the identification and characterisation of prokaryotes. *Current Opinion in Microbiology*, *6*(3), 302–309. https://doi.org/10.1016/S1369-5274(03)00054-7

Wain, J., Diep, T. S., Ho, V. A., Walsh, A. M., Hoa, N. T. T., Parry, C. M., & White, N. J. (1998). Quantitation of bacteria in blood of typhoid fever patients and relationship between counts and clinical features, transmissibility, and antibiotic resistance. *J Clin Microbiol*, *36*(6), 1683–1687.

Wain, J., & Mavrogiorgou, E. (2013). Next-generation sequencing in clinical microbiology. *Expert Rev Mol Diagn*, *13*(3), 225–227. https://doi.org/10.1586/erm.13.8

Warhurst, G., Maddi, S., Dunn, G., Ghrew, M., Chadwick, P., Alexander, P., … Dark, P. (2015). Diagnostic accuracy of SeptiFast multi-pathogen real-time PCR in the setting of suspected healthcare-associated bloodstream infection. *Intensive Care Med*, *41*(1), 86–93. https://doi.org/10.1007/s00134-014-3551-x

Washington, J. A. (2012). *Laboratory procedures in clinical microbiology*. Springer Science & Business Media.

Wei, Y.-X., Zhang, Z.-Y., Liu, C., Zhu, Y.-Z., Zhu, Y.-Q., Zheng, H., … Guo, X.-K. (2010). Complete Genome Sequence of Bifidobacterium longum JDM301. *J Bacteriol*, *192*(15), 4076–4077. https://doi.org/10.1128/JB.00538-10

Weiss, S., Amir, A., Hyde, E. R., Metcalf, J. L., Song, S. J., & Knight, R. (2014). Tracking down the sources of experimental contamination in microbiome studies. *Genome Biol*, *15*(12), 564.

Wellinghausen, N., Kochem, A. J., Disque, C., Muhl, H., Gebert, S., Winter, J., … Sakka, S. G. (2009). Diagnosis of bacteremia in whole-blood samples by use of a commercial universal 16S rRNA gene-based PCR and sequence analysis. *J Clin Microbiol*, *47*(9), 2759–2765. https://doi.org/10.1128/JCM.00567-09

Wiesinger-Mayr, H., Jordana-Lluch, E., Martró, E., Schoenthaler, S., & Noehammer, C. (2011). Establishment of a semi-automated pathogen DNA isolation from whole blood and comparison with commercially available kits. *J Microbiol Methods*, *85*(3), 206–213.

Wilcox, M. H., Gerding, D. N., Poxton, I. R., Kelly, C., Nathan, R., Birch, T., … Lee, C. (2017). Bezlotoxumab for prevention of recurrent Clostridium difficile infection. *New England Journal of Medicine*, *376*(4), 305–317.

Wilcox, M. H., Shetty, N., Fawley, W. N., Shemko, M., Coen, P., Birtles, a, … Wren, M. W. D. (2012). Changing epidemiology of Clostridium difficile infection following the introduction of a national ribotyping-based surveillance scheme in England. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *55*(8), 1056–1063. https://doi.org/10.1093/cid/cis614

Wilson, M. L., & Gaido, L. (2004). Laboratory diagnosis of urinary tract infections in adult patients. *Clinical Infectious Diseases*, *38*(8), 1150–1158.

Wojewoda, C. M., Sercia, L., Navas, M., Tuohy, M., Wilson, D., Hall, G. S., … Richter, S. S. (2013). Evaluation of the Verigene Gram-positive blood culture nucleic acid test for rapid detection of bacteria and resistance determinants. *J*

Clin Microbiol, *51*(7), 2072–2076.

Wood, D. E., & Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol*, *15*(3), R46. https://doi.org/10.1186/gb-2014-15-3-r46

Yagupsky, P., & Nolte, F. S. (1990). Quantitative aspects of septicemia. *Clinical Microbiology Reviews*, *3*(3), 269–279.

Yigit, E., Hernandez, D. I., Trujillo, J. T., Dimalanta, E., & Bailey, C. D. (2014). Genome and metagenome sequencing: Using the human methyl-binding domain to partition genomic DNA derived from plant tissues. *Applications in Plant Sciences*, *2*(11), 1400064.

Zaman, Z., Fogazzi, G. B., Garigali, G., Croci, M. D., Bayer, G., & Kránicz, T. (2010). Urine sediment analysis: analytical and diagnostic performance of sediMAX®—a new automated microscopy image-based urine sediment analyser. *Clinica Chimica Acta*, *411*(3), 147–154.

Zhang, X., Niu, S., & Zhang, L. (2014). Antimicrobial susceptibilities and clinical characterization of Pseudomonas aeruginosa isolates from urinary tract infections. *Urol Int*, *93*(4), 464–469. https://doi.org/10.1159/000358493

Zhao, S., Tyson, G. H., Chen, Y., Li, C., Mukherjee, S., Young, S., … McDermott, P. F. (2016). Whole-genome sequencing analysis accurately predicts antimicrobial resistance phenotypes in Campylobacter spp. *Applied and Environmental Microbiology*, *82*(2), 459–466.

Zhou, L., & Pollard, A. J. (2012). A novel method of selective removal of human DNA improves PCR sensitivity for detection of Salmonella Typhi in blood samples. *BMC Infect Dis*, *12*, 164. https://doi.org/10.1186/1471-2334-12-164

Ziegler, I., Fagerström, A., Strålin, K., & Mölling, P. (2016). Evaluation of a Commercial Multiplex PCR Assay for Detection of Pathogen DNA in Blood from Patients with Suspected Sepsis. *PLoS One*, *11*(12), e0167883.

Zimlichman, E., Henderson, D., Tamir, O., Franz, C., Song, P., Yamin, C. K., … Bates, D. W. (2013). Health care–associated infections: a meta-analysis of costs and financial impact on the US health care system. *JAMA Internal Medicine*, *173*(22), 2039–2046.

Ackermann, R. J., & Monroe, P. W. (1996). Bacteremic urinary tract infection in older people. *Journal of the American Geriatrics Society*, *44*(8), 927–933. Retrieved from

http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med4&NEWS=N&AN=8708302

Afshari, A., Schrenzel, J., Ieven, M., & Harbarth, S. (2012). Bench-to-bedside review: Rapid molecular diagnostics for bloodstream infection--a new frontier? *Crit Care*, *16*(3), 222. https://doi.org/10.1186/cc11202

Akiyoshi, D. E., Dilo, J., Pearson, C., Chapman, S., Tumwine, J., & Tzipori, S. (2003). Characterization of Cryptosporidium meleagridis of human origin passaged through different host species. *Infection and Immunity*, *71*(4), 1828–1832. https://doi.org/10.1128/IAI.71.4.1828-1832.2003

Angelakis, E., Bachar, D., Henrissat, B., Armougom, F., Audoly, G., Lagier, J. C., … Raoult, D. (2016). Glycans affect DNA extraction and induce substantial differences in gut metagenomic studies. *Sci Rep*, *6*, 26276. https://doi.org/10.1038/srep26276

Arroyo, L. G., Rousseau, J., Willey, B. M., Low, D. E., Staempfli, H., McGeer, A., & Weese, J. S. (2005). Use of a selective enrichment broth to recover Clostridium difficile from stool swabs stored under different conditions. *J Clin Microbiol*, *43*(10), 5341–5343.

Ashton, P. M., Nair, S., Dallman, T., Rubino, S., Rabsch, W., Mwaigwisya, S., … O'grady, J. (2015). MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nat Biotechnol*, *33*(3), 296–300.

Bader, O., Weig, M., Taverne-Ghadwal, L., Lugert, R., Gross, U., & Kuhns, M. (2011). Improved clinical laboratory identification of human pathogenic yeasts by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clinical Microbiology and Infection*, *17*(9), 1359–1365.

Baltekin, Ö., Boucharin, A., Tano, E., Andersson, D. I., & Elf, J. (2017). Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proceedings of the National Academy of Sciences*, 201708558. https://doi.org/10.1073/pnas.1708558114

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Prjibelski, A. D. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477.

Baquero, F. (2017). Emerging Concepts and Strategies in Clinical Microbiology and

Infectious Diseases. *Clinical Infectious Diseases*, *65*, 1.

Baraban, L., Bertholle, F., Salverda, M. L. M., Bremond, N., Panizza, P., Baudry, J., … Bibette, J. (2011). Millifluidic droplet analyser for microbiology. *Lab on a Chip*, *11*(23), 4057. https://doi.org/10.1039/c1lc20545e

Barbosa, T. M., Serra, C. R., La Ragione, R. M., Woodward, M. J., & Henriques, A. O. (2005). Screening for Bacillus isolates in the broiler gastrointestinal tract. *Applied and Environmental Microbiology*, *71*(2), 968–978.

Benagli, C., Rossi, V., Dolina, M., Tonolla, M., Petrini, O., Mellmann, A., … Moehring, R. W. (2013). Evaluation of new preanalysis sample treatment tools and DNA isolation protocols to improve bacterial pathogen detection in whole blood. *J Clin Microbiol*, *6*(1), 2629–2631. https://doi.org/10.1001/jama.2013.3231

Bernardo, K., Krut, O., Wiegmann, K., Kreder, D., Micheli, M., Schäfer, R., … Meyer, H. E. (2000). Purification and characterization of a magnesium-dependent neutral sphingomyelinase from bovine brain. *Journal of Biological Chemistry*, *275*(11), 7641–7647.

Bidet, P., Lalande, V., Salauze, B., Burghoffer, B., Avesani, V., Delmée, M., … Petit, J.-C. (2000). Comparison of PCR-ribotyping, arbitrarily primed PCR, and pulsed-field gel electrophoresis for typing Clostridium difficile. *J Clin Microbiol*, *38*(7), 2484–2487.

Bilecen, K., Yaman, G., Ciftci, U., & Laleli, Y. R. (2015). Performances and reliability of Bruker Microflex LT and Vitek MS MALDI-TOF mass spectrometry systems for the identification of clinical microorganisms. *BioMed Research International*, *2015*.

Bizzini, A., & Greub, G. (2010). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification. *Clinical Microbiology and Infection*, *16*(11), 1614–1619.

Bochud, P.-Y., Bonten, M., Marchetti, O., & Calandra, T. (2004). Antimicrobial therapy for patients with severe sepsis and septic shock: an evidence-based review. *Critical Care Medicine*, *32*(11), S495–S512.

Boedicker, J. Q., Li, L., Kline, T. R., & Ismagilov, R. F. (2008). Detecting bacteria and determining their susceptibility to antibiotics by stochastic confinement in nanoliter droplets using plug-based microfluidics. *Lab on a Chip*, *8*(8), 1265. https://doi.org/10.1039/b804911d

Bradley, P., Gordon, N. C., Walker, T. M., Dunn, L., Heys, S., Huang, B., … De Cesare, M. (2015). Rapid antibiotic-resistance predictions from genome sequence data for Staphylococcus aureus and Mycobacterium tuberculosis. *Nat Commun*, *6*, 10063.

Britton, R. A., & Young, V. B. (2014). Role of the intestinal microbiota in resistance to colonization by Clostridium difficile. *Gastroenterology*, *146*(6), 1547–1553.

Broeren, M. A. C., Bahçeci, S., Vader, H. L., & Arents, N. L. A. (2011). Screening for urinary tract infection with the Sysmex UF-1000i urine flow cytometer. *J Clin Microbiol*, *49*(3), 1025–1029.

Buchanan, A. G. (1984). Selective enrichment broth culture for detection of Clostridium difficile and associated cytotoxin. *J Clin Microbiol*, *20*(1), 74–76.

Burillo, A., Rodríguez-Sánchez, B., Ramiro, A., Cercenado, E., Rodríguez-Créixems, M., & Bouza, E. (2014). Gram-Stain Plus MALDI-TOF MS (Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry) for a Rapid Diagnosis of Urinary Tract Infection. *PLoS ONE*, *9*(1), e86915. https://doi.org/10.1371/journal.pone.0086915

Burke, R., McKenna, J., Cox, C., … P. C.-J. of, & 2016, undefined. (n.d.). A comparison of different pre-lysis methods and extraction kits for recovery of Streptococcus agalacticae (Lancefield group B Streptococcus) DNA from whole blood. *Elsevier*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0167701216302275

Burnham, C.-A. D., Leeds, J., Nordmann, P., O'Grady, J., & Patel, J. (2017). Diagnosing antimicrobial resistance. *Nature Reviews Microbiology*, *15*(11), 697.

Burns, J. L., Saiman, L., Whittier, S., Krzewinski, J., Liu, Z., Larone, D., … Jones, R. N. (2001). Comparison of two commercial systems (Vitek and MicroScan-WalkAway) for antimicrobial susceptibility testing of Pseudomonas aeruginosa isolates from cystic fibrosis patients. *Diagnostic Microbiology and Infectious Disease*, *39*(4), 257–260. https://doi.org/10.1016/S0732-8893(01)00234-6

Carrara, L., Navarro, F., Turbau, M., Seres, M., Moran, I., Quintana, I., … Cordon, O. (2013). Molecular diagnosis of bloodstream infections with a new dual-priming oligonucleotide-based multiplex PCR assay. *Journal of Medical Microbiology*, *62*(11), 1673–1679.

Cetinkaya, Y., Falk, P., & Mayhall, C. G. (2000). Vancomycin-resistant enterococci. *Clinical Microbiology Reviews*, *13*(4), 686–707.

Chang, S.-S., Hsieh, W.-H., Liu, T.-S., Lee, S.-H., Wang, C.-H., Chou, H.-C., … Lee, C.-C. (2013). Multiplex PCR system for rapid detection of pathogens in patients with presumed sepsis–a systemic review and meta-analysis. *PLoS One*, *8*(5), e62323.

Chankhamhaengdecha, S., Hadpanus, P., Aroonnual, A., Ngamwongsatit, P., Chotiprasitsakul, D., Chongtrakool, P., & Janvilisri, T. (2013). Evaluation of multiplex PCR with enhanced spore germination for detection of clostridium difficile from stool samples of the hospitalized patients. *BioMed Research International*, *2013*. https://doi.org/10.1155/2013/875437

Chen, C. H., Lu, Y., Sin, M. L. Y., Mach, K. E., Zhang, D. D., Gau, V., … Wong, P. K. (2010). Antimicrobial susceptibility testing using high surface-to-volume ratio microchannels. *Analytical Chemistry*, *82*(xx), 1012–1019. https://doi.org/10.1021/ac9022764

Cho, I., & Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. *Nature Reviews. Genetics*, *13*(4), 260.

Choi, J., Yoo, J., Lee, M., Kim, E.-G., Lee, J. S., Lee, S., … Kwon, S. (2014). A rapid antimicrobial susceptibility test based on single-cell morphological analysis. *Science Translational Medicine*, *6*(267), 267ra174. https://doi.org/10.1126/scitranslmed.3009650

Clements, A. C. A., Magalhães, R. J. S., Tatem, A. J., Paterson, D. L., & Riley, T. V. (2010). Clostridium difficile PCR ribotype 027: assessing the risks of further worldwide spread. *The Lancet Infectious Diseases*, *10*(6), 395–404.

Cohen, J. (2002). The immunopathogenesis of sepsis. *Nature*, *420*(6917), 885–891. https://doi.org/10.1038/nature01326

Cohen, J., Vincent, J.-L., Adhikari, N. K. J., Machado, F. R., Angus, D. C., Calandra, T., … Opal, S. (2015). Sepsis: a roadmap for future research. *The Lancet Infectious Diseases*, *15*(5), 581–614.

Cohen, S. H., Gerding, D. N., Johnson, S., Kelly, C. P., Loo, V. G., McDonald, L. C., … Wilcox, M. H. (2010). Clinical practice guidelines for Clostridium difficile infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). *Infect Control Hosp Epidemiol*, *31*(5), 431–455. https://doi.org/10.1086/651706

Consortium, H. M. P. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, *486*(7402), 207–214.

Cooper, M. A., & Shlaes, D. (2011). Fix the antibiotics pipeline. *Nature*, *472*(7341), 32.

Daniels, R. (2011). Surviving the first hours in sepsis: getting the basics right (an intensivist's perspective). *J Antimicrob Chemother*, *66 Suppl 2*, ii11-23. https://doi.org/10.1093/jac/dkq515

Dark, P., Blackwood, B., Gates, S., McAuley, D., Perkins, G. D., McMullan, R., … Warhurst, G. (2015). Accuracy of LightCycler((R)) SeptiFast for the detection and identification of pathogens in the blood of patients with suspected sepsis: a systematic review and meta-analysis. *Intensive Care Med*, *41*(1), 21–33. https://doi.org/10.1007/s00134-014-3553-8

Dark, P., Dunn, G., Chadwick, P., Young, D., Bentley, A., Carlson, G., & Warhurst, G. (2011). The clinical diagnostic accuracy of rapid detection of healthcare-associated bloodstream infection in intensive care using multipathogen real-time PCR technology. *BMJ Open*, *1*(1), e000181. https://doi.org/10.1136/bmjopen-2011-000181

Davies, S. C., & Gibbens, N. (2013). UK five year antimicrobial resistance strategy 2013 to 2018. *London: Department of Health*.

De Angelis, G., Posteraro, B., De Carolis, E., Menchinelli, G., Franceschi, F., Tumbarello, M., … Sanguinetti, M. (2018). T2Bacteria magnetic resonance assay for the rapid detection of ESKAPEc pathogens directly in whole blood. *Journal of Antimicrobial Chemotherapy*, *73*, iv20-iv26. https://doi.org/10.1093/jac/dky049

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., … Lawley, T. D. (2012). The Clostridium difficile spo0A gene is a persistence and transmission factor. *Infect Immun*, *80*(8), 2704–2711.

Dellinger, R. P., Levy, M. M., Rhodes, A., Annane, D., Gerlach, H., Opal, S. M., … Campaign, S. S. (2013). Surviving Sepsis Campaign: International Guidelines for Management of Severe Sepsis and Septic Shock, 2012. *Intensive Care Medicine*, *39*(2), 165–228. https://doi.org/DOI 10.1007/s00134-012-2769-8

Delmont, T. O., Robe, P., Clark, I., Simonet, P., & Vogel, T. M. (2011). Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J Microbiol Methods*, *86*(3), 397–400.

Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E. A., & Crook, D. W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nature*

Reviews. Genetics, *13*(9), 601.

Dingle, K. E., Didelot, X., Ansari, M. A., Eyre, D. W., Vaughan, A., Griffiths, D., … Crook, D. W. (2013). Recombinational Switching of the Clostridium difficile S-Layer and a Novel Glycosylation Gene Cluster Revealed by Large-Scale Whole-Genome Sequencing. *The Journal of Infectious Diseases*, *207*(4), 675–686. https://doi.org/10.1093/infdis/jis734

Dodémont, M., De Mendonça, R., Nonhoff, C., Roisin, S., & Denis, O. (2014). Performance of the Verigene Gram-negative blood culture assay for rapid detection of bacteria and resistance determinants. *J Clin Microbiol*, *52*(8), 3085–3087.

