Investigating a potential infectious cause of death in Sudden Unexpected Death in Infancy and Childhood using 16S rRNA gene sequencing

Lily Gates

A thesis submitted
for the degree of
Doctor of Philosphy
of
University College London

Great Ormond Street Institute of Child Health
University College London

June 2022

Declaration

The following people contributed to the work presented in this thesis.

The Histopathology department at Great Ormond Street Hospital formalin-fixed and paraffin embedded all tissue samples. Aimee Avery, Talisa Mistry, and Olumide Ogunbiyi assisted with the identification of suitable study cases for both the frozen and formalin-fixed paraffin-embedded (FFPE) material from Sudden Unexpected Death in Infancy and Childhood (SUDIC) cases and the collection and cutting of archival FFPE material. Victoria Bryant assisted with the identification of Sudden Unexplained Death in Childhood (SUDC) cases that held consent for scientific research. Diane Rampling identified chorioamnionitis human placenta cases and controls.

FFPE mouse placenta samples used in the FFPE optimisation were kindly donated by Dr. Natalie Suff. Mice used in the study of post-mortem (PM) bacterial translocation were culled and donated to this study by Dr. Diana Matei.

Bioinformatic processing of sequencing data was performed using VSEARCH pipelines obtained from Dr. Ronan Doyle. The pipeline was modified to assign OTUs based on the hypervariable region(s) sequenced.

Study design was performed by Lily Gates with assistance and consultation from Prof. Nigel Klein, Prof. Neil Sebire, and Dr. Dagmar Alber.

I, Lily Gates, confirm that all other work presented in this thesis is my own. Where information has been derived from other sources I confirm that this has been indicated in the relevant section of the thesis.

Abstract

Sudden Unexpected Death in Infancy (SUDI) is the most common cause of post-neonatal infant death in the developed world. Following thorough investigation approximately 40% of cases remain unexplained. It has been hypothesised that a subset of these currently unexplained infant deaths are due to infection that could not be diagnosed at PM examination due to limitations in current diagnostic techniques. The aim of this study was to optimise the 16S rRNA sequencing technique for post-mortem (PM) tissue samples from both SUDI cases and their childhood equivalent (collectively referred to as SUDIC) to investigate whether this technique can reliably identify infected tissues and assist in the PM diagnosis of infection.

This thesis investigated several aspects of the PM process in order to determine the use of the 16S rRNA gene sequencing technique in routine clinical diagnosis of infection. The process of PM bacterial translocation was investigated using the 16S rRNA gene sequencing technique in two animal models over a 14-day study period in order to determine the effects this process may have on bacteria identified within internal organs in a PM setting. This investigation found no evidence for significant PM bacterial translocation in the mouse model. In the piglet model, low-levels of bacteria were detected in some tissues, but there was no evidence for significant translocation from the gastrointestinal (GI) tract or nasal cavity.

With this insight into the PM process, a 16S rRNA gene sequencing method was successfully optimised for use in both FFPE and frozen tissues from archived cases of SUDI. These investigations included examination of the effects of different sequencing protocols and tissue processing.

The optimised 16S rRNA gene sequencing method was then performed on PM tissues from SUDIC cases. These cases were assigned to one of three groups depending on the final cause of death (CoD) as recorded on the autopsy report: (i) infectious, (ii) explained non-infectious, and (iii) unexplained. Results showed that the technique could successfully distinguish between explained non-infectious and infectious cases of SUDIC. The method correctly identified the causative organism in all cases of infectious SUDIC due to a bacterial infection where the pathogen had been identified at PM investigation. In four cases of infectious SUDIC where the causative organism was not identified at PM investigation using current routine techniques, a dominant bacterial pathogen was identified in this study using the

sequencing technique. Using control groups a threshold was defined to identify infected PM tissue samples. When applied to the unexplained cohort four cases of SUDIC were identified where overwhelming bacterial colonisation of PM tissues was observed. Considering clinical history and other PM findings, it is possible that these deaths had a potentially unidentified infectious cause.

The results of this thesis support the further evaluation and use of 16S rRNA gene sequencing in PM tissues from cases of SUDIC. The technique has the potential to identify infectious agents which are undetected using current techniques. Further work should increase the study cohort and perform this investigation prospectively as a part of routine clinical diagnosis of SUDIC.

Impact statement

The research undertaken in this thesis spans several fields including pathology, microbiology, molecular biology, forensics, and paediatrics. The findings may be of interest to academic researchers, microbiologsts, pathologists, as well as the wider public especially the SUDIC community.

To begin, this thesis shows that PM bacterial translocation is minimal when the cadaver is stored in refrigerated conditions prior to autopsy when investigating bacterial presence using the 16S ribosomal RNA (rRNA) gene sequencing technique. In some PM tissues, low levels of bacteria were identified, but these all showed mixed bacterial communities of physiological flora. These results may be of interest to clinical pathologists as they may contribute to the interpretation of PM microbiology results, as well as forensic pathologists who seek to understand the effects of the cause and manner of death on PM microbial communities.

This thesis also shows successful application the 16S rRNA gene sequencing to PM tissues from cases of SUDIC. Results demonstrate high specificity and sensitivity of this method when identifying cases of SUDIC with an infectious cause. These results provide the foundation for further investigations into the use of this technique in routine clinical testing in order to diagnose infection as a cause of death in SUDIC. Further investigations could be performed in future research projects or translational clinical studies.

The use of this technique in routine PM investigation of SUDIC has the potential to lead to an improved diagnosis of infection in SUDIC. This study highlighted cases of SUDIC that could potentially have occurred due to infection, reducing the proportion of cases that currently remain unexplained. These findings are of interest to academic researchers, clinical pathologists, patients and their families.

The methods described in this thesis would be of interest to academics. In particular the successful optimisation of DNA extraction from FFPE material in order

to perform 16S rRNA gene sequencing technique in this tissue. The typically damaged nucleic acids recovered from such material have been previously difficult to perform such investigations on. This would benefit many retrospective studies that wish to extract nucleic acids for molecular investigation.

Lastly and importantly, identifying cases of SUDIC that have an infectious cause that have been previously misclassified as unexplained allows for several changes. Improving the diagnosis of SUDIC and identifying cases with an infectious cause of death that had been previously unexplained, better research and investigation into SIDS will be possible. Understanding infections that are resulting in SUDIC may allow for the development interventions or screening methods which could save future lives. This project allows for a greater understanding of the aetiology of SIDS, which is a key aim of The Lullaby Trust who supported this project. It is important to reduce the burden of unexplained death on families experiencing this tragedy by offering a diagnosis other than SUDIC where possible.

Acknowledgements

Firstly, I would like to thank my supervisors for allowing me to take on this project. Your support, guidance, and expert knowledge has really helped shape this project. Nigel, although these past few years have been tough, you have still taken the time for weekly check-ins with the lab and always with positive words and a smile on your face. Neil, thank you for always making this work feel achievable. Each time we met and I was overwhelmed with data you always helped me reframe it and I always left feeling like everything was under control. Dagmar, thank you for being there with support and advice on a day-to-day basis and for always challenging me to view things from a different perspective. Your endless comments on my work really brought this study together. Thank you for believing in me and helping me to believe in myself.

A large portion of this work was performed with assistance from the GOSH histopathology team who were always so welcoming and friendly. A special thanks to Aimee Avery, Ciaran Hutchinson, Olumide Ogunbiyi, Talisa Mistry, and Toby Hunt for always making it a joy to pop over to collect samples and have a chat. Thank you to Ronan Doyle for helping me get up and running with all things bioinformatics and for reminding me that this will not last forever! Thank you to the Klein Lab and everyone in Office 646 for your support.

Thank you to The Lullaby Trust without whom this research would not be possible. It is a privilege to be involved in this area of research and to work alongside such an amazing charity. I would like to extend my thanks to the families who consented to research and allowed us to use such precious samples. Without you, this work would not have been possible.

I am very grateful for my wonderful friends and family who have managed to keep me sane and focused over the past few years and helped me carry on even in the most difficult of times. Thank you for your constant love and support.

A special thank you to Kerry-Anne Kite who helped me get through the days I wanted to give up. Thank you for all the cups of tea and for making me laugh every single day. I feel incredibly lucky to have met you on this journey and appreciate your friendship more than you will ever know.

Last, but by no means least, my family deserve my endless gratitude. Without your constant love, support, and guidance none of this would have been possible. Thank you to my mum and dad for building such strong foundations for me in my early years and thank you to my sister, Katie, for being my biggest supporter and always reminding me how I can achieve anything I put my mind to.

Contents

Abbreviations

ANKRD1 Ankyrin Repeat Domain 1

APOC4 Apolipoprotein C4

GAPDH Glyceraldehyde-3-Phosphate Dehydrogenase

SFTPC Surfactant Protein C

bp base pair

cDNA complementary DNA

CMV cytomegalovirus

CoD cause of death

CoNS coagulase-negative staphylococci

CSF cerebrospinal fluid

CT cyclic threshold

D0 day 0

D10 day 10

D14 day 14

D3 day 3

D7 day 7

eC explained case

FFPE formalin-fixed paraffin-embedded

GAS Group A *Streptococcus*

GI gastrointestinal

GOSH Great Ormond Street Hospital

GP general practitioner

HMC Her Majesty's Coroner

iC infectious case

ICD-10 International Classification of Diseases 10

mRNA messenger ribonucleic acid

NCBI National Center for Biotechnology Information

nf-H_2O nuclease-free H_2O

NGS next-generation sequencing

NTC no template control

OTU operational taxonomic unit

PBS phosphate-buffered saline

PM post-mortem

PMI post-mortem interval

PVC polyvinyl chloride

qPCR quantitative-PCR

RFU relative fluorescence units

rRNA ribosomal RNA

RSV respiratory syncytial virus

RT-qPCR reverse transcriptase quantitative polymerase chain reaction

SEA Staphylococcal enterotoxin A

SEB Staphylococcal enterotoxin B

SEC Staphylococcal enterotoxin C

SIDS Sudden Infant Death Syndrome

SNP single nucleotide polymorphism

SUDC Sudden Unexplained Death in Childhood

SUDI Sudden Unexpected Death in Infancy

SUDIC Sudden Unexpected Death in Infancy and Childhood

TSST-1 toxic shock syndrome toxin 1

uC unexplained case

UCSC University of California Santa Cruz

URT upper respiratory tract

UV ultraviolet

List of Figures

List of Tables

Introduction

1.1 Introduction and definitions

The sudden and unexpected death of a seemingly healthy child is arguably one of the most traumatic experiences a person can encounter. As will be discussed in greater depth throughout this thesis, many of these unexpected deaths share death circumstances, clinical presentation, and autopsy findings.

The thorough investigation of unexpected death in infants and children results in two death cohorts as shown in Figure 1.1. Explained sudden deaths are those with an identified CoD. In the case of an unexplained deaths the CoD remains unknown.

A lack of standardised terminology to describe these deaths, as well as variances in definitions used by different institutions, results in difficulties comparing statistics and characteristics across populations [1]. For the purpose of this thesis the following definitions will be used:

Sudden Unexpected Death in Infancy (SUDI) - The sudden and unexpected death of an apparently healthy infant occurring between 7 and 365 days of life where the cause is not immediately obvious.

Sudden Unexpected Death in Infancy and Childhood (SUDIC) - The sudden and unexpected death of an apparently healthy infant (7 - 365 days) or child (1 - 16 years) where the cause is not immediately obvious.

Sudden Infant Death Syndrome (SIDS) - A subset of SUDI cases where the cause of death cannot be determined following thorough investigation.

Sudden Unexplained Death in Childhood (SUDC) - The childhood equivalent of SIDS. A subset of unexpected death in childhood where the cause of death cannot be determined following thorough investigation.

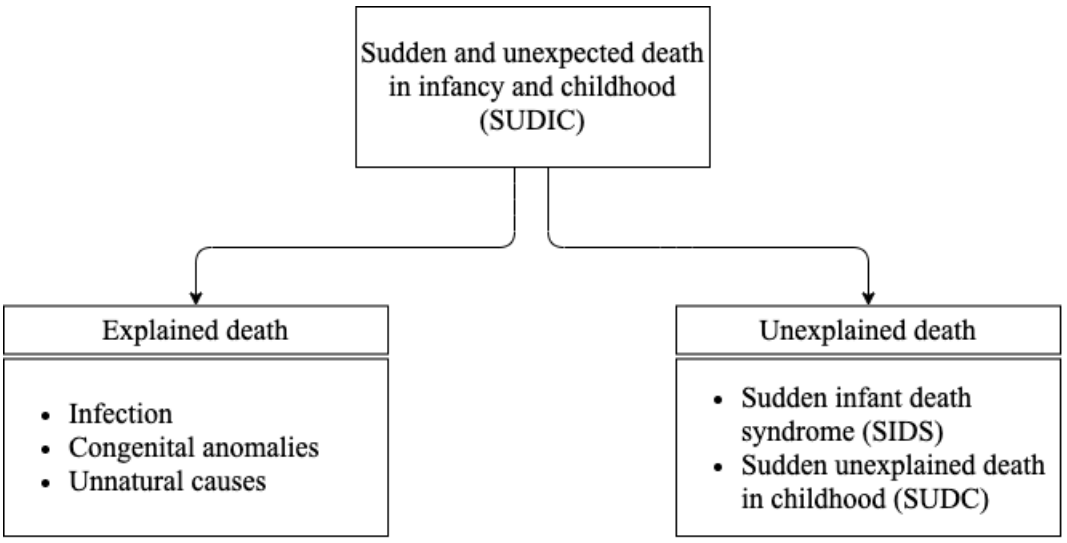

FIGURE 1.1: Flowchart showing death classification in cases of SUDIC. Infection, congenital anomalies, and unnatural deaths are a few examples of the cause of explained death.

1.2 Explained infant death

Explained SUDI account for approximately 60% of all sudden and unexpected infant deaths. The commonest CoD are shown in Figure 1.2. These include infection, cardiac anomalies, and metabolic disorders. Explained SUDI are more common in the first two months of life and these infants often display higher rates of neonatal problems and hospital admissions prior to death [2, 3].

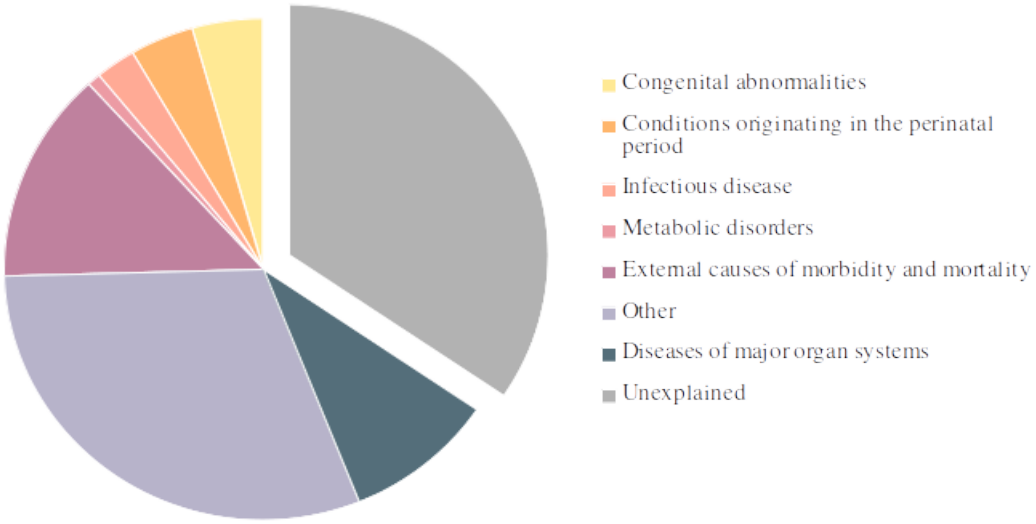

FIGURE 1.2: Final cause of death for SUDI cases occurring in the post-neonatal period (2018). Dataset obtained from the Office of National Statistics [4]

1.3 Unexplained infant death

Unexplainable infant deaths that are believed to have a natural cause are classified as Sudden Infant Death Syndrome (SIDS). This term was first introduced in 1969 to highlight a unique group of SUDI. This subset of deaths shares common clinical, epidemiological, and pathological features but following extensive investigation the CoD remains inconclusive [5]. By definition, SIDS is a diagnosis of exclusion and therefore requires all other possibilities to be explored and eliminated before the diagnosis can be made. This 'syndrome' affects infants worldwide and in 2018 there were approximately 200 cases in England and Wales [4].

Whilst not widely acknowledged, SUDI has a childhood equivalent, and together these cases are collectively referred to as SUDIC. Although a much rarer occurrence, the incidence of such deaths is a concern that must be addressed. In 2017, there were 42 cases of SUDC in England and Wales [6]. There is currently a lack of research and literature in this field and further investigations are required to deepen the understanding of this condition to provide explanation and closure to bereaved families. Therefore, although this PhD thesis focuses mainly on SUDI, as part of the investigation cases of sudden death in childhood will be included in the study cohort.

1.3.1 Incidence, distribution and risk factors

The rates of SIDS vary across the world and range from 0.09 to 0.3 per 1,000 live births [7, 8]. Countries with the lowest rates of SUDI include Austria, Denmark, the Netherlands, Sweden, and Norway, with higher rates observed in Western Europe, Australia, New Zealand, and the United States of America [9]. A study investigating international trends in SUDI from 1969 to 2012 found that countries with low overall infant mortality rates had higher rates of SUDI and SIDS [9]. Historically, most countries displayed a higher rate of incidence prior to the initiation of sleep awareness campaigns launched in the 1990s which promoted a supine sleeping position [9].

Northern European countries such as Denmark, Norway, Sweden, and Finland have consistently shown the lowest rates of SIDS worldwide [7]. These extremely low rates, and their ability to sustain them, has prompted much interest into their pre- and post-natal practices. In recent years there has been particular interest in the use of baby boxes to sleep newborns with many media outlets linking the practice to SIDS reduction [10, 11]. However, professionals warned against their use due to lack of scientific evidence to suggest these cardboard baby boxes are a safer alternative to traditional cots or Moses baskets as they are not required to meet safety regulations [12].

It is important that the comparison of SUDI and SIDS rates worldwide is performed with caution due to the different coding practices used in different regions. Even with the implementation of the International Classification of Diseases 10 (ICD-10) coding system there is still great variability in their usage. For example, a study carried out by Taylor *et al.* investigated death classification of unexplained deaths in eight countries [13]. The study found that the classification under which unexplained infant deaths are registered differs greatly between countries. In Japan of all unexplained infant deaths 44.8% were classified as '*sudden death, unknown*' and 32.6% classified SIDS. However in Australia the '*sudden death, unknown*' classification was not used in any case and SIDS was used in

61.6% of cases [13]. Using the SIDS classification to calculate SIDS rates would therefore give some countries an underestimated rate making worldwide comparisons difficult.

Within the UK, the rate of SIDS has drastically decreased over the last two decades (shown in Figure 1.3), largely owing to reduction campaigns aimed to increase awareness and limit infant exposure to risk factors, which can be found in Table 1.1. The 'Back-to-Sleep' campaign, launched in 1991, successfully reduced the rates of SIDS from a rate of 2.26 per 1,000 live births in 1986 to 0.66 per 1,000 live births in 1993 [14]. Over the following decade, as more risk factors were identified and focus was shifted to parental and caregiver education resulting in a consistent decline through the early noughties. However, in recent years the decline has stagnated and from 2014 to 2018, SIDS rates have remained consistently between 0.28 to 0.31 per 1,000 live births [6].

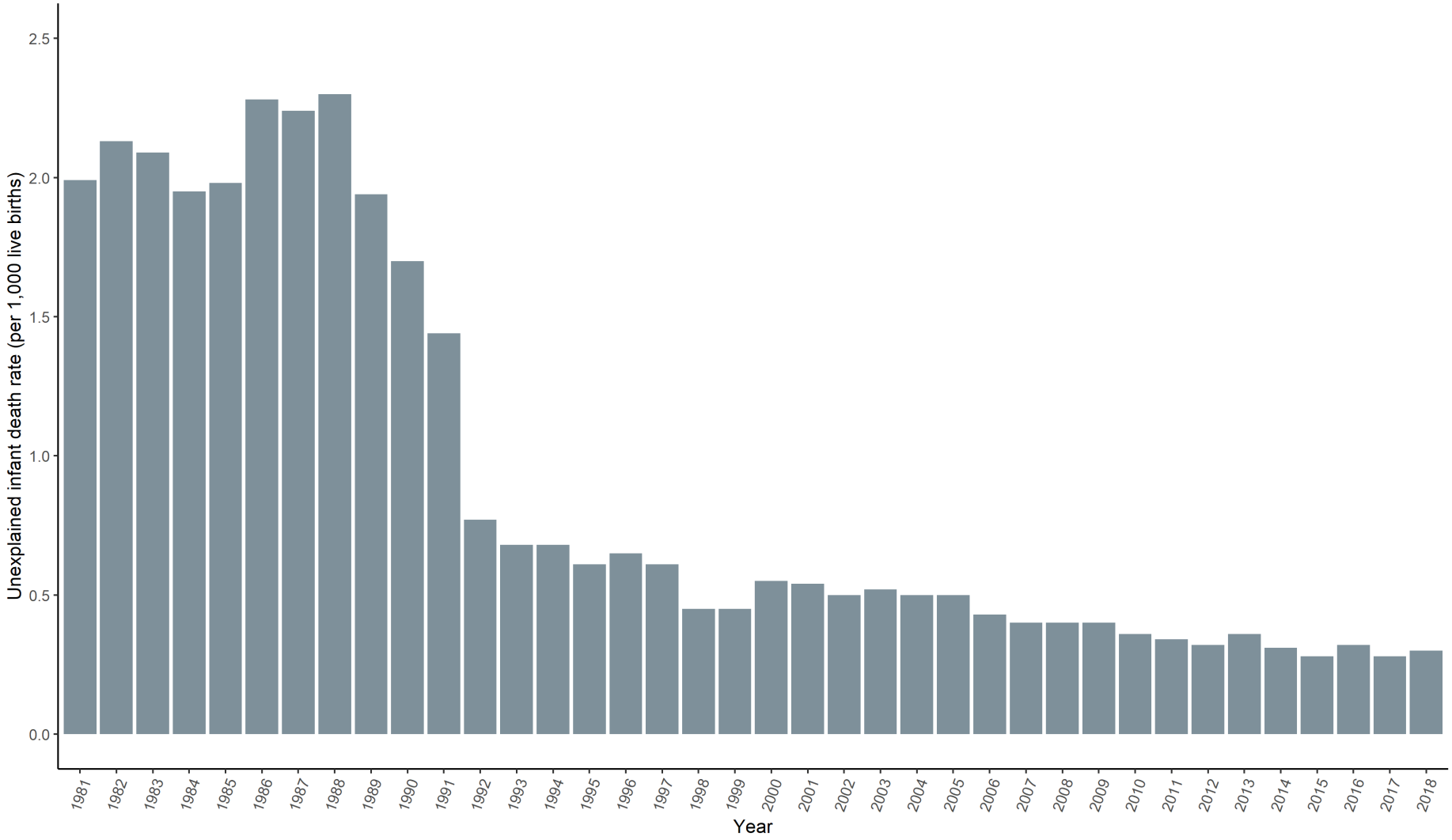

FIGURE 1.3: Declining unexplained infant death rates in England and Wales (1981-2018). Data obtained from ONS [15, 16]

Table 1.1: Summary of risk and protective factors for SIDS

Factor	Increased risk	Decreased risk	Reference
Intrinsic			
	Male sex		
	Preterm birth		
	Low birth weight		
	Young maternal age		[6]
	Single parent family		
	Low socioeconomic class		[17]
	Low education levels		[17]
Extrinsic			
Sleep position	Prone	Supine	
	Side sleeping		
Sleep environment	Bed sharing	Room sharing	
	Use of pillows, blankets, toys and cot bumpers	Waterproof mattress covers	
	Second-hand mattress use	Dummy use	
Temperature	Overheating	Room temperature between 16°C and 20°C	[18–22]
	Swaddling and overwrapping	Well-fitting baby sleep bag	
Drug/alcohol exposure	Pre-natal parental smoking and		[23, 24]
	Second-hand smoke exposure		[17, 25–28]
	Drug or alcohol use by caregiver		
Feed		Breastfeeding	[29–31]
Immunisations		All immunisations up-to-date	

1.4 Investigating SIDS aetiology

Unravelling the risk factors and implementing prevention campaigns to increase awareness and reduce the exposure of infants has proven successful in lowering the rates of SIDS. Although many risk factors have been identified, these merely increase susceptibility to, but do not cause, SIDS. With the causes still poorly understood it is becoming increasingly difficult to reduce the incidence further.

Over the last decade, with SIDS incidence stabilising, there has been considerable interest in uncovering the potential causes. Proposed theories often examine a small subset of SIDS cases within an institution or region and attempt to identify similarities between them, based on the underlying mechanism being investigated. Theories are often supported by theoretical literature with most work focused on explaining SIDS as a disease of immature cardiorespiratory control [32], failure of arousal from sleep [33], obstructive sleep apnoea [34], 'rebreathing' of carbon dioxide resulting in hypoxia [35], and infection [36]. Unfortunately, research tends to group SIDS cases together and treat them as a result of a disease with a single underlying mechanism, rather than an umbrella term which encompasses a range of underlying mechanisms. It is highly unlikely that these unexplained deaths that share apparent clinical features and PM findings can be explained by a single unifying cause, and perhaps it is only these features that they have in common.

The triple risk hypothesis is a widely accepted multi-factorial theory that describes SIDS as a complex multi-factorial disease encompassing a range of different conditions (Figure 1.4). The theory proposes that to succumb to SIDS an infant must be: (i) vulnerable, (ii) at a critical stage in development, and (iii) exposed to an exogenous stressor [37]. The model states that there is no single characteristic or event that results in a sudden infant death, but rather an interaction between several risk factors. In the case where all three criteria are met in unison; infants are unable to respond appropriately to the supposed stressor resulting in a sudden death. It is important to note that if any of these criteria are met in isolation, they may not be sufficient to cause death.

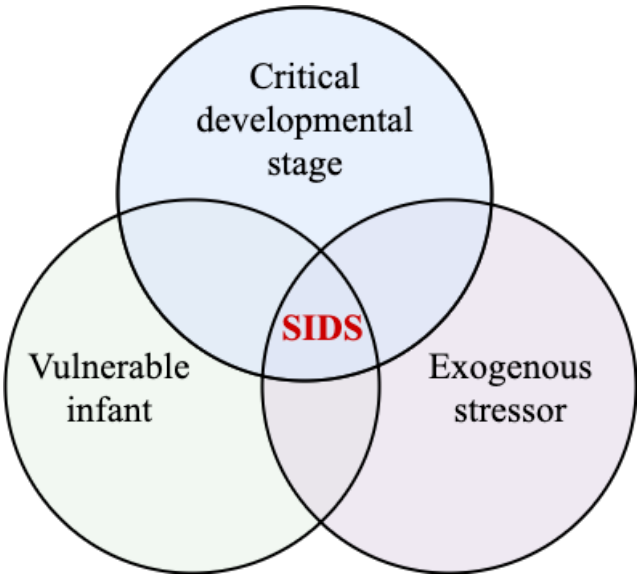

FIGURE 1.4: The Triple Risk hypothesis schematic proposed by Filiano *et al.* [37]. If all three criteria are met in unison an infant can succumb to SIDS.

Although this model attempts to enhance the understanding of the fatal interactions that precede SIDS, it does not provide insight into underlying mechanisms. It does, however, provide fundamental reasoning and a foundation upon which other theories can begin to develop. Given the common characteristics of these deaths such as the age distribution, the infants immature immune system, and increased likelihood of belonging to a family of low socio-economic status, it has been theorised that these deaths could be caused by an undiagnosed infection.

1.5 Infection as a cause of SIDS

The last few decades have seen much discussion around the role of undetected infection as a cause of death for a subset of SIDS. Many researchers have cross-examined risk factors, epidemiological features, clinical presentation, and issues with current diagnostic tools, all of which fit an infectious model for SIDS.

1.5.1 Clinical presentation and epidemiology

The shared circumstances of death in a large proportion of SIDS cases are an important feature to consider. Many SIDS cases occur during a period of sleep, where the infant is often found lifeless by the caregiver in the same position they had been placed. Many SIDS infants are found with a high core temperature and parents have described finding their child "drenched in sweat and buried under blankets" and with "hair matted with dried sweat" [19]. In the 24 hours preceding death, a large proportion of SIDS infants are reported to be sweatier than usual and not feeding as well [3, 19]. A higher proportion of SIDS infants are diagnosed by health professionals as having a minor illness in the week prior to death compared to healthy controls [3]. Changes in infant behaviour reported by parents such as irritability, loss of appetite and drowsiness prior to death are all suggestive of an infective aetiology of SIDS.

SIDS is more prevalent amongst infants between one and four months of age which spans a vulnerable period of development in terms of the infantile immune system. Serum levels of maternally-transferred immunoglobulin G (IgG) that occurs *in utero* begin to wane after birth. The maternally-transferred antibodies reach a nadir between 3 to 6 months of age, at which point the immature infantile immune system cannot produce their own antibodies. The leads to a period of low immunoglobulin levels placing the infant at an increased risk of infection (shown in Figure 1.5). This immune immaturity is more pronounce in premature infants

putting them at a heightened risk of infection [38]. Premature infants are also at an increased risk of SIDS [39].

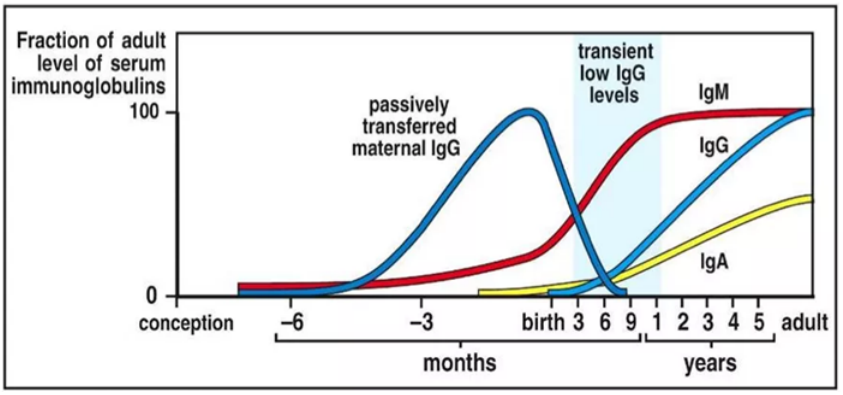

FIGURE 1.5: Infant immunoglobulin levels from conception through the first few years of life. The infant is protected by passively transferred maternal IgG for the first few months of life, after which these levels wane and reach a nadir at 3-6 months. Source: *Immunobiology: The Immune System in Health and Disease, 6th edition* [40]

Breastfeeding provides infants with maternally-derived antibodies and other bioactive compounds which reduce the incidence of infection [41]. The delivery of these protective cellular and chemical components to infants through breast milk has the potential to bridge the gap of waning immunity that occurs between 3 and 6 months as described above. Long-term exclusive breastfeeding has shown to have greater protective effects against infections, and this has also shown to provide and increased protection against SIDS. These similarities in protection against infection and SIDS point to a possible similar role.

The prone sleeping position has been shown to alter the nasopharyngeal flora, increase bacterial colonisation of the upper respiratory tract (URT), and increase the temperature of the airways [42–46]. Colonisation of the URT with toxin-producing bacterial species, together with the increased temperature provides optimal conditions for the production of toxins and this has been hypothesised to result in SUDI [46, 47]. Sleeping in the prone position could result in the ingestion and/or

inhalation of bacteria present on the sleep surface [48, 49]. This could explain the increased risk of SIDS when sleeping on contaminated sleep surfaces such as the parental bed, sofa, and second-hand mattresses [50]. Nasopharyngeal swabs collected from infants after a period of sleep have shown increased rate of detection of *Haemophilus influenzae* and *Neisseria* spp. in prone sleepers with a URT infection than in infants sleeping in the supine position [43]. In a study investigating bacterial presence on mattresses, alternative mattress coverings such as cotton or polyester have shown a significant increase in growth and survival of bacteria such as *Staphylococcus aureus* when compared to waterproof coverings such as polyvinyl chloride (PVC) [51, 52].

The risk factors of SIDS are also similar to those of proven respiratory tract infections in children [53]. This similarity between infection and SIDS is not only observed in risk factors and death circumstances, but also in findings at autopsy.

1.5.2 Autopsy findings

Externally, SIDS infants appear well-nourished and normally developed. In cases where resuscitation or other medical procedures are performed external findings correspond accordingly.

Internally, petechial haemorrhages are often observed within the heart, lung, and thymic tissue and have been found in a greater proportion of SIDS infants than infants dying from other causes [54, 55]. A 2008 study examined autopsy reports from 174 SIDS infants and 67 non-SIDS deaths to examine the presence or absence of intrathoracic petechiae [56]. Petechiae were present in the heart, lung, and thymic tissue in 62.7% of SIDS infants compared to 26.8% of non-SIDS deaths [56]. Petechiae can be indicative of infection. However, petechiae are a common finding in many courses of infant death as they may occur due to a pre-terminal event and therefore are non-specific. In SIDS deaths, the lungs are typically heavy and congested with an abnormal accumulation of fluid which is occasionally also found in the mouth and nares [21]. The respiratory tract is often inflamed, with

purulent mucus found in the upper airways [21]. The lungs commonly show signs of mild inflammation with the detection of inflammatory cells [21, 57]. However, although these findings are suggestive of a mild infection, these findings alone are not sufficient to certify fatal infection as the CoD.

1.5.3 Bacteria and bacterial toxins

Together with autopsy findings suggestive of infection, bacteria have also been consistently identified in PM samples from SIDS infants. A recent study investigated the presence of bacteria in heart-blood, spleen, and CSF samples collected at autopsy as reported in PM reports [58]. *Staphylococcus aureus* was cultured in 10.76% of PM tissues collected from SIDS infants compared to 18.75% of those with an infectious origin and 0% of those whose death was accidental [58]. In a large retrospective analysis of bacteriology results obtained from 470 SUDI cases, significantly more pathogens known to cause septicaemia in young children were isolated in SIDS infants than in those with a non-infectious cause of death [59]. The greatest proportion of *Staphylococcus aureus* and *Escherichia coli* isolates were yielded from unexplained deaths compared to those with an infectious cause and accidental deaths [59]. A study carried out by Gilbert *et al.* found it was more likely to find coliforms in the respiratory tract of SIDS infants compared to healthy live infants [60]. However, in a more recent study which compared PM tissues from SIDS infants to those from SUDI infants with an identified cause of death, no significant difference in the identification of coliforms was found [58].

Throughout the 1990s there was heightened interest in the role of bacterial toxins in SIDS. Morris *et al.* proposed the common bacterial hypothesis in 1987 suggesting that the bacterial toxins produced by commensal nasopharyngeal flora can have lethal effects in vulnerable infants following a transient viral infection [47]. They argued that if these bacterial toxins were absorbed by infants in the period between waning of maternal antibodies and the production of infant antibodies, the infant would experience a fatal outcome [47, 61]. It has been shown that there is increased

nasopharyngeal carriage of *Staphylococcus* spp. and *Streptococcus* spp. in SIDS infants compared with healthy age-matched controls [61, 62], and this carriage was shown to be inversely associated with age [43].

The prevalence of staphylococcal enterotoxins toxic shock syndrome toxin 1 (TSST-1), Staphylococcal enterotoxin A (SEA), Staphylococcal enterotoxin B (SEB) and Staphylococcal enterotoxin C (SEC) has been investigated in serum, lung, and kidney samples collected from SIDS deaths. At least one staphylococcal toxin was identified in 53% of SIDS cases investigated in a case series of Scottish infants [63]. In another cohort of 30 Australian SIDS infants and 19 infants who died from other causes, TSST-1 was identified in 53% SIDS compared to 16% non-SIDS deaths using flow cytometry assays. The deaths occurring in the non-SIDS group were due to pneumonia and complications of cystic fibrosis [63]. Similar ratios were observed in other studies carried out using immunohistochemical techniques where TSST-1 was identified in 18% of SIDS infants compared to 6% non-SIDS deaths and SEC was identified in 36% SIDS infants compared to 12% non-SIDS deaths [61, 64]. A more recent study by Highet *et al.* examined the presence of *Staphylococcus aureus* and associated enterotoxins in faecal samples from SIDS infants compared to age- and gender-matched live controls using PCR methods. They found a higher proportion of *Staphylococcus aureus* in SIDS infants compared to controls (68.4% vs. 40.5%) as well as more staphylococcal enterotoxin genes (43.8% vs. 21.5%) [65]. However, it is important to note that the comparison group were live infants. Results from these studies show that although colonisation of infants with *Staphylococcus* spp. is common, it appears to be present in a higher proportion of SIDS infants.

1.5.4 Viruses

The ability to investigate the presence of viruses in PM tissues has improved drastically in recent years due to the emergence of molecular diagnostic techniques such as PCR. In the past there has been a lot of uncertainty surrounding the

role of viruses due to difficulties in culturing virus from PM tissues given body decomposition and low viral loads [62].

Many SIDS are believed to have a mild viral infection prior to death. However, given the high frequency of viral infections in young children the detection of virus does not mean that it was the CoD [66, 67]. The presence of viral respiratory infection in children is more common in winter months and has a similar seasonal distribution as has been observed in SIDS [68]. However, some studies have reported that in recent years the seasonal distribution of SIDS has dissipated [69]. The increased risk of SIDS in household with smokers and an increased number of siblings also fits well with viral respiratory infection as a potential cause [70].

Infection with cytomegalovirus (CMV) may result in asymptomatic infection with higher mortality in immunocompromised individuals including premature and neonatal infants [71]. Enteroviruses and parechoviruses are a frequent cause of infection in children although they typically result in mild illness that passes within a few days. However, in vulnerable individuals, particularly infants less than three months old, they are known to cause more severe disease [72]. A recent case report has shown enterovirus and parechovirus identified in the blood of a five-month old infant who died suddenly and exhibited the common features of SIDS [73].

Another potential involvement of viruses in SUDIC is their role in facilitating superimposed secondary bacterial infection. These secondary bacterial infections occur when a bacterial pathogen invades when the host immune system is unable to respond appropriately due to the primary infection. Therefore, although the viral infection may not be solely responsible for the death, it may weaken the immune system rendering the infant unable to respond to the exposure. Primary viral infection has been linked to severe bacterial infections including bacteraemia, pneumonia, and urinary tract infection. Studies have shown a correlation between invasive pneumococcal disease and meningococcal bacteraemia with respiratory syncytial virus (RSV), influenza virus, and human metapneumovirus infection in children [74, 75]. Viral infection is known to damage ciliated cells which may reduce the body's ability to clear bacteria from the upper airways [**?**]. Viruses such as

RSV have also been shown to enhance the adherence of known pathogenic bacteria to airway epithelial cells [76, 77]. *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Neisseria meningitidis*, *Haemophilus influenzae*, and *Klebsiella pneumoniae* are often isolated from secondary bacterial infections [78], all of which have been linked to SIDS.

Following current recommendations, viral investigation is not a mandatory component of routine investigation into SUDIC. However, given the high frequency of viral infection in infants and young children it is important that their role in SUDIC is not overlooked.

1.6 Interpreting post-mortem results and diagnosing infection in SUDIC

1.6.1 Current procedures

Diagnosing infection in cases of SUDIC is particularly difficult due to their unexpected nature [79]. The diagnosis requires definitive evidence of an infection, ideally with both histological and microbiological evidence together with supporting clinical history.

Many histological changes are common observations in SUDI, but without supporting evidence from other PM tests these results are difficult to interpret. For example, the presence of petechial haemorrhages is common in SIDS cases but their presence alone is not indicative of infection and often interpreted as a product of the PM process [54, 56]. Similarly, if microbes are identified in PM tissues, without clear histological evidence of infection, the interpretation of results is difficult. It is generally accepted that a pure culture must be obtained from a tissue site for it to have caused an infection [79]. Currently, the diagnosis of infection requires agreement of microbiological and histological evidence, both of which present with major difficulties in interpretation.

In most of the microbiology diagnostic laboratories worldwide, the gold standard procedure for identification of infectious microbes is live bacterial culture. This method involves culturing patient samples on a range of selective media and incubating these under specific conditions. The selective media and culture conditions allow the growth of target bacterial or fungal pathogens. This method is highly selective and can only provide qualitative results [80]. It is also associated with limitations such as an increased likelihood of identifying readily cultivable bacteria [81], the inability to culture all bacterial species under laboratory conditions [82, 83], long incubation periods [84], skilled persons or high-cost equipment for bacterial identification. This method also requires live bacteria, meaning that if a

patient has been given empiric antimicrobial therapy, the likelihood of recovery of live bacteria is greatly reduced.

The positive culture of an organism requires interpretation by the investigating team and as there are currently no guidelines on the interpretation of PM microbiology results, the interpretation is often varied [85, 86]. One of the difficulties in determining the relevance of bacteria isolated from samples collected at PM is the number of circumstances from which a positive microbial culture can arise [87]. The acquisition of a positive PM bacterial culture could represent the following: (i) a true ante-mortem infection by the identified pathogen, (ii) contamination of the sample during sample handling, or (iii) bacterial translocation that has occurred during the time between death and PM examination (the post-mortem interval (PMI)). These possible sources of false-positive results that may occur in a PM setting makes it difficult to interpret results, particularly in cases where a positive culture is obtained without supporting clinical or histological evidence. Although contamination during sample handling can be limited by ensuring sterile technique, PM bacterial translocation is a natural process that occurs due to body decomposition and is currently not well understood. This phenomena will be discussed in greater depth in Chapter 3.

Due to the inability of current microbiological techniques to distinguish between commensal contamination and true infection, as well as the issue with culturing bacteria under laboratory conditions, the use of alternative techniques should be explored. Although current techniques may be suitable for the diagnosis of infection in live patients who usually present with symptoms suggestive of infection, in retrospective diagnosis of SUDIC where there is no obvious indication of the CoD there is a demand for more detailed investigation. In the investigation of SUDIC the time constraint present in the rapid diagnosis of a live patient who requires imminent treatment is removed meaning alternative, more detailed methods could serve useful given their increased sensitivity and the ability to identify a larger proportion of bacteria. Molecular techniques such as PCR target specific bacteria which is sufficient in cases where a bacterial pathogen is suspected or symptoms

are present. However, this method cannot provide an unbiased assessment of bacteria present within tissues. This leads to the purpose of this study, investigating the use of 16S rRNA gene sequencing for identification of infection in SUDIC.

1.6.2 16S rRNA gene sequencing

The 16S rRNA is a key component of prokaryotic ribosomes which are the fundamental machinery for protein synthesis. The gene was first implemented for the taxonomic classification of prokaryotes in 1977 due to its favourable characteristics including its slow rates of evolution, its universal presence in bacterial and archaeal species and the presence of hyper-variable regions that act as fingerprints and enable the identification of bacteria present. Through PCR methods, the gene has been characterised in many different bacterial species and the nucleotide sequenced decoded, resulting in the generation of detailed reference databases including Greengenes [88], SILVA [89], and the Ribosomal Database Project (RDP) [90].

The 16S rRNA gene spans approximately 1500 bps and is composed of nine hypervariable regions (V1 to V9), separated by highly conserved regions as depicted in Figure 1.6. These conserved regions provide a suitable anchoring site for primer binding to allow amplification of a large range of bacterial species. The unique sequences found in the hypervariable regions allows for the identification of the organisms through comparison to reference databases.

FIGURE 1.6: Diagram of the 16S rRNA gene. This gene spans approximately 1500 bps and is composed of nine hyper-variable regions (shown in orange) which are separated by highly conserved regions (shown in grey).

The 16S rRNA gene has been targeted using PCR methods such as quantitative-PCR (qPCR) and has been used in clinical studies where it has shown to identify bacterial pathogens in a range of clinical samples where bacterial culture was negative [91–95]. With the advancement of molecular techniques and emergence of next-generation sequencing (NGS) technologies, it is now possible to obtain millions of sequences from multiple samples simultaneously within a relatively short time frame. The ability of machines such as the Illumina MiSeq to generate upwards of 10 million reads per sequencing run has reduced the cost of generating sequencing data which currently stands at approximately 25 pence per megabase [96]. The technique allows extensive and in-depth characterization of microbial communities, not only providing a list of microbial members but also quantitative analysis enabling calculations of relative abundances within a population. It has also been shown to possess greater resolution and accuracy in identifying microbial pathogens. In advanced centres such as the diagnostic laboratories in GOSH, 16S Sanger sequencing can be requested by the pathologist in special cases although it is not part of routine protocol. However, this method only allows sequencing of a single amplicon at any one time and cannot provide results in polymicrobial samples [97]. When comparing Sanger sequencing to NGS, NGS has shown to provide more accurate, higher resolution results when applied to clinical specimens [98, 99].