Dolinger, D. L., & Jacobs, A. A. (2015). Molecular diagnostics and active screening for health care-associated infections: stepping-up the game. *Laboratory Medicine*, *42*(5), 267–272.

Donaldson, G. P., Lee, S. M., & Mazmanian, S. K. (2016). Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology*, *14*(1), 20–32.

Driks, A. (1999). Bacillus subtilis spore coat. *Microbiology and Molecular Biology Reviews*, *63*(1), 1–20.

Dunne Jr., W. M., Westblade, L. F., & Ford, B. (2012). Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. *Eur J Clin Microbiol Infect Dis*, *31*(8), 1719–1726. https://doi.org/10.1007/s10096-012-1641-7

Dutta, S., Chatterjee, A., Dutta, P., Rajendran, K., Roy, S., Pramanik, K. C., & Bhattacharya, S. K. (2001). Sensitivity and performance characteristics of a direct PCR with stool samples in comparison to conventional techniques for diagnosis of Shigella and enteroinvasive Escherichia coli infection in children with acute diarrhoea in Calcutta, India. *Journal of Medical Microbiology*, *50*(8), 667–674.

Edwards, A. N., Suárez, J. M., & McBride, S. M. (2013). Culturing and maintaining Clostridium difficile in an anaerobic environment. *Journal of Visualized Experiments: JoVE*, (79).

Eliopoulos, G. M., & Gold, H. S. (2001). Vancomycin-resistant enterococci: mechanisms and clinical observations. *Clinical Infectious Diseases*, *33*(2), 210–219.

Ellepola, A. N., & Morrison, C. J. (2005). Laboratory diagnosis of invasive

candidiasis. *J Microbiol*, *43 Spec No*, 65–84.

Ellington, M. J., Ekelund, O., Aarestrup, F. M., Canton, R., Doumith, M., Giske, C., … Hopkins, K. L. (2017). The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. *Clinical Microbiology and Infection*, *23*(1), 2–22.

Emmadi, R., Boonyaratanakornkit, J. B., Selvarangan, R., Shyamala, V., Zimmer, B. L., Williams, L., … Wilson, J. A. A. (2011). Molecular methods and platforms for infectious diseases testing: a review of FDA-approved and cleared assays. *The Journal of Molecular Diagnostics*, *13*(6), 583–604.

Eyre, D. W., Cule, M. L., Wilson, D. J., Griffiths, D., Vaughan, A., O'connor, L., … Finney, J. M. (2013). Diverse sources of C. difficile infection identified on whole-genome sequencing. *New England Journal of Medicine*, *369*(13), 1195–1205.

Eyre, D. W., Fawley, W. N., Best, E. L., Griffiths, D., Stoesser, N. E., Crook, D. W., … Wilcox, M. H. (2013). Comparison of multilocus variable-number tandem-repeat analysis and whole-genome sequencing for investigation of Clostridium difficile transmission. *J Clin Microbiol*, *51*(12), 4141–4149. https://doi.org/10.1128/jcm.01095-13

Fair, R. J., & Tor, Y. (2014). Antibiotics and bacterial resistance in the 21st century. *Perspectives in Medicinal Chemistry*, (6), 25–64. https://doi.org/10.4137/PMC.S14459

Falbo, R., Sala, M. R., Signorelli, S., Venturi, N., Signorini, S., & Brambilla, P. (2012). Bacteriuria screening by automated whole-field-image-based microscopy reduces the number of necessary urine cultures. *J Clin Microbiol*, *50*(4), 1427–1429.

Feehery, G. R., Yigit, E., Oyola, S. O., Langhorst, B. W., Schmidt, V. T., Stewart, F. J., … Pradhan, S. (2013). A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. *PLoS One*, *8*(10), e76096. https://doi.org/10.1371/journal.pone.0076096

Felt, J. R., Yurkovich, C., Garshott, D. M., Kamat, D., Farooqi, A., Fribley, A. M., … Hebert, K. (2017). The Utility of Real-Time Quantitative Polymerase Chain Reaction Genotype Detection in the Diagnosis of Urinary Tract Infections in Children. *Clin Pediatr (Phila)*, *56*(10), 912–919. https://doi.org/10.1177/0009922817706144

Fenollar, F., & Raoult, D. (2007). Molecular diagnosis of bloodstream infections

caused by non-cultivable bacteria. *International Journal of Antimicrobial Agents*, *30*, 7–15.

Ferreira, L., Sánchez-Juanes, F., González-Ávila, M., Cembrero-Fuciños, D., Herrero-Hernández, A., González-Buitrago, J. M., & Muñoz-Bellido, J. L. (2010). Direct identification of urinary tract pathogens from urine samples by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol*, *48*(6), 2110–2115.

Ferrer, R., Martin-Loeches, I., Phillips, G., Osborn, T. M., Townsend, S., Dellinger, R. P., … Levy, M. M. (2014). Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. *Critical Care Medicine*, *42*(8), 1749–1755.

Fichtel, J., Köster, J., Rullkötter, J., & Sass, H. (2007). Spore dipicolinic acid contents used for estimating the number of endospores in sediments. *FEMS Microbiology Ecology*, *61*(3), 522–532.

Fleischmann, C., Scherag, A., Adhikari, N. K. J., Hartog, C. S., Tsaganos, T., Schlattmann, P., … Reinhart, K. (2016). Assessment of global incidence and mortality of hospital-treated sepsis. Current estimates and limitations. *American Journal of Respiratory and Critical Care Medicine*, *193*(3), 259–272.

Fournier, P.-E., Drancourt, M., Colson, P., Rolain, J.-M., La Scola, B., & Raoult, D. (2013). Modern clinical microbiology: new challenges and solutions. *Nature Reviews Microbiology*, *11*(8), 574–585.

Fournier, P., Drancourt, M., … P. C.-N. R., & 2013, undefined. (n.d.). Modern clinical microbiology: new challenges and solutions. *Nature.Com*. Retrieved from https://www.nature.com/articles/nrmicro3068

Fox, J. L. (2014). Technology comes to typing. *Nat Biotechnol*, *32*(11), 1081–1084. https://doi.org/10.1038/nbt.3067

Foxman, B. (2010). The epidemiology of urinary tract infection. *Nature Reviews Urology*, *7*(12), 653.

Frith, M. C., Hamada, M., & Horton, P. (2010). Parameters for accurate genome alignment. *BMC Bioinformatics*, *11*(1), 80.

Fukushima, H., Tsunomori, Y., & Seki, R. (2003). Duplex real-time SYBR green PCR assays for detection of 17 species of food-or waterborne pathogens in stools. *J Clin Microbiol*, *41*(11), 5134–5146.

Gadsby, N. J., Russell, C. D., McHugh, M. P., Mark, H., Conway Morris, A., Laurenson, I. F., … Templeton, K. E. (2016). Comprehensive molecular testing for respiratory pathogens in community-acquired pneumonia. *Clinical Infectious Diseases*, *62*(7), 817–823.

Gaillot, O., Di Camillo, P., Berche, P., Courcol, R., & Savage, C. (1999). Comparison of CHROMagar Salmonella medium and Hektoen enteric agar for isolation of salmonellae from stool samples. *J Clin Microbiol*, *37*(3), 762–765.

Gangaram, R., Ojwang, P. J., Moodley, J., & Maharaj, D. (2005). The accuracy of urine dipsticks as a screening test for proteinuria in hypertensive disorders of pregnancy. *Hypertension in Pregnancy*, *24*(2), 117–123. https://doi.org/10.1081/PRG-200059849

Garalde, D. R., Snell, E. A., Jachimowicz, D., Sipos, B., Lloyd, J. H., Bruce, M., … Warland, A. (2018). Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods*.

García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L. M., Götz, S., Tarazona, S., … Conesa, A. (2012). Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics*, *28*(20), 2678–2679.

Garnacho-Montero, J., Gutierrez-Pizarraya, A., Escoresca-Ortega, A., Corcia-Palomo, Y., Fernandez-Delgado, E., Herrera-Melero, I., … Marquez-Vacaro, J. A. (2014). De-escalation of empirical therapy is associated with lower mortality in patients with severe sepsis and septic shock. *Intensive Care Med*, *40*(1), 32–40. https://doi.org/10.1007/s00134-013-3077-7

Gebert, S., Siegel, D., & Wellinghausen, N. (2008). Rapid detection of pathogens in blood culture bottles by real-time PCR in conjunction with the pre-analytic tool MolYsis. *Journal of Infection*, *57*(4), 307–316.

Gerner-Smidt, P., Hise, K., Kincaid, J., Hunter, S., Rolando, S., Hyytiä-Trees, E., … Taskforce, P. (2006). PulseNet USA: a five-year update. *Foodbourne Pathogens & Disease*, *3*(1), 9–19.

Gerritzen, A., Wittke, J.-W., & Wolff, D. (2011). Rapid and Sensitive Detection of Shiga Toxin-Producing Escherichia Coli Directly From Stool Samples by Real-Time PCR in Comparison to Culture, Enzyme Immunoassay, and Vero Cell Cytotoxicity Assay. *Clinical Laboratory*, *57*(11), 993.

Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B., Jorden, J. R., & Chiodini, R. J. (2015). Changes in 16S RNA gene microbial community profiling by

concentration of prokaryotic DNA. *J Microbiol Methods*, *119*, 239–242.

Goldberg, B., Sichtig, H., Geyer, C., Ledeboer, N., & Weinstock, G. M. (2015). Making the leap from research laboratory to clinic: challenges and opportunities for next-generation sequencing in infectious disease diagnostics. *MBio*, *6*(6), e01888-15.

Goldenberg, S. D., & French, G. L. (2011). Diagnostic testing for Clostridium difficile: a comprehensive survey of laboratories in England. *Journal of Hospital Infection*, *79*(1), 4–7.

Goldenberg, S. D., Price, N. M., Tucker, D., Wade, P., & French, G. L. (2011). Mandatory reporting and improvements in diagnosing Clostridium difficile infection: an incompatible dichotomy? *Journal of Infection*, *62*(5), 363–370.

Goodwin, S., Wappel, R., & McCombie, W. R. (2017). 1D Genome Sequencing on the Oxford Nanopore MinION. *Current Protocols in Human Genetics*, 18.11. 1-18.11. 14.

Goswitz, J. J., Willard, K. E., Eastep, S. J., Shanholtzer, C. J., Olson, M. L., Pinnell, M., … Peterson, L. R. (1993). Utility of slide centrifuge Gram's stain versus quantitative culture for diagnosis of urinary tract infection. *American Journal of Clinical Pathology*, *99*(2), 132–136.

Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., … Chiu, C. Y. (2015). Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med*, *7*, 99. https://doi.org/10.1186/s13073-015-0220-9

Grice, E. A., Kong, H. H., Renaud, G., Young, A. C., Bouffard, G. G., Blakesley, R. W., … Segre, J. A. (2008). A diversity profile of the human skin microbiota. *Genome Res*, *18*(7), 1043–1050. https://doi.org/10.1101/gr.075549.107

Grumaz, S., Stevens, P., Grumaz, C., Decker, S. O., Weigand, M. A., Hofer, S., … Sohn, K. (2016). Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Med*, *8*(1), 73. https://doi.org/10.1186/s13073-016-0326-8

Guido, M., Tumolo, M. R., De Donno, A., Verri, T., Serio, F., Bagordo, F., & Zizza, A. (2016). In vitro diagnosis of sepsis: a review. *PATHOLOGY AND LABORATORY MEDICINE INTERNATIONAL*, *8*, 1–14.

Gurtler, V. (1993). Typing of Clostridium difficile strains by PCR-amplification of variable length 16S-23S rDNA spacer regions. *Journal of General Microbiology*,

139(12), 3089–3097. https://doi.org/10.1099/00221287-139-12-3089

Gyarmati, P., Kjellander, C., Aust, C., Song, Y., Öhrmalm, L., & Giske, C. G. (2016). Metagenomic analysis of bloodstream infections in patients with acute leukemia and therapy-induced neutropenia. *Sci Rep*, *6*.

Haag, H., Locher, F., & Nolte, O. (2013). Molecular diagnosis of microbial aetiologies using SepsiTest™ in the daily routine of a diagnostic laboratory. *Diagnostic Microbiology and Infectious Disease*, *76*(4), 413–418.

Hansen, W. L. J., Bruggeman, C. A., & Wolffs, P. F. G. (2009). Evaluation of new preanalysis sample treatment tools and DNA isolation protocols to improve bacterial pathogen detection in whole blood. *J Clin Microbiol*, *47*(8), 2629–2631.

Harris, D. M., & Hata, D. J. (2013). Rapid identification of bacteria and Candida using PNA-FISH from blood and peritoneal fluid cultures: a retrospective clinical study. *Annals of Clinical Microbiology and Antimicrobials*, *12*(1), 2.

Hasman, H., Saputra, D., Sicheritz-Ponten, T., Lund, O., Svendsen, C. A., Frimodt-Moller, N., & Aarestrup, F. M. (2014). Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. *J Clin Microbiol*, *52*(1), 139–146. https://doi.org/10.1128/JCM.02452-13

Hay, A., Birnie, K., Busby, J., Delaney, B., Downing, H., Dudley, J., … Hollingworth, W. (2016). The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute i.

Hayden, R. T., Clinton, L. K., Hewitt, C., Koyamatsu, T., Sun, Y., Jamison, G., … Bankowski, M. J. (2016). Rapid antimicrobial susceptibility testing using forward laser light scatter technology. *J Clin Microbiol*, *54*(11), 2701–2706.

He, M., Miyajima, F., Roberts, P., Ellison, L., Pickard, D. J., Martin, M. J., … Lawley, T. D. (2013). Emergence and global spread of epidemic healthcare-associated Clostridium difficile. *Nat Genet*, *45*(1), 109–113. https://doi.org/10.1038/ng.2478

Hevia, A., Delgado, S., Margolles, A., & Sanchez, B. (2015). Application of density gradient for the isolation of the fecal microbial stool component and the potential use thereof. *Sci Rep*, *5*, 16807. https://doi.org/10.1038/srep16807

Hewitt, F. C., Guertin, S. L., Ternus, K. L., Schulte, K., & Kadavy, D. R. (2017). Toward Rapid Sequenced-Based Detection and Characterization of Causative Agents of Bacteremia. *BioRxiv*, 162735.

Hill, J. T., Tran, K.-D. T., Barton, K. L., Labreche, M. J., & Sharp, S. E. (2014). Evaluation of the nanosphere Verigene BC-GN assay for direct identification of gram-negative bacilli and antibiotic resistance markers from positive blood cultures and potential impact for more-rapid antibiotic interventions. *J Clin Microbiol*, *52*(10), 3805–3807.

Holt, K. E., Phan, M. D., Baker, S., Duy, P. T., Nga, T. V, Nair, S., … Wain, J. (2011). Emergence of a globally dominant IncHI1 plasmid type associated with multiple drug resistant typhoid. *PLoS Negl Trop Dis*, *5*(7), e1245. https://doi.org/10.1371/journal.pntd.0001245

Horz, H.-P., Scheer, S., Huenger, F., Vianna, M. E., & Conrads, G. (2008). Selective isolation of bacterial DNA from human clinical specimens. *J Microbiol Methods*, *72*(1), 98–102.

Horz, H. P., Scheer, S., Vianna, M. E., & Conrads, G. (2010). New methods for selective isolation of bacterial DNA from human clinical specimens. *Anaerobe*, *16*(1), 47–53. https://doi.org/10.1016/j.anaerobe.2009.04.009

Hosokawa, M., Nishikawa, Y., Kogawa, M., & Takeyama, H. (2017). Massively parallel whole genome amplification for single-cell sequencing using droplet microfluidics. *Sci Rep*, *7*(1), 5199.

Huang, T. D., Bogaerts, P., Ghilani, E., Heinrichs, A., Gavage, P., Roisin, S., … Glupczynski, Y. (2015). Multicentre evaluation of the Check-Direct CPE(R) assay for direct screening of carbapenemase-producing Enterobacteriaceae from rectal swabs. *J Antimicrob Chemother*, *70*(6), 1669–1673. https://doi.org/10.1093/jac/dkv009

Huber, C. A., Foster, N. F., Riley, T. V, & Paterson, D. L. (2013). Challenges for standardization of Clostridium difficile typing methods. *J Clin Microbiol*, *51*(9), 2810–2814.

Ip, C. L. C., Loose, M., Tyson, J. R., de Cesare, M., Brown, B. L., Jain, M., … Urban, J. M. (2015). MinION Analysis and Reference Consortium: Phase 1 data release and analysis. *F1000Res*, *4*.

Ipe, D. S., Sundac, L., Benjamin Jr, W. H., Moore, K. H., & Ulett, G. C. (2013). Asymptomatic bacteriuria: prevalence rates of causal microorganisms, etiology of infection in different patient populations, and recent advances in molecular detection. *FEMS Microbiol Lett*, *346*(1), 1–10.

Iyengar, V., Albaugh, G. P., Lohani, A., & Nair, P. P. (1991). Human stools as a

source of viable colonic epithelial cells. *Faseb J*, *5*(13), 2856–2859.

Jain, M., Fiddes, I. T., Miga, K. H., Olsen, H. E., Paten, B., & Akeson, M. (2015). Improved data analysis for the MinION nanopore sequencer. *Nat Methods*, *12*(4), 351–356.

Jain, M., Koren, S., Quick, J., Rand, A. C., Sasani, T. A., Tyson, J. R., … Malla, S. (2017). Nanopore sequencing and assembly of a human genome with ultra-long reads. *BioRxiv*, 128835.

Jorgensen, J. H., & Ferraro, M. J. (2009). Antimicrobial susceptibility testing: a review of general principles and contemporary practices. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *49*(11), 1749–1755. https://doi.org/10.1086/647952

Juul, S., Izquierdo, F., Hurst, A., Dai, X., Wright, A., Kulesha, E., … Turner, D. J. (2015). What9s in my pot? Real-time species identification on the MinION. *BioRxiv*, 30742.

Kachrimanidou, M., & Malisiovas, N. (2011). Clostridium difficile infection: a comprehensive review. *Critical Reviews in Microbiology*, *37*(3), 178–187.

Karlsson, E., Larkeryd, A., Sjodin, A., Forsman, M., & Stenberg, P. (2015). Scaffolding of a bacterial genome using MinION nanopore sequencing. *Sci Rep*, *5*, 11996. https://doi.org/10.1038/srep11996

Kellogg, J. A., Manzella, J. P., & Bankert, D. A. (2000). Frequency of low-level bacteremia in children from birth to fifteen years of age. *J Clin Microbiol*, *38*(6), 2181–2185.