When targeting the 16S rRNA gene, the choice of hyper-variable region in characterising microbial communities is an important consideration. As described above, the 16S rRNA gene is composed of nine hyper-variable regions. Current next generation sequencing technologies, such as the Illumina MiSeq, are able to sequence up to 2 x 250 bp paired-end reads. As no single region of the 16S rRNA gene is able to distinguish between all bacterial species, it is important for the region of choice to be a balance between required classification power and obtainable sequence length. A study carried out by Chakravorty *et al.* investigated the ability of each region to distinguish between 110 common human bacterial pathogens to identify which region or combination of regions would be the most useful in clinical investigation [100]. They found the V3 region optimal for distinguishing between

all 110 bacteria to genera level classification. They identified a region between positions 456 to 479 containing the most SNPs to enable high resolution identification. Similar identification power was observed in the V2 region, but the longer fragment length (105 bps compared to 64 bps) made it slightly less favourable [100].

The use of 16S rRNA gene targeted NGS for application in the clinic to aid in the diagnosis of infection and enhance patient management has been investigated, mainly exploring the replacement of current culture techniques. The introduction of this technique into diagnostic procedures removes the need for viable bacteria, meaning the technique could be used in cases where patients have received antimicrobial therapy prior to sample collection [91]. In studies carried out on blood, synovial fluid, CSF, urine, swabs and tissue samples 16S rRNA sequencing detected more organisms than were detected using bacterial culture [91, 101, 102]. NGS of the 16S rRNA gene sequencing has been shown to be of particular use in cases with bacterial co-infection and polymicrobial specimens [91]. When compared to bacterial culture, 16S rRNA gene sequencing has shown higher specificity and sensitivity when detecting clinically relevant bacterial pathogens such as *Streptococcus* spp., *Pseudomonas* spp. and *Haemophilus* spp. [101, 103]. It has also been suggested that this method may enable an enhanced understanding of pathogenic bacteria and help to identify novel pathogens causing severe infection in infants [104].

As 16S sequencing has a greater sensitivity and provides a less bias, more in-depth analysis into bacterial communities compared to traditional culture-only methods, this thesis assessed its use as a potential diagnostic tool of bacterial infection in SUDIC.

1.7 Research aims and objectives

As has been discussed throughout this introduction SIDS is a diagnosis of exclusion and its cause is unknown. Risk factors, epidemiological features, clinical presentation together with limitations in current PM diagnostic tools fit an infectious model of SIDS. Therefore the aim of this study was

To investigate the effectiveness of 16S rRNA gene sequencing in detecting potential infection-causing bacteria in cases of SUDIC for the diagnosis of infection

In order to achieve this aim, two other aspects had to first be investigated.

1. As a large proportion of archived material worldwide is stored as formalin-fixed paraffin-embedded (FFPE) material, it was important to optimise the 16S rRNA sequencing technique following this type of tissue processing. This will be optimised in Chapter 4.

2. The interpretation of PM microbiological results is difficult. To assist in results interpretation post-mortem bacterial translocation had to be addressed. Post-mortem bacterial translocation is the spread of bacteria that is believed to occur from sites such as the gut to other tissues in the time between death and post-mortem investigation. This spread can cause difficulty when interpreting post-mortem bacterial culture results as it is often difficult to determine whether identification of a bacteria was due to an infection, or whether the bacteria had spread to these sites after death. This phenomenon was assessed using 16S sequencing in two animal models to characterise the post-mortem process. With this knowledge, implementation of the 16S rRNA gene sequencing technique into routine clinical diagnostics would allow a more reliable retrospective diagnosis of ante-mortem infection. This will be investigated in Chapter 3.

These aims of this thesis are summarised in Figure 1.7.

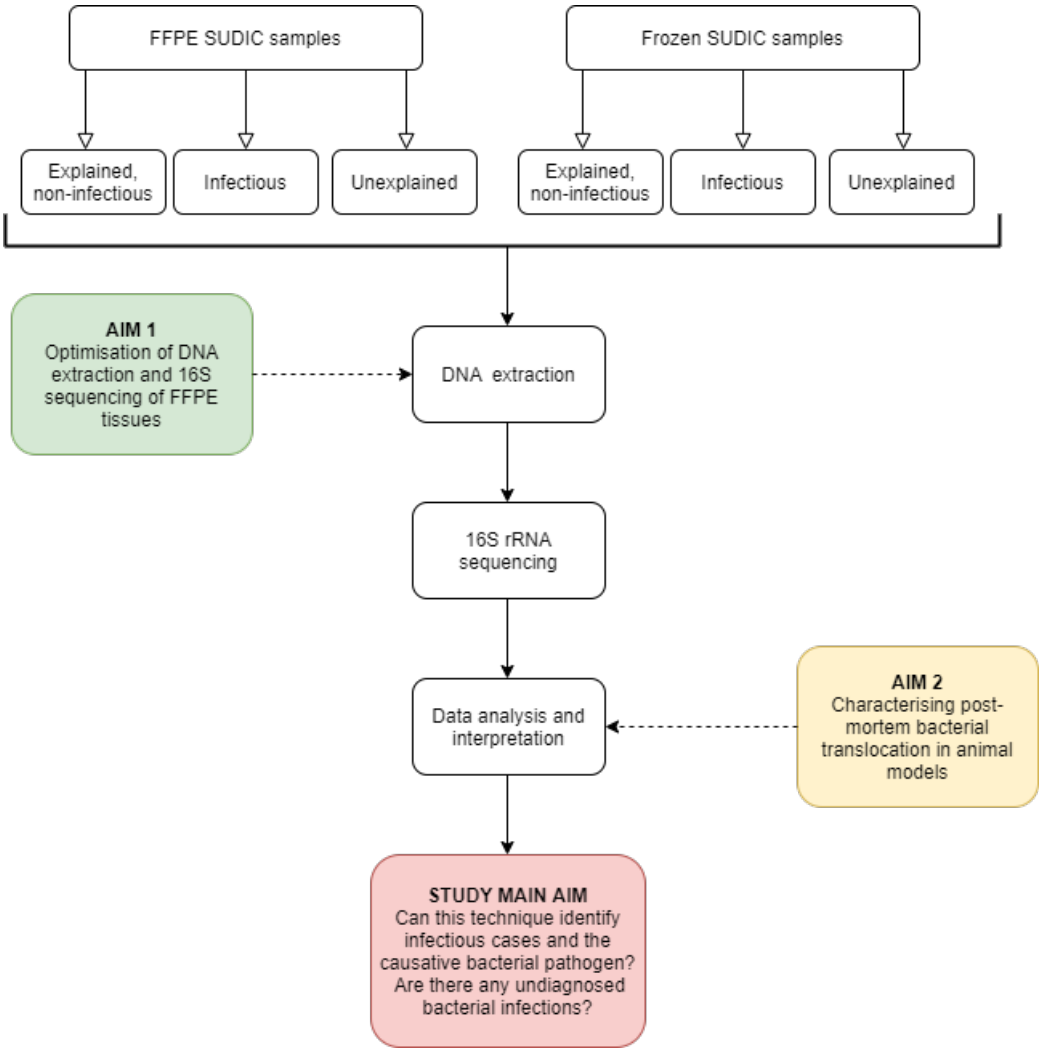

FIGURE 1.7: Thesis aims

Methodology

2.1 Sample collection

2.1.1 Harvesting mouse tissue samples

Surplus nine-week-old C57BL/6 female mice had been housed for approximately one week to allow for acclimatisation and standardisation of the microbiota for another study. Mice were culled on day 0 according to Schedule 1 of the Animals Scientific Procedures Act 1986 by dislocation of the neck and donated for use in this study. Culled mice were placed into airtight freezer bags and kept on ice until the first sample was collected.

A total of 15 mice were kept in three cages (five mice per cage). Tissue samples were collected from one mouse per cage at five time points over a two-week period (three mice sampled per time point) and mice were stored in an airtight freezer bag at 4°C until sampling. Samples were collected on day 0 (D0), day 3 (D3), day 7 (D7), day 10 (D10) and day 14 (D14) PM. The first sample was collected within an hour of death and mice were sampled in quick succession. Prior to dissection, mice were pinned to a dissection board and doused with 70% ethanol. An incision was made down the ventral midline using sterile dissection scissors and the skin was pulled back using sterile forceps. The abdominal membrane was cut from the ribcage upwards keeping the lower region intact whilst removing organs from the thoracic cavity to limit cross contamination. A 25-gauge needle was used to obtain a blood sample via cardiac puncture. On D3 it was apparent that the blood had accumulated in the thoracic cavity with little to no sample available from the heart, so a 'leaked' blood sample was taken. Samples were obtained in the following order: blood, heart, lungs, liver, spleen and the lower GI tract. A new

set of sterile scissors and forceps were used to extract each organ. All samples were placed into sterile 1.5mL Eppendorf tubes and immediately snap frozen followed by storage at -80°C until DNA extraction.

2.1.2 Obtaining piglet tissue samples

Three 1-month-old Welsh piglets were obtained from the Royal Veterinary College (RVC) (n=1) and a local farm (n=2). The piglets were received and samples were obtained no more than 3 hours PM. Tissue samples were collected from the heart, lung and liver using a new sterile 16-gauge Quick-Core Biopsy Needle (Cook Medical) for each sample. Prior to sampling, piglets were cleaned and the area of needle insertion was doused with 70% ethanol. Nasal and rectal swabs were also collected from each piglet at every time point to examine bacterial communities from which post-mortem bacterial translocation could potentially originate. Swabs were collected using polyurethane cellular foam dry swabs (MWE Medical Wire). Tissue samples and swabs were placed into sterile 1.5mL Eppendorf tubes, immediately snap frozen, and stored at -80°C until DNA extraction. Sampling was performed on D0, D3, D7, D10 and D14. Each piglet was stored in a zipped body bag at 4°C between sampling points.

2.1.2.1 Piglet tissue identification using RT-PCR

Tissue type confirmation was performed using tissue-specific reverse transcriptase quantitative polymerase chain reaction (RT-qPCR). Nucleic acids were extracted from tissue samples as described in Section 2.2. Differentially expressed messenger ribonucleic acid (mRNA) was identified using the University of California Santa Cruz (UCSC) Genome Browser followed by primer design using the National Center for Biotechnology Information (NCBI) Primer-BLAST tool. Primers were designed to span exon-exon junctions as to not detect genomic DNA. Target genes along with primer sequences are shown in table 2.1.

Frozen pig heart, lung, liver, larynx, and trachea were obtained from Samples for Schools (www.samples-for-schools.co.uk). These samples had been taken from pigs under 12 months of age within 24 hours of slaughter and immediately frozen. Approximately a $1mm^2$ section was cut from each organ and DNA extraction was performed as described in Section 2.2.

The reverse transcriptase reaction was performed on DNA extracts using the ProtoScript II First Strand cDNA synthesis kit (New England Biolabs). 2μL of total nucleic acid was added to 2μL of oligo $d(T)_{23}VN$ oligonucleotide and made up to a total reaction volume of 8μL using 6μL of nuclease-free H_2O (nf-H_2O). The reaction was completed following the manufacturers Standard Protocol. No-RT control reactions consisting of all components of the reaction without the reverse transcriptase enzyme were used to monitor potential products arising from genomic DNA.

The qPCR reaction was performed as follows: 10μL Bioline SensiFAST SYBR No-ROX master mix (Bioline), 0.5μM forward primer, 0.5μM reverse primer, 2μL cDNA template and 6μL nf-H_2O to make a final volume of 20μL. The reactions were cycled on a BioRad CFX96 real time system with the following cycling parameters: an initial 3 minutes at 95°C followed by 45 cycles of 95°C for 5 seconds and 65°C for 30 seconds. This was followed by the addition of a melt curve to monitor non-specific amplicons. Each reaction contained a positive control and a no template control (NTC).

Table 2.1: Primer sequences for piglet tissue identification

Primer name	Target gene	Primer sequence (5' to 3')
Liver FP	APOC4	TGGTTCCGGGCAAGATGAAG
Liver RP		GCTATGGGCCTTGTTCAGGA
Heart FP	ANKRD1	TGCTCGGGATAAGTTGCTCA
Heart RP		GGCGCCATACGTAATCAGGA
Lung FP	SFTPC	GGTCTATGACTACCAGCGGC
Lung RP		CAGCTTAGAGGTAGGCGTCG
Control FP	GAPDH	ACACTCACTCTTCTACCTTTG
Control RP		CAAATTCATTGTCGTACCAG

2.1.3 FFPE and frozen Mouse placenta samples

Mouse placenta samples were kindly donated by Dr. Natalie Suff (UCLH). Briefly, 20μL of *Escherichia coli* or phosphate-buffered saline (PBS) was delivered into the vagina of mice anaesthetised with isoflane and placenta samples were collected within 24 hours of administration. Two placenta samples were fixed in 10% buffered formalin and embedded into a block of paraffin. The remaining samples were frozen as controls.

2.1.4 FFPE Human placenta samples

Local Research Ethics Committee approval for both assay development on surplus pathological specimens and collection and sampling of surplus placental material was in place for this study. FFPE placenta samples with chorioamnionitis diagnosis and negative controls were obtained were anonymised and chorioamnionitis status was obtained.

2.1.5 FFPE and Frozen SUDIC samples

SUDIC cases were selected from the archive of PM tissue at Great Ormond Street Hospital (GOSH). In all cases, consent for scientific research had been given by the parents for retention and use of tissue for research approved by the Research Ethics Committee. All cases selected for this study underwent a comprehensive autopsy between 2010 and 2019 and were investigated on behalf of Her Majesty's Coroner (HMC). Cases were selected based on their final diagnosis and grouped into one of three groups:

1. Explained, non-infectious death
2. Infectious death
3. Unexplained death

Both frozen and FFPE materials were obtained. Frozen material collected for this study included heart, kidney, liver, spleen, and muscle. FFPE material was collected for select cases and included heart, kidney, liver, spleen, muscle, and lung samples. Each collected sample was subject to DNA extraction and 16S sequencing analysis as described below.

2.2 DNA extraction

2.2.1 Fresh material

DNA was extracted from fresh tissue samples measuring approximately 1mm^2. Upon collection the tissue was digested at 56°C overnight in a mixture of 180μL ATL lysis buffer (Qiagen) and 20μL of Proteinase K (Qiagen). The DNA was then extracted using the QIAamp DNA Mini Kit (Qiagen) as per the manufacturers protocol. An initial lysing step was added following tissue digestion. This step consisted of adding 1/6 Ribolysing Matrix B (MP Biomedicals, Cambridge, UK) to the digested material and bead beating for one minute at 50 oscillations per second. A negative control undergoing the same process as tissue samples, without the tissue was included for each DNA extraction.

2.2.2 Frozen material

A 1mm^2 section of each frozen tissue was cut using a sterile scalpel. This was done in a sterile Petri dish kept on dry ice. The tissue was then digested overnight and DNA extraction was carried out as described in Section 2.2.1.

2.2.3 Blood samples

DNA was extracted from 200μL of each blood sample. In cases where 200μL was not available, the total volume was made up to 200μL using ultraviolet (UV)-treated AE buffer (Qiagen). The DNA extraction was carried out using the QIAamp DNA Mini Kit (Qiagen) as per the manufacturers' protocol with the addition of the initial lysing step detailed in Section 2.2.1.

2.2.4 Formalin-fixed paraffin-embedded material

Once the FFPE blocks had been selected, scrolls were cut in 10μm sections using a microtome. The first few scrolls were discarded of to further minimise surface contamination. Before each block was cut, the microtome was decontaminated using absolute ethanol and DNA-away Surface Decontaminant (Thermo Scientific) to prevent cross contamination and a new sterile blade and forceps were used for each block. The scrolls were placed into sterile 1.5mL Eppendorf tubes (Eppendorf). For each set of blocks cut, a negative wax control was obtained to account for any bacterial contamination that may arise from the paraffin. This was obtained by cutting a 10μm scroll from a part of the block without any sample.

Deparaffinisation of the FFPE tissues was carried out in one of the following ways:

2.2.4.1 Deparaffinisation using xylene

Deparaffinisation using xylene was performed by adding 1mL of xylene to 10μm rolls of each sample in a 1.5mL Eppendorf tube, vortexing for 30 seconds and then centrifuging at 12,000pm for one minute. The xylene was then carefully removed taking care not to disturb the pellet. Two absolute ethanol washes were then performed as follows: add 1mL of absolute ethanol, vortex for 30 seconds, and centrifuge at 12,000rpm for five minutes.

2.2.4.2 Deparaffinisation using Mineral Oil

Deparaffinisation using mineral oil was performed following a published protocol [105]. Briefly, 300μL of mineral was added to eight 10μM sections of sample and vortexed vigorously for 15 seconds, followed by incubation at 90°C for 20 minutes.

2.2.4.3 Deparaffinisation using the Qiagen Deparaffinisation Solution

Tissues were deparaffinised using the deparaffinisation solution (Qiagen) as described in the manufacturers' protocol.

Once the tissue had been deparaffinised, the DNA was extracted using the QIAamp DNA FFPE Tissue kit (Qiagen) following the manufacturers' protocol.

2.3 Quantification and Quality Control methods for nucleic acids

2.3.1 Qubit dsDNA High Sensitivity (HS) Assay

The concentration of double-stranded DNA extracted from each sample was estimated using the Qubit dsDNA HS Kit (ThermoFisher Scientific) on the Qubit Fluorometer 2.0 (ThermoFisher Scientific) following the manufacturer's protocol.

2.3.2 TapeStation DNA Assay

The concentration and lengths of the DNA extracts were analysed using either the High Sensitivity D1000 ScreenTape (Agilent) or the Genomic DNA Analysis ScreenTape (Agilent) on the 2200 TapeStation Instrument (Agilent). This was carried out as per manufacturers' protocol and results were analysed using the TapeStation software.

Table 2.2: Tissue-Tek VIP Processing schedule

Stage	Solution	Concentration (%)	Time (minutes)	Temperature
1	Formal Saline	10	30	-
2	Alcoholic Formalin	50	30	-
3	Alcohol	70	60	-
4	Alcohol	95	60	-
5	Alcohol (IMS)	100	60	-
6	Alcohol (IMS)	100	60	-
7	Alcohol (IMS)	100	60	-
8	Xylene	100	60	-
9	Xylene	100	60	-
10	Xylene	100	60	-
11	Wax	-	60	60
12	Wax	-	60	60
13	Wax	-	60	60
14	Wax	-	60	60

2.3.3 Nanodrop Assay

The DNA concentration and 260/280 ratio were analysed using the NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific). This was carried out as described in the manufacturer's protocol.

2.4 Formalin-fixation and paraffin-embedding of fresh material

Lung samples were collected from mice as described in Section 2.1.1. All reagents were loaded onto the machine and tissues were fixed overnight on the Tissue-Tek VIP 5 Processor according to the processing schedule shown in Table 2.2. Following fixation, samples were embedded into a paraffin block and stored at room temperature until further analysis.

2.5 PCR amplification

2.5.1 Bacterial 16S rRNA gene PCR

PCR of the 16S gene was carried out using the Taq PCR core kit (Qiagen) and primer sequences shown in Table 2.3. The reaction mixture was as follows: 5μL 10x PCR buffer, 1μL $MgCl_2$, 10μL Q solution, 1μL dNTPs, 0.5μM forward primer, 0.5μM reverse primer, 0.25μL Taq polymerase, 5μL DNA template and nf-H_2O to make the final reaction volume up to 50μL. The following cycling parameters were used: initial denaturation at 95°C for 3 minutes followed by 30 cycles of denaturation at 95°C for 30 seconds, annealing at 54°C, extension at 72°C for 90 seconds and a final extension at 72°C for 10 minutes.

2.5.2 B-actin PCR

Amplification of the β-actin gene was performed in both human and mouse tissue using primers designed using the NCBI Primer-BLAST tool. The FASTA sequence to which the primers were designed for was NCBI Reference Sequence: NC_000007.14 for the human β-actin gene and NCBI Reference Sequence: NC_000071.7 for the mouse NCBI β-actin gene. Primer sequences are shown in Table 2.5 and Table 2.4.

The PCR reaction was performed using the Taq PCR Core kit with the following reaction mixture: The reaction mixture was as follows: 6μL 10x PCR buffer, 1μL dNTPs, 0.2μM forward primer, 0.2μM reverse primer, 5μL DNA template, and 36.75μL to make the final reaction volume to 50μL. The following cycling parameters were used: initial denaturation at 95°C for 3 minutes followed by 30

Table 2.3: Primer sequences targeting the V3-V4 and V4 regions of the bacterial 16S rRNA gene

Primer name	Primer sequence (5' to 3')	Product (bp)	Reference
V3-V4 FP	TCCTACGGGAGGCAGCAGT	466	[106]
V3-V4 RP	GGACTACCAGGGTATCTAATCCTGTT	466	[106]
V4 FP	GTGCCAGCMGCCGCGGTAA	292	[107]
V4 RP	GGACTACHVGGGTWTCTAAT	292	[107]

cycles of denaturation at 95°C for 30 seconds, annealing at 58°C, extension at 72°C for 90 seconds and a final extension at 72°C for 10 minutes.

Table 2.4: Primer sequences targetting the β-actin gene in mice

Primer name	Primer sequence (5' to 3')	Product (bp)
Mouse FP	CTCTGTGGCTTTGCTGGGTG	-
Mouse RP 1	CCTGTAGCCCTCCCACTAGAT	114
Mouse RP 2	AGTGTTAGTGCAGGCCAACTT	217
Mouse RP 3	CTTCAGACTAGCAGAGGGGCTTG	297
Mouse RP 4	CAAAACCAGGGAAAAGATGCCC	385
Mouse RP 5	ACGACTGACAAAGGCTGGAG	488

Table 2.5: Primer sequences targetting the human β-actin gene

Primer name	Primer sequence (5' to 3')	Product (bp)
Human FP	CCGAAGCGAGACCATAGGAC	-
Human RP 1	CAAAGCAGCAACTGCCCAAT	123
Human RP 2	CCTTCGGCTGGAAAACAAGC	245
Human RP 3	AAGCAATTCTCTGTATGCACATC	319
Human RP 4	TCTGGTGACGAAGAGCAGTT	440

2.6 Agarose gel electrophoresis

A 1.5% (w/v) agarose gel was prepared in a horizontal gel tank containing 10μL of GelRed Nucleic Acid Stain (Biotium) per 100mL gel (1:10,000 dilution). Amplified products were loaded into the gel alongside 10μL of Quick-load Purple 1Kb DNA ladder (New England Biolabs) and the gel was electrophoresed at 110V using UltraPure TBE buffer (Invitrogen) until the dye reached the bottom edge of the gel. The bands were visualised under UV light.

2.7 16S rRNA Gene Sequencing

2.7.1 Sample preparation

DNA templates were amplified for sequencing using the Taq PCR Core Kit (Qiagen) and region-specific primers with adapters and indexes attached (adapted from [107]). For sequencing of the V3-V4 region, a combination of F1 forward primers (Table 2.8) and R1 reverse primers (Table 2.10) were used. For sequencing of the V4 region, a combination of F2 forward primers (Table 2.9) and R1 reverse primers (Table 2.10) were used. Figure 2.1 shows how the sequencing primers bind to the 16S gene in order to amplify the desired region. A ZymoBIOMICS Microbial Community DNA Standard (catalogue #6305, Zymo Research) of known bacterial composition was amplified and sequenced alongside samples in order to assess bias and sequencing error rates. The theoretical composition of the Microbial Community DNA Standard is shown in Table 2.6. Negative controls were also amplified and sequenced and consisted of negative controls from DNA extraction as well as no-template controls which consisted of nf-H_2O.

FIGURE 2.1: Binding of F1, F2 and R sequencing primers to the V3 to V4 region of the 16S rRNA gene. The grey depicts the conserved regions to which the primers bind. The orange depicts the variable regions that are unique to each bacterial species.

Amplification was carried out using the following mixture: 5μL PCR 10x buffer, 1μL dNTPs, 0.5μM forward primer, 0.5μM reverse primer, 0.5mM $MgCl_2$, 10μL Q solution, 0.25μL Taq polymerase, 5μL DNA template and 25.25μL nf-H_2O to

TABLE 2.6: Theoretical composition of the 16S gene in the ZymoBIOMICS Microbial Community DNA Standard

Bacterial species	16S theoretical composition (%)
Pseudomonas aeruginosa	4.2
Escherichia coli	10.1
Salmonella enterica	10.4
Lactobacillus fermentum	18.4
Enterococcus faecalis	9.9
Staphylococcus aureus	15.5
Listeria monocytogenes	14.1
Bacillus subtilis	17.4

make the solution up to a total volume of 50μL. Each sample received a different primer pair (shown in Figure 2.2). Reactions were amplified on a thermocycler with the following conditions: initial denaturation at 95 °C for 3 minutes, 30 cycles at 95°C for 30 seconds, annealing at 54°C for 90 seconds for amplification of the V3-V4 region and 56°C for amplification of the V4 region and a final extension at 72°C for 10 minutes.

Resulting amplicons were cleaned up using 0.7x AMPure XP Beads (Beckman Coulter) to DNA (35μL beads added to 50μL DNA). DNA was eluted in 50μL AE buffer (Qiagen). Sample concentration was measured using the Qubit dsDNA HS Assay (described in Section 2.3.1) and each sample normalised to the desired final library concentration using nf-H_2O. The samples were then pooled to give a final library for sequencing.

The concentration of the final equimolar library was measured using three methods. The library was quantified using the Qubit as described in Section 2.3.1. The Tapestation was performed as described in Section 2.3.2. The NEBNext Library Quant Kit for Illumina (New England Biolabs) was used to quantify the library and each standard and the library was run in triplicate, following the manufacturers protocol. An average of the measured concentrations by each method was

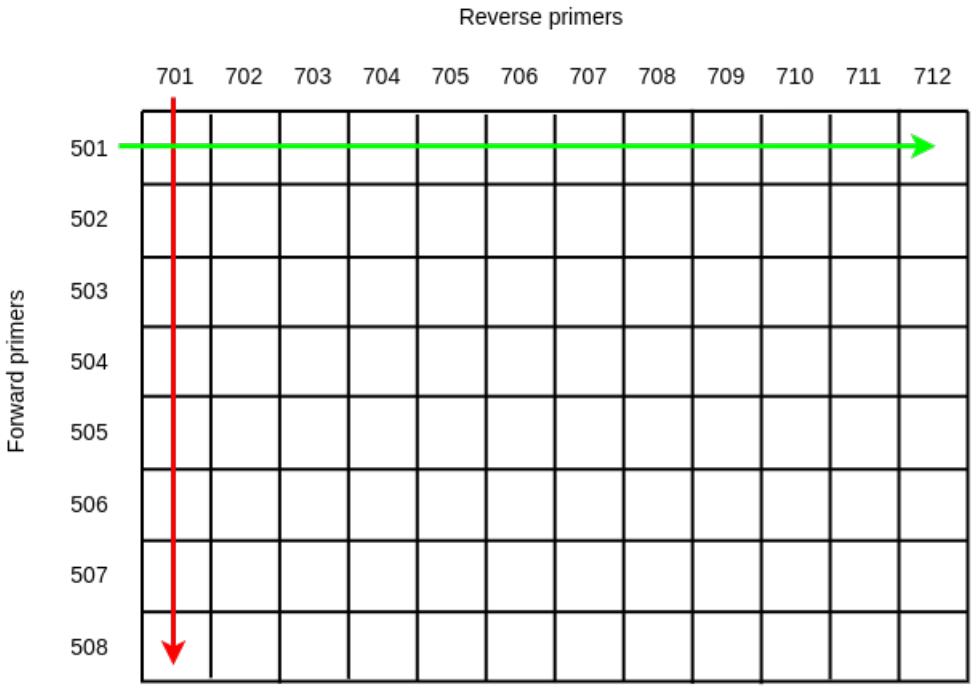

FIGURE 2.2: Diagram demonstrating how each well on a 96-well plate receives a different primer pair for sequencing. Forward primers (501-508) are loaded in each well on the plate in a horizontal line. Reverse primers (701 to 712) are loaded into each well on the plate in a vertical line. Each well receives a different primer pair i.e. the first well will receive 501 and 701, the second well will receive 501 and 702, etc.

taken and the library was diluted and denatured accordingly.

2.7.2 Library denaturation and loading

The equimolar library and the PhiX control (Illumina) were denatured by equal amounts of library and 0.2N NaOH, briefly vortexing and centrifuging at 12,000rpm for 1 minute. The final library was loaded into a Mi-Seq 500-V2 cartridge (Illumina) at a concentration of 3.6pM and the PhiX control loaded at a final concentration of 1.2pM. Custom sequencing primers (Table 2.7) were spiked into the cartridge at a concentration of 0.5μM. The cartridge was then loaded onto the MiSeq platform (Illumina) for sequencing.

Table 2.7: Custom sequencing primers for Illumina MiSeq platform

Primer name	Primer sequence (5' to 3')
Illumina V3F Read I	TATGGGTAATTGGCCTACGGGNGGCWGCAG
Illumina V4F Read I	TATGGTAATTGTGTGCCAGCMGCCGCGGTAA
Illumina V4R Read II	AGTCAGTCAGCCGGACTACHVGGGTWTCTAAT
Illumina V3-V4 Index	ATTAGAWACCCBDGTAGTCCGGCTGACTGACT

Table 2.8: Barcoded forward F1 primers for amplification of the V3-V4 region of the bacterial 16S rRNA gene

	Primer sequence (5' to 3')				
Primer name	Adaptor	Barcode	PAD	Linker	16S Primer
F1-A501	AATGATACGGCGACCACCGAGATCTACAC	ATCGTACG	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-A502	AATGATACGGCGACCACCGAGATCTACAC	ACTATCTG	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-A503	AATGATACGGCGACCACCGAGATCTACAC	TAGCGAGT	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-A504	AATGATACGGCGACCACCGAGATCTACAC	CTGCGTGT	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-A505	AATGATACGGCGACCACCGAGATCTACAC	TCATCGAG	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-A506	AATGATACGGCGACCACCGAGATCTACAC	CGTGAGTG	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-A507	AATGATACGGCGACCACCGAGATCTACAC	GGATATCT	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-A508	AATGATACGGCGACCACCGAGATCTACAC	GACACCGT	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-B501	AATGATACGGCGACCACCGAGATCTACAC	CTACTATA	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-B502	AATGATACGGCGACCACCGAGATCTACAC	CGTTACTA	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-B503	AATGATACGGCGACCACCGAGATCTACAC	AGAGTCAC	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-B504	AATGATACGGCGACCACCGAGATCTACAC	TACGAGAC	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-B505	AATGATACGGCGACCACCGAGATCTACAC	ACGTCTCG	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-B506	AATGATACGGCGACCACCGAGATCTACAC	TCGACGAG	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-B507	AATGATACGGCGACCACCGAGATCTACAC	GATCGTGT	TATGGTAATT	GG	CCTACGGGNGGCWGCAG
F1-B508	AATGATACGGCGACCACCGAGATCTACAC	GTCAGATA	TATGGTAATT	GG	CCTACGGGNGGCWGCAG

Table 2.9: Barcoded forward F2 primers for amplification of the V4 region of the bacterial 16S rRNA gene

	Primer sequence (5' to 3')				
Primer name	Adaptor	Barcode	PAD	Linker	16S Primer
F2-A501	AATGATACGGCGACCACCGAGATCTACAC	ATCGTACG	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-A502	AATGATACGGCGACCACCGAGATCTACAC	ACTATCTG	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-A503	AATGATACGGCGACCACCGAGATCTACAC	TAGCGAGT	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-A504	AATGATACGGCGACCACCGAGATCTACAC	CTGCGTGT	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-A505	AATGATACGGCGACCACCGAGATCTACAC	TCATCGAG	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-A506	AATGATACGGCGACCACCGAGATCTACAC	CGTGAGTG	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-A507	AATGATACGGCGACCACCGAGATCTACAC	GGATATCT	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-A508	AATGATACGGCGACCACCGAGATCTACAC	GACACCGT	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-B501	AATGATACGGCGACCACCGAGATCTACAC	CTACTATA	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-B502	AATGATACGGCGACCACCGAGATCTACAC	CGTTACTA	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-B503	AATGATACGGCGACCACCGAGATCTACAC	AGAGTCAC	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-B504	AATGATACGGCGACCACCGAGATCTACAC	TACGAGAC	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-B505	AATGATACGGCGACCACCGAGATCTACAC	ACGTCTCG	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-B506	AATGATACGGCGACCACCGAGATCTACAC	TCGACGAG	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-B507	AATGATACGGCGACCACCGAGATCTACAC	GATCGTGT	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA
F2-B508	AATGATACGGCGACCACCGAGATCTACAC	GTCAGATA	TATGGTAATT	GT	GTGCCAGCMGCCGCGGTAA

TABLE 2.10: Barcoded i7-indexed primers for amplification of the V3-V4 region of the bacteria 16S gene

	Primer sequence (5' to 3')				
Primer name	Adaptor	Barcode	PAD	Linker	16S Primer
R701	CAAGCAGAAGACGGCATACGAGAT	AACTCTCG	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R702	CAAGCAGAAGACGGCATACGAGAT	ACTATGTC	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R703	CAAGCAGAAGACGGCATACGAGAT	AGTAGCGT	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R704	CAAGCAGAAGACGGCATACGAGAT	CAGTGAGT	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R705	CAAGCAGAAGACGGCATACGAGAT	CGTACTCA	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R706	CAAGCAGAAGACGGCATACGAGAT	CTACGCAG	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R707	CAAGCAGAAGACGGCATACGAGAT	GGAGACTA	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R708	CAAGCAGAAGACGGCATACGAGAT	GTCGCTCG	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R709	CAAGCAGAAGACGGCATACGAGAT	GTCGTAGT	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R710	CAAGCAGAAGACGGCATACGAGAT	TAGCAGAC	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R711	CAAGCAGAAGACGGCATACGAGAT	TCATAGAC	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT
R712	CAAGCAGAAGACGGCATACGAGAT	TCGCTATA	AGTCAGTCAG	CC	GGACTACHVGGGTWTCTAAT

2.7.3 Sequencing data processing and statistical analysis

Sequencing data was demultiplexed on the MiSeq platform and FASTq files containing sequences for each sample were obtained. Prior to selection of the appropriate bioinformatic software to process the data and assign OTUs, taxonomic assignment with two commonly used open-source software packages were compared: Mothur [110] and VSEARCH [109].

The taxonomic assignment using the Mothur platform utilises the RDP Naive Bayesian Classifier which is a leave-one-out validation method. Put simply, this method looks for patterns within the reference database (or 'training set') to classify each query sequence. Although this method is thorough in its approach, it is also computationally expensive and can therefore be slow.

VSEARCH allows utilisation of the SINTAX (simple non-Bayesian taxonomy) classification method which predicts taxonomy based on k-mer similarity to identify the top-hit in the reference database. As this method does not require training, it is less computationally intensive. A random subsample of 32 8-mers are taken from the query sequence and compared to the reference database. This is repeated for 100 iterations and the resulting top-hit is reported. This method is faster than the RDP Naive Bayesian Classifier and therefore it is often utilised for larger datasets.

Comparison of the two methods found that there were no differences in the quality of data following initial processing of raw sequencing data using the Mothur and VSEARCH pipelines. The VSEARCH pipeline was more easily modified, significantly faster and less computationally intensive than Mothur. For the large datasets generated in this thesis, VSEARCH was used for raw sequence processing. OTU assignment and taxonomic classification was superior using Mothur compared to VSEARCH. Using the Mothur software package, more sequences were classified to genera with higher confidence levels. Therefore, raw sequencing data was processed using VSEARCH followed by taxonomic classification using

Mothur. This is outlined in Figure 2.3 and the final data analysis pipeline is described below.

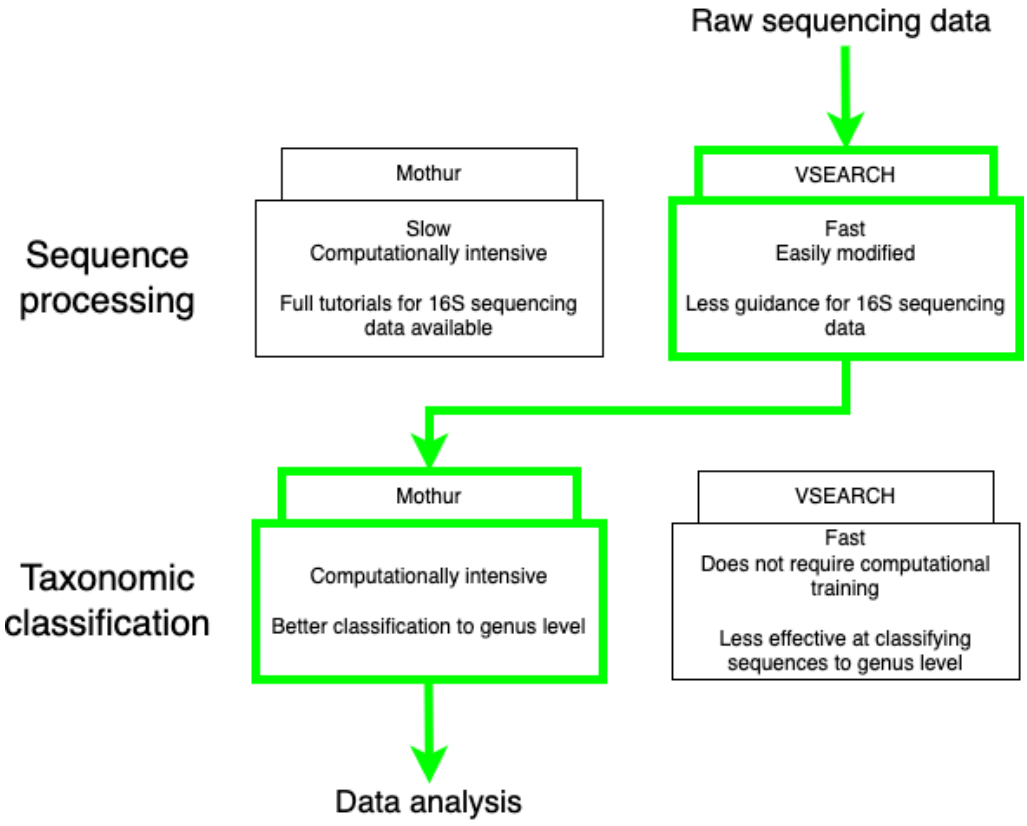

FIGURE 2.3: Bioinformatic pipeline for data analysis. Sequencing data was analysed using VSEARCH for raw seqeuencing processing and quality filtering, followed by Mothur for taxonomic classification

Paired-end 250bp reads were merged using FLASH v1.2.11 [108] and quality filtering was performed using VSEARCH [109]. Quality filtering consisted of removing any sequences shorter than the target region and removing any singletons. Taxonomic classification was performed using the Mothur software [110] with the RDP reference database (version 14) [111]. A similarity cut-off of 97% was implemented and any sequences that aligned to a non-bacterial domain were removed. Output files included an OTU table and taxonomy table which were used for analysis as described below.

The OTU and taxonomy table were imported into RStudio and merged with a metadata file containing all sample information using the *phyloseq* package [112].

Table 2.11: Significance levels as determined by P-values

P-value	**Significance**	**Symbol**
<0.0001	Extremely significant	****
0.0001 to 0.001	Extremely significant	***
0.001 to 0.01	Very significant	**
0.01 to 0.05	Significant	*
≤0.05	Not significant	ns

Negative controls including both negative DNA extraction controls and no template controls were assessed for contamination. For each sequencing run the mean number of reads in these controls were calculated and can be found in the respective results sections. Samples with sequencing reads within two standard deviations from the mean number of reads were considered negative. The mock community of known theoretical composition was assessed for any unexpected or spurious OTUs. If these were present their abundance was used as a cut-off for false sequence generation.

Data was analysed in RStudio using a range of packages. Packages used included *vegan*, *data.table*, *dplyr*, *ggplot2*, and *phyloseq*. Abundance data was imported into GraphPad Prism (version 9) where data was visualised and statistical analysis was performed. Significance was determine based on the P values shown in Table 2.11.

Investigating bacterial translocation in animal models

Results presented in this chapter were published in April 2021.
Lily Gates, Nigel J. Klein, Neil J. Sebire, and Dagmar G. Alber. Characterising Post-mortem Bacterial Translocation Under Clinical Conditions Using 16S rRNA Gene Sequencing in Two Animal Models. *Frontiers in Microbiology*, 2021. ISSN 1664302X. doi: 10.3389/fmicb.2021.649312. [113].

3.1 Introduction

As discussed in Section 1.6, the interpretation of routine PM microbiology results is difficult in cases of SUDIC given the often unspecific clinical history and histological findings. When a positive PM microbiology result is obtained from visceral tissues that are believed to be sterile in life, it is often difficult to determine the relevance of results. A positive PM microbiology result could signify one of the following: (i) a true infection that occurred during life, (ii) contamination during sample collection and processing, or (iii) contamination due to the natural decay process of the body through a process known as PM bacterial translocation [87, 114, 115].

PM bacterial translocation is defined as the migration of viable bacteria from highly colonised bodily sites, such as the GI tract, to extra-intestinal tissues after death [114]. Following death, mechanisms that protect the host from overwhelming translocation of microbes from the GI tract to extra-intestinal tissue cease to exist. In a healthy living host physiological mechanisms such as the presence of tight junctions between epithelial cells, the mucosal barrier, and the immune

system prevent bacterial translocation. After death bacteria overgrow and with the deterioration of normal chemical and physical barriers, the bacteria have been hypothesised to escape the GI tract and migrate to surrounding tissues and blood. When a sample is then collected at autopsy a positive bacterial culture may obtained as a result of this bacterial spread.

Few studies have investigated the phenomenon of PM bacterial translocation, and those that have typically investigate bacterial invasion under ambient conditions for the purpose of forensic investigation. Heimesaat *et al.* investigated PM translocation in a mouse model kept under ambient conditions using conventional bacterial culture [116]. Results from this study showed that extra-intestinal sites including cardiac blood, kidney, liver, mesenteric lymph nodes, and spleen were invaded by intestinal bacteria as soon as 5 minutes PM [116]. Another study tracked fluorescently-labelled *Staphylococcus aureus* following intranasal inoculation of a mouse model found that these bacteria could be detected in all organs one hour PM [117]. This animal model was also kept under ambient conditions following sacrifice.

There is a current lack of understanding of the PM bacterial translocation process in a clinical setting where bodies are refrigerated during the PMI [114, 115, 118, 119]. This lack of knowledge causes difficulty when interpreting positive PM microbiology results as the relevance of the result must be determined. A positive PM results poses the question *is the cultured organism causative of the fatal event or due to sample contamination?* Unlike contamination during sampling, putrefaction is a natural process that can be limited but not completely controlled [115]. Routine procedures such as refrigeration of the body soon after death aims to limit decomposition, but the extent to which this prevents bacterial translocation has yet to be studied.

PM bacterial translocation has not yet been characterised under conditions mimicking standard clinical practice using contemporary molecular approaches. Essential to the interpretation of PM microbiological results of SUDI, is a clear understanding of the natural bacterial translocation that occurs PM in this clinical

setting. The hypothesis of this study was that although PM bacterial translocation occurs, it is not extensive and can be defined using 16S rRNA gene sequencing.

This study will use 16S rRNA gene sequencing to enable an identification of the viable bacteria present in the tissues and an unbiased quantitative determination of the bacterial community structure. Through this investigation a baseline sequencing signal will be established for the PM process in animal models free of known infection.

3.2 Chapter aims

Aim To characterise PM bacterial translocation under controlled conditions in two model organisms using the 16S rRNA gene sequencing method and assess the implications it may have on the interpretation of PM microbiology results. This study used two separate animal models; a mouse model for ease of sampling, followed by a piglet model given its anatomical similarities to humans and ability to perform repeated sampling.

3.3 Methods

3.3.1 16S rRNA gene sequencing

DNA extracts were amplified for sequencing as described in Section 2.7.1. A post-PCR concentration of <0.5ng/μL was considered negative. Tissue samples were amplified thrice before being considered negative.

Samples collected in this study were sequenced on four separate MiSeq sequencing runs according to the animal model and sample type i.e. 1. mouse tissue, 2. piglet tissue, 3. mouse GI tract and 4. piglet nasal and rectal swabs. The GI tract samples and swabs were normalised to 1.0ng/μL and pooled into an equimolar solution. As the tissue samples had a lower bacterial biomass and therefore amplified a lower concentration of DNA, they were normalised and pooled to an equimolar library with a lower final concentration of 0.2ng/μL. Equimolar libraries were loaded onto the MiSeq platform as described in Section 2.7.