Kelly, MD, C. P., & LaMont, MD, J. T. (1998). *CLOSTRIDIUM DIFFICILE* INFECTION. *Annual Review of Medicine*, *49*(1), 375–390. https://doi.org/10.1146/annurev.med.49.1.375

Kiełbasa, S. M., Wan, R., Sato, K., Horton, P., & Frith, M. C. (2011). Adaptive seeds tame genomic sequence comparison. *Genome Res*, *21*(3), 487–493.

Killgore, G., Thompson, A., Johnson, S., Brazier, J., Kuijper, E., Pepin, J., … van den Berg, R. J. (2008). Comparison of seven techniques for typing international epidemic strains of Clostridium difficile: restriction endonuclease analysis, pulsed-field gel electrophoresis, PCR-ribotyping, multilocus sequence typing, multilocus variable-number tandem-repeat an. *J Clin Microbiol*, *46*(2), 431–437.

Knetsch, C. W., Lawley, T. D., Hensgens, M. P., Corver, J., Wilcox, M. W., & Kuijper, E. J. (2013). Current application and future perspectives of molecular typing

methods to study Clostridium difficile infections. *Euro Surveill*, *18*(4), 20381.

Kohn, A. (1960). Lysis of frozen and thawed cells of Escherichia coli by lysozyme, and their conversion into spheroplasts. *J Bacteriol*, *79*(5), 697.

Kollef, M. H. (2000). Inadequate antimicrobial treatment: an important determinant of outcome for hospitalized patients. *Clin Infect Dis*, *31 Suppl 4*, S131-8. https://doi.org/10.1086/314079

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/gr.215087.116

Köser, C. U., Ellington, M. J., Cartwright, E. J. P., Gillespie, S. H., Brown, N. M., Farrington, M., … Parkhill, J. (2012). Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathogens*, *8*(8), e1002824.

Krcmery, S., Hromec, J., & Demesova, D. (2001). Treatment of lower urinary tract infection in pregnancy. *International Journal of Antimicrobial Agents*, *17*(4), 279–282.

Kuczynski, J., Lauber, C. L., Walters, W. A., Parfrey, L. W., Clemente, J. C., Gevers, D., & Knight, R. (2012). Experimental and analytical tools for studying the human microbiome. *Nature Reviews Genetics*, *13*(1), 47–58.

Kuhn, C., Disque, C., Muhl, H., Orszag, P., Stiesch, M., & Haverich, A. (2011). Evaluation of commercial universal rRNA gene PCR plus sequencing tests for identification of bacteria and fungi associated with infectious endocarditis. *J Clin Microbiol*, *49*(8), 2919–2923. https://doi.org/10.1128/jcm.00830-11

Kuijper, E. J., Coignard, B., & Tüll, P. (2006). Emergence of Clostridium difficile-associated disease in North America and Europe. *Clinical Microbiology and Infection*, *12*(s6), 2–18.

Kumar, A., Ellis, P., Arabi, Y., Roberts, D., Light, B., Parrillo, J. E., … Chateau, D. (2009). Initiation of inappropriate antimicrobial therapy results in a fivefold reduction of survival in human septic shock. *Chest*, *136*(5), 1237–1248. https://doi.org/10.1378/chest.09-0087

Kumar, A., Roberts, D., Wood, K. E., Light, B., Parrillo, J. E., Sharma, S., … Cheang, M. (2006). Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit Care*

Med, *34*(6), 1589–1596. https://doi.org/10.1097/01.CCM.0000217961.75225.E9

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.

Lawley, T. D., Clare, S., Walker, A. W., Goulding, D., Stabler, R. A., Croucher, N., … Dougan, G. (2009). Antibiotic treatment of clostridium difficile carrier mice triggers a supershedder state, spore-mediated transmission, and severe disease in immunocompromised hosts. *Infect Immun*, *77*(9), 3661–3669. https://doi.org/10.1128/iai.00558-09

Leffler, D. A., & Lamont, J. T. (2015). Clostridium difficile infection. *New England Journal of Medicine*, *372*(16), 1539–1548.

Lehmann, L. E., Hauser, S., Malinka, T., Klaschik, S., Weber, S. U., Schewe, J.-C., … Book, M. (2011). Rapid qualitative urinary tract infection pathogen identification by SeptiFast® real-time PCR. *PLoS One*, *6*(2), e17146.

Lehmann, L. E., Hunfeld, K.-P., Steinbrucker, M., Brade, V., Book, M., Seifert, H., … Stüber, F. (2010). Improved detection of blood stream pathogens by real-time PCR in severe sepsis. *Intensive Care Medicine*, *36*(1), 49–56.

Lessa, F. C., Mu, Y., Bamberg, W. M., Beldavs, Z. G., Dumyati, G. K., Dunn, J. R., … Phipps, E. C. (2015). Burden of Clostridium difficile infection in the United States. *New England Journal of Medicine*, *372*(9), 825–834.

Li, B., Qiu, Y., Glidle, A., McIlvenna, D., Luo, Q., Cooper, J., … Yin, H. (2014). Gradient microfluidics enables rapid bacterial growth inhibition testing. *Analytical Chemistry*, *86*(6), 3131–3137. https://doi.org/10.1021/ac5001306

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv Preprint ArXiv:1303.3997*.

Li, H. (2016). Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics*, *32*(14), 2103–2110.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079.

Liesenfeld, O., Lehman, L., Hunfeld, K. P., & Kost, G. (2014). Molecular diagnosis of sepsis: New aspects and recent developments. *Eur J Microbiol Immunol (Bp)*, *4*(1), 1–25. https://doi.org/10.1556/EuJMI.4.2014.1.1

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., …

Ecker, J. R. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, *462*(7271), 315–322. https://doi.org/10.1038/nature08514

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., … Law, M. (2012). Comparison of next-generation sequencing systems. *BioMed Research International*, *2012*.

Livermore, D. M., & Wain, J. (2013). Revolutionising bacteriology to improve treatment outcomes and antibiotic stewardship. *Infect Chemother*, *45*(1), 1–10. https://doi.org/10.3947/ic.2013.45.1.1

Lofgren, E. T., Cole, S. R., Weber, D. J., Anderson, D. J., & Moehring, R. W. (2014). Hospital-Acquired Clostridium difficile Infections Estimating All-Cause Mortality and Length of Stay. *Epidemiology (Cambridge, Mass.)*, *25*(4), 570.

Loman, N. J., Quick, J., & Simpson, J. T. (2015). A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat Methods*, *12*(8), 733–735. https://doi.org/10.1038/nmeth.3444

Loman, N. J., & Quinlan, A. R. (2014). Poretools: a toolkit for analyzing nanopore sequence data. *Bioinformatics*, *30*(23), 3399–3401.

Longtin, Y., Trottier, S., Brochu, G., Paquet-Bolduc, B., Garenc, C., Loungnarath, V., … Longtin, J. (2012). Impact of the type of diagnostic assay on Clostridium difficile infection and complication rates in a mandatory reporting program. *Clinical Infectious Diseases*, *56*(1), 67–73.

Loonen, A. J., Bos, M. P., van Meerbergen, B., Neerken, S., Catsburg, A., Dobbelaer, I., … van den Brule, A. J. (2013). Comparison of pathogen DNA isolation methods from large volumes of whole blood to improve molecular diagnosis of bloodstream infections. *PLoS One*, *8*(8), e72349. https://doi.org/10.1371/journal.pone.0072349

Loose, M. W. (2017). The potential impact of nanopore sequencing on human genetics. *Human Molecular Genetics*, *26*(R2), R202–R207.

López-Pérez, M., Gonzaga, A., Martin-Cuadrado, A. B., Onyshchenko, O., Ghavidel, A., Ghai, R., & Rodriguez-Valera, F. (2012). Genomes of surface isolates of Alteromonas macleodii: The life of a widespread marine opportunistic copiotroph. *Scientific Reports*, *2*. https://doi.org/10.1038/srep00696

Lucke, K., Hombach, M., Hug, M., & Pfyffer, G. E. (2010). Rapid detection of methicillin-resistant Staphylococcus aureus (MRSA) in diverse clinical specimens by the BD GeneOhm MRSA assay and comparison with culture. *J*

Clin Microbiol, *48*(3), 981–984.

Maddocks, S., Olma, T., & Chen, S. (2002). Comparison of CHROMagar Salmonella medium and xylose-lysine-desoxycholate and Salmonella-Shigella agars for isolation of Salmonella strains from stool samples. *J Clin Microbiol*, *40*(8), 2999–3003.

Mancini, N., Carletti, S., Ghidoli, N., Cichero, P., Burioni, R., & Clementi, M. (2010). The era of molecular and other non-culture-based methods in diagnosis of sepsis. *Clin Microbiol Rev*, *23*(1), 235–251. https://doi.org/10.1128/CMR.00043-09

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., … Marteau, P. (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut*, *55*(2), 205–211.

Manoni, F., Fornasiero, L., Ercolin, M., Tinello, A., Ferrian, M., Hoffer, P., … Gessoni, G. (2009). Cutoff values for bacteria and leukocytes for urine flow cytometer Sysmex UF-1000i in urinary tract infections. *Diagnostic Microbiology and Infectious Disease*, *65*(2), 103–107.

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, *24*(3), 133–141.

Marler, L. M., Siders, J. A., Wolters, L. C., Pettigrew, Y., Skitt, B. L., & Allen, S. D. (1992). Comparison of five cultural procedures for isolation of clostridium difficile from stools. *Journal of Clinical Microbiology*.

McArthur, A. G., Waglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A. J., … Ejim, L. (2013). The comprehensive antibiotic resistance database. *Antimicrobial Agents and Chemotherapy*, *57*(7), 3348–3357.

McCann, C. D., & Jordan, J. A. (2014). Evaluation of MolYsis Complete5 DNA extraction method for detecting Staphylococcus aureus DNA from whole blood in a sepsis model using PCR/pyrosequencing. *J Microbiol Methods*, *99*, 1–7. https://doi.org/10.1016/j.mimet.2014.01.013

McPherson, D., Griffiths, C., Williams, M., Baker, A., Klodawski, E., Jacobson, B., & Donaldson, L. (2013). Sepsis-associated mortality in England: an analysis of multiple cause of death data from 2001 to 2010. *BMJ Open*, *3*(8). https://doi.org/10.1136/bmjopen-2013-002586

Mellmann, A., Cloud, J., Maier, T., Keckevoet, U., Ramminger, I., Iwen, P., … Harmsen, D. (2008). Evaluation of matrix-assisted laser desorption ionization-

time-of-flight mass spectrometry in comparison to 16S rRNA gene sequencing for species identification of nonfermenting bacteria. *Journal of Clinical Microbiology*, *46*(6), 1946–1954. https://doi.org/10.1128/JCM.00157-08

Melnikov, A., Galinsky, K., Rogov, P., Fennell, T., Van Tyne, D., Russ, C., … Neafsey, D. E. (2011). Hybrid selection for sequencing pathogen genomes from clinical samples. *Genome Biol*, *12*(8), R73. https://doi.org/10.1186/gb-2011-12-8-r73

Melzer, M., & Welch, C. (2013). Outcomes in UK patients with hospital-acquired bacteraemia and the risk of catheter-associated urinary tract infections. *Postgraduate Medical Journal*, *89*(1052), 329–334.

Menozzi, M. G., Eigner, U., Covan, S., Rossi, S., Somenzi, P., Dettori, G., … Fahr, A. M. (2006). Two-center collaborative evaluation of performance of the BD phoenix automated microbiology system for identification and antimicrobial susceptibility testing of gram-negative bacteria. *Journal of Clinical Microbiology*, *44*(11), 4085–4094. https://doi.org/10.1128/JCM.00614-06

Moir, A., Corfe, B. M., & Behravan, J. (2002). Spore germination. *Cellular and Molecular Life Sciences*, *59*(3), 403–409.

Morris, K., Wilson, C., & Wilcox, M. H. (2012). Evaluation of chromogenic meticillin-resistant Staphylococcus aureus media: sensitivity versus turnaround time. *Journal of Hospital Infection*, *81*(1), 20–24.

Muhl, H., Kochem, A. J., Disque, C., & Sakka, S. G. (2010). Activity and DNA contamination of commercial polymerase chain reaction reagents for the universal 16S rDNA real-time polymerase chain reaction detection of bacterial pathogens in blood. *Diagn Microbiol Infect Dis*, *66*(1), 41–49. https://doi.org/10.1016/j.diagmicrobio.2008.07.011

Murray, C., Adeyiga, O., Owsley, K., & Di Carlo, D. (2015). Research highlights: microfluidic analysis of antimicrobial susceptibility. *Lab on a Chip*, *15*(5), 1226–1229. https://doi.org/10.1039/c5lc90017d

Mwaigwisya, S., Assiri, R. A., & O'Grady, J. (2015). Emerging commercial molecular tests for the diagnosis of bloodstream infection. *Expert Rev Mol Diagn*, *15*(5), 681–692. https://doi.org/10.1586/14737159.2015.1029459

Naccache, S. N., Federman, S., Veeraraghavan, N., Zaharia, M., Lee, D., Samayoa, E., … Enge, B. (2014). A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples.

Genome Res, *24*(7), 1180–1192.

Nair, S., Ashton, P., Doumith, M., Connell, S., Painset, A., Mwaigwisya, S., … Day, M. (2016). WGS for surveillance of antimicrobial resistance: a pilot study to detect the prevalence and mechanism of resistance to azithromycin in a UK population of non-typhoidal Salmonella. *Journal of Antimicrobial Chemotherapy*, *71*(12), 3400–3408.

Neely, L. A., Audeh, M., Phung, N. A., Min, M., Suchocki, A., Plourde, D., … Lowery, T. J. (2013). T2 magnetic resonance enables nanoparticle-mediated rapid detection of candidemia in whole blood. *Sci Transl Med*, *5*(182), 182ra54. https://doi.org/10.1126/scitranslmed.3005377

Norling, M., Karlsson-Lindsjö, O. E., Gourlé, H., Bongcam-Rudloff, E., & Hayer, J. (2016). MetLab: An in silico experimental design, simulation and analysis tool for viral metagenomics studies. *PLoS ONE*, *11*(8). https://doi.org/10.1371/journal.pone.0160334

O'Neill, J. (2014). Antimicrobial Resistance : Tackling a crisis for the health and wealth of nations. *Review on Antimicrobial Resistance*, (December), 1–16. https://doi.org/10.1038/510015a

Opota, O., Jaton, K., & Greub, G. (2015). Microbial diagnosis of bloodstream infection: Towards molecular diagnosis directly from blood. *Clinical Microbiology and Infection*. https://doi.org/10.1016/j.cmi.2015.02.005

Oyola, S. O., Gu, Y., Manske, M., Otto, T. D., O'Brien, J., Alcock, D., … Quail, M. A. (2013). Efficient depletion of host DNA contamination in malaria clinical sequencing. *J Clin Microbiol*, *51*(3), 745–751. https://doi.org/10.1128/jcm.02507-12

Pancholi, P., Kelly, C., Raczkowski, M., & Balada-Llasat, J. M. (2012). Detection of toxigenic Clostridium difficile: comparison of the cell culture neutralization, Xpert C. difficile, Xpert C. difficile/Epi, and Illumigene C. difficile assays. *J Clin Microbiol*, *50*(4), 1331–1335.

Paredes-Sabja, D., Shen, A., & Sorg, J. A. (2014). Clostridium difficile spore biology: sporulation, germination, and spore structural proteins. *Trends in Microbiology*, *22*(7), 406–416.

Patel, H. D., Livsey, S. A., Swann, R. A., & Bukhari, S. S. (2005). Can urine dipstick testing for urinary tract infection at point of care reduce laboratory workload? *Journal of Clinical Pathology*, *58*(9), 951–954.

Patel, J. B. (2001). 16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. *Molecular Diagnosis*, *6*(4), 313–321.

Paulino, L. C., Tseng, C. H., Strober, B. E., & Blaser, M. J. (2006). Molecular analysis of fungal microbiota in samples from healthy human skin and psoriatic lesions. *J Clin Microbiol*, *44*(8), 2933–2941. https://doi.org/10.1128/jcm.00785-06

Penders, J., Vink, C., Driessen, C., London, N., Thijs, C., & Stobberingh, E. E. (2005). Quantification of Bifidobacterium spp., Escherichia coli and Clostridium difficile in faecal samples of breast-fed and formula-fed infants by real-time PCR. *FEMS Microbiol Lett*, *243*(1), 141–147. https://doi.org/10.1016/j.femsle.2004.11.052

Perry-O'Keefe, H., Rigby, S., Oliveira, K., Sørensen, D., Stender, H., Coull, J., & Hyldig-Nielsen, J. J. (2001). Identification of indicator microorganisms using a standardized PNA FISH method. *Journal of Microbiological Methods*, *47*(3), 281–292. https://doi.org/10.1016/S0167-7012(01)00303-7

Pezzlo, M. (1988). Detection of urinary tract infections by rapid methods. *Clinical Microbiology Reviews*, *1*(3), 268–280.

Poretsky, R., Rodriguez-R, L. M., Luo, C., Tsementzi, D., & Konstantinidis, K. T. (2014). Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. *PLoS One*, *9*(4), e93827.

Putnam, L. R., Howard, W. J., Pfaller, M. A., Koontz, F. P., & Jones, R. N. (1997). Accuracy of the vitek system for antimicrobial susceptibility testing enterobacteriaceae bloodstream infection isolates: Use of "direct" inoculation from bactec 9240 blood culture bottles. *Diagnostic Microbiology and Infectious Disease*, *28*(2), 101–104. https://doi.org/10.1016/S0732-8893(97)00014-X

Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., … Mikhail, A. (2016). Real-time, portable genome sequencing for Ebola surveillance. *Nature*, *530*(7589), 228–232.

Quick, J., Quinlan, A. R., & Loman, N. J. (2014). A reference bacterial genome dataset generated on the MinION portable single-molecule nanopore sequencer. *Gigascience*, *3*, 22. https://doi.org/10.1186/2047-217x-3-22

Raoult, D., Fournier, P. E., & Drancourt, M. (2004). What does the future hold for clinical microbiology? *Nature Reviews Microbiology*, *2*(2), 151–159. https://doi.org/Doi 10.1038/Nrmicro820

Ray, B., & Speck, M. L. (1973). Freeze-injury in bacteria. *CRC Critical Reviews in Clinical Laboratory Sciences*, *4*(2), 161–213.

Reece, A., Xia, B., Jiang, Z., Noren, B., McBride, R., & Oakey, J. (2016). Microfluidic techniques for high throughput single cell analysis. *Current Opinion in Biotechnology*. https://doi.org/10.1016/j.copbio.2016.02.015

Resistance, R. on A. (2016). *Tackling drug-resistant infections globally: final report and recommendations*. Review on Antimicrobial Resistance.

Rhoads, D. D., Sintchenko, V., Rauch, C. A., & Pantanowitz, L. (2014). Clinical microbiology informatics. *Clin Microbiol Rev*, *27*(4), 1025–1047. https://doi.org/10.1128/cmr.00049-14

Rhodes, A., Evans, L. E., Alhazzani, W., Levy, M. M., Antonelli, M., Ferrer, R., … Dellinger, R. P. (2017). Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. *Intensive Care Medicine*, *43*(3), 304–377. https://doi.org/10.1007/s00134-017-4683-6

Richards, M., Knox, J., Elliott, B., Mackin, K., Lyras, D., Waring, L. J., & Riley, T. V. (2011). Severe infection with Clostridium difficile PCR ribotype 027 acquired in Melbourne, Australia. *Medical Journal of Australia*, *194*(7), 369–371.

Riley, T. V, Brazier, J. S., Hassan, H., Williams, K., & Phillips, K. D. (1987). Comparison of alcohol shock enrichment and selective enrichment for the isolation of Clostridium difficile. *Epidemiology & Infection*, *99*(2), 355–359.

Rivers, E., Nguyen, B., Havstad, S., Ressler, J., Muzzin, A., Knoblich, B., … Early Goal-Directed Therapy Collaborative, G. (2001). Early goal-directed therapy in the treatment of severe sepsis and septic shock. *N Engl J Med*, *345*(19), 1368–1377. https://doi.org/10.1056/NEJMoa010307

Roberts, K. B. (2011). Urinary tract infection: clinical practice guideline for the diagnosis and management of the initial UTI in febrile infants and children 2 to 24 months. Am Acad Pediatrics.

Rogol, M., Shpak, B., Rothman, D., & Sechter, I. (1985). Enrichment medium for isolation of Campylobacter jejuni-Campylobacter coli. *Applied and Environmental Microbiology*, *50*(1), 125–126.

Rose, G., Wooldridge, D. J., Anscombe, C., Mee, E. T., Misra, R. V., & Gharbia, S. (2015). Challenges of the Unknown: Clinical Application of Microbial Metagenomics. *International Journal of Genomics*, *2015*. https://doi.org/10.1155/2015/292950

Sabat, A. J., Budimir, A., Nashev, D., Sá-Leão, R., Van Dijl, J. M., Laurent, F., … Markers, E. S. G. of E. (2013). Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill*, *18*(4), 20380.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., … Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*. https://doi.org/10.1126/science.2448875

Salipante, S. J., Sengupta, D. J., Hoogestraat, D. R., Cummings, L. A., Bryant, B. H., Natividad, C., … Hoffman, N. G. (2013). Molecular diagnosis of Actinomadura madurae infection by 16S rRNA deep sequencing. *J Clin Microbiol*, *51*(12), 4262–4265. https://doi.org/10.1128/jcm.02227-13

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., … Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol*, *12*(1), 87.

Schappert, S. M., & Rechtsteiner, E. A. (2011). Ambulatory medical care utilization estimates for 2007. *Vital and Health Statistics. Series 13, Data from the National Health Survey*, (169), 1–38.

Schlaberg, R., Chiu, C. Y., Miller, S., Procop, G. W., Weinstock, G., Committee, P. P., … Pathologists, M. R. C. of the C. of A. (2017). Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Archives of Pathology and Laboratory Medicine*, *141*(6), 776–786.

Schmidt, K., Mwaigwisya, S., Crossman, L. C., Doumith, M., Munroe, D., Pires, C., … Wain, J. (2016). Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *Journal of Antimicrobial Chemotherapy*, *72*(1), 104–114.

Schmieder, R., & Edwards, R. (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One*, *6*(3), e17288.

Schmiemann, G., Kniehl, E., Gebhardt, K., Matejczyk, M. M., & Hummers-Pradier, E. (2010). The diagnosis of urinary tract infection: a systematic review. *Deutsches Ärzteblatt International*, *107*(21), 361.

Schoepp, N. G., Khorosheva, E. M., Schlappi, T. S., Curtis, M. S., Humphries, R. M., Hindler, J. A., & Ismagilov, R. F. (2016). Digital Quantification of DNA

Replication and Chromosome Segregation Enables Determination of Antimicrobial Susceptibility after only 15 Minutes of Antibiotic Exposure. *Angewandte Chemie - International Edition*, *55*(33), 9557–9561. https://doi.org/10.1002/anie.201602763

Schröder, U.-C., Bokeloh, F., O'Sullivan, M., Glaser, U., Wolf, K., Pfister, W., … Neugebauer, U. (2015). Rapid, culture-independent, optical diagnostics of centrifugally captured bacteria from urine samples. *Biomicrofluidics*, *9*(4), 44118.

Schulte, W., Bernhagen, J., & Bucala, R. (2013). Cytokines in sepsis: potent immunoregulators and potential therapeutic targets—an updated view. *Mediators of Inflammation*, *2013*.

Semeniuk, H., & Church, D. (1999). Evaluation of the leukocyte esterase and nitrite urine dipstick screening tests for detection of bacteriuria in women with suspected uncomplicated urinary tract infections. *Journal of Clinical Microbiology*, *37*(9), 3051–3052.

Sender, R., Fuchs, S., & Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol*, *14*(8), e1002533. https://doi.org/10.1371/journal.pbio.1002533

Seng, P., Abat, C., Rolain, J. M., Colson, P., Lagier, J.-C., Gouriet, F., … Raoult, D. (2013). Identification of rare pathogenic bacteria in a clinical microbiology laboratory: impact of MALDI-TOF mass spectrometry. *J Clin Microbiol*, JCM. 00492-13.

Shigemura, K., Shirakawa, T., Okada, H., Tanaka, K., Kamidono, S., Arakawa, S., & Gotoh, A. (2005). Rapid detection and differentiation of Gram-negative and Gram-positive pathogenic bacteria in urine using TaqMan probe. *Clinical and Experimental Medicine*, *4*(4), 196–201.

Shrestha, N. K., Tuohy, M. J., Hall, G. S., Isada, C. M., & Procop, G. W. (2002). Rapid identification of Staphylococcus aureus and the mecA gene from BacT/ALERT blood culture bottles by using the LightCycler system. *J Clin Microbiol*, *40*(7), 2659–2661.

Simner, P. J., Uhl, J. R., Hall, L., Weber, M. M., Walchak, R. C., Buckwalter, S., & Wengenack, N. L. (2013). Broad-range direct detection and identification of fungi by use of the PLEX-ID PCR-electrospray ionization mass spectrometry (ESIMS) system. *Journal of Clinical Microbiology*, *51*(6), 1699–1706. https://doi.org/10.1128/JCM.03282-12

Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., … Coopersmith, C. M. (2016). The third international consensus definitions for sepsis and septic shock (sepsis-3). *Jama*, *315*(8), 801–810.

Sixt, B. S., Siegl, A., Müller, C., Watzka, M., Wultsch, A., Tziotis, D., … Horn, M. (2013). Metabolic features of Protochlamydia amoebophila elementary bodies–a link between activity and infectivity in Chlamydiae. *PLoS Pathogens*, *9*(8), e1003553.

Smith, A. M., Jain, M., Mulroney, L., Garalde, D. R., & Akeson, M. (2017). Reading canonical and modified nucleotides in 16S ribosomal RNA using nanopore direct RNA sequencing. *BioRxiv*, 132274.

Sparbier, K., Weller, U., Boogen, C., & Kostrzewa, M. (2012). Rapid detection of Salmonella sp. by means of a combination of selective enrichment broth and MALDI-TOF MS. *European Journal of Clinical Microbiology & Infectious Diseases*, *31*(5), 767–773.

Springer, J., Loeffler, J., Heinz, W., Schlossnagel, H., Lehmann, M., Morton, O., … Kurzai, O. (2011). Pathogen-specific DNA enrichment does not increase sensitivity of PCR for diagnosis of invasive aspergillosis in neutropenic patients. *J Clin Microbiol*, *49*(4), 1267–1273. https://doi.org/10.1128/jcm.01679-10

Stamm, W. E., & Norrby, S. R. (2001). Urinary Tract Infections: Disease Panorama and Challenges. *The Journal of Infectious Diseases*, *183*(s1), S1–S4. https://doi.org/10.1086/318850

Stamper, P. D., Babiker, W., Alcabasa, R., Aird, D., Wehrlin, J., Ikpeama, I., … Carroll, K. C. (2009). Evaluation of a new commercial TaqMan PCR assay for direct detection of the clostridium difficile toxin B gene in clinical stool specimens. *J Clin Microbiol*, *47*(12), 3846–3850. https://doi.org/10.1128/jcm.01490-09

Sterry-Blunt, R. E., Randall, K. S., Doughton, M. J., Aliyu, S. H., & Enoch, D. A. (2015). Screening urine samples for the absence of urinary tract infection using the sediMAX automated microscopy analyser. *Journal of Medical Microbiology*, *64*(6), 605–609.

Stevenson, M., Pandor, A., James, M.-S., Rafia, R., Uttley, L., Stevens, J., … McMullan, R. (2016). Sepsis: the LightCycler SeptiFast Test MGRADE®, SepsiTest™ and IRIDICA BAC BSI assay for rapidly identifying bloodstream bacteria and fungi–a systematic review and economic evaluation. *Health*

Technology Assessment, *20*(46), 1–246.

Stewart, E. J. (2012). Growing unculturable bacteria. *J Bacteriol*, *194*(16), 4151–4160.

Stonehouse, M. J., Cota-Gomez, A., Parker, S. K., Martin, W. E., Hankin, J. A., Murphy, R. C., … Vasil, A. I. (2002). A novel class of microbial phosphocholine-specific phospholipases C. *Molecular Microbiology*, *46*(3), 661–676.

Stoneking, L. R., Patanwala, A. E., Winkler, J. P., Fiorello, A. B., Lee, E. S., Olson, D. P., & Wolk, D. M. (2013). Would earlier microbe identification alter antibiotic therapy in bacteremic emergency department patients? *J Emerg Med*, *44*(1), 1–8. https://doi.org/10.1016/j.jemermed.2012.02.036

Surawicz, C. M., Brandt, L. J., Binion, D. G., Ananthakrishnan, A. N., Curry, S. R., Gilligan, P. H., … Zuckerbraun, B. S. (2013). Guidelines for diagnosis, treatment, and prevention of Clostridium difficile infections. *The American Journal of Gastroenterology*, *108*(4), 478–498.

Telenti, A., Honoré, N., Bernasconi, C., March, J., Ortega, A., Heym, B., … Cole, S. T. (1997). Genotypic assessment of isoniazid and rifampin resistance in Mycobacterium tuberculosis: A blind study at reference laboratory level. *Journal of Clinical Microbiology*, *35*(3), 719–723.

Tenover, F. C., Novak-Weekley, S., Woods, C. W., Peterson, L. R., Davis, T., Schreckenberger, P., … Nomura, J. H. (2010). Impact of strain type on detection of toxigenic Clostridium difficile: comparison of molecular diagnostic and enzyme immunoassay approaches. *J Clin Microbiol*, *48*(10), 3719–3724.

Tessari, A., Osti, N., & Scarin, M. (2015). Screening of presumptive urinary tract infections by the automated urine sediment analyser sediMAX. *Clinical Chemistry and Laboratory Medicine (CCLM)*, *53*(s2), s1503–s1508.

Thoendel, M., Jeraldo, P., Greenwood-Quaintance, K. E., Yao, J., Chia, N., Hanssen, A. D., … Patel, R. (2017). Impact of Contaminating DNA in Whole-Genome Amplification Kits Used for Metagenomic Shotgun Sequencing for Infection Diagnosis. *J Clin Microbiol*, *55*(6), 1789–1801.

Thoendel, M., Jeraldo, P. R., Greenwood-Quaintance, K. E., Yao, J. Z., Chia, N., Hanssen, A. D., … Patel, R. (2016). Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. *J Microbiol Methods*, *127*, 141–145.

Tille, P. (2013). *Bailey & Scott's Diagnostic Microbiology-E-Book*. Elsevier Health

Sciences.

Tissari, P., Zumla, A., Tarkka, E., Mero, S., Savolainen, L., Vaara, M., … Gant, V. (2010). Accurate and rapid identification of bacterial species from positive blood cultures with a DNA-based microarray platform: an observational study. *The Lancet*, *375*(9710), 224–230. https://doi.org/10.1016/S0140-6736(09)61569-5

Titball, R. W., Leslie, D. L., Harvey, S., & Kelly, D. (1991). Hemolytic and sphingomyelinase activities of Clostridium perfringens alpha-toxin are dependent on a domain homologous to that of an enzyme from the human arachidonic acid pathway. *Infect Immun*, *59*(5), 1872–1874.

Torio, C. M. (2013). *National inpatient hospital costs: the most expensive conditions by payer, 2011. Healthcare Cost and Utilization Project Statistical Brief# 160*.

Treangen, T. J., & Salzberg, S. L. (2011). Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nature Reviews Genetics*, *13*(1), nrg3117.

Tyson, G. H., McDermott, P. F., Li, C., Chen, Y., Tadesse, D. A., Mukherjee, S., … Loneragan, G. H. (2015). WGS accurately predicts antimicrobial resistance in Escherichia coli. *Journal of Antimicrobial Chemotherapy*, *70*(10), 2763–2769.

van der Zwet, W. C., Hessels, J., Canbolat, F., & Deckers, M. M. L. (2010). Evaluation of the Sysmex UF-1000i® urine flow cytometer in the diagnostic work-up of suspected urinary tract infection in a Dutch general hospital. *Clinical Chemistry and Laboratory Medicine*, *48*(12), 1765–1771.

Vandecandelaere, I., Nercessian, O., Segaert, E., Achouak, W., Mollica, A., Faimali, M., … Vandamme, P. (2008). Alteromonas genovensis sp. nov., isolated from a marine electroactive biofilm and emended description of Alteromonas macleodii Baumann et al. 1972 (Approved Lists 1980). *International Journal of Systematic and Evolutionary Microbiology*, *58*(11), 2589–2596. https://doi.org/10.1099/ijs.0.65691-0

Viala, C., Le Monnier, A., Maataoui, N., Rousseau, C., Collignon, A., & Poilane, I. (2012). Comparison of commercial molecular assays for toxigenic Clostridium difficile detection in stools: BD GeneOhm Cdiff, XPert C. difficile and illumigene C. difficile. *J Microbiol Methods*, *90*(2), 83–85.

Vince, A., Lepej, S. Ž., Baršić, B., Dušek, D., Mitrović, Z., Serventi-Seiwerth, R., & Labar, B. (2008). LightCycler SeptiFast assay as a tool for the rapid diagnosis of sepsis in patients during antimicrobial therapy. *Journal of Medical Microbiology*,

57(10), 1306–1307.

Vincent, J.-L., Sakr, Y., Sprung, C. L., Ranieri, V. M., Reinhart, K., Gerlach, H., … Payen, D. (2006). Sepsis in European intensive care units: results of the SOAP study. *Critical Care Medicine*, *34*(2), 344–353.

Votintseva, A. A., Bradley, P., Pankhurst, L., del Ojo Elias, C., Loose, M., Nilgiriwala, K., … Walker, T. M. (2017). Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. *J Clin Microbiol*, *55*(5), 1285–1298.

Wagner, M., Horn, M., & Daims, H. (2003). Fluorescence in situ hybridisation for the identification and characterisation of prokaryotes. *Current Opinion in Microbiology*, *6*(3), 302–309. https://doi.org/10.1016/S1369-5274(03)00054-7

Wain, J., Diep, T. S., Ho, V. A., Walsh, A. M., Hoa, N. T. T., Parry, C. M., & White, N. J. (1998). Quantitation of bacteria in blood of typhoid fever patients and relationship between counts and clinical features, transmissibility, and antibiotic resistance. *J Clin Microbiol*, *36*(6), 1683–1687.

Wain, J., & Mavrogiorgou, E. (2013). Next-generation sequencing in clinical microbiology. *Expert Rev Mol Diagn*, *13*(3), 225–227. https://doi.org/10.1586/erm.13.8

Warhurst, G., Maddi, S., Dunn, G., Ghrew, M., Chadwick, P., Alexander, P., … Dark, P. (2015). Diagnostic accuracy of SeptiFast multi-pathogen real-time PCR in the setting of suspected healthcare-associated bloodstream infection. *Intensive Care Med*, *41*(1), 86–93. https://doi.org/10.1007/s00134-014-3551-x

Washington, J. A. (2012). *Laboratory procedures in clinical microbiology*. Springer Science & Business Media.

Wei, Y.-X., Zhang, Z.-Y., Liu, C., Zhu, Y.-Z., Zhu, Y.-Q., Zheng, H., … Guo, X.-K. (2010). Complete Genome Sequence of Bifidobacterium longum JDM301. *J Bacteriol*, *192*(15), 4076–4077. https://doi.org/10.1128/JB.00538-10

Weiss, S., Amir, A., Hyde, E. R., Metcalf, J. L., Song, S. J., & Knight, R. (2014). Tracking down the sources of experimental contamination in microbiome studies. *Genome Biol*, *15*(12), 564.

Wellinghausen, N., Kochem, A. J., Disque, C., Muhl, H., Gebert, S., Winter, J., … Sakka, S. G. (2009). Diagnosis of bacteremia in whole-blood samples by use of a commercial universal 16S rRNA gene-based PCR and sequence analysis. *J Clin Microbiol*, *47*(9), 2759–2765. https://doi.org/10.1128/JCM.00567-09

Wiesinger-Mayr, H., Jordana-Lluch, E., Martró, E., Schoenthaler, S., & Noehammer, C. (2011). Establishment of a semi-automated pathogen DNA isolation from whole blood and comparison with commercially available kits. *J Microbiol Methods*, *85*(3), 206–213.

Wilcox, M. H., Gerding, D. N., Poxton, I. R., Kelly, C., Nathan, R., Birch, T., … Lee, C. (2017). Bezlotoxumab for prevention of recurrent Clostridium difficile infection. *New England Journal of Medicine*, *376*(4), 305–317.

Wilcox, M. H., Shetty, N., Fawley, W. N., Shemko, M., Coen, P., Birtles, a, … Wren, M. W. D. (2012). Changing epidemiology of Clostridium difficile infection following the introduction of a national ribotyping-based surveillance scheme in England. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *55*(8), 1056–1063. https://doi.org/10.1093/cid/cis614

Wilson, M. L., & Gaido, L. (2004). Laboratory diagnosis of urinary tract infections in adult patients. *Clinical Infectious Diseases*, *38*(8), 1150–1158.