3.3.2 Data analysis

Mouse GI tract samples were randomly subsampled at 20,000 reads. This subsampling resulted in the exclusion of a single GI tract sample (C1.D14.A) due to a low number of sequencing reads. Piglet nasal and rectal swabs were randomly subsampled at 100,000 and 14,000 reads, respectively. Two samples (nasal swab P3.D10.right and rectal swab P2.D3) were excluded from downstream analysis due to low quality sequencing reads. Differences in the total number of generated sequencing reads that passed quality filtering were statistically compared between sample type and organism sampled using an independent t-test. Piglet tissues were subsampled to 500 reads prior to β-diversity calculation. Changes in community structure in the piglet tissues (β-diversity) was assessed with a non-metric multi-dimensional scaling (NMDS) plot based on Bray-Curtis dissimilarities between the samples. The effects of sampling day on bacterial community changes

were measured by performing a permutational multivariate analysis of variance (PERMANOVA) test on Bray-Curtis dissimilarity measures using the *phyloseq* package [112]. A Kruskal-Wallis test by ranks was performed on communities at different time points to compare taxonomic differences at family level classification using the stats package [120]. For statistical tests a $P<0.05$ was considered significant. Figures were generated using the ggplot2 package [121] and Graphpad Prism (version 9) [122].

3.4 Results

3.4.1 Validation of sequencing technique

Validation of the sequencing technique was required before this study could be performed. This consisted of sequencing a commercial bacterial community of known composition in order to determine how accurately the technique could estimate the community structure. The true abundance of 16S rRNA genes from each bacteria within the known community was compared to those estimated by the sequencing run following data processing and cleaning. Relative abundances of each sequenced mock community together with the true composition can be found in Figure 3.1 and the average relative abundances can be found in Table 3.1.

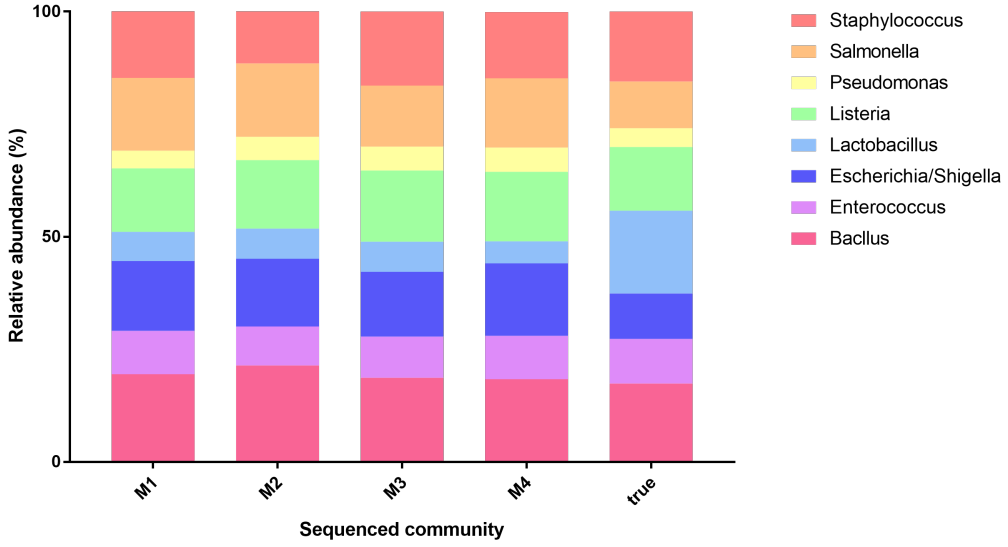

FIGURE 3.1: Relative abundance plots showing the bacterial genera identified in the four sequenced mock communities (M1-M4) compared to the true composition (true)

The sequencing and bioinformatic protocol successfully classified the known bacterial constituents to genus level. There was no significant difference between the proportional abundance of each Genera from all four sequenced mock communities (P=0.990). In all four mock communities the Lactobacillus genera was under represented, with an average estimated percentage of 6.2% compared to the actual composition of 18.4%. Other bacterial members were estimated at a slightly increased relative abundance. There was no significant difference between the mean relative abundance of the mock communities and the actual composition of the mock communities (P=0.993). These results demonstrate the ability of the technique to successfully identify all 16S rRNA genes within the mock community and validates the use of this technique for the remainder of this study.

Table 3.1: Comparison of actual composition vs. mock estimation

Genera	Actual composition (%)	Average mock estimation (%)	Standard deviation
Bacillus	17.4	19.5	±1.2
Enterococcus	9.9	9.3	±0.4
Escherichia/Shigella	10.1	15.3	±0.6
Lactobacillus	18.4	6.2	±0.8
Listeria	14.1	15.1	±0.6
Pseudomonas	4.2	5	±0.6
Salmonella	10.4	15.4	±1.1
Staphylococcus	15.5	14.4	±1.7

3.4.2 Mouse model

3.4.2.1 Changes in the GI tract microbiome

To begin the investigation of the PM process in the mouse model, the gut microbiome in the mouse GI tract samples were compared. Samples collected from the GI tract generated an average of 190,007 sequencing reads (range 9,986 - 410,623). The gut microbiome in the three mice sampled at each time point showed no significant differences based on Bray-Curtis distances as was expected due to their inbred nature and controlled diet and housing environment (P=0.07). The samples from mice sampled at the same time points were averaged and the top 10 bacterial families at each time point are shown in Figure 3.2.

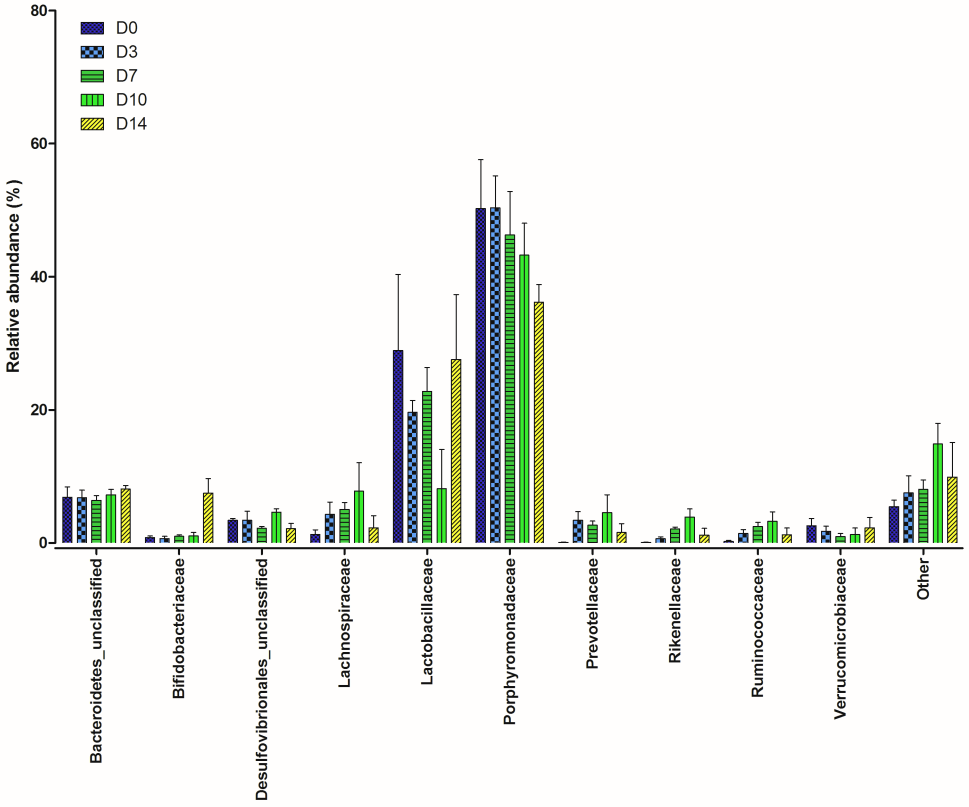

FIGURE 3.2: Changes in the relative abundance of the top 10 bacterial families present within the mouse gastrointestinal (GI) tract samples collected at five time points post-mortem.

The β-diversity of each community was calculated using Bray-Curtis distances. The gut microbiome at each time point was compared to the microbiota on D0 to track community changes over the study period. There were no significant changes observed in the gut microbiota from D0 to D7 (P=0.096). Significant changes were, however, observed in the bacterial communities on D10 (P=0.007) and D14 (P=0.037). To investigate which bacterial families were responsible for these differences a Kruskal-Wallis test was performed. There were no significant changes in the dominant families over the two-week study period. Significant differences were observed in Bifidobacteriaceae (P=0.013) and Rikenellaceae (P=0.041). The changes in relative abundance in these two bacterial families can be seen in Figure 3.3A and Figure 3.3B. At each time point from D0 to D10 Bifidobacteriaceae represented less than 1% of the total bacterial community. This was followed by a sharp increase occurring between D10 and D14 where their relative abundance increased to 7.2%. Rikenellaceae was also present at less than 1% in the starting community on D0. There a was a steady increase to D10 where the relative abundance peaked at an average of 3.9% relative abundance, which then declined and on D14 Rikenellaceae represented just 0.6% of the bacterial community. These results show that over the two-week study period there were only slight changes detected in the mouse gut microbiome. The main bacterial members present in the gut microbiome remained stable over the study period.

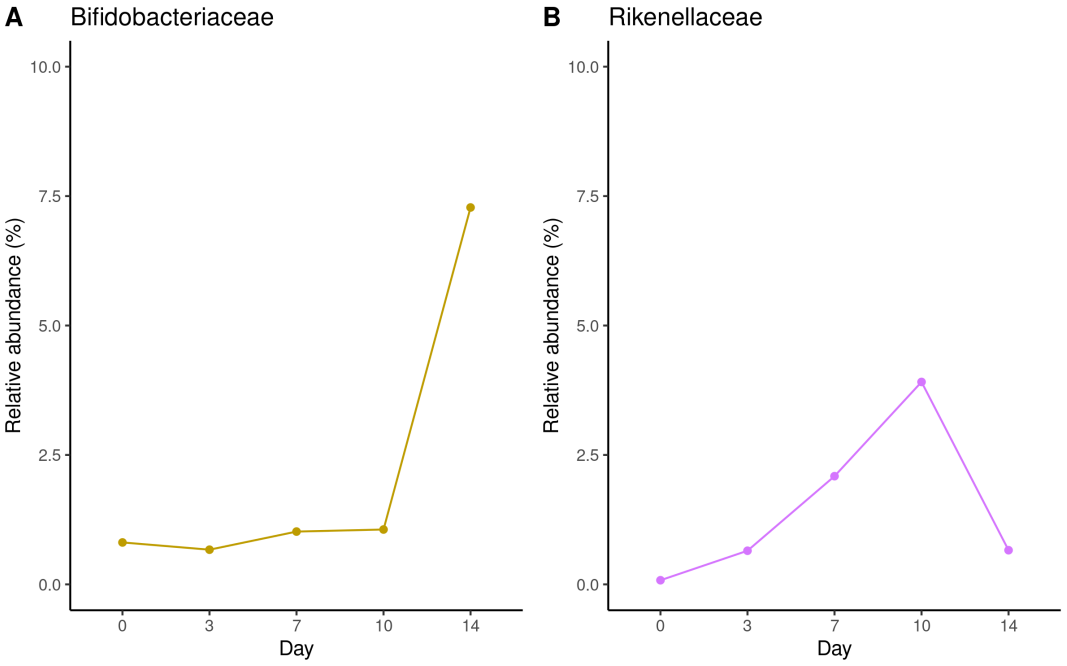

FIGURE 3.3: Relative change of (A) Bifidobacteriaceae and (B) Rikenellaceae in mouse GI tract samples over the two-week study period

3.4.2.2 Tissue samples

A total of 75 mouse tissue samples were obtained. Of these 75 tissue samples, 4 (5%) were considered as positive for the 16S rRNA gene using the quality filtering threshold described in the Section 3.3.2. Table 3.2 shows the positive samples. Positive tissue samples had an average of 3,267 sequencing reads (range 642 – 10,839). The number of sequencing reads in the tissue samples were significantly lower than those in the GI tract samples (P=0.004). The number of positive samples did not increase with time PM, nor did they correlate with cage or tissue type.

In order to assess whether the bacteria detected in the positive tissues could have originated in the GI tract and translocated during the PMI, the OTUs identified in the tissues were compared to those identified in the GI tract. Two of the four positive samples (C1.D3.Liver and C2.D14.Blood) did not share any OTUs with the GI tract samples collected from mice housed in the same cage at the same or previous time points. Sample C2.D3.Liver shared all OTUs with the corresponding GI tract sample. These bacterial families included Enterobacteriaceae, Lactobacillaceae, and Porphyromonadaceae. These families were of high abundance in the GI tract samples. Sample C2.D10.Heart shared four of 11 OTUs with the corresponding GI tract sample. These OTUs were members of the Entercoccaceae, Enterobacteriaceae and Staphylococcaceae bacterial families.

Three of the four positive tissue samples were dominated by either Enterobacteriaceae or *Entercoccus*, which represented >45% of each positive sample. As the read numbers were low, a specific qPCR was carried out to identify whether these reads were true positives. These bacteria were also identified in three GI tract samples from C2 at low relative abundances (between 5% and 11% of the total community). These samples were used as positive controls. Two of the three tissues were positive for Enterobacteriaceae, but generated a very weak signal with a cyclic threshold (CT) value >34. The positive control samples had an average CT value of 27. These results suggest that in the mouse model there is little to no

TABLE 3.2: Mouse tissues positive for the 16S rRNA gene. Table headers describe the cage from which the sample was collected (C1, C2, C3) and the day on which the sample was collected (D0, D3, D7, D10, D14). P=positive for 16S rRNA gene following sequencing and quality filtering. N=negative for 16S rRNA gene after sequencing and quality filtering.

	D0			D3			D7			D10			D14		
	C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
Blood	N	N	N	N	N	N	N	N	N	N	N	N	N	**P**	N
Heart	N	N	N	N	N	N	N	N	N	N	**P**	N	N	N	N
Liver	N	N	N	**P**	**P**	N	N	N	N	N	N	N	N	N	N
Lung	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Spleen	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N

bacterial translocation observed from the GI tract to the heart, blood, lung, liver, or spleen that can be detected using the 16S rRNA gene sequencing technique over the two-week study period.

3.4.3 Piglet model

3.4.3.1 Rectal swabs

Rectal swabs were collected from the piglet carcasses to assess the gut microbiome without perforating the GI tract lining which would potentially affect bacterial translocation. The bacteria within the rectal cavities of the piglet carcasses were significantly different in each piglet as was expected due to differences in genetics, diet, and environmental conditions (P=0.018). Rectal swabs generated an average of 224,936 (range 139,068 - 321,549). The relative abundance of bacterial families in each piglet are shown in Figure 3.4. The rectal swab community in each piglet progressed in a different manner over the two-week period. In P1, the bacterial community in the rectal cavity remained relatively stable from D0 to D10. On D14, an increase in Pseudomonadaceae and Moraxellaceae was observed. The dominant families in the rectal cavities in P2 were Enterobacteriaceae and Moraxcellaceae on D0 and D3 which progressed to Bacteroidaceae and Veillonellaceae on D10 and D14. In P3, Porphyromonadaceae and Veillonellaceae dominated on D0, but communities from D3 to D14 were dominated by Clostridiaceae and Pseudomonadaceae.

3.4.3.2 Nasal swabs

Nasal swabs were obtained from both the left and right nares of each piglet at each sampling time point. Nasal swab samples generated an average of 239,446 sequencing reads. Prior to comparative analysis, 100,000 sequences were randomly subsampled from each swab resulting in the loss of a single right nasal sample collected from P3 on D10 due to a low number of generated sequencing reads. As duplicate samples were collected, an average of the bacterial members in paired samples was calculated to give an overall view of the nasal cavity.

To analyse community changes in more detail, bacterial families were visualised on stacked bar plots showing their relative abundance (Figure 3.5). Despite clear

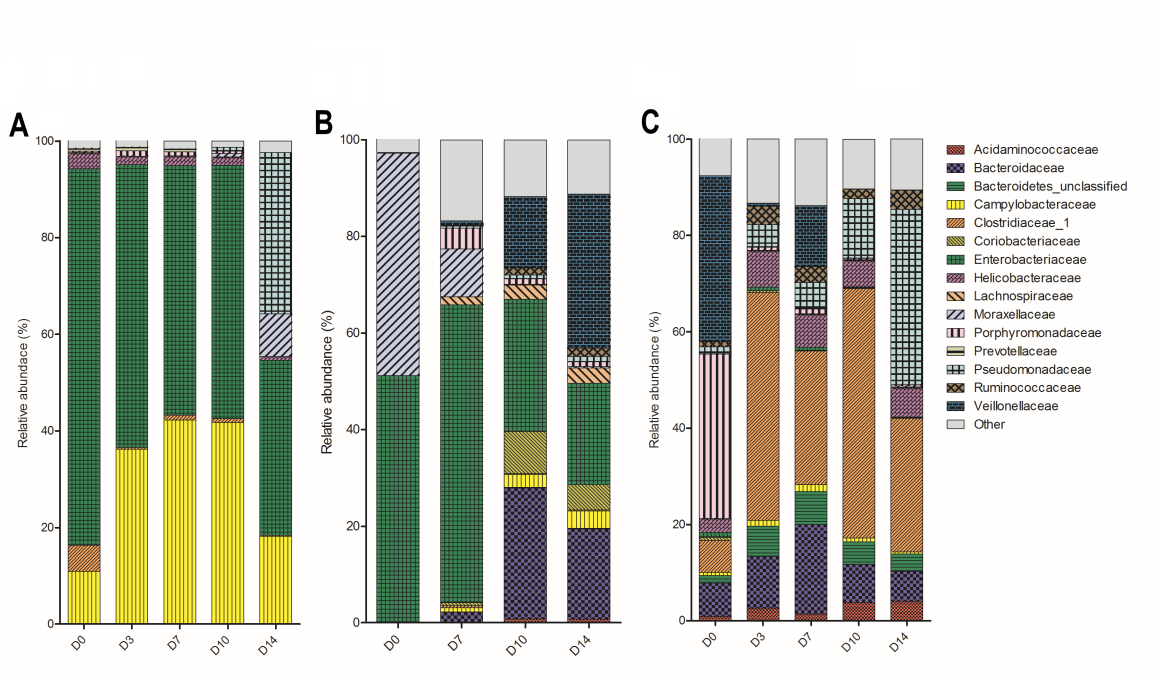

FIGURE 3.4: Changes in the relative abundance of the top 10 bacterial families in piglet rectal swabs collected from (A) P1, (B) P2, and (C) P3 over the study period. The x-axis describes the day on which the sample was collected (D0, D3, D7, D10 or D14).

differences between the communities at early time points from the three piglets, bacterial communities progressed in a similar fashion. On D0 the dominant bacterial families included Enterobacteriaceae, Flavobacteriaceae, Moraxellaceae, and Pasteurellaceae. From D3 to D10 the communities in the different piglets began to converge and by D10 and D14 there was no significant different between the communities in the nasal cavities of all three piglets (P=0.33). The similarity between communities was due to the gradual domination of Pseudomonadaceae in all communities that occurred over the study period. Pseudomonadaceae bacteria began to emerge on D7 and their relative abundance progressively increased until D14 where they represented >90% of the bacteria within the nasal cavity of each piglet.

3.4.3.3 Development of a method for piglet tissues identification

As repeated sampling of the piglet carcasses was performed through the skin using needle biopsies, it was important to determine whether the correct tissue had been obtained for analysis. To do this, a tissue-specific reverse-transcriptase qPCR was designed using differentially expressed genes in piglet heart, lung, and liver tissue. Primer pairs targeted lung-expressed gene Surfactant Protein C (*SFTPC*), heart-expressed gene Ankyrin Repeat Domain 1 (*ANKRD1*), liver-expressed gene Apolipoprotein C4 (*APOC4*), and universally-expressed gene Glyceraldehyde-3-Phosphate Dehydrogenase (*GAPDH*) as a positive control and measure of amplifiable DNA.

Piglet tissues of known identity were used to validate this method. Tissues included heart, liver, lung, kidney, and trachea; the latter two being used as negative controls. RNA was extracted from each tissue and reverse-transcribed to produce complementary DNA (cDNA). Each cDNA extract underwent PCR amplification with all four primer pairs.

Amplification curves and melt peaks obtained from each primer pair are shown in Figure 3.6. The coloured peak in each graph shows the correct target tissue. For

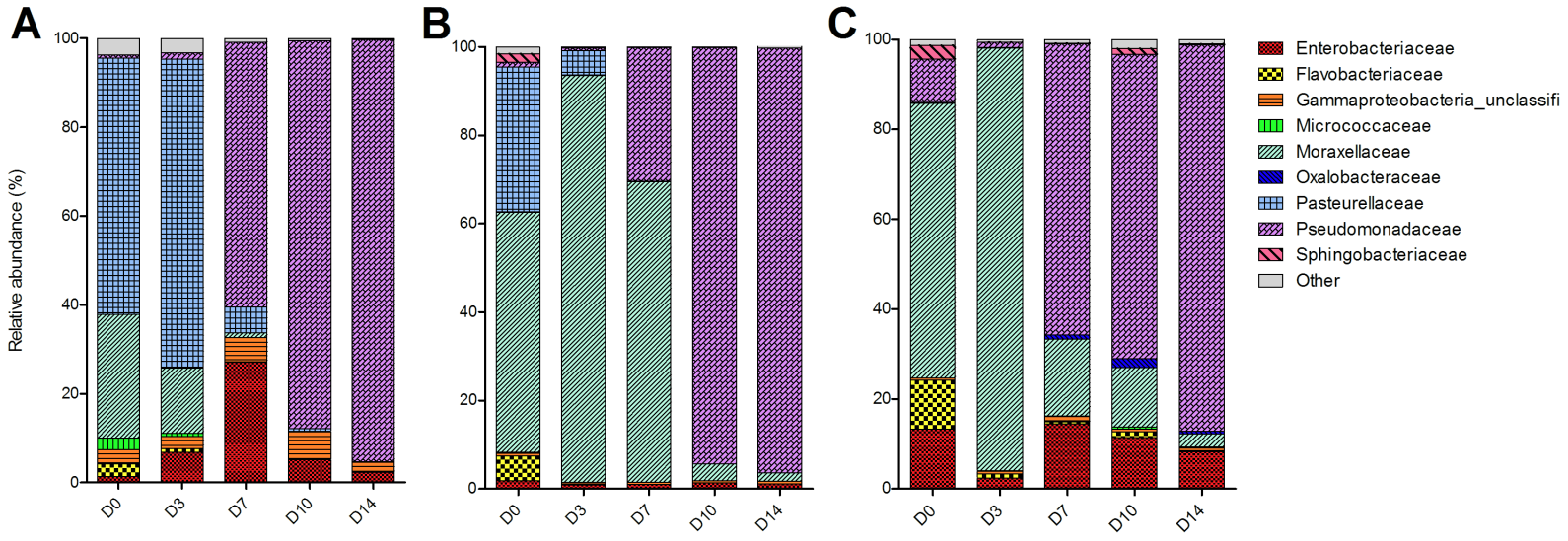

FIGURE 3.5: Changes in the relative abundance of the top 10 bacterial families in piglet nasal swabs collected from (A) P1, (B) P2, and (C) P3 over the study period. The x-axis describes the day on which the sample was collected (D0, D3, D7, D10 or D14).