Wojewoda, C. M., Sercia, L., Navas, M., Tuohy, M., Wilson, D., Hall, G. S., … Richter, S. S. (2013). Evaluation of the Verigene Gram-positive blood culture nucleic acid test for rapid detection of bacteria and resistance determinants. *J Clin Microbiol*, *51*(7), 2072–2076.

Wood, D. E., & Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol*, *15*(3), R46. https://doi.org/10.1186/gb-2014-15-3-r46

Yagupsky, P., & Nolte, F. S. (1990). Quantitative aspects of septicemia. *Clinical Microbiology Reviews*, *3*(3), 269–279.

Yigit, E., Hernandez, D. I., Trujillo, J. T., Dimalanta, E., & Bailey, C. D. (2014). Genome and metagenome sequencing: Using the human methyl-binding domain to partition genomic DNA derived from plant tissues. *Applications in Plant Sciences*, *2*(11), 1400064.

Zaman, Z., Fogazzi, G. B., Garigali, G., Croci, M. D., Bayer, G., & Kránicz, T. (2010). Urine sediment analysis: analytical and diagnostic performance of sediMAX®—a new automated microscopy image-based urine sediment analyser. *Clinica Chimica Acta*, *411*(3), 147–154.

Zhang, X., Niu, S., & Zhang, L. (2014). Antimicrobial susceptibilities and clinical characterization of Pseudomonas aeruginosa isolates from urinary tract

infections. *Urol Int*, *93*(4), 464–469. https://doi.org/10.1159/000358493

Zhao, S., Tyson, G. H., Chen, Y., Li, C., Mukherjee, S., Young, S., … McDermott, P. F. (2016). Whole-genome sequencing analysis accurately predicts antimicrobial resistance phenotypes in Campylobacter spp. *Applied and Environmental Microbiology*, *82*(2), 459–466.

Zhou, L., & Pollard, A. J. (2012). A novel method of selective removal of human DNA improves PCR sensitivity for detection of Salmonella Typhi in blood samples. *BMC Infect Dis*, *12*, 164. https://doi.org/10.1186/1471-2334-12-164

Ziegler, I., Fagerström, A., Strålin, K., & Mölling, P. (2016). Evaluation of a Commercial Multiplex PCR Assay for Detection of Pathogen DNA in Blood from Patients with Suspected Sepsis. *PLoS One*, *11*(12), e0167883.

Zimlichman, E., Henderson, D., Tamir, O., Franz, C., Song, P., Yamin, C. K., … Bates, D. W. (2013). Health care–associated infections: a meta-analysis of costs and financial impact on the US health care system. *JAMA Internal Medicine*, *173*(22), 2039–2046.

Ackermann, R. J., & Monroe, P. W. (1996). Bacteremic urinary tract infection in older people. *Journal of the American Geriatrics Society*, *44*(8), 927–933. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med4&NEWS=N&AN=8708302

Afshari, A., Schrenzel, J., Ieven, M., & Harbarth, S. (2012). Bench-to-bedside review: Rapid molecular diagnostics for bloodstream infection--a new frontier? *Crit Care*, *16*(3), 222. https://doi.org/10.1186/cc11202

Akiyoshi, D. E., Dilo, J., Pearson, C., Chapman, S., Tumwine, J., & Tzipori, S. (2003). Characterization of Cryptosporidium meleagridis of human origin passaged through different host species. *Infection and Immunity*, *71*(4), 1828–1832. https://doi.org/10.1128/IAI.71.4.1828-1832.2003

Angelakis, E., Bachar, D., Henrissat, B., Armougom, F., Audoly, G., Lagier, J. C., … Raoult, D. (2016). Glycans affect DNA extraction and induce substantial differences in gut metagenomic studies. *Sci Rep*, *6*, 26276. https://doi.org/10.1038/srep26276

Arroyo, L. G., Rousseau, J., Willey, B. M., Low, D. E., Staempfli, H., McGeer, A., & Weese, J. S. (2005). Use of a selective enrichment broth to recover Clostridium

difficile from stool swabs stored under different conditions. *J Clin Microbiol*, *43*(10), 5341–5343.

Ashton, P. M., Nair, S., Dallman, T., Rubino, S., Rabsch, W., Mwaigwisya, S., … O'grady, J. (2015). MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nat Biotechnol*, *33*(3), 296–300.

Bader, O., Weig, M., Taverne-Ghadwal, L., Lugert, R., Gross, U., & Kuhns, M. (2011). Improved clinical laboratory identification of human pathogenic yeasts by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clinical Microbiology and Infection*, *17*(9), 1359–1365.

Baltekin, Ö., Boucharin, A., Tano, E., Andersson, D. I., & Elf, J. (2017). Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proceedings of the National Academy of Sciences*, 201708558. https://doi.org/10.1073/pnas.1708558114

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Prjibelski, A. D. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477.

Baquero, F. (2017). Emerging Concepts and Strategies in Clinical Microbiology and Infectious Diseases. *Clinical Infectious Diseases*, *65*, 1.

Baraban, L., Bertholle, F., Salverda, M. L. M., Bremond, N., Panizza, P., Baudry, J., … Bibette, J. (2011). Millifluidic droplet analyser for microbiology. *Lab on a Chip*, *11*(23), 4057. https://doi.org/10.1039/c1lc20545e

Barbosa, T. M., Serra, C. R., La Ragione, R. M., Woodward, M. J., & Henriques, A. O. (2005). Screening for Bacillus isolates in the broiler gastrointestinal tract. *Applied and Environmental Microbiology*, *71*(2), 968–978.

Benagli, C., Rossi, V., Dolina, M., Tonolla, M., Petrini, O., Mellmann, A., … Moehring, R. W. (2013). Evaluation of new preanalysis sample treatment tools and DNA isolation protocols to improve bacterial pathogen detection in whole blood. *J Clin Microbiol*, *6*(1), 2629–2631. https://doi.org/10.1001/jama.2013.3231

Bernardo, K., Krut, O., Wiegmann, K., Kreder, D., Micheli, M., Schäfer, R., … Meyer, H. E. (2000). Purification and characterization of a magnesium-dependent neutral sphingomyelinase from bovine brain. *Journal of Biological Chemistry*,

275(11), 7641–7647.

Bidet, P., Lalande, V., Salauze, B., Burghoffer, B., Avesani, V., Delmée, M., … Petit, J.-C. (2000). Comparison of PCR-ribotyping, arbitrarily primed PCR, and pulsed-field gel electrophoresis for typing Clostridium difficile. *J Clin Microbiol*, *38*(7), 2484–2487.

Bilecen, K., Yaman, G., Ciftci, U., & Laleli, Y. R. (2015). Performances and reliability of Bruker Microflex LT and Vitek MS MALDI-TOF mass spectrometry systems for the identification of clinical microorganisms. *BioMed Research International*, *2015*.

Bizzini, A., & Greub, G. (2010). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification. *Clinical Microbiology and Infection*, *16*(11), 1614–1619.

Bochud, P.-Y., Bonten, M., Marchetti, O., & Calandra, T. (2004). Antimicrobial therapy for patients with severe sepsis and septic shock: an evidence-based review. *Critical Care Medicine*, *32*(11), S495–S512.

Boedicker, J. Q., Li, L., Kline, T. R., & Ismagilov, R. F. (2008). Detecting bacteria and determining their susceptibility to antibiotics by stochastic confinement in nanoliter droplets using plug-based microfluidics. *Lab on a Chip*, *8*(8), 1265. https://doi.org/10.1039/b804911d

Bradley, P., Gordon, N. C., Walker, T. M., Dunn, L., Heys, S., Huang, B., … De Cesare, M. (2015). Rapid antibiotic-resistance predictions from genome sequence data for Staphylococcus aureus and Mycobacterium tuberculosis. *Nat Commun*, *6*, 10063.

Britton, R. A., & Young, V. B. (2014). Role of the intestinal microbiota in resistance to colonization by Clostridium difficile. *Gastroenterology*, *146*(6), 1547–1553.

Broeren, M. A. C., Bahçeci, S., Vader, H. L., & Arents, N. L. A. (2011). Screening for urinary tract infection with the Sysmex UF-1000i urine flow cytometer. *J Clin Microbiol*, *49*(3), 1025–1029.

Buchanan, A. G. (1984). Selective enrichment broth culture for detection of Clostridium difficile and associated cytotoxin. *J Clin Microbiol*, *20*(1), 74–76.

Burillo, A., Rodríguez-Sánchez, B., Ramiro, A., Cercenado, E., Rodríguez-Créixems, M., & Bouza, E. (2014). Gram-Stain Plus MALDI-TOF MS (Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry) for a Rapid Diagnosis of Urinary Tract Infection. *PLoS ONE*, *9*(1), e86915.

https://doi.org/10.1371/journal.pone.0086915

Burke, R., McKenna, J., Cox, C., … P. C.-J. of, & 2016, undefined. (n.d.). A comparison of different pre-lysis methods and extraction kits for recovery of Streptococcus agalacticae (Lancefield group B Streptococcus) DNA from whole blood. *Elsevier*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0167701216302275

Burnham, C.-A. D., Leeds, J., Nordmann, P., O'Grady, J., & Patel, J. (2017). Diagnosing antimicrobial resistance. *Nature Reviews Microbiology*, *15*(11), 697.

Burns, J. L., Saiman, L., Whittier, S., Krzewinski, J., Liu, Z., Larone, D., … Jones, R. N. (2001). Comparison of two commercial systems (Vitek and MicroScan-WalkAway) for antimicrobial susceptibility testing of Pseudomonas aeruginosa isolates from cystic fibrosis patients. *Diagnostic Microbiology and Infectious Disease*, *39*(4), 257–260. https://doi.org/10.1016/S0732-8893(01)00234-6

Carrara, L., Navarro, F., Turbau, M., Seres, M., Moran, I., Quintana, I., … Cordon, O. (2013). Molecular diagnosis of bloodstream infections with a new dual-priming oligonucleotide-based multiplex PCR assay. *Journal of Medical Microbiology*, *62*(11), 1673–1679.

Cetinkaya, Y., Falk, P., & Mayhall, C. G. (2000). Vancomycin-resistant enterococci. *Clinical Microbiology Reviews*, *13*(4), 686–707.

Chang, S.-S., Hsieh, W.-H., Liu, T.-S., Lee, S.-H., Wang, C.-H., Chou, H.-C., … Lee, C.-C. (2013). Multiplex PCR system for rapid detection of pathogens in patients with presumed sepsis–a systemic review and meta-analysis. *PLoS One*, *8*(5), e62323.

Chankhamhaengdecha, S., Hadpanus, P., Aroonnual, A., Ngamwongsatit, P., Chotiprasitsakul, D., Chongtrakool, P., & Janvilisri, T. (2013). Evaluation of multiplex PCR with enhanced spore germination for detection of clostridium difficile from stool samples of the hospitalized patients. *BioMed Research International*, *2013*. https://doi.org/10.1155/2013/875437

Chen, C. H., Lu, Y., Sin, M. L. Y., Mach, K. E., Zhang, D. D., Gau, V., … Wong, P. K. (2010). Antimicrobial susceptibility testing using high surface-to-volume ratio microchannels. *Analytical Chemistry*, *82*(xx), 1012–1019. https://doi.org/10.1021/ac9022764

Cho, I., & Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. *Nature Reviews. Genetics*, *13*(4), 260.

Choi, J., Yoo, J., Lee, M., Kim, E.-G., Lee, J. S., Lee, S., … Kwon, S. (2014). A rapid antimicrobial susceptibility test based on single-cell morphological analysis. *Science Translational Medicine*, *6*(267), 267ra174. https://doi.org/10.1126/scitranslmed.3009650

Clements, A. C. A., Magalhães, R. J. S., Tatem, A. J., Paterson, D. L., & Riley, T. V. (2010). Clostridium difficile PCR ribotype 027: assessing the risks of further worldwide spread. *The Lancet Infectious Diseases*, *10*(6), 395–404.

Cohen, J. (2002). The immunopathogenesis of sepsis. *Nature*, *420*(6917), 885–891. https://doi.org/10.1038/nature01326

Cohen, J., Vincent, J.-L., Adhikari, N. K. J., Machado, F. R., Angus, D. C., Calandra, T., … Opal, S. (2015). Sepsis: a roadmap for future research. *The Lancet Infectious Diseases*, *15*(5), 581–614.

Cohen, S. H., Gerding, D. N., Johnson, S., Kelly, C. P., Loo, V. G., McDonald, L. C., … Wilcox, M. H. (2010). Clinical practice guidelines for Clostridium difficile infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). *Infect Control Hosp Epidemiol*, *31*(5), 431–455. https://doi.org/10.1086/651706

Consortium, H. M. P. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, *486*(7402), 207–214.

Cooper, M. A., & Shlaes, D. (2011). Fix the antibiotics pipeline. *Nature*, *472*(7341), 32.

Daniels, R. (2011). Surviving the first hours in sepsis: getting the basics right (an intensivist's perspective). *J Antimicrob Chemother*, *66 Suppl 2*, ii11-23. https://doi.org/10.1093/jac/dkq515

Dark, P., Blackwood, B., Gates, S., McAuley, D., Perkins, G. D., McMullan, R., … Warhurst, G. (2015). Accuracy of LightCycler((R)) SeptiFast for the detection and identification of pathogens in the blood of patients with suspected sepsis: a systematic review and meta-analysis. *Intensive Care Med*, *41*(1), 21–33. https://doi.org/10.1007/s00134-014-3553-8

Dark, P., Dunn, G., Chadwick, P., Young, D., Bentley, A., Carlson, G., & Warhurst, G. (2011). The clinical diagnostic accuracy of rapid detection of healthcare-associated bloodstream infection in intensive care using multipathogen real-time PCR technology. *BMJ Open*, *1*(1), e000181. https://doi.org/10.1136/bmjopen-2011-000181

Davies, S. C., & Gibbens, N. (2013). UK five year antimicrobial resistance strategy 2013 to 2018. *London: Department of Health*.

De Angelis, G., Posteraro, B., De Carolis, E., Menchinelli, G., Franceschi, F., Tumbarello, M., … Sanguinetti, M. (2018). T2Bacteria magnetic resonance assay for the rapid detection of ESKAPEc pathogens directly in whole blood. *Journal of Antimicrobial Chemotherapy*, *73*, iv20-iv26. https://doi.org/10.1093/jac/dky049

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., … Lawley, T. D. (2012). The Clostridium difficile spo0A gene is a persistence and transmission factor. *Infect Immun*, *80*(8), 2704–2711.

Dellinger, R. P., Levy, M. M., Rhodes, A., Annane, D., Gerlach, H., Opal, S. M., … Campaign, S. S. (2013). Surviving Sepsis Campaign: International Guidelines for Management of Severe Sepsis and Septic Shock, 2012. *Intensive Care Medicine*, *39*(2), 165–228. https://doi.org/DOI 10.1007/s00134-012-2769-8

Delmont, T. O., Robe, P., Clark, I., Simonet, P., & Vogel, T. M. (2011). Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J Microbiol Methods*, *86*(3), 397–400.

Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E. A., & Crook, D. W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nature Reviews. Genetics*, *13*(9), 601.

Dingle, K. E., Didelot, X., Ansari, M. A., Eyre, D. W., Vaughan, A., Griffiths, D., … Crook, D. W. (2013). Recombinational Switching of the Clostridium difficile S-Layer and a Novel Glycosylation Gene Cluster Revealed by Large-Scale Whole-Genome Sequencing. *The Journal of Infectious Diseases*, *207*(4), 675–686. https://doi.org/10.1093/infdis/jis734

Dodémont, M., De Mendonça, R., Nonhoff, C., Roisin, S., & Denis, O. (2014). Performance of the Verigene Gram-negative blood culture assay for rapid detection of bacteria and resistance determinants. *J Clin Microbiol*, *52*(8), 3085–3087.

Dolinger, D. L., & Jacobs, A. A. (2015). Molecular diagnostics and active screening for health care-associated infections: stepping-up the game. *Laboratory Medicine*, *42*(5), 267–272.

Donaldson, G. P., Lee, S. M., & Mazmanian, S. K. (2016). Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology*, *14*(1), 20–32.

Driks, A. (1999). Bacillus subtilis spore coat. *Microbiology and Molecular Biology Reviews*, *63*(1), 1–20.

Dunne Jr., W. M., Westblade, L. F., & Ford, B. (2012). Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. *Eur J Clin Microbiol Infect Dis*, *31*(8), 1719–1726. https://doi.org/10.1007/s10096-012-1641-7

Dutta, S., Chatterjee, A., Dutta, P., Rajendran, K., Roy, S., Pramanik, K. C., & Bhattacharya, S. K. (2001). Sensitivity and performance characteristics of a direct PCR with stool samples in comparison to conventional techniques for diagnosis of Shigella and enteroinvasive Escherichia coli infection in children with acute diarrhoea in Calcutta, India. *Journal of Medical Microbiology*, *50*(8), 667–674.

Edwards, A. N., Suárez, J. M., & McBride, S. M. (2013). Culturing and maintaining Clostridium difficile in an anaerobic environment. *Journal of Visualized Experiments: JoVE*, (79).

Eliopoulos, G. M., & Gold, H. S. (2001). Vancomycin-resistant enterococci: mechanisms and clinical observations. *Clinical Infectious Diseases*, *33*(2), 210–219.

Ellepola, A. N., & Morrison, C. J. (2005). Laboratory diagnosis of invasive candidiasis. *J Microbiol*, *43 Spec No*, 65–84.

Ellington, M. J., Ekelund, O., Aarestrup, F. M., Canton, R., Doumith, M., Giske, C., … Hopkins, K. L. (2017). The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. *Clinical Microbiology and Infection*, *23*(1), 2–22.

Emmadi, R., Boonyaratanakornkit, J. B., Selvarangan, R., Shyamala, V., Zimmer, B. L., Williams, L., … Wilson, J. A. A. (2011). Molecular methods and platforms for infectious diseases testing: a review of FDA-approved and cleared assays. *The Journal of Molecular Diagnostics*, *13*(6), 583–604.

Eyre, D. W., Cule, M. L., Wilson, D. J., Griffiths, D., Vaughan, A., O'connor, L., … Finney, J. M. (2013). Diverse sources of C. difficile infection identified on whole-genome sequencing. *New England Journal of Medicine*, *369*(13), 1195–1205.

Eyre, D. W., Fawley, W. N., Best, E. L., Griffiths, D., Stoesser, N. E., Crook, D. W., … Wilcox, M. H. (2013). Comparison of multilocus variable-number tandem-repeat analysis and whole-genome sequencing for investigation of Clostridium

difficile transmission. *J Clin Microbiol*, *51*(12), 4141–4149. https://doi.org/10.1128/jcm.01095-13

Fair, R. J., & Tor, Y. (2014). Antibiotics and bacterial resistance in the 21st century. *Perspectives in Medicinal Chemistry*, (6), 25–64. https://doi.org/10.4137/PMC.S14459

Falbo, R., Sala, M. R., Signorelli, S., Venturi, N., Signorini, S., & Brambilla, P. (2012). Bacteriuria screening by automated whole-field-image-based microscopy reduces the number of necessary urine cultures. *J Clin Microbiol*, *50*(4), 1427–1429.

Feehery, G. R., Yigit, E., Oyola, S. O., Langhorst, B. W., Schmidt, V. T., Stewart, F. J., … Pradhan, S. (2013). A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. *PLoS One*, *8*(10), e76096. https://doi.org/10.1371/journal.pone.0076096

Felt, J. R., Yurkovich, C., Garshott, D. M., Kamat, D., Farooqi, A., Fribley, A. M., … Hebert, K. (2017). The Utility of Real-Time Quantitative Polymerase Chain Reaction Genotype Detection in the Diagnosis of Urinary Tract Infections in Children. *Clin Pediatr (Phila)*, *56*(10), 912–919. https://doi.org/10.1177/0009922817706144

Fenollar, F., & Raoult, D. (2007). Molecular diagnosis of bloodstream infections caused by non-cultivable bacteria. *International Journal of Antimicrobial Agents*, *30*, 7–15.

Ferreira, L., Sánchez-Juanes, F., González-Ávila, M., Cembrero-Fuciños, D., Herrero-Hernández, A., González-Buitrago, J. M., & Muñoz-Bellido, J. L. (2010). Direct identification of urinary tract pathogens from urine samples by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol*, *48*(6), 2110–2115.

Ferrer, R., Martin-Loeches, I., Phillips, G., Osborn, T. M., Townsend, S., Dellinger, R. P., … Levy, M. M. (2014). Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. *Critical Care Medicine*, *42*(8), 1749–1755.

Fichtel, J., Köster, J., Rullkötter, J., & Sass, H. (2007). Spore dipicolinic acid contents used for estimating the number of endospores in sediments. *FEMS Microbiology Ecology*, *61*(3), 522–532.

Fleischmann, C., Scherag, A., Adhikari, N. K. J., Hartog, C. S., Tsaganos, T., Schlattmann, P., … Reinhart, K. (2016). Assessment of global incidence and mortality of hospital-treated sepsis. Current estimates and limitations. *American Journal of Respiratory and Critical Care Medicine*, *193*(3), 259–272.

Fournier, P.-E., Drancourt, M., Colson, P., Rolain, J.-M., La Scola, B., & Raoult, D. (2013). Modern clinical microbiology: new challenges and solutions. *Nature Reviews Microbiology*, *11*(8), 574–585.

Fournier, P., Drancourt, M., … P. C.-N. R., & 2013, undefined. (n.d.). Modern clinical microbiology: new challenges and solutions. *Nature.Com*. Retrieved from https://www.nature.com/articles/nrmicro3068

Fox, J. L. (2014). Technology comes to typing. *Nat Biotechnol*, *32*(11), 1081–1084. https://doi.org/10.1038/nbt.3067

Foxman, B. (2010). The epidemiology of urinary tract infection. *Nature Reviews Urology*, *7*(12), 653.

Frith, M. C., Hamada, M., & Horton, P. (2010). Parameters for accurate genome alignment. *BMC Bioinformatics*, *11*(1), 80.

Fukushima, H., Tsunomori, Y., & Seki, R. (2003). Duplex real-time SYBR green PCR assays for detection of 17 species of food-or waterborne pathogens in stools. *J Clin Microbiol*, *41*(11), 5134–5146.

Gadsby, N. J., Russell, C. D., McHugh, M. P., Mark, H., Conway Morris, A., Laurenson, I. F., … Templeton, K. E. (2016). Comprehensive molecular testing for respiratory pathogens in community-acquired pneumonia. *Clinical Infectious Diseases*, *62*(7), 817–823.

Gaillot, O., Di Camillo, P., Berche, P., Courcol, R., & Savage, C. (1999). Comparison of CHROMagar Salmonella medium and Hektoen enteric agar for isolation of salmonellae from stool samples. *J Clin Microbiol*, *37*(3), 762–765.

Gangaram, R., Ojwang, P. J., Moodley, J., & Maharaj, D. (2005). The accuracy of urine dipsticks as a screening test for proteinuria in hypertensive disorders of pregnancy. *Hypertension in Pregnancy*, *24*(2), 117–123. https://doi.org/10.1081/PRG-200059849

Garalde, D. R., Snell, E. A., Jachimowicz, D., Sipos, B., Lloyd, J. H., Bruce, M., … Warland, A. (2018). Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods*.

García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L. M., Götz, S., Tarazona,

S., … Conesa, A. (2012). Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics*, *28*(20), 2678–2679.

Garnacho-Montero, J., Gutierrez-Pizarraya, A., Escoresca-Ortega, A., Corcia-Palomo, Y., Fernandez-Delgado, E., Herrera-Melero, I., … Marquez-Vacaro, J. A. (2014). De-escalation of empirical therapy is associated with lower mortality in patients with severe sepsis and septic shock. *Intensive Care Med*, *40*(1), 32–40. https://doi.org/10.1007/s00134-013-3077-7

Gebert, S., Siegel, D., & Wellinghausen, N. (2008). Rapid detection of pathogens in blood culture bottles by real-time PCR in conjunction with the pre-analytic tool MolYsis. *Journal of Infection*, *57*(4), 307–316.

Gerner-Smidt, P., Hise, K., Kincaid, J., Hunter, S., Rolando, S., Hyytiä-Trees, E., … Taskforce, P. (2006). PulseNet USA: a five-year update. *Foodbourne Pathogens & Disease*, *3*(1), 9–19.

Gerritzen, A., Wittke, J.-W., & Wolff, D. (2011). Rapid and Sensitive Detection of Shiga Toxin-Producing Escherichia Coli Directly From Stool Samples by Real-Time PCR in Comparison to Culture, Enzyme Immunoassay, and Vero Cell Cytotoxicity Assay. *Clinical Laboratory*, *57*(11), 993.

Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B., Jorden, J. R., & Chiodini, R. J. (2015). Changes in 16S RNA gene microbial community profiling by concentration of prokaryotic DNA. *J Microbiol Methods*, *119*, 239–242.

Goldberg, B., Sichtig, H., Geyer, C., Ledeboer, N., & Weinstock, G. M. (2015). Making the leap from research laboratory to clinic: challenges and opportunities for next-generation sequencing in infectious disease diagnostics. *MBio*, *6*(6), e01888-15.

Goldenberg, S. D., & French, G. L. (2011). Diagnostic testing for Clostridium difficile: a comprehensive survey of laboratories in England. *Journal of Hospital Infection*, *79*(1), 4–7.

Goldenberg, S. D., Price, N. M., Tucker, D., Wade, P., & French, G. L. (2011). Mandatory reporting and improvements in diagnosing Clostridium difficile infection: an incompatible dichotomy? *Journal of Infection*, *62*(5), 363–370.

Goodwin, S., Wappel, R., & McCombie, W. R. (2017). 1D Genome Sequencing on the Oxford Nanopore MinION. *Current Protocols in Human Genetics*, 18.11. 1-18.11. 14.

Goswitz, J. J., Willard, K. E., Eastep, S. J., Shanholtzer, C. J., Olson, M. L., Pinnell,

M., … Peterson, L. R. (1993). Utility of slide centrifuge Gram's stain versus quantitative culture for diagnosis of urinary tract infection. *American Journal of Clinical Pathology*, *99*(2), 132–136.

Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., … Chiu, C. Y. (2015). Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med*, *7*, 99. https://doi.org/10.1186/s13073-015-0220-9

Grice, E. A., Kong, H. H., Renaud, G., Young, A. C., Bouffard, G. G., Blakesley, R. W., … Segre, J. A. (2008). A diversity profile of the human skin microbiota. *Genome Res*, *18*(7), 1043–1050. https://doi.org/10.1101/gr.075549.107

Grumaz, S., Stevens, P., Grumaz, C., Decker, S. O., Weigand, M. A., Hofer, S., … Sohn, K. (2016). Next-generation sequencing diagnostics of bacteremia in septic patients. *Genome Med*, *8*(1), 73. https://doi.org/10.1186/s13073-016-0326-8

Guido, M., Tumolo, M. R., De Donno, A., Verri, T., Serio, F., Bagordo, F., & Zizza, A. (2016). In vitro diagnosis of sepsis: a review. *PATHOLOGY AND LABORATORY MEDICINE INTERNATIONAL*, *8*, 1–14.

Gurtler, V. (1993). Typing of Clostridium difficile strains by PCR-amplification of variable length 16S-23S rDNA spacer regions. *Journal of General Microbiology*, *139*(12), 3089–3097. https://doi.org/10.1099/00221287-139-12-3089

Gyarmati, P., Kjellander, C., Aust, C., Song, Y., Öhrmalm, L., & Giske, C. G. (2016). Metagenomic analysis of bloodstream infections in patients with acute leukemia and therapy-induced neutropenia. *Sci Rep*, *6*.

Haag, H., Locher, F., & Nolte, O. (2013). Molecular diagnosis of microbial aetiologies using SepsiTest™ in the daily routine of a diagnostic laboratory. *Diagnostic Microbiology and Infectious Disease*, *76*(4), 413–418.

Hansen, W. L. J., Bruggeman, C. A., & Wolffs, P. F. G. (2009). Evaluation of new preanalysis sample treatment tools and DNA isolation protocols to improve bacterial pathogen detection in whole blood. *J Clin Microbiol*, *47*(8), 2629–2631.

Harris, D. M., & Hata, D. J. (2013). Rapid identification of bacteria and Candida using PNA-FISH from blood and peritoneal fluid cultures: a retrospective clinical study. *Annals of Clinical Microbiology and Antimicrobials*, *12*(1), 2.

Hasman, H., Saputra, D., Sicheritz-Ponten, T., Lund, O., Svendsen, C. A., Frimodt-Moller, N., & Aarestrup, F. M. (2014). Rapid whole-genome sequencing for

detection and characterization of microorganisms directly from clinical samples. *J Clin Microbiol*, *52*(1), 139–146. https://doi.org/10.1128/JCM.02452-13

Hay, A., Birnie, K., Busby, J., Delaney, B., Downing, H., Dudley, J., … Hollingworth, W. (2016). The Diagnosis of Urinary Tract infection in Young children (DUTY): a diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute i.

Hayden, R. T., Clinton, L. K., Hewitt, C., Koyamatsu, T., Sun, Y., Jamison, G., … Bankowski, M. J. (2016). Rapid antimicrobial susceptibility testing using forward laser light scatter technology. *J Clin Microbiol*, *54*(11), 2701–2706.

He, M., Miyajima, F., Roberts, P., Ellison, L., Pickard, D. J., Martin, M. J., … Lawley, T. D. (2013). Emergence and global spread of epidemic healthcare-associated Clostridium difficile. *Nat Genet*, *45*(1), 109–113. https://doi.org/10.1038/ng.2478

Hevia, A., Delgado, S., Margolles, A., & Sanchez, B. (2015). Application of density gradient for the isolation of the fecal microbial stool component and the potential use thereof. *Sci Rep*, *5*, 16807. https://doi.org/10.1038/srep16807

Hewitt, F. C., Guertin, S. L., Ternus, K. L., Schulte, K., & Kadavy, D. R. (2017). Toward Rapid Sequenced-Based Detection and Characterization of Causative Agents of Bacteremia. *BioRxiv*, 162735.

Hill, J. T., Tran, K.-D. T., Barton, K. L., Labreche, M. J., & Sharp, S. E. (2014). Evaluation of the nanosphere Verigene BC-GN assay for direct identification of gram-negative bacilli and antibiotic resistance markers from positive blood cultures and potential impact for more-rapid antibiotic interventions. *J Clin Microbiol*, *52*(10), 3805–3807.

Holt, K. E., Phan, M. D., Baker, S., Duy, P. T., Nga, T. V, Nair, S., … Wain, J. (2011). Emergence of a globally dominant IncHI1 plasmid type associated with multiple drug resistant typhoid. *PLoS Negl Trop Dis*, *5*(7), e1245. https://doi.org/10.1371/journal.pntd.0001245

Horz, H.-P., Scheer, S., Huenger, F., Vianna, M. E., & Conrads, G. (2008). Selective isolation of bacterial DNA from human clinical specimens. *J Microbiol Methods*, *72*(1), 98–102.

Horz, H. P., Scheer, S., Vianna, M. E., & Conrads, G. (2010). New methods for selective isolation of bacterial DNA from human clinical specimens. *Anaerobe*, *16*(1), 47–53. https://doi.org/10.1016/j.anaerobe.2009.04.009

Hosokawa, M., Nishikawa, Y., Kogawa, M., & Takeyama, H. (2017). Massively parallel whole genome amplification for single-cell sequencing using droplet microfluidics. *Sci Rep*, *7*(1), 5199.

Huang, T. D., Bogaerts, P., Ghilani, E., Heinrichs, A., Gavage, P., Roisin, S., … Glupczynski, Y. (2015). Multicentre evaluation of the Check-Direct CPE(R) assay for direct screening of carbapenemase-producing Enterobacteriaceae from rectal swabs. *J Antimicrob Chemother*, *70*(6), 1669–1673. https://doi.org/10.1093/jac/dkv009

Huber, C. A., Foster, N. F., Riley, T. V, & Paterson, D. L. (2013). Challenges for standardization of Clostridium difficile typing methods. *J Clin Microbiol*, *51*(9), 2810–2814.

Ip, C. L. C., Loose, M., Tyson, J. R., de Cesare, M., Brown, B. L., Jain, M., … Urban, J. M. (2015). MinION Analysis and Reference Consortium: Phase 1 data release and analysis. *F1000Res*, *4*.

Ipe, D. S., Sundac, L., Benjamin Jr, W. H., Moore, K. H., & Ulett, G. C. (2013). Asymptomatic bacteriuria: prevalence rates of causal microorganisms, etiology of infection in different patient populations, and recent advances in molecular detection. *FEMS Microbiol Lett*, *346*(1), 1–10.

Iyengar, V., Albaugh, G. P., Lohani, A., & Nair, P. P. (1991). Human stools as a source of viable colonic epithelial cells. *Faseb J*, *5*(13), 2856–2859.

Jain, M., Fiddes, I. T., Miga, K. H., Olsen, H. E., Paten, B., & Akeson, M. (2015). Improved data analysis for the MinION nanopore sequencer. *Nat Methods*, *12*(4), 351–356.

Jain, M., Koren, S., Quick, J., Rand, A. C., Sasani, T. A., Tyson, J. R., … Malla, S. (2017). Nanopore sequencing and assembly of a human genome with ultra-long reads. *BioRxiv*, 128835.

Jorgensen, J. H., & Ferraro, M. J. (2009). Antimicrobial susceptibility testing: a review of general principles and contemporary practices. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *49*(11), 1749–1755. https://doi.org/10.1086/647952

Juul, S., Izquierdo, F., Hurst, A., Dai, X., Wright, A., Kulesha, E., … Turner, D. J. (2015). What9s in my pot? Real-time species identification on the MinION. *BioRxiv*, 30742.

Kachrimanidou, M., & Malisiovas, N. (2011). Clostridium difficile infection: a

comprehensive review. *Critical Reviews in Microbiology*, *37*(3), 178–187.

Karlsson, E., Larkeryd, A., Sjodin, A., Forsman, M., & Stenberg, P. (2015). Scaffolding of a bacterial genome using MinION nanopore sequencing. *Sci Rep*, *5*, 11996. https://doi.org/10.1038/srep11996

Kellogg, J. A., Manzella, J. P., & Bankert, D. A. (2000). Frequency of low-level bacteremia in children from birth to fifteen years of age. *J Clin Microbiol*, *38*(6), 2181–2185.

Kelly, MD, C. P., & LaMont, MD, J. T. (1998). *CLOSTRIDIUM DIFFICILE* INFECTION. *Annual Review of Medicine*, *49*(1), 375–390. https://doi.org/10.1146/annurev.med.49.1.375

Kiełbasa, S. M., Wan, R., Sato, K., Horton, P., & Frith, M. C. (2011). Adaptive seeds tame genomic sequence comparison. *Genome Res*, *21*(3), 487–493.

Killgore, G., Thompson, A., Johnson, S., Brazier, J., Kuijper, E., Pepin, J., … van den Berg, R. J. (2008). Comparison of seven techniques for typing international epidemic strains of Clostridium difficile: restriction endonuclease analysis, pulsed-field gel electrophoresis, PCR-ribotyping, multilocus sequence typing, multilocus variable-number tandem-repeat an. *J Clin Microbiol*, *46*(2), 431–437.

Knetsch, C. W., Lawley, T. D., Hensgens, M. P., Corver, J., Wilcox, M. W., & Kuijper, E. J. (2013). Current application and future perspectives of molecular typing methods to study Clostridium difficile infections. *Euro Surveill*, *18*(4), 20381.

Kohn, A. (1960). Lysis of frozen and thawed cells of Escherichia coli by lysozyme, and their conversion into spheroplasts. *J Bacteriol*, *79*(5), 697.

Kollef, M. H. (2000). Inadequate antimicrobial treatment: an important determinant of outcome for hospitalized patients. *Clin Infect Dis*, *31 Suppl 4*, S131-8. https://doi.org/10.1086/314079

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/gr.215087.116

Köser, C. U., Ellington, M. J., Cartwright, E. J. P., Gillespie, S. H., Brown, N. M., Farrington, M., … Parkhill, J. (2012). Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathogens*, *8*(8), e1002824.

Krcmery, S., Hromec, J., & Demesova, D. (2001). Treatment of lower urinary tract

infection in pregnancy. *International Journal of Antimicrobial Agents*, *17*(4), 279–282.

Kuczynski, J., Lauber, C. L., Walters, W. A., Parfrey, L. W., Clemente, J. C., Gevers, D., & Knight, R. (2012). Experimental and analytical tools for studying the human microbiome. *Nature Reviews Genetics*, *13*(1), 47–58.

Kuhn, C., Disque, C., Muhl, H., Orszag, P., Stiesch, M., & Haverich, A. (2011). Evaluation of commercial universal rRNA gene PCR plus sequencing tests for identification of bacteria and fungi associated with infectious endocarditis. *J Clin Microbiol*, *49*(8), 2919–2923. https://doi.org/10.1128/jcm.00830-11

Kuijper, E. J., Coignard, B., & Tüll, P. (2006). Emergence of Clostridium difficile-associated disease in North America and Europe. *Clinical Microbiology and Infection*, *12*(s6), 2–18.