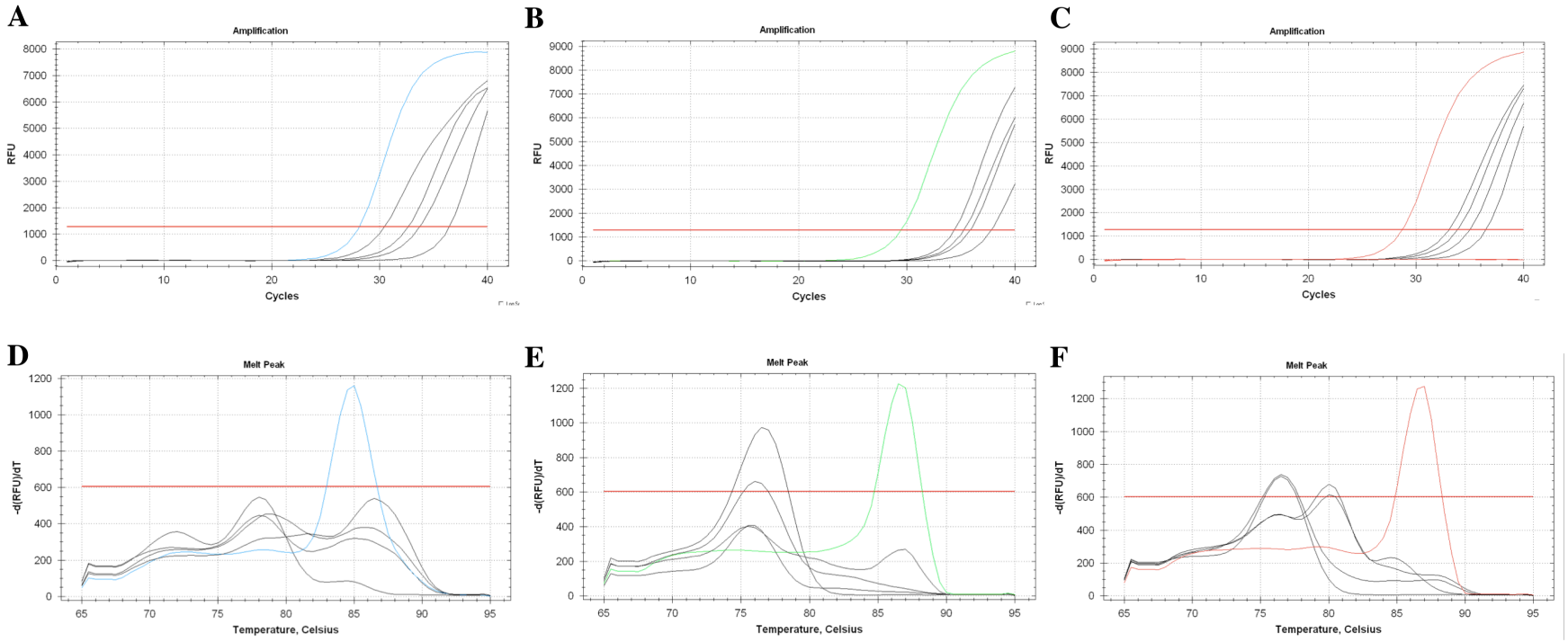

FIGURE 3.6: Amplification curves (A-C) and melt peaks (D-F) produced by differentially expressed mRNA RT-qPCR. The coloured curves show the target tissue- blue = lung, green = liver, red = heart.

Table 3.3: Tissue type identified using the designed RT-qPCR method versus the actual tissue type

DNA extract	Tissue identified	Actual tissue	Agreement
1	Liver	Liver	✓
2	Lung	Lung	✓
3	Other	Trachea	✓
4	Heart	Heart	✓
5	Heart	Heart	✓
6	Other	Kidney	✓
7	Liver	Liver	✓
8	Heart	Heart	✓

each pair, the target tissue crossed the CT between 26 and 30 cycles. The control tissues also crossed the CT, but at more than 30 cycles.

The melt peak was analysed to assess whether the qPCR reaction produced a single specific product and to assess primer specificity. DNA extracted from the target tissue showed a single melt peak, representing amplification of a single amplicon. Melt peaks produced from DNA extracted from other tissues demonstrated multiple peaks of low relative fluorescence units (RFU) suggesting non-specific binding of the primers to other nucleic acid material or primer-dimer formation.

To confirm whether this method could be successful in identifying tissue type when processing DNA extracts from unknown tissues, extracts were processed in a blind fashion. In this investigation eight randomly selected DNA extracts from pig heart, liver, lung, kidney, and trachea were subject to PCR amplification with the four primer sets. All eight DNA extracts were identified as the correct tissues (shown in Table 3.3). In tissue samples where the amplification curves and melt peaks were not present with heart, liver, or lung primer pairs but the positive control GAPDH gene was successful, the tissue was identified as other. This investigation demonstrated that this designed method for identification of correct tissue types was sufficient for use on needle biopsy samples collected from the piglet.

3.4.3.4 Tissue-specific RT-PCR for tissue identification

A total of 45 tissue samples were collected (15 heart, 15 lung, and 15 liver samples) from three piglet carcasses. Following sample collection and nucleic acid extraction, a specific qPCR targetting differentially expressed mRNA was performed on each extract to confirm the tissue type. The successfully obtained tissues are shown in Table 3.4. The success rate of sampling the heart, the lung, and the liver using the needle biopsies was 87%, 73% and 87% respectively.

TABLE 3.4: Successfully obtained piglet tissues by needle biopsy sampling

	Piglet 1					Piglet 2					Piglet 3				
	Day					Day					Day				
Tissue	0	3	7	10	14	0	3	7	10	14	0	3	7	10	14
Heart	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓
Lung	✓	✗	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗	✓	✗	✓
Liver	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗	✓	✓

3.4.3.5 Tissue samples

A total of 35 piglet tissue samples were successfully obtained. Of these, 29 samples were considered Sequencing data were obtained for 29 of these 40 tissue samples, as shown in Table 3.4. Sequencing reads in the piglet tissues were higher than those obtained from the mouse tissues with an average of 30,960 reads from tissues across all three piglets. These read numbers were, however, significantly lower than those found in the nasal and rectal swabs (P=0.0001). Variation amongst the reads obtained from each piglet was observed, but there were no associations with sampling day or tissue type.

Bacterial families identified in the piglet tissues were assessed using β-diversity measures. Figure 3.7 shows a non-metric multi-dimensional scaling (NMDS) plot based on calculated Bray-Curtis distances of β-diversity. Each point on the NMDS

	D0			D3			D7			D10			D14		
	P1	P2	P3	P1	P2	P3	P1	P2	P3	P1	P2	P3	P1	P2	P3
Heart	x	P	P	P	P	x	P	P	P	N	N	N	P	N	P
Liver	P	N	P	N	P	N	P	P	P	N	P	N	P	N	P
Lung	P	x	P	x	P	x	P	P	N	P	P	N	P	P	P

plot represents a single tissue sample and the closer together the points lie, the more similar the bacterial community. This plot shows that the bacterial communities within tissue samples cluster based on the piglet rather than with timepoints or tissue sample type. Tissues collected from P1 were mainly dominated by Clostridiaceae (70%, 7/10), P2 tissues were dominated by Moraxellaceae (40%, 4/10) and P3 tissues by Lactobacilliaceae (44%, 4/9).

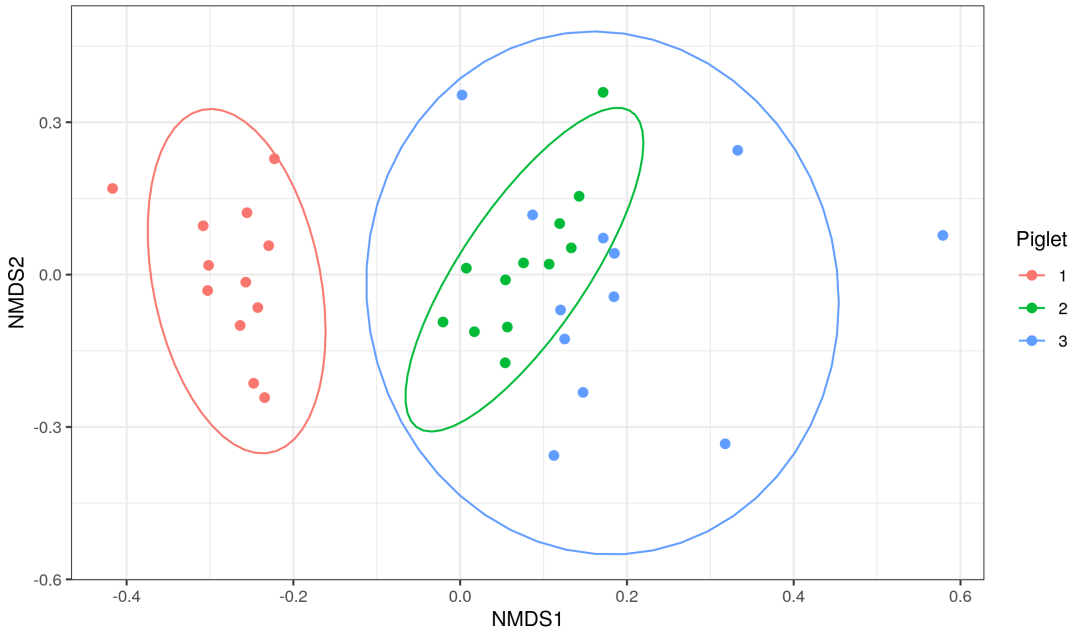

FIGURE 3.7: Non-metric multi-dimensional scaling (NMDS) plot based on Bray-Curtis dissimilarity index of piglet tissue samples. Each point represents an individual piglet tissue sample. Ellipses are based on 95% confidence intervals. Colours represent the piglet from which the sample was collected.

To determine whether translocation occurs from sites of high colonisation to otherwise sterile tissues, OTUs present within the tissues were cross-checked with those identified in nasal and rectal swabs from the same piglet to identify if this was a possible source of translocation. Of the 29 sequenced samples, 15 (52%) did not share any OTUs with either rectal or nasal swabs. There were 10/29 (34%) tissue samples that had OTUs that were also present in the rectal swabs (Figure 3.8A). Of the 10 samples that shared at least one OTU with the rectal swabs, the shared OTUs were dominant in 3/10 (30%). In comparison to the nasal swabs, 8/29 (28%) of the tissue samples shared at least one OTU with the nasal cavity, of which 1/8 (13%) shared their dominant OTU (Figure 3.8B).

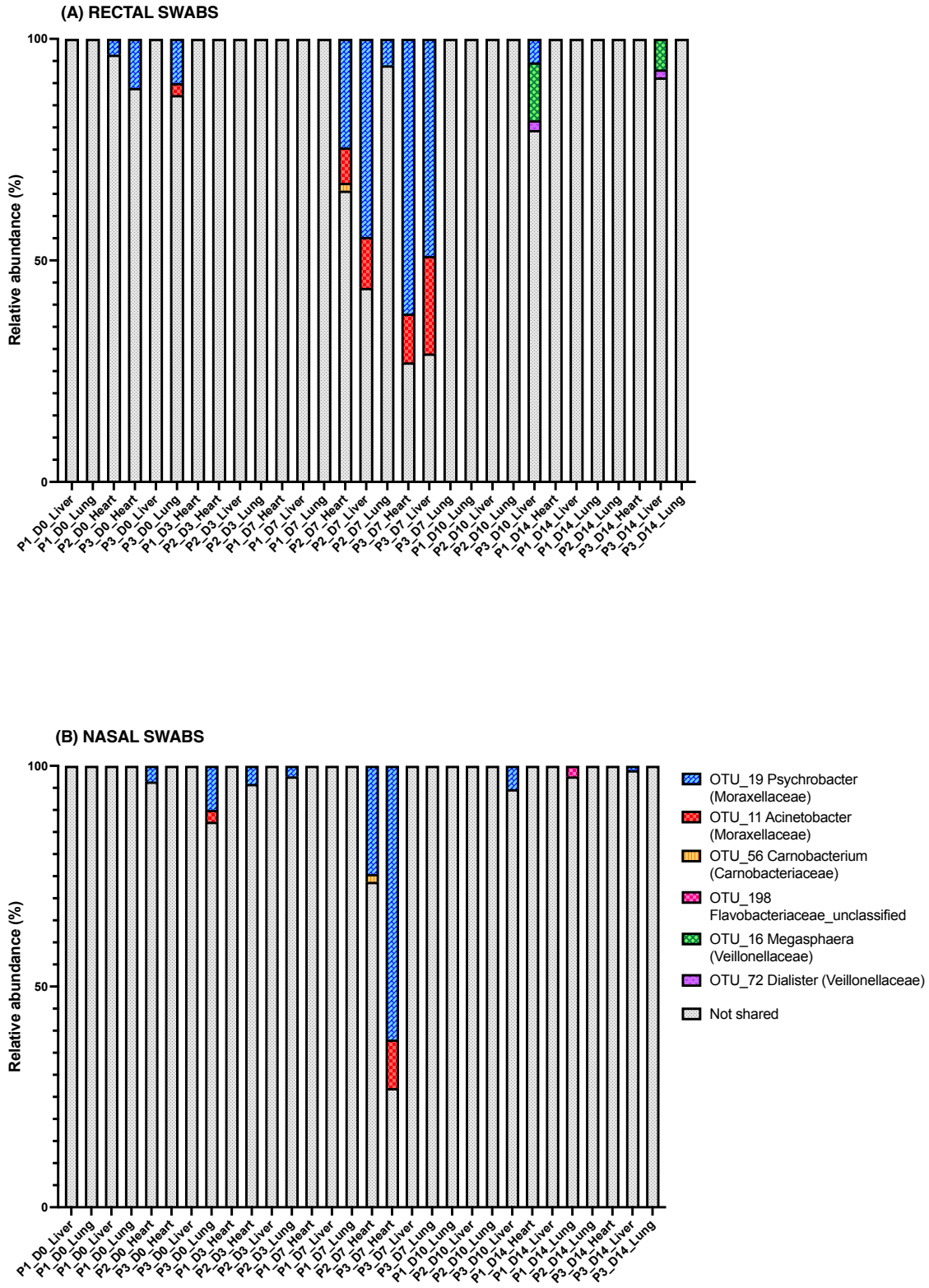

FIGURE 3.8: OTUs shared between positive piglet tissue samples and (A) rectal swabs or (B) nasal swabs collected from the piglet on the same sampling day

According to these results PM bacterial translocation from the rectal or nasal cavity may occur in the piglet model but this translocation is minimal and cannot be detected in a large proportion of tissue samples. In samples where an OTUs was detected in both the tissue sample and the rectal or nasal cavity, the OTU was not dominating in a majority of samples suggesting that these highly colonised sites were not the source of bacterial spread.

3.5 Discussion

This investigation failed to demonstrate significant and consistent PM bacterial translocation in two animal models. Although bacteria were detected within some tissues in both models, the study did not find evidence for true bacterial translocation from the GI tract or nasal cavity. These data do not support the concept of significant PM translocation as part of the normal PM process in these animal models.

Significant changes were observed in the mouse GI tract in the Bifidobacteriaceae and Ruminococcaceae families over the study period, both of which increased in relative abundance over the study period. These bacterial families consist largely of anaerobic bacterial species which are expected to proliferate during body decomposition when oxygen sources are low and nutrients limited. Although there was a significant change in these two families, the relative stability of the microbiome demonstrates the ability of body refrigeration to slow the decay process. These results support findings of another study performed by Tuomisto *et al.* who used real-time qPCR to track changes in six main bacterial families in the stool of autopsy cases [123]. The study found no significant changes in these bacteria up to five days PM [123]. A study that investigated the PM microbiome of internal organs in a rabbit model kept at 4°C found no significant differences in the bacterial communities in the caecum up to 48 hours PM [124]. The bacterial communities present in the rectal cavity of the piglets showed inter-piglet variation reflecting observations in humans where the microbiome between individuals differs depending on a number of factors including genetics, age, diet, and environment [125]. No trends were observed in changes over the study period, possibly due to the low sample size. The bacterial communities in P1 remained stable for the first 10 days PM reflecting the results of our mouse study. In P2 although a change was observed between the communities within the first three days PM, there is arguably little difference observed over the remianer of the study period. The biggest shift

was observed in Veillonellaceae, which are a bacterial family consisting of anaerobic and microaerophilic bacterial species, which may explain their ability to thrive in the conditions within the carcass as the time PM increases. Microbial succession has been shown in various studies to be directly influenced by oxygen availability, with a longer PMI associated with a shift from aerobic to anaerobic bacteria at the completion of the bloat stage of decomposition [124, 126–128].

The exponential increase of Pseudomonadaceae bacteria in the piglet nasal swabs is likely due to the PM conditions within the body bag. As decomposition progresses, it is known that the body begins to purge fluids [129]. This loss of fluid from the carcass resulted in the body bag becoming damp, which we observed on sampling at D7. These conditions could provide suitable conditions for the proliferation of the Pseudomonadaceae family. Pseudomonadaceae have been found on porcine carcasses and are associated with food spoilage given their ability to proliferate at low temperatures [130]. As the nasal cavity is exposed to the body bag it is possible that this is the source of colonisation. Pseudomonadaceae also possess a flagella enabling motility which further increases the probablility of colonisaing the nasal cavity. It is therefore likely that these findings are a potential contaminant of surroundings. Interestingly, a recent study by Hurtado *et al.* investigated the impact of PMI on microbiological results obtained at autopsy and found an increase in the identification of the *Pseudomonas* genus as PMI increased when using molecular analysis [131].

The mouse model used in this study failed to demonstrate ubiquitous PM translocation using the 16S rRNA gene sequencing method. Although Enterobacteriaceae and *Enterococcus* were detected by sequencing, the use of the highly sensitive qPCR assay and the late fluorescent signal suggest that the bacteria are present at very low numbers. If this were due to bacterial translocation PM from the GI tract it is unlikely that it would be seen in these few tissues at seemingly random time points. This observation could therefore represent PM bacterial translocation, but given the low read numbers these findings are likely due to contamination or background signal. These results are reflected in a similar study performed on

rabbit carcasses which investigated the effects of temperature on the PM microbiome. In this study bacterial DNA was too low to be quantified in a majority of PM lung and kidney samples due to low bacterial load [124]. The study also obtained low sequencing reads for these visceral tissues compared to those obtained from the caecum and ileum and hypothesised that the low bacterial DNA present in these samples resulted in non-specific amplification of host DNA that then failed to align to the reference database [124].The results of this study are also in line with previous work that has shown a poor correlation between PMI and positive bacterial culture [119, 132].

A larger number of positive tissues and higher sequencing read numbers were obtained in the piglet model compared to the mouse model. The similarity of bacterial families shared between tissues collected from the same piglet suggest that these bacterial families originate from the piglet itself, rather than the external environment. The lack of concordance between bacterial families found in the tissues and both rectal and nasal swabs suggest it is unlikely that the identified bacteria originate from these sites. Bacterial families such as Moraxellaceae, Clostridiaceae, Flavobacteriaceae, and Peptostreptococcaceae have been found in high abundance on the skin, tonsil, and nasal cavities of piglets in various studies [133–135]. These bacterial families were detected in tissues in the present study. It is therefore possible that these positive tissues have arisen from these sites, perhaps as a result of the sampling method used.

The vast number of positive cultures that are typically detected in human SUDI PM examinations are reflected in positive piglet tissues in this study; >80% of human PM tissues from SUDI autopsies generate positive culture results [85, 119]. However, in these human tissues a majority of bacteria identified are enteric bacteria [136], which differs from the bacterial species identified in this study. This may be due to differences in the PM translocation process or perhaps contamination occurs at a higher rate during human autopsy and tissue processing. The technique used in this study could also introduce differences in results as for a positive bacterial culture a single viable bacteria needs to be present within the

sample. Examining bacterial components using molecular targets such as the 16S rRNA gene allows a more accurate evaluation of bacterial communities.

3.5.1 Study limitations

The lack of skin sampling performed could have enhanced the interpretation of bacteria within tissues. The characterisation of the skin microbiome at puncture sites on the piglet carcass may have assisted in hypothesising whether the bacterial species identified within the tissue samples had been introduced via the sampling technique.

It is also important to note that this study was performed in healthy animals who were free of known infection. If an infection were to have been present in these animal models the conclusions of this study may have been different. This may be due to a the increased invasive abilities of pathogenic bacteria which may favour their entry at mucosal surfaces such as the GI tract. The presence of pathogenic bacteria may also disrupt the homeostatic balance of commensal microbes or damage to tissue epithelium compromising barrier function, further facilitating their invasion across the GI tract.

In this study the time period between death and sampling was also controlled in the model organisms and initial sampling was performed no longer than three hours PM after which carcasses were immediately refrigerated. In cases of SUDIC, often the time between death and body identification cannot be determined due to death often occurring during a period of sleep.

Finally, although this study could not detect PM bacterial translocation in these model organisms, the PM process in human subjects may differ. Given the small sample size and the use of animal models, no clinical conclusions can be made. This study does however provide an insight into the bacterial changes that occur PM under a clinical setting and acts as a proof-of-principle study into the use of 16S rRNA gene sequencing under these circumstances. The inability to detect

ubiquitous PM bacterial translocation in these animal models suggest this tool may be useful in PM diagnosis of infection.

Optimisation of Nucleic Acid extraction from FFPE tissues

4.1 Introduction to FFPE material

FFPE tissues make up the largest proportion of archived clinical samples worldwide. This method of specimen preservation has been used for over a hundred years due to its ability to fix tissues in their current state and maintain their structural integrity for many years [137]. Histopathology laboratories routinely process FFPE samples for pathological diagnosis as it allows for thin sections of tissues to be cut, mounted on slides, and stained for microscopic examination (overview of process shown in Figure 4.1).