Kumar, A., Ellis, P., Arabi, Y., Roberts, D., Light, B., Parrillo, J. E., … Chateau, D. (2009). Initiation of inappropriate antimicrobial therapy results in a fivefold reduction of survival in human septic shock. *Chest*, *136*(5), 1237–1248. https://doi.org/10.1378/chest.09-0087

Kumar, A., Roberts, D., Wood, K. E., Light, B., Parrillo, J. E., Sharma, S., … Cheang, M. (2006). Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit Care Med*, *34*(6), 1589–1596. https://doi.org/10.1097/01.CCM.0000217961.75225.E9

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.

Lawley, T. D., Clare, S., Walker, A. W., Goulding, D., Stabler, R. A., Croucher, N., … Dougan, G. (2009). Antibiotic treatment of clostridium difficile carrier mice triggers a supershedder state, spore-mediated transmission, and severe disease in immunocompromised hosts. *Infect Immun*, *77*(9), 3661–3669. https://doi.org/10.1128/iai.00558-09

Leffler, D. A., & Lamont, J. T. (2015). Clostridium difficile infection. *New England Journal of Medicine*, *372*(16), 1539–1548.

Lehmann, L. E., Hauser, S., Malinka, T., Klaschik, S., Weber, S. U., Schewe, J.-C., … Book, M. (2011). Rapid qualitative urinary tract infection pathogen identification by SeptiFast® real-time PCR. *PLoS One*, *6*(2), e17146.

Lehmann, L. E., Hunfeld, K.-P., Steinbrucker, M., Brade, V., Book, M., Seifert, H., …

Stüber, F. (2010). Improved detection of blood stream pathogens by real-time PCR in severe sepsis. *Intensive Care Medicine*, *36*(1), 49–56.

Lessa, F. C., Mu, Y., Bamberg, W. M., Beldavs, Z. G., Dumyati, G. K., Dunn, J. R., … Phipps, E. C. (2015). Burden of Clostridium difficile infection in the United States. *New England Journal of Medicine*, *372*(9), 825–834.

Li, B., Qiu, Y., Glidle, A., McIlvenna, D., Luo, Q., Cooper, J., … Yin, H. (2014). Gradient microfluidics enables rapid bacterial growth inhibition testing. *Analytical Chemistry*, *86*(6), 3131–3137. https://doi.org/10.1021/ac5001306

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv Preprint ArXiv:1303.3997*.

Li, H. (2016). Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics*, *32*(14), 2103–2110.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079.

Liesenfeld, O., Lehman, L., Hunfeld, K. P., & Kost, G. (2014). Molecular diagnosis of sepsis: New aspects and recent developments. *Eur J Microbiol Immunol (Bp)*, *4*(1), 1–25. https://doi.org/10.1556/EuJMI.4.2014.1.1

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., … Ecker, J. R. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, *462*(7271), 315–322. https://doi.org/10.1038/nature08514

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., … Law, M. (2012). Comparison of next-generation sequencing systems. *BioMed Research International*, *2012*.

Livermore, D. M., & Wain, J. (2013). Revolutionising bacteriology to improve treatment outcomes and antibiotic stewardship. *Infect Chemother*, *45*(1), 1–10. https://doi.org/10.3947/ic.2013.45.1.1

Lofgren, E. T., Cole, S. R., Weber, D. J., Anderson, D. J., & Moehring, R. W. (2014). Hospital-Acquired Clostridium difficile Infections Estimating All-Cause Mortality and Length of Stay. *Epidemiology (Cambridge, Mass.)*, *25*(4), 570.

Loman, N. J., Quick, J., & Simpson, J. T. (2015). A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat Methods*, *12*(8), 733–735. https://doi.org/10.1038/nmeth.3444

Loman, N. J., & Quinlan, A. R. (2014). Poretools: a toolkit for analyzing nanopore

sequence data. *Bioinformatics*, *30*(23), 3399–3401.

Longtin, Y., Trottier, S., Brochu, G., Paquet-Bolduc, B., Garenc, C., Loungnarath, V., … Longtin, J. (2012). Impact of the type of diagnostic assay on Clostridium difficile infection and complication rates in a mandatory reporting program. *Clinical Infectious Diseases*, *56*(1), 67–73.

Loonen, A. J., Bos, M. P., van Meerbergen, B., Neerken, S., Catsburg, A., Dobbelaer, I., … van den Brule, A. J. (2013). Comparison of pathogen DNA isolation methods from large volumes of whole blood to improve molecular diagnosis of bloodstream infections. *PLoS One*, *8*(8), e72349. https://doi.org/10.1371/journal.pone.0072349

Loose, M. W. (2017). The potential impact of nanopore sequencing on human genetics. *Human Molecular Genetics*, *26*(R2), R202–R207.

López-Pérez, M., Gonzaga, A., Martin-Cuadrado, A. B., Onyshchenko, O., Ghavidel, A., Ghai, R., & Rodriguez-Valera, F. (2012). Genomes of surface isolates of Alteromonas macleodii: The life of a widespread marine opportunistic copiotroph. *Scientific Reports*, *2*. https://doi.org/10.1038/srep00696

Lucke, K., Hombach, M., Hug, M., & Pfyffer, G. E. (2010). Rapid detection of methicillin-resistant Staphylococcus aureus (MRSA) in diverse clinical specimens by the BD GeneOhm MRSA assay and comparison with culture. *J Clin Microbiol*, *48*(3), 981–984.

Maddocks, S., Olma, T., & Chen, S. (2002). Comparison of CHROMagar Salmonella medium and xylose-lysine-desoxycholate and Salmonella-Shigella agars for isolation of Salmonella strains from stool samples. *J Clin Microbiol*, *40*(8), 2999–3003.

Mancini, N., Carletti, S., Ghidoli, N., Cichero, P., Burioni, R., & Clementi, M. (2010). The era of molecular and other non-culture-based methods in diagnosis of sepsis. *Clin Microbiol Rev*, *23*(1), 235–251. https://doi.org/10.1128/CMR.00043-09

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., … Marteau, P. (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut*, *55*(2), 205–211.

Manoni, F., Fornasiero, L., Ercolin, M., Tinello, A., Ferrian, M., Hoffer, P., … Gessoni, G. (2009). Cutoff values for bacteria and leukocytes for urine flow cytometer Sysmex UF-1000i in urinary tract infections. *Diagnostic Microbiology*

and Infectious Disease, *65*(2), 103–107.

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, *24*(3), 133–141.

Marler, L. M., Siders, J. A., Wolters, L. C., Pettigrew, Y., Skitt, B. L., & Allen, S. D. (1992). Comparison of five cultural procedures for isolation of clostridium difficile from stools. *Journal of Clinical Microbiology*.

McArthur, A. G., Waglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A. J., … Ejim, L. (2013). The comprehensive antibiotic resistance database. *Antimicrobial Agents and Chemotherapy*, *57*(7), 3348–3357.

McCann, C. D., & Jordan, J. A. (2014). Evaluation of MolYsis Complete5 DNA extraction method for detecting Staphylococcus aureus DNA from whole blood in a sepsis model using PCR/pyrosequencing. *J Microbiol Methods*, *99*, 1–7. https://doi.org/10.1016/j.mimet.2014.01.013

McPherson, D., Griffiths, C., Williams, M., Baker, A., Klodawski, E., Jacobson, B., & Donaldson, L. (2013). Sepsis-associated mortality in England: an analysis of multiple cause of death data from 2001 to 2010. *BMJ Open*, *3*(8). https://doi.org/10.1136/bmjopen-2013-002586

Mellmann, A., Cloud, J., Maier, T., Keckevoet, U., Ramminger, I., Iwen, P., … Harmsen, D. (2008). Evaluation of matrix-assisted laser desorption ionization-time-of-flight mass spectrometry in comparison to 16S rRNA gene sequencing for species identification of nonfermenting bacteria. *Journal of Clinical Microbiology*, *46*(6), 1946–1954. https://doi.org/10.1128/JCM.00157-08

Melnikov, A., Galinsky, K., Rogov, P., Fennell, T., Van Tyne, D., Russ, C., … Neafsey, D. E. (2011). Hybrid selection for sequencing pathogen genomes from clinical samples. *Genome Biol*, *12*(8), R73. https://doi.org/10.1186/gb-2011-12-8-r73

Melzer, M., & Welch, C. (2013). Outcomes in UK patients with hospital-acquired bacteraemia and the risk of catheter-associated urinary tract infections. *Postgraduate Medical Journal*, *89*(1052), 329–334.

Menozzi, M. G., Eigner, U., Covan, S., Rossi, S., Somenzi, P., Dettori, G., … Fahr, A. M. (2006). Two-center collaborative evaluation of performance of the BD phoenix automated microbiology system for identification and antimicrobial susceptibility testing of gram-negative bacteria. *Journal of Clinical Microbiology*, *44*(11), 4085–4094. https://doi.org/10.1128/JCM.00614-06

Moir, A., Corfe, B. M., & Behravan, J. (2002). Spore germination. *Cellular and Molecular Life Sciences*, *59*(3), 403–409.

Morris, K., Wilson, C., & Wilcox, M. H. (2012). Evaluation of chromogenic meticillin-resistant Staphylococcus aureus media: sensitivity versus turnaround time. *Journal of Hospital Infection*, *81*(1), 20–24.

Muhl, H., Kochem, A. J., Disque, C., & Sakka, S. G. (2010). Activity and DNA contamination of commercial polymerase chain reaction reagents for the universal 16S rDNA real-time polymerase chain reaction detection of bacterial pathogens in blood. *Diagn Microbiol Infect Dis*, *66*(1), 41–49. https://doi.org/10.1016/j.diagmicrobio.2008.07.011

Murray, C., Adeyiga, O., Owsley, K., & Di Carlo, D. (2015). Research highlights: microfluidic analysis of antimicrobial susceptibility. *Lab on a Chip*, *15*(5), 1226–1229. https://doi.org/10.1039/c5lc90017d

Mwaigwisya, S., Assiri, R. A., & O'Grady, J. (2015). Emerging commercial molecular tests for the diagnosis of bloodstream infection. *Expert Rev Mol Diagn*, *15*(5), 681–692. https://doi.org/10.1586/14737159.2015.1029459

Naccache, S. N., Federman, S., Veeraraghavan, N., Zaharia, M., Lee, D., Samayoa, E., … Enge, B. (2014). A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. *Genome Res*, *24*(7), 1180–1192.

Nair, S., Ashton, P., Doumith, M., Connell, S., Painset, A., Mwaigwisya, S., … Day, M. (2016). WGS for surveillance of antimicrobial resistance: a pilot study to detect the prevalence and mechanism of resistance to azithromycin in a UK population of non-typhoidal Salmonella. *Journal of Antimicrobial Chemotherapy*, *71*(12), 3400–3408.

Neely, L. A., Audeh, M., Phung, N. A., Min, M., Suchocki, A., Plourde, D., … Lowery, T. J. (2013). T2 magnetic resonance enables nanoparticle-mediated rapid detection of candidemia in whole blood. *Sci Transl Med*, *5*(182), 182ra54. https://doi.org/10.1126/scitranslmed.3005377

Norling, M., Karlsson-Lindsjö, O. E., Gourlé, H., Bongcam-Rudloff, E., & Hayer, J. (2016). MetLab: An in silico experimental design, simulation and analysis tool for viral metagenomics studies. *PLoS ONE*, *11*(8). https://doi.org/10.1371/journal.pone.0160334

O'Neill, J. (2014). Antimicrobial Resistance : Tackling a crisis for the health and

wealth of nations. *Review on Antimicrobial Resistance*, (December), 1–16. https://doi.org/10.1038/510015a

Opota, O., Jaton, K., & Greub, G. (2015). Microbial diagnosis of bloodstream infection: Towards molecular diagnosis directly from blood. *Clinical Microbiology and Infection*. https://doi.org/10.1016/j.cmi.2015.02.005

Oyola, S. O., Gu, Y., Manske, M., Otto, T. D., O'Brien, J., Alcock, D., … Quail, M. A. (2013). Efficient depletion of host DNA contamination in malaria clinical sequencing. *J Clin Microbiol*, *51*(3), 745–751. https://doi.org/10.1128/jcm.02507-12

Pancholi, P., Kelly, C., Raczkowski, M., & Balada-Llasat, J. M. (2012). Detection of toxigenic Clostridium difficile: comparison of the cell culture neutralization, Xpert C. difficile, Xpert C. difficile/Epi, and Illumigene C. difficile assays. *J Clin Microbiol*, *50*(4), 1331–1335.

Paredes-Sabja, D., Shen, A., & Sorg, J. A. (2014). Clostridium difficile spore biology: sporulation, germination, and spore structural proteins. *Trends in Microbiology*, *22*(7), 406–416.

Patel, H. D., Livsey, S. A., Swann, R. A., & Bukhari, S. S. (2005). Can urine dipstick testing for urinary tract infection at point of care reduce laboratory workload? *Journal of Clinical Pathology*, *58*(9), 951–954.

Patel, J. B. (2001). 16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. *Molecular Diagnosis*, *6*(4), 313–321.

Paulino, L. C., Tseng, C. H., Strober, B. E., & Blaser, M. J. (2006). Molecular analysis of fungal microbiota in samples from healthy human skin and psoriatic lesions. *J Clin Microbiol*, *44*(8), 2933–2941. https://doi.org/10.1128/jcm.00785-06

Penders, J., Vink, C., Driessen, C., London, N., Thijs, C., & Stobberingh, E. E. (2005). Quantification of Bifidobacterium spp., Escherichia coli and Clostridium difficile in faecal samples of breast-fed and formula-fed infants by real-time PCR. *FEMS Microbiol Lett*, *243*(1), 141–147. https://doi.org/10.1016/j.femsle.2004.11.052

Perry-O'Keefe, H., Rigby, S., Oliveira, K., Sørensen, D., Stender, H., Coull, J., & Hyldig-Nielsen, J. J. (2001). Identification of indicator microorganisms using a standardized PNA FISH method. *Journal of Microbiological Methods*, *47*(3), 281–292. https://doi.org/10.1016/S0167-7012(01)00303-7

Pezzlo, M. (1988). Detection of urinary tract infections by rapid methods. *Clinical Microbiology Reviews*, *1*(3), 268–280.

Poretsky, R., Rodriguez-R, L. M., Luo, C., Tsementzi, D., & Konstantinidis, K. T. (2014). Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. *PLoS One*, *9*(4), e93827.

Putnam, L. R., Howard, W. J., Pfaller, M. A., Koontz, F. P., & Jones, R. N. (1997). Accuracy of the vitek system for antimicrobial susceptibility testing enterobacteriaceae bloodstream infection isolates: Use of “direct” inoculation from bactec 9240 blood culture bottles. *Diagnostic Microbiology and Infectious Disease*, *28*(2), 101–104. https://doi.org/10.1016/S0732-8893(97)00014-X

Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., … Mikhail, A. (2016). Real-time, portable genome sequencing for Ebola surveillance. *Nature*, *530*(7589), 228–232.

Quick, J., Quinlan, A. R., & Loman, N. J. (2014). A reference bacterial genome dataset generated on the MinION portable single-molecule nanopore sequencer. *Gigascience*, *3*, 22. https://doi.org/10.1186/2047-217x-3-22

Raoult, D., Fournier, P. E., & Drancourt, M. (2004). What does the future hold for clinical microbiology? *Nature Reviews Microbiology*, *2*(2), 151–159. https://doi.org/Doi 10.1038/Nrmicro820

Ray, B., & Speck, M. L. (1973). Freeze-injury in bacteria. *CRC Critical Reviews in Clinical Laboratory Sciences*, *4*(2), 161–213.

Reece, A., Xia, B., Jiang, Z., Noren, B., McBride, R., & Oakey, J. (2016). Microfluidic techniques for high throughput single cell analysis. *Current Opinion in Biotechnology*. https://doi.org/10.1016/j.copbio.2016.02.015

Resistance, R. on A. (2016). *Tackling drug-resistant infections globally: final report and recommendations*. Review on Antimicrobial Resistance.

Rhoads, D. D., Sintchenko, V., Rauch, C. A., & Pantanowitz, L. (2014). Clinical microbiology informatics. *Clin Microbiol Rev*, *27*(4), 1025–1047. https://doi.org/10.1128/cmr.00049-14

Rhodes, A., Evans, L. E., Alhazzani, W., Levy, M. M., Antonelli, M., Ferrer, R., … Dellinger, R. P. (2017). Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. *Intensive Care Medicine*, *43*(3), 304–377. https://doi.org/10.1007/s00134-017-4683-6

Richards, M., Knox, J., Elliott, B., Mackin, K., Lyras, D., Waring, L. J., & Riley, T. V.

(2011). Severe infection with Clostridium difficile PCR ribotype 027 acquired in Melbourne, Australia. *Medical Journal of Australia*, *194*(7), 369–371.

Riley, T. V, Brazier, J. S., Hassan, H., Williams, K., & Phillips, K. D. (1987). Comparison of alcohol shock enrichment and selective enrichment for the isolation of Clostridium difficile. *Epidemiology & Infection*, *99*(2), 355–359.

Rivers, E., Nguyen, B., Havstad, S., Ressler, J., Muzzin, A., Knoblich, B., … Early Goal-Directed Therapy Collaborative, G. (2001). Early goal-directed therapy in the treatment of severe sepsis and septic shock. *N Engl J Med*, *345*(19), 1368–1377. https://doi.org/10.1056/NEJMoa010307

Roberts, K. B. (2011). Urinary tract infection: clinical practice guideline for the diagnosis and management of the initial UTI in febrile infants and children 2 to 24 months. Am Acad Pediatrics.

Rogol, M., Shpak, B., Rothman, D., & Sechter, I. (1985). Enrichment medium for isolation of Campylobacter jejuni-Campylobacter coli. *Applied and Environmental Microbiology*, *50*(1), 125–126.

Rose, G., Wooldridge, D. J., Anscombe, C., Mee, E. T., Misra, R. V., & Gharbia, S. (2015). Challenges of the Unknown: Clinical Application of Microbial Metagenomics. *International Journal of Genomics*, *2015*. https://doi.org/10.1155/2015/292950

Sabat, A. J., Budimir, A., Nashev, D., Sá-Leão, R., Van Dijl, J. M., Laurent, F., … Markers, E. S. G. of E. (2013). Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill*, *18*(4), 20380.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., … Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*. https://doi.org/10.1126/science.2448875

Salipante, S. J., Sengupta, D. J., Hoogestraat, D. R., Cummings, L. A., Bryant, B. H., Natividad, C., … Hoffman, N. G. (2013). Molecular diagnosis of Actinomadura madurae infection by 16S rRNA deep sequencing. *J Clin Microbiol*, *51*(12), 4262–4265. https://doi.org/10.1128/jcm.02227-13

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., … Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol*, *12*(1), 87.

Schappert, S. M., & Rechtsteiner, E. A. (2011). Ambulatory medical care utilization estimates for 2007. *Vital and Health Statistics. Series 13, Data from the National Health Survey*, (169), 1–38.