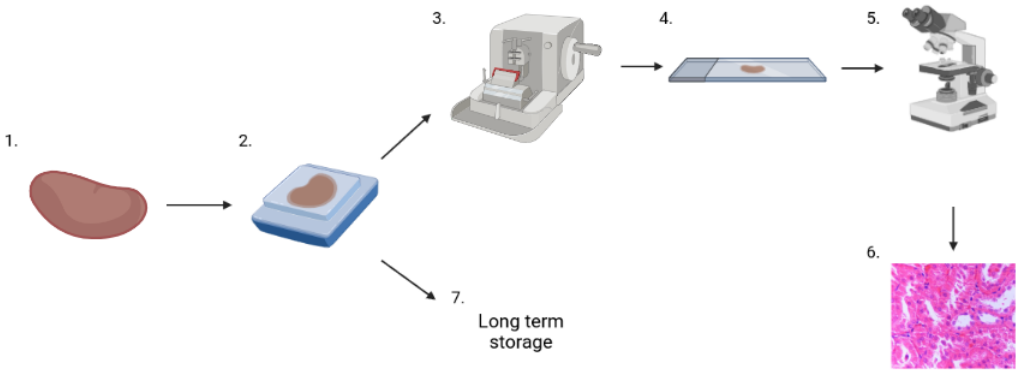

FIGURE 4.1: Laboratory use of FFPE tissues. (1) A tissue sample is collected, (2) fixed with formalin and infiltrated with paraffin which provides support to the tissue. (3) Using a microtome thin sections are cut, (4) mounted onto glass slides and stained so they can be (5) microscopically examined. (6) This allows cell morphology to be examined. (7) The FFPE blocks can then be stored long term.

As these samples are routinely collected and stored for long periods of time, application of molecular methods on these tissues would be ideal. Unfortunately, although FFPE tissues allow for examination of cell and tissue morphology, the processing of the tissue compromises the state of molecular components making their use in downstream molecular techniques difficult [138, 139]. This is due to the adverse effects formalin has on cellular components. On the addition of the formalin solution, the formaldehyde permeates the cell and causes cross-linking between nucleic acids and proteins. The low pH of the formalin solution also damages the quality of nucleic acid by weakening the phosphodiester bonds resulting in random, and often heavy fragmentation of the DNA [139]. Although modifications such as deamination and oxidation of bases can be treated, complete fragmentation of DNA is irreversible. This limits the amount of repair that can be done to this damaged DNA.

Multiple studies have attempted to maximise the quality, quantity, and fragment length of DNA extracted from FFPE tissues [137, 140–144], but all have struggled to consistently recover DNA fragments >300 bps that are suitable for molecular applications such as DNA amplification. With the development of more sophisticated techniques, the interest in using FFPE tissues for molecular diagnostics has grown. Given the vast number of samples stored in institutions worldwide and the ease of both storage and shipping compared to frozen specimens, optimisation of a DNA extraction method to retrieve DNA of sufficient quality to perform genetic investigations would provide researchers with an invaluable resource.

DNA sequencing methods have been successfully applied to FFPE tissues to examine genetic changes occurring in cancer progression. These studies have been performed on samples obtained during tumour resection and applied whole exome sequencing techniques to identify SNPs in cancer driving genes [144, 145], which requires short DNA fragments <150 bps. Although these techniques have been successfully applied to these samples, these tissues are obtained during surgery where samples are fresh and it is likely that they will undergo processing soon after collection to achieve a rapid clinical diagnosis. This is often not the case with

PM tissues where tissues may be at the early stages of decomposition on collection and not processed immediately depending on the demands placed on diagnostic laboratories for surgical diagnostics. In a study performed by Zhi *et al.* DNA was extracted from 65 FFPE PM lung tissues collected at autopsy and a PCR was performed targeting two amplicon lengths of the human β-globin gene. The shorter 110bp amplicon could only be amplified in 15/65 (23%) FFPE PM lung tissue samples, and the longer 250bp amplicon was successfully amplified in 1/65 (1.5%) [146]. The authors suggested that DNA extracted from FFPE PM tissues may be more degraded than DNA from surgical specimens as they are often left in formalin solution for an extended period of time which may result in greater DNA damage [146].

Studying the bacterial 16S rRNA gene in FFPE material has been performed in a few published studies. Two case reports have been published which performed Sanger sequencing of the 16S rRNA gene to identify infection-causing bacteria in PM tissues [147, 148]. Both studies targeted 16S rRNA gene regions of multiple lengths, but neither could obtain amplifiable DNA >300 bps in length. Jobbagy *et al.* amplified five regions of the 16S rRNA gene and performed Sanger sequencing on each amplicon. The sequenced fragments were aligned to a 16S rRNA gene reference to obtain a to obtained a near-full length sequence of the 16S rRNA gene for bacterial identification. Watanabe *et al.* performed 16S Sanger sequencing on PM FFPE liver samples from a sepsis patient. In the investigation three regions of the 16S rRNA gene were targeted (143bp, 510bp, and 1484bp). Only the shortest fragment of 143bp could be successfully amplified for sequencing. Using a combination of the Sanger sequencing results and immunohistochemistry the causative organism was successfully identified.

NGS of the 16S rRNA gene has been performed on FFPE material in two studies. Both of these studies investigated the microbiota present in surgical FFPE intestinal samples [149, 150]. These studies targeted the V5-V6 region (264 bps) and the V4 region (291 bps). These samples were surgical and had a high bacterial load.

To date there are no studies that have successfully amplified regions of the 16S rRNA gene >300 bps from PM tissue. As discussed in Section 1.6.2, the choice of region is important as longer regions contain a higher number of SNPs that allow bacterial members to be identified to a lower taxonomic rank. For example the V3-V4 region has more SNPs than V4 region alone which are unique to bacteria genera and therefore different members of the same bacterial family can be distinguished from one another. For this reason, the method used to extract DNA from FFPE material was optimised in order obtain fragments > 300 bps.

After development of an extraction method was complete, 16S rRNA gene sequencing of both the V3-V4 and V4 region was performed on DNA extracts from frozen material in order to identify differences in the ability to classify bacteria to each taxonomic level using these sequences. Then, as FFPE tissues would be used in lieu of frozen tissues where the latter were unavailable, it was important to establish whether the different processing of the material still would allow detection of potential infection-causing bacteria.

4.2 Chapter aims

Aim To optimise a protocol to extract high quality DNA from FFPE tissue that is of sufficient length for amplification of specified regions the 16S gene.

In order to meet this aim, this study will first compare:

1. Commercially available nucleic acid extraction kits
2. Deparaffinisation methods
3. Starting material

In order to assess the feasibility of DNA sequencing and output sequencing results this study will compare

1. Targetting of the V3-V4 and the V4 region of the 16S rRNA gene
2. The use of FFPE material compared to fresh-frozen material

4.3 Results

4.3.1 Comparison of DNA extraction kits

Three commercially available DNA extraction kits for FFPE material were compared using *Escherichia coli* infected mouse placenta samples (MPL1 and MPL2) that were left over from another study and kindly provided by Dr. Natalie Suff, Institute of Women Health, UCLH. The kits used in the comparison were the QIAamp DNA FFPE Tissue kit (kit Q), the High Pure FFPET DNA Isolation kit (kit R) and the AllPrep DNA/RNA FFPE kit (kit QAP). Frozen mouse placenta samples (both infected and uninfected) were used as controls for DNA quality. To monitor potential contamination during sample processing, wax controls were obtained from each block. These controls consisted of three 10-micron sections of wax containing no tissue sample. The workflow is outlined in Figure 4.2.

The DNA concentration of the extracts from each kit were measured using the Qubit dsDNA HS assay (results shown in Figure 4.3). Negative wax control samples were negative for DNA (Qubit reading of <0.05ng/μL). Kit Q outperformed kit QAP and kit R by extracting the highest concentration of DNA for both samples.

The average fragment length extracted from each sample was estimated using the TapeStation D1000 High Sensitivity (HS) ScreenTape. DNA fragments extracted from both samples using kit Q and kit QAP and sample 1 extracted using kit R were too large to be estimated using this kit suggesting that the extracted DNA had lengths >1500bps. Extraction of sample 2 using kit R showed highly degraded fragments with an average fragment length of 205bps.

To assess whether the DNA extracts were suitable for amplification, DNA extract concentrations were normalised to 2.14ng/μL using nf-H_2O. Extracts were subjected to qPCR using primers targetting a region of the β-actin housekeeping gene spanning 114 base pairs. The lower the CT value (the number of cycles at which

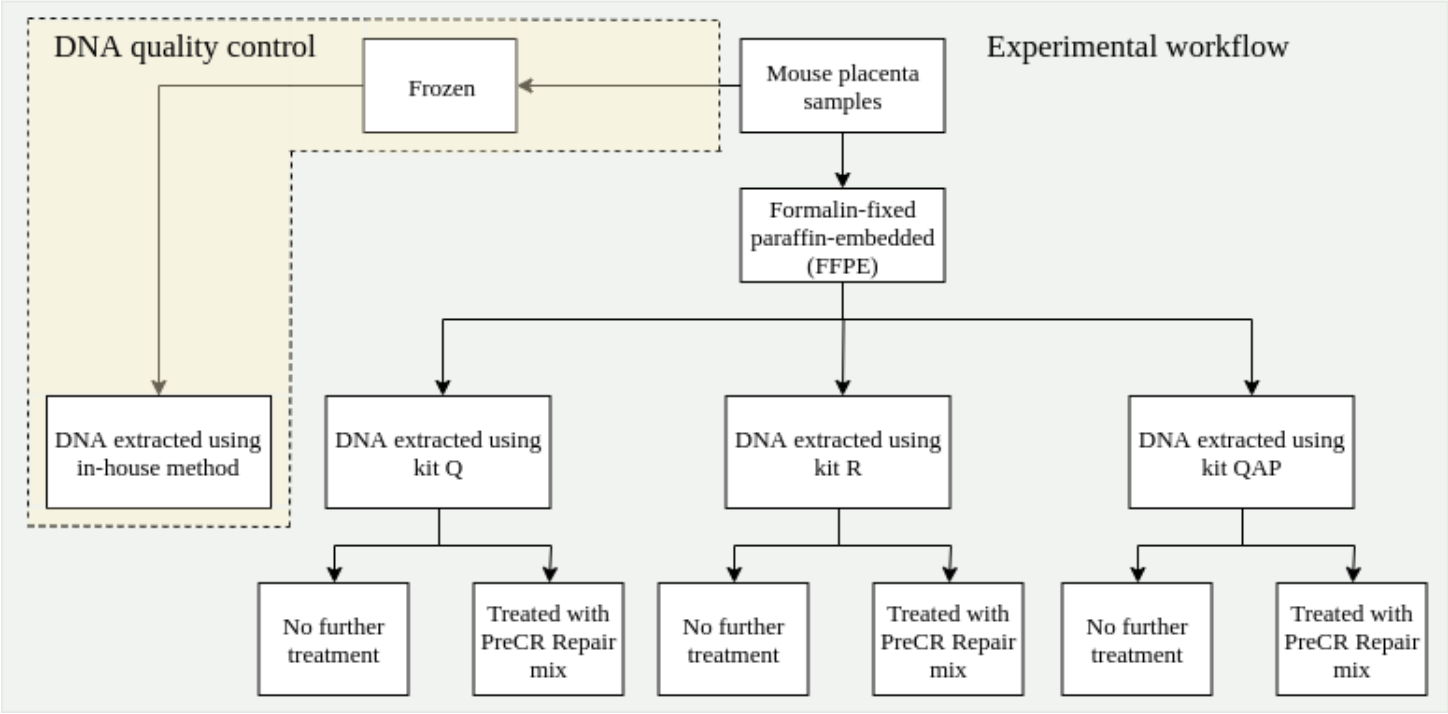

FIGURE 4.2: Workflow for DNA extraction kit comparison

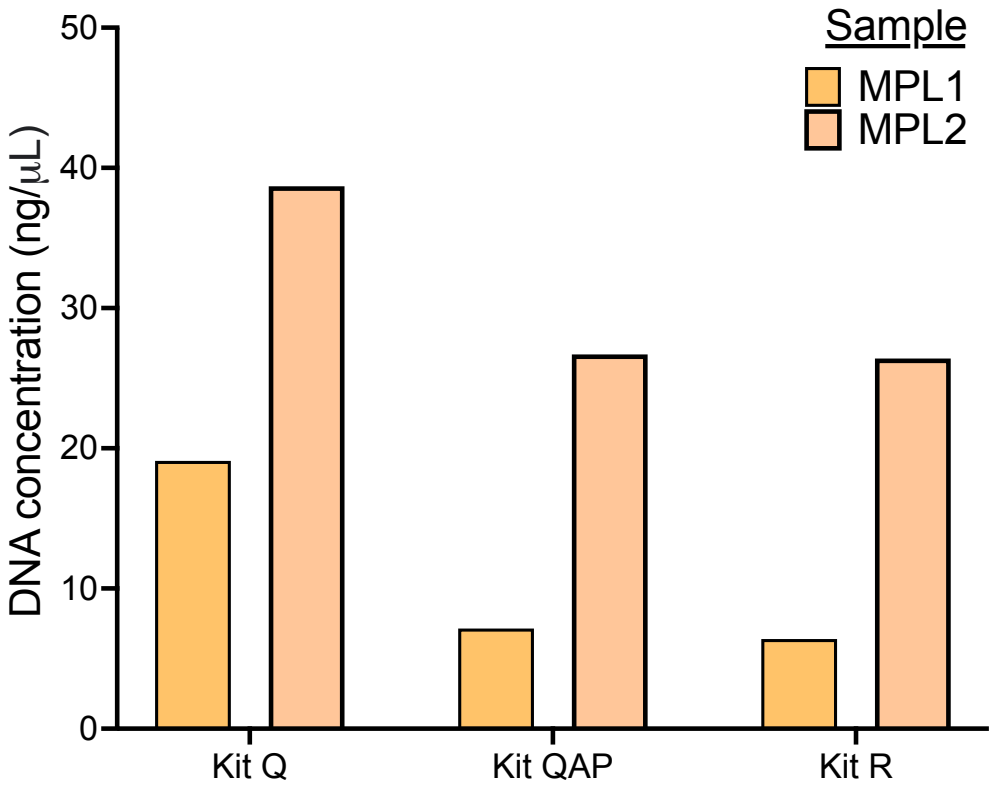

FIGURE 4.3: DNA concentrations extracted from mouse placenta samples (MPL1 and MPL2) with each DNA extraction kit. Q = Qiagen, QAP = Qiagen AllPrep, R = Roche

the fluorescence emitted by the amplicon passes the threshold), the more DNA is amplified. CT values for FFPE DNA extracts along with frozen controls are shown in Figure 4.4. DNA extracts from the frozen mouse placentas showed relatively consistent CT values and crossing the CT between 22 and 24 cycles. DNA extracts from FFPE tissues had significantly higher CT values, crossing the CT between 26 and 39 cycles.

The DNA isolated from the FFPE tissues with each kit were then treated with the PreCR repair mix which contains enzymes that repair damage in the DNA template that would otherwise interfere with PCR. The treated DNA samples were amplified alongside the untreated equivalent. The treated extracts had a CT value between 0.18 and 0.74 cycles lower than the untreated sample meaning there was only slightly more amplifiable DNA. These differences showed no statistical significance.

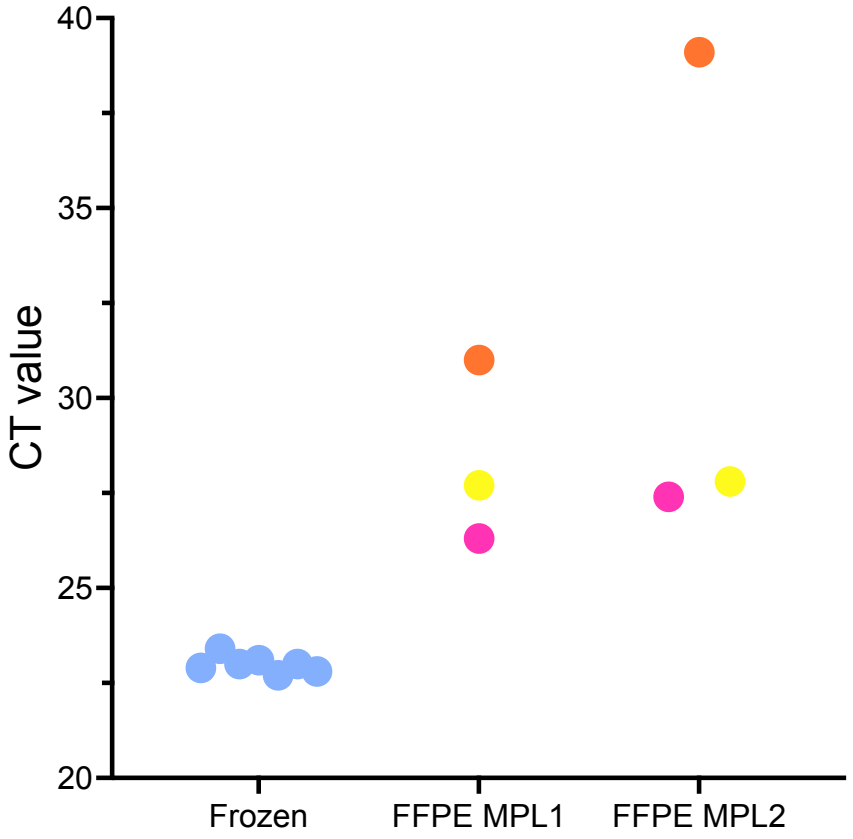

FIGURE 4.4: Cyclic threshold (CT) values of FFPE versus frozen mouse placenta samples when amplifying the housekeeping β-actin gene. Each frozen sample (blue) is collected from a different mouse. The CT values for DNA extracted by each kit is shown in a different colour (pink - kit Q, yellow - kit QAP, and orange- kit R).

A summary of results obtained from the extraction kit comparison is shown in Table 4.1.

Table 4.1: Summary of results from the DNA extraction kit comparison

Criteria			Sample 1			Sample 2		
			Kit Q	Kit R	Kit QAP	Kit Q	Kit R	Kit QAP
DNA Quality	Concentration (ng/μL)	Qubit assay	19.11	7.14	6.4	38.7	26.7	26.4
	Purity (260/280 ratio)	Nanodrop assay	2.13	1.74	2.17	2.10	1.61	2.11
	Fragment length	TapeStation assay	>1500bp	208bp	>1500bp	>1500bp	>1500bp	>1500bp
DNA amplification	qPCR	B-actin gene	Positive	Negative	Positive	Positive	Positive	Positive

4.3.2 Comparison of input material

Further optimisation of the method involved comparing varying amounts of starting material. The selected extraction kit from Section 4.3.1 (kit Q) was used with 6, 8 and 10 10-micron sections of mouse lung tissue samples (ML1, ML2 and ML3). These samples were collected and processed as described in the methods Section 2.4. The concentration, fragment length, and the ability to amplify DNA were investigated. The workflow is summarised in Figure 4.5.

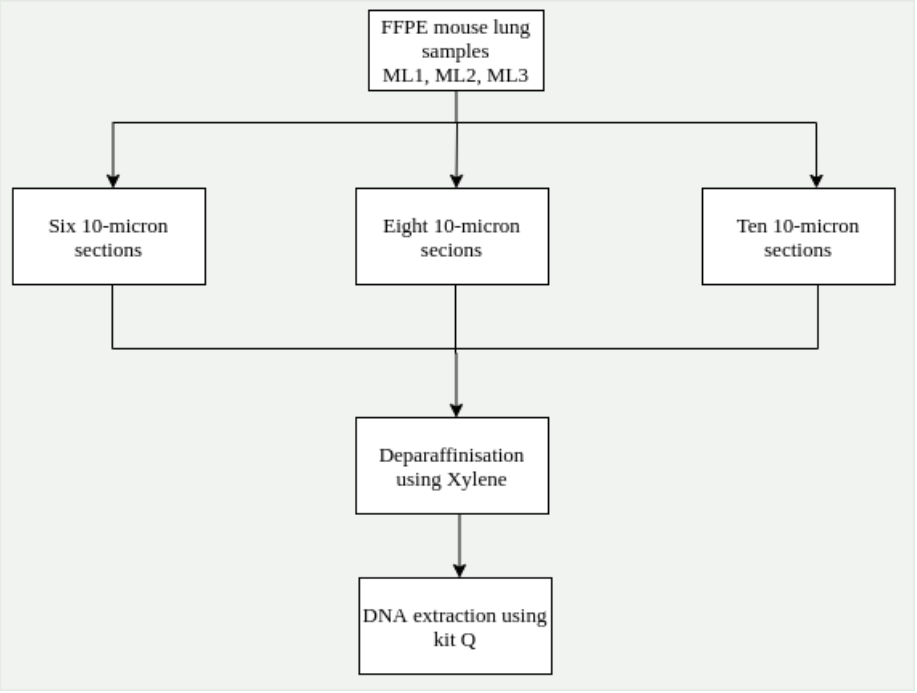

FIGURE 4.5: Workflow for comparison of input material

DNA concentrations of each extract is shown in Figure 4.6. The use of 8 sections yielded the highest concentration of DNA from all three samples. DNA extraction from eight 10-micron sections was higher than DNA extracted from six 10-micron sections, but was not significantly different from the concentration of DNA extracted from ten 10-micron sections. During the deparaffinisation step of ten 10-micron sections, the higher amount of paraffin due to more sections being

used could not be completely dissolved by the xylene resulting in blocking of the DNA extraction column and a reduction in the quality of the final sample.