Schlaberg, R., Chiu, C. Y., Miller, S., Procop, G. W., Weinstock, G., Committee, P. P., … Pathologists, M. R. C. of the C. of A. (2017). Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Archives of Pathology and Laboratory Medicine*, *141*(6), 776–786.

Schmidt, K., Mwaigwisya, S., Crossman, L. C., Doumith, M., Munroe, D., Pires, C., … Wain, J. (2016). Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *Journal of Antimicrobial Chemotherapy*, *72*(1), 104–114.

Schmieder, R., & Edwards, R. (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One*, *6*(3), e17288.

Schmiemann, G., Kniehl, E., Gebhardt, K., Matejczyk, M. M., & Hummers-Pradier, E. (2010). The diagnosis of urinary tract infection: a systematic review. *Deutsches Ärzteblatt International*, *107*(21), 361.

Schoepp, N. G., Khorosheva, E. M., Schlappi, T. S., Curtis, M. S., Humphries, R. M., Hindler, J. A., & Ismagilov, R. F. (2016). Digital Quantification of DNA Replication and Chromosome Segregation Enables Determination of Antimicrobial Susceptibility after only 15 Minutes of Antibiotic Exposure. *Angewandte Chemie - International Edition*, *55*(33), 9557–9561. https://doi.org/10.1002/anie.201602763

Schröder, U.-C., Bokeloh, F., O'Sullivan, M., Glaser, U., Wolf, K., Pfister, W., … Neugebauer, U. (2015). Rapid, culture-independent, optical diagnostics of centrifugally captured bacteria from urine samples. *Biomicrofluidics*, *9*(4), 44118.

Schulte, W., Bernhagen, J., & Bucala, R. (2013). Cytokines in sepsis: potent immunoregulators and potential therapeutic targets—an updated view. *Mediators of Inflammation*, *2013*.

Semeniuk, H., & Church, D. (1999). Evaluation of the leukocyte esterase and nitrite urine dipstick screening tests for detection of bacteriuria in women with suspected uncomplicated urinary tract infections. *Journal of Clinical Microbiology*, *37*(9), 3051–3052.

Sender, R., Fuchs, S., & Milo, R. (2016). Revised Estimates for the Number of

Human and Bacteria Cells in the Body. *PLoS Biol*, *14*(8), e1002533. https://doi.org/10.1371/journal.pbio.1002533

Seng, P., Abat, C., Rolain, J. M., Colson, P., Lagier, J.-C., Gouriet, F., … Raoult, D. (2013). Identification of rare pathogenic bacteria in a clinical microbiology laboratory: impact of MALDI-TOF mass spectrometry. *J Clin Microbiol*, JCM. 00492-13.

Shigemura, K., Shirakawa, T., Okada, H., Tanaka, K., Kamidono, S., Arakawa, S., & Gotoh, A. (2005). Rapid detection and differentiation of Gram-negative and Gram-positive pathogenic bacteria in urine using TaqMan probe. *Clinical and Experimental Medicine*, *4*(4), 196–201.

Shrestha, N. K., Tuohy, M. J., Hall, G. S., Isada, C. M., & Procop, G. W. (2002). Rapid identification of Staphylococcus aureus and the mecA gene from BacT/ALERT blood culture bottles by using the LightCycler system. *J Clin Microbiol*, *40*(7), 2659–2661.

Simner, P. J., Uhl, J. R., Hall, L., Weber, M. M., Walchak, R. C., Buckwalter, S., & Wengenack, N. L. (2013). Broad-range direct detection and identification of fungi by use of the PLEX-ID PCR-electrospray ionization mass spectrometry (ESIMS) system. *Journal of Clinical Microbiology*, *51*(6), 1699–1706. https://doi.org/10.1128/JCM.03282-12

Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., … Coopersmith, C. M. (2016). The third international consensus definitions for sepsis and septic shock (sepsis-3). *Jama*, *315*(8), 801–810.

Sixt, B. S., Siegl, A., Müller, C., Watzka, M., Wultsch, A., Tziotis, D., … Horn, M. (2013). Metabolic features of Protochlamydia amoebophila elementary bodies–a link between activity and infectivity in Chlamydiae. *PLoS Pathogens*, *9*(8), e1003553.

Smith, A. M., Jain, M., Mulroney, L., Garalde, D. R., & Akeson, M. (2017). Reading canonical and modified nucleotides in 16S ribosomal RNA using nanopore direct RNA sequencing. *BioRxiv*, 132274.

Sparbier, K., Weller, U., Boogen, C., & Kostrzewa, M. (2012). Rapid detection of Salmonella sp. by means of a combination of selective enrichment broth and MALDI-TOF MS. *European Journal of Clinical Microbiology & Infectious Diseases*, *31*(5), 767–773.

Springer, J., Loeffler, J., Heinz, W., Schlossnagel, H., Lehmann, M., Morton, O., …

Kurzai, O. (2011). Pathogen-specific DNA enrichment does not increase sensitivity of PCR for diagnosis of invasive aspergillosis in neutropenic patients. *J Clin Microbiol*, *49*(4), 1267–1273. https://doi.org/10.1128/jcm.01679-10

Stamm, W. E., & Norrby, S. R. (2001). Urinary Tract Infections: Disease Panorama and Challenges. *The Journal of Infectious Diseases*, *183*(s1), S1–S4. https://doi.org/10.1086/318850

Stamper, P. D., Babiker, W., Alcabasa, R., Aird, D., Wehrlin, J., Ikpeama, I., … Carroll, K. C. (2009). Evaluation of a new commercial TaqMan PCR assay for direct detection of the clostridium difficile toxin B gene in clinical stool specimens. *J Clin Microbiol*, *47*(12), 3846–3850. https://doi.org/10.1128/jcm.01490-09

Sterry-Blunt, R. E., Randall, K. S., Doughton, M. J., Aliyu, S. H., & Enoch, D. A. (2015). Screening urine samples for the absence of urinary tract infection using the sediMAX automated microscopy analyser. *Journal of Medical Microbiology*, *64*(6), 605–609.

Stevenson, M., Pandor, A., James, M.-S., Rafia, R., Uttley, L., Stevens, J., … McMullan, R. (2016). Sepsis: the LightCycler SeptiFast Test MGRADE®, SepsiTest™ and IRIDICA BAC BSI assay for rapidly identifying bloodstream bacteria and fungi–a systematic review and economic evaluation. *Health Technology Assessment*, *20*(46), 1–246.

Stewart, E. J. (2012). Growing unculturable bacteria. *J Bacteriol*, *194*(16), 4151–4160.

Stonehouse, M. J., Cota-Gomez, A., Parker, S. K., Martin, W. E., Hankin, J. A., Murphy, R. C., … Vasil, A. I. (2002). A novel class of microbial phosphocholine-specific phospholipases C. *Molecular Microbiology*, *46*(3), 661–676.

Stoneking, L. R., Patanwala, A. E., Winkler, J. P., Fiorello, A. B., Lee, E. S., Olson, D. P., & Wolk, D. M. (2013). Would earlier microbe identification alter antibiotic therapy in bacteremic emergency department patients? *J Emerg Med*, *44*(1), 1–8. https://doi.org/10.1016/j.jemermed.2012.02.036

Surawicz, C. M., Brandt, L. J., Binion, D. G., Ananthakrishnan, A. N., Curry, S. R., Gilligan, P. H., … Zuckerbraun, B. S. (2013). Guidelines for diagnosis, treatment, and prevention of Clostridium difficile infections. *The American Journal of Gastroenterology*, *108*(4), 478–498.

Telenti, A., Honoré, N., Bernasconi, C., March, J., Ortega, A., Heym, B., … Cole, S.

T. (1997). Genotypic assessment of isoniazid and rifampin resistance in Mycobacterium tuberculosis: A blind study at reference laboratory level. *Journal of Clinical Microbiology*, *35*(3), 719–723.

Tenover, F. C., Novak-Weekley, S., Woods, C. W., Peterson, L. R., Davis, T., Schreckenberger, P., … Nomura, J. H. (2010). Impact of strain type on detection of toxigenic Clostridium difficile: comparison of molecular diagnostic and enzyme immunoassay approaches. *J Clin Microbiol*, *48*(10), 3719–3724.

Tessari, A., Osti, N., & Scarin, M. (2015). Screening of presumptive urinary tract infections by the automated urine sediment analyser sediMAX. *Clinical Chemistry and Laboratory Medicine (CCLM)*, *53*(s2), s1503–s1508.

Thoendel, M., Jeraldo, P., Greenwood-Quaintance, K. E., Yao, J., Chia, N., Hanssen, A. D., … Patel, R. (2017). Impact of Contaminating DNA in Whole-Genome Amplification Kits Used for Metagenomic Shotgun Sequencing for Infection Diagnosis. *J Clin Microbiol*, *55*(6), 1789–1801.

Thoendel, M., Jeraldo, P. R., Greenwood-Quaintance, K. E., Yao, J. Z., Chia, N., Hanssen, A. D., … Patel, R. (2016). Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. *J Microbiol Methods*, *127*, 141–145.

Tille, P. (2013). *Bailey & Scott's Diagnostic Microbiology-E-Book*. Elsevier Health Sciences.

Tissari, P., Zumla, A., Tarkka, E., Mero, S., Savolainen, L., Vaara, M., … Gant, V. (2010). Accurate and rapid identification of bacterial species from positive blood cultures with a DNA-based microarray platform: an observational study. *The Lancet*, *375*(9710), 224–230. https://doi.org/10.1016/S0140-6736(09)61569-5

Titball, R. W., Leslie, D. L., Harvey, S., & Kelly, D. (1991). Hemolytic and sphingomyelinase activities of Clostridium perfringens alpha-toxin are dependent on a domain homologous to that of an enzyme from the human arachidonic acid pathway. *Infect Immun*, *59*(5), 1872–1874.

Torio, C. M. (2013). *National inpatient hospital costs: the most expensive conditions by payer, 2011. Healthcare Cost and Utilization Project Statistical Brief# 160*.

Treangen, T. J., & Salzberg, S. L. (2011). Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nature Reviews Genetics*, *13*(1), nrg3117.

Tyson, G. H., McDermott, P. F., Li, C., Chen, Y., Tadesse, D. A., Mukherjee, S., …

Loneragan, G. H. (2015). WGS accurately predicts antimicrobial resistance in Escherichia coli. *Journal of Antimicrobial Chemotherapy*, *70*(10), 2763–2769.

van der Zwet, W. C., Hessels, J., Canbolat, F., & Deckers, M. M. L. (2010). Evaluation of the Sysmex UF-1000i® urine flow cytometer in the diagnostic work-up of suspected urinary tract infection in a Dutch general hospital. *Clinical Chemistry and Laboratory Medicine*, *48*(12), 1765–1771.

Vandecandelaere, I., Nercessian, O., Segaert, E., Achouak, W., Mollica, A., Faimali, M., … Vandamme, P. (2008). Alteromonas genovensis sp. nov., isolated from a marine electroactive biofilm and emended description of Alteromonas macleodii Baumann et al. 1972 (Approved Lists 1980). *International Journal of Systematic and Evolutionary Microbiology*, *58*(11), 2589–2596. https://doi.org/10.1099/ijs.0.65691-0

Viala, C., Le Monnier, A., Maataoui, N., Rousseau, C., Collignon, A., & Poilane, I. (2012). Comparison of commercial molecular assays for toxigenic Clostridium difficile detection in stools: BD GeneOhm Cdiff, XPert C. difficile and illumigene C. difficile. *J Microbiol Methods*, *90*(2), 83–85.

Vince, A., Lepej, S. Ž., Baršić, B., Dušek, D., Mitrović, Z., Serventi-Seiwerth, R., & Labar, B. (2008). LightCycler SeptiFast assay as a tool for the rapid diagnosis of sepsis in patients during antimicrobial therapy. *Journal of Medical Microbiology*, *57*(10), 1306–1307.

Vincent, J.-L., Sakr, Y., Sprung, C. L., Ranieri, V. M., Reinhart, K., Gerlach, H., … Payen, D. (2006). Sepsis in European intensive care units: results of the SOAP study. *Critical Care Medicine*, *34*(2), 344–353.

Votintseva, A. A., Bradley, P., Pankhurst, L., del Ojo Elias, C., Loose, M., Nilgiriwala, K., … Walker, T. M. (2017). Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. *J Clin Microbiol*, *55*(5), 1285–1298.

Wagner, M., Horn, M., & Daims, H. (2003). Fluorescence in situ hybridisation for the identification and characterisation of prokaryotes. *Current Opinion in Microbiology*, *6*(3), 302–309. https://doi.org/10.1016/S1369-5274(03)00054-7

Wain, J., Diep, T. S., Ho, V. A., Walsh, A. M., Hoa, N. T. T., Parry, C. M., & White, N. J. (1998). Quantitation of bacteria in blood of typhoid fever patients and relationship between counts and clinical features, transmissibility, and antibiotic resistance. *J Clin Microbiol*, *36*(6), 1683–1687.

Wain, J., & Mavrogiorgou, E. (2013). Next-generation sequencing in clinical microbiology. *Expert Rev Mol Diagn*, *13*(3), 225–227. https://doi.org/10.1586/erm.13.8

Warhurst, G., Maddi, S., Dunn, G., Ghrew, M., Chadwick, P., Alexander, P., … Dark, P. (2015). Diagnostic accuracy of SeptiFast multi-pathogen real-time PCR in the setting of suspected healthcare-associated bloodstream infection. *Intensive Care Med*, *41*(1), 86–93. https://doi.org/10.1007/s00134-014-3551-x

Washington, J. A. (2012). *Laboratory procedures in clinical microbiology*. Springer Science & Business Media.

Wei, Y.-X., Zhang, Z.-Y., Liu, C., Zhu, Y.-Z., Zhu, Y.-Q., Zheng, H., … Guo, X.-K. (2010). Complete Genome Sequence of Bifidobacterium longum JDM301. *J Bacteriol*, *192*(15), 4076–4077. https://doi.org/10.1128/JB.00538-10

Weiss, S., Amir, A., Hyde, E. R., Metcalf, J. L., Song, S. J., & Knight, R. (2014). Tracking down the sources of experimental contamination in microbiome studies. *Genome Biol*, *15*(12), 564.

Wellinghausen, N., Kochem, A. J., Disque, C., Muhl, H., Gebert, S., Winter, J., … Sakka, S. G. (2009). Diagnosis of bacteremia in whole-blood samples by use of a commercial universal 16S rRNA gene-based PCR and sequence analysis. *J Clin Microbiol*, *47*(9), 2759–2765. https://doi.org/10.1128/JCM.00567-09

Wiesinger-Mayr, H., Jordana-Lluch, E., Martró, E., Schoenthaler, S., & Noehammer, C. (2011). Establishment of a semi-automated pathogen DNA isolation from whole blood and comparison with commercially available kits. *J Microbiol Methods*, *85*(3), 206–213.

Wilcox, M. H., Gerding, D. N., Poxton, I. R., Kelly, C., Nathan, R., Birch, T., … Lee, C. (2017). Bezlotoxumab for prevention of recurrent Clostridium difficile infection. *New England Journal of Medicine*, *376*(4), 305–317.

Wilcox, M. H., Shetty, N., Fawley, W. N., Shemko, M., Coen, P., Birtles, a, … Wren, M. W. D. (2012). Changing epidemiology of Clostridium difficile infection following the introduction of a national ribotyping-based surveillance scheme in England. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *55*(8), 1056–1063. https://doi.org/10.1093/cid/cis614

Wilson, M. L., & Gaido, L. (2004). Laboratory diagnosis of urinary tract infections in adult patients. *Clinical Infectious Diseases*, *38*(8), 1150–1158.

Wojewoda, C. M., Sercia, L., Navas, M., Tuohy, M., Wilson, D., Hall, G. S., … Richter, S. S. (2013). Evaluation of the Verigene Gram-positive blood culture nucleic acid test for rapid detection of bacteria and resistance determinants. *J Clin Microbiol*, *51*(7), 2072–2076.

Wood, D. E., & Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol*, *15*(3), R46. https://doi.org/10.1186/gb-2014-15-3-r46

Yagupsky, P., & Nolte, F. S. (1990). Quantitative aspects of septicemia. *Clinical Microbiology Reviews*, *3*(3), 269–279.

Yigit, E., Hernandez, D. I., Trujillo, J. T., Dimalanta, E., & Bailey, C. D. (2014). Genome and metagenome sequencing: Using the human methyl-binding domain to partition genomic DNA derived from plant tissues. *Applications in Plant Sciences*, *2*(11), 1400064.

Zaman, Z., Fogazzi, G. B., Garigali, G., Croci, M. D., Bayer, G., & Kránicz, T. (2010). Urine sediment analysis: analytical and diagnostic performance of sediMAX®—a new automated microscopy image-based urine sediment analyser. *Clinica Chimica Acta*, *411*(3), 147–154.

Zhang, X., Niu, S., & Zhang, L. (2014). Antimicrobial susceptibilities and clinical characterization of Pseudomonas aeruginosa isolates from urinary tract infections. *Urol Int*, *93*(4), 464–469. https://doi.org/10.1159/000358493

Zhao, S., Tyson, G. H., Chen, Y., Li, C., Mukherjee, S., Young, S., … McDermott, P. F. (2016). Whole-genome sequencing analysis accurately predicts antimicrobial resistance phenotypes in Campylobacter spp. *Applied and Environmental Microbiology*, *82*(2), 459–466.

Zhou, L., & Pollard, A. J. (2012). A novel method of selective removal of human DNA improves PCR sensitivity for detection of Salmonella Typhi in blood samples. *BMC Infect Dis*, *12*, 164. https://doi.org/10.1186/1471-2334-12-164

Ziegler, I., Fagerström, A., Strålin, K., & Mölling, P. (2016). Evaluation of a Commercial Multiplex PCR Assay for Detection of Pathogen DNA in Blood from Patients with Suspected Sepsis. *PLoS One*, *11*(12), e0167883.

Zimlichman, E., Henderson, D., Tamir, O., Franz, C., Song, P., Yamin, C. K., … Bates, D. W. (2013). Health care–associated infections: a meta-analysis of costs and financial impact on the US health care system. *JAMA Internal Medicine*, *173*(22), 2039–2046.

Website references:

DH 2012
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215135/dh_133016.pdf

House of Parliament, 2017
http://researchbriefings.parliament.uk/ResearchBriefing/Summary/POST-PN-0545

PHE, 2010
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/342707/CDRN_ERS_NHS_Trusts_V3_Mar_2010.pdf

PHE, 2014
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/619774/Urinary_tract_infection_UTI_guidance_summary.pdf

PHE 2017
https://www.gov.uk/government/publications/typhoid-and-paratyphoid-public-health-operational-guidelines

WHO 2014
http://apps.who.int/iris/bitstream/10665/112738/1/9789240692671_eng.pdf