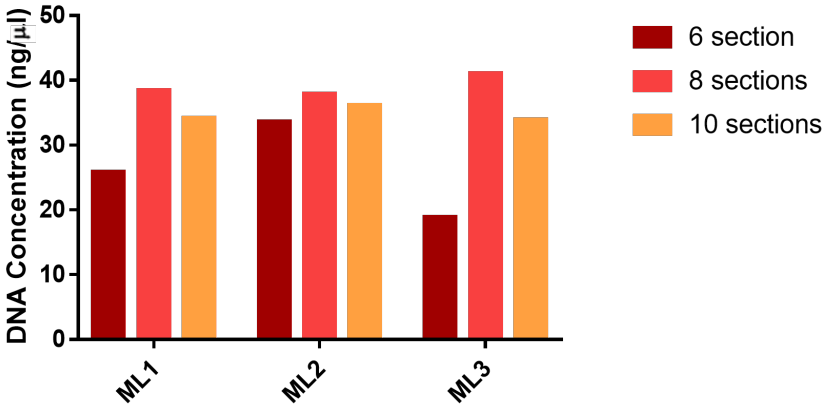

FIGURE 4.6: DNA concentration of extracts using varying input material as measured using the Qubit ($n = 3$)

In order to assess whether the extracted DNA was suitable for amplification, extracts were subject to PCR amplification using β-actin primers designed to amplify multiple product lengths (114 bp, 297 bp, 386bp and 488bp). The PCR product was subject to agarose gel electrophoresis and visualised under UV light. All four product lengths were amplified in each DNA extract (representative sample amplification is shown in Figure 4.7).

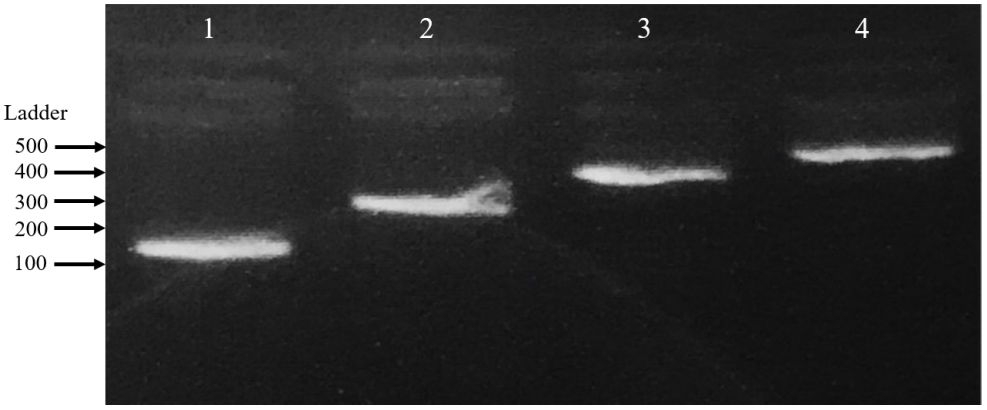

FIGURE 4.7: Gel electrophoresis of β-actin amplicons from FFPE mouse lung DNA extracts. The molecular weight size marker (DNA ladder) is shown on the left hand side of the figure and numbers represent the length in base pairs. Product sizes are as follows (1) 114bp, (2) 297bp, (3) 386bp, and (4) 488bp.

4.3.3 Comparison of deparaffinisation methods

To optimise the nucleic acid extraction further, three methods to deparaffinise the tissue were compared: (1) paraffin removal using xylene followed by the addition of absolute ethanol, (2) a commercially available deparaffinisation solution, and (3) the use of mineral oil as described in a recently published article [105].

Archived clinical FFPE placenta samples from the GOSH histopathology archive were collected for this part of the study ($n = 10$). This study cohort included placenta samples from cases with a histological diagnosis of chorioamnionitis ($n = 5$) and cases with no pathological diagnosis ($n = 5$). In cases with chorioamnionitis, bacteria were expected to be present in the placenta and therefore the 16S gene could be targeted using PCR methods. Archived placenta samples were selected as they provided a more realistic representation of routinely processed samples, as opposed to tissue samples described in Section 4.3.2 where tissues were embedded specifically with the purpose of optimising nucleic acid extraction.

The human placenta samples were larger than the mouse lung samples used previously and were embedded into larger wax blocks, giving a lower tissue to wax ratio. For this reason, the decision was made to use one and three 10-micron sections to allow for efficient paraffin removal and extraction. The different deparaffinisation techniques were used to remove the wax from the tissues prior to DNA extraction using kit Q. As anticipated, the use of three 10-micron sections provided too much tissue to be successfully digested using kit Q. Therefore one 10-micron section was selected for downstream application.

DNA concentration and fragment length were measured and the DNA was subject to amplification of the human β-actin gene. The estimated DNA concentrations for each deparaffinisation method are shown in Figure 4.8. The xylene and deparaffinisation solution were more consistent than the mineral oil method shown by the interquartile range. Although the mineral oil method had the highest average concentration of DNA extracted, it also had the largest range and extracted $<$1.0ng/μL in 3 out of the 10 samples investigated. The maximum concentration

the deparaffinisation method extracted was 26.5ng/μL, which was the lowest of all methods. Using these results, the traditional xylene method of deparaffinisation was selected.

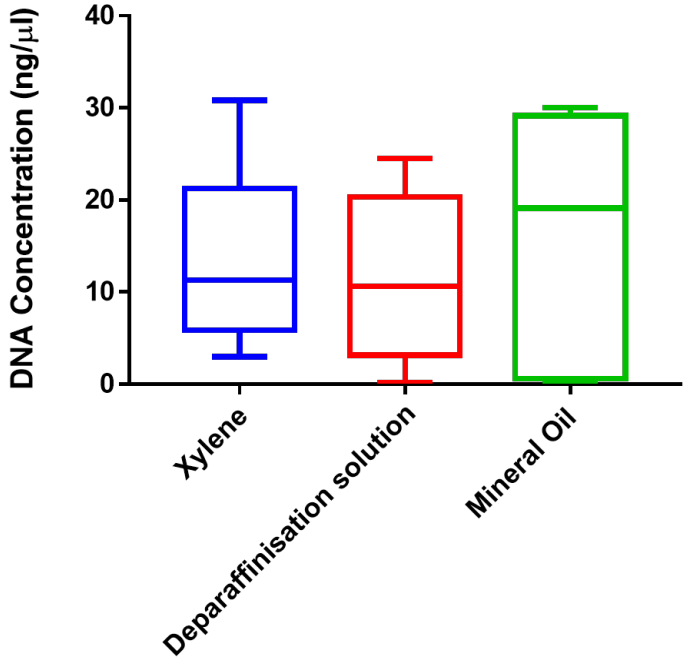

FIGURE 4.8: Double stranded DNA concentration measured using the Qubit extracted from FFPE human placenta samples (n = 10) using different deparaffinisation methods

Both the β-actin and 16S genes in the human placenta samples were targeted for amplification as a test of DNA quality. B-actin primers that amplified four different amplicon lengths were designed to provide an indication of the length of DNA that was suitable for amplification. Both the V4 region (product length of 292bps) and the V3-V4 region (product length of 466bps) of the 16S gene were also targeted to determine which fragment size was suitable for amplification. Amplified products were visualised on an agarose gel and results are summarised in Table 4.2.

The 440bp product was successfully amplified in 8 of the 10 human placenta DNA extracts. In one placenta sample, only the 114bp product was successfully amplified, but both the V4 and V3-V4 region were amplified successfully.

Table 4.2: Human placenta DNA isolate amplification success

Sample name	B-actin product length				16S rRNA gene region	
	114bp	297bp	385bp	440bp	V4	V3-V4
HPL1	✓	✓	✓	✓	✓	✓
HPL2	✓	✓	✓	✓	✓	✓
HPL3	✓	✓	✓	✓	✓	✓
HPL4	✓	✓	✓	✓	✓	✓
HPL5	✓	✓	✗	✗	✓	✓
HPL6	✓	✓	✓	✓	✓	✓
HPL7	✓	✓	✓	✓	✓	✓
HPL8	✓	✗	✗	✗	✓	✓
HPL9	✓	✓	✓	✓	✓	✓
HPL10	✓	✓	✓	✓	✓	✓

4.3.4 Comparison of the V3-V4 region and the V4 region of the 16S rRNA gene

As discussed in Section 1.6.2, the choice of target region of the 16S rRNA gene is important and can often introduce different results depending on which region is selected. The effects of FFPE processing and the subsequent damage to the DNA may limit the region choice for the investigation of bacterial species within these tissues. For this reason and as proof of principle, the V3-V4 region and the V4 region were amplified and sequenced in frozen SUDIC tissues as both regions should be readily amplifiable in these tissues.

Archived SUDIC samples that had undergone comprehensive autopsy at GOSH were identified using the GOSH PM database and used for this part of the study ($n = 12$). The study cohort included samples collected from six patients. These six SUDIC patients included two explained cases (eCs), two infectious cases (iCs) and two unexplained cases (uCs). Tissue samples included heart ($n = 5$), kidney ($n = 2$), liver ($n = 3$) and spleen ($n = 2$) tissue. Both the V3-V4 and the V4 region of the 16S rRNA gene were amplified and sequenced in all samples. Five negative extraction controls and two blank wax controls were sequenced to control for potential contamination and false-positive sequence generation. These negative extraction controls generated an average of 36 (range 0 - 149) and 663 (range 1 - 2412) reads on the V3-V4 and V4 sequencing run respectively, compared to 19,861 and 32,979 reads for the V3-V4 and V4 regions amplified from the positive control (DNA standard of known composition).

The V3-V4 region sequencing run generated an average of 605 reads per sample (range 2 - 3844) and the V4 sequencing run generated an average of 1,297 reads per sample (range 6 - 7,330). More sequencing reads were generated from the V4 sequencing run (Figure 4.9).

A total of four samples did not pass quality filtering (three did not pass on either the V4 or V3-V4 sequencing run and one did not pass on the V3-V4 sequencing

run only). These samples were therefore excluded from comparative abundance analysis. These samples included a spleen and heart sample from uC5, and a spleen and liver sample from eC4.

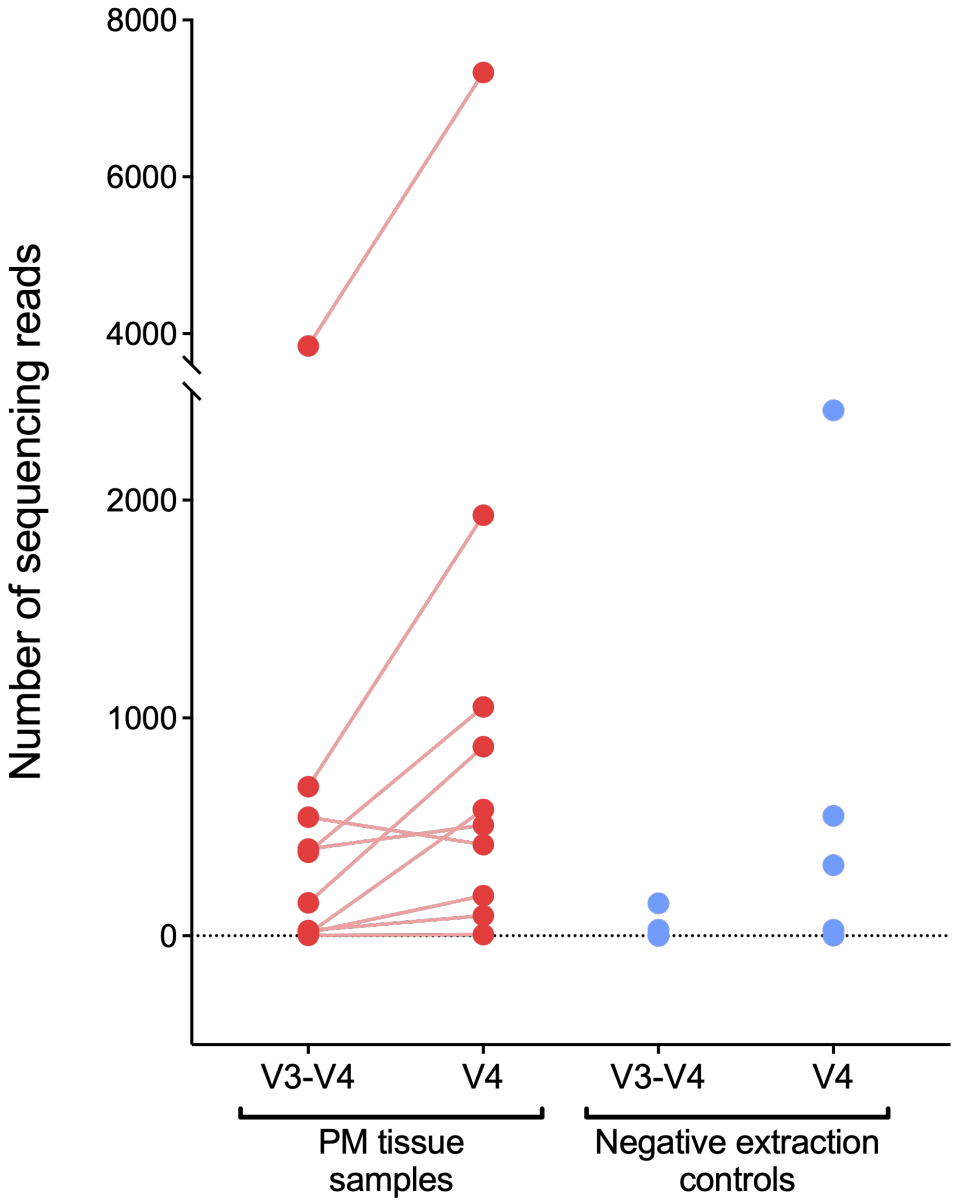

FIGURE 4.9: Total number of sequencing reads in each PM tissue sample when sequencing the V3-V4 and V4 region of the 16S rRNA gene. Red dots represent PM tissue samples. Each sample pair is joined with a line. Blue dots represent negative DNA extraction controls.

Comparison of bacterial members identified at family level classification within the matched samples is shown in Figure 4.10 A-F. In each tissue sample, whether the V3-V4 or the V4 region was sequenced, the same top bacterial families were identified. When assessing the classification at the lower taxonomic rank of genera, more bacterial members were unable to be classified. In particular, the V4 region consistently failed to classify bacteria in the *Escherichia*/*Shigella* genera, which were identified using the V3-V4 region of the 16S rRNA gene. Therefore, family level classification gave more comparable results between the two tested regions.

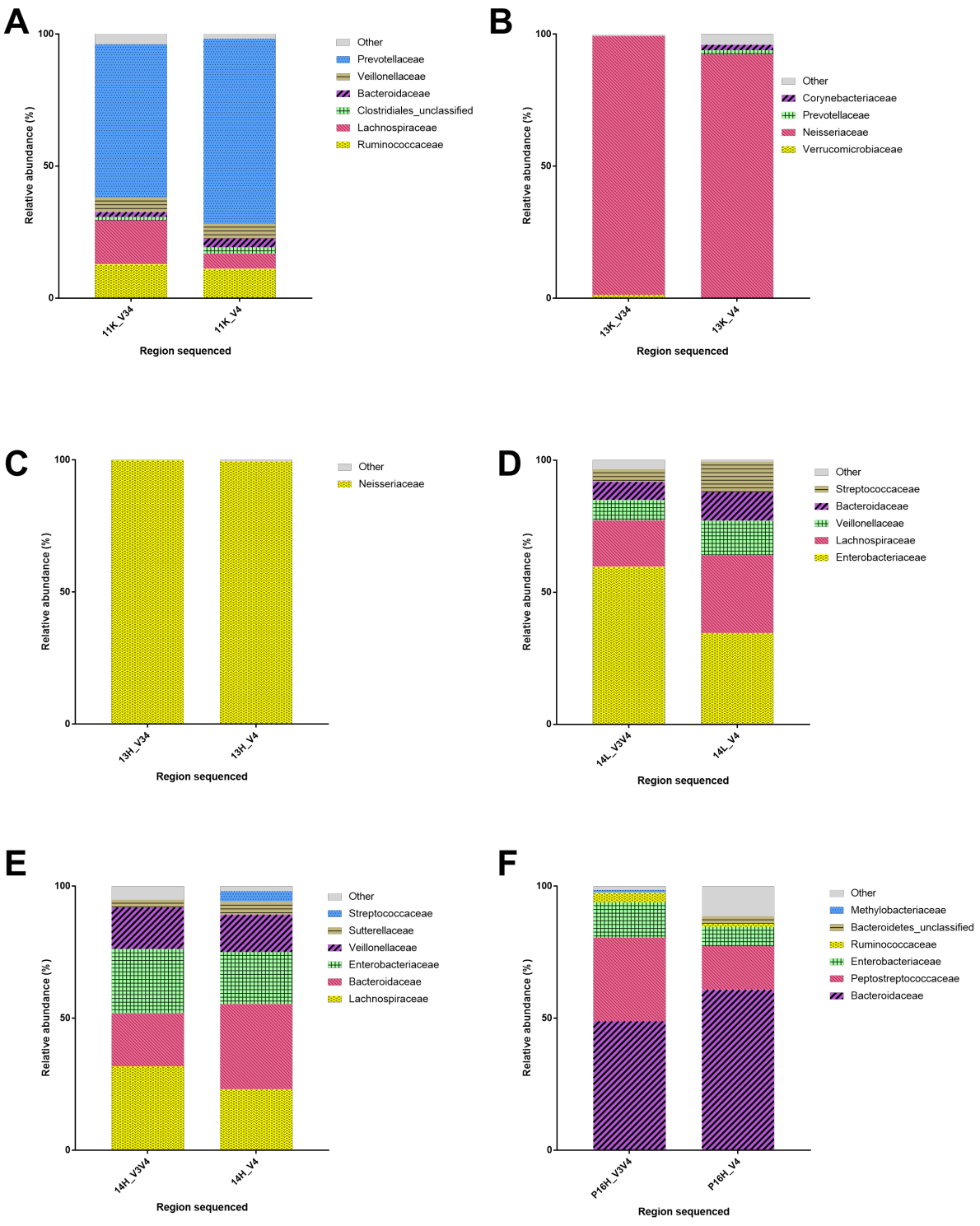

FIGURE 4.10: Relative abundance plots of bacterial families identified when sequencing the V3-V4 and the V4 region in matched frozen tissue samples as follows (A) eC3 kidney, (B) iC2 kidney, (C) iC2 heart, (D) iC8 liver, (E) iC8 heart, (F) eC4 heart. In each plot the first bar shows results from sequencing the V3-V4 region and the second bar shows results from sequencing the V4 region.

4.3.5 Comparison of FFPE material and fresh frozen material

After investigating the effects that the use of the shorter V4 region had on bacterial identification compared to the longer V3-V4 region, the next stage was to investigate whether tissue processing would affect bacterial identification.

Fifteen matched FFPE and fresh-frozen tissue samples (a total of 30 samples) were obtained for this part of the study. These samples were collected from 11 cases of SUDIC; four unexplained cases (uCs), four explained cases (eCs) and three infectious cases (iCs). The samples included 11 heart samples, five kidney samples, seven liver samples and six spleen samples. When amplified, 14 (47%) samples were considered positive for the 16S gene and were subsequently sequenced.

After sequence processing, the negative extraction controls were analysed to track any potential contamination. Negative extraction controls processed alongside frozen material had an average of 720 sequencing reads (range 0 - 2486). FFPE negative extraction controls had an average of 5,058 reads (range 2215 - 7821). Burkholderia was identified in all FFPE negative extraction controls at relative abundances between 61% and 99% as shown in Figure 4.11. Given their persistent presence and higher relative abundance in all FFPE negative extraction controls these bacteria were considered contaminants and therefore removed from all FFPE patient samples.

Following sequence processing and quality filtering, 24 tissue samples (12 pairs) were used for downstream analysis. The 12 positive tissue samples pairs included six heart samples, two liver samples, two spleen samples, and two kidney samples.

In 10/12 (83%) of the tissue sample pairs there was little concordance between the frozen and the FFPE tissues (representative plots in Figure 4.12). Of these ten tissue samples, eight were from patients whose final diagnosis was either explained or unexplained. The explained cases included final cause of deaths such as asthma, head injury, and cardiac arrest. The remaining two samples were a heart and liver

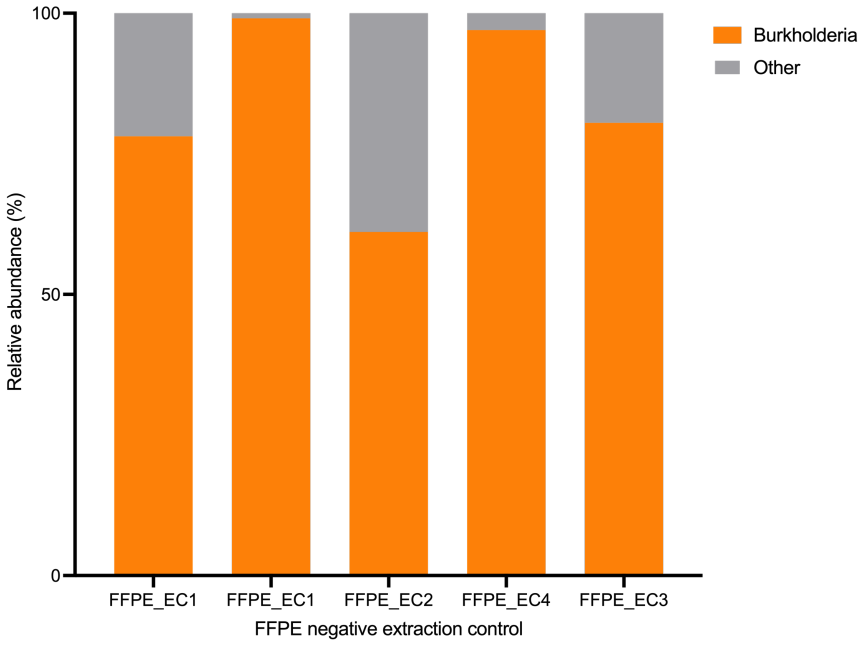

FIGURE 4.11: Relative abundance plots of bacterial families identified in FFPE negative extraction controls when sequencing the V4 region of the 16S rRNA gene. Burkholderia were identified in all five negative extraction controls at >60%

sample from a patient whose final cause of death was deemed to be a lung infection. In this patient lung samples were not obtained for this comparative investigation. Bacteria observed in these tissues included Bacteroidaceae, Enterobacteriaceae, Lachnospiraceae, Streptococcaaceae, Sutterellaceae, and Veillonellaceae all present at <35%.

The remaining two sample pairs were collected from a patient who had a known infection. The tissue samples showed concordance in the bacteria identified in the frozen and FFPE material. The relative abundance plots of the matched samples from the heart and kidney are shown in Figure 4.13A and Figure 4.13B. In both samples the same dominant bacteria was found in the FFPE and frozen samples at relative abundances >70%. Other bacteria identified in these samples included Bacillaceae, Corynebacteriaceae and Prevotellaceae in frozen and Bacillaceae, Moraxellaceae, Pseudomonadaceae, Rhodobacteraceae and Sphingobacteriaceae, but all were at relative abundances <25%.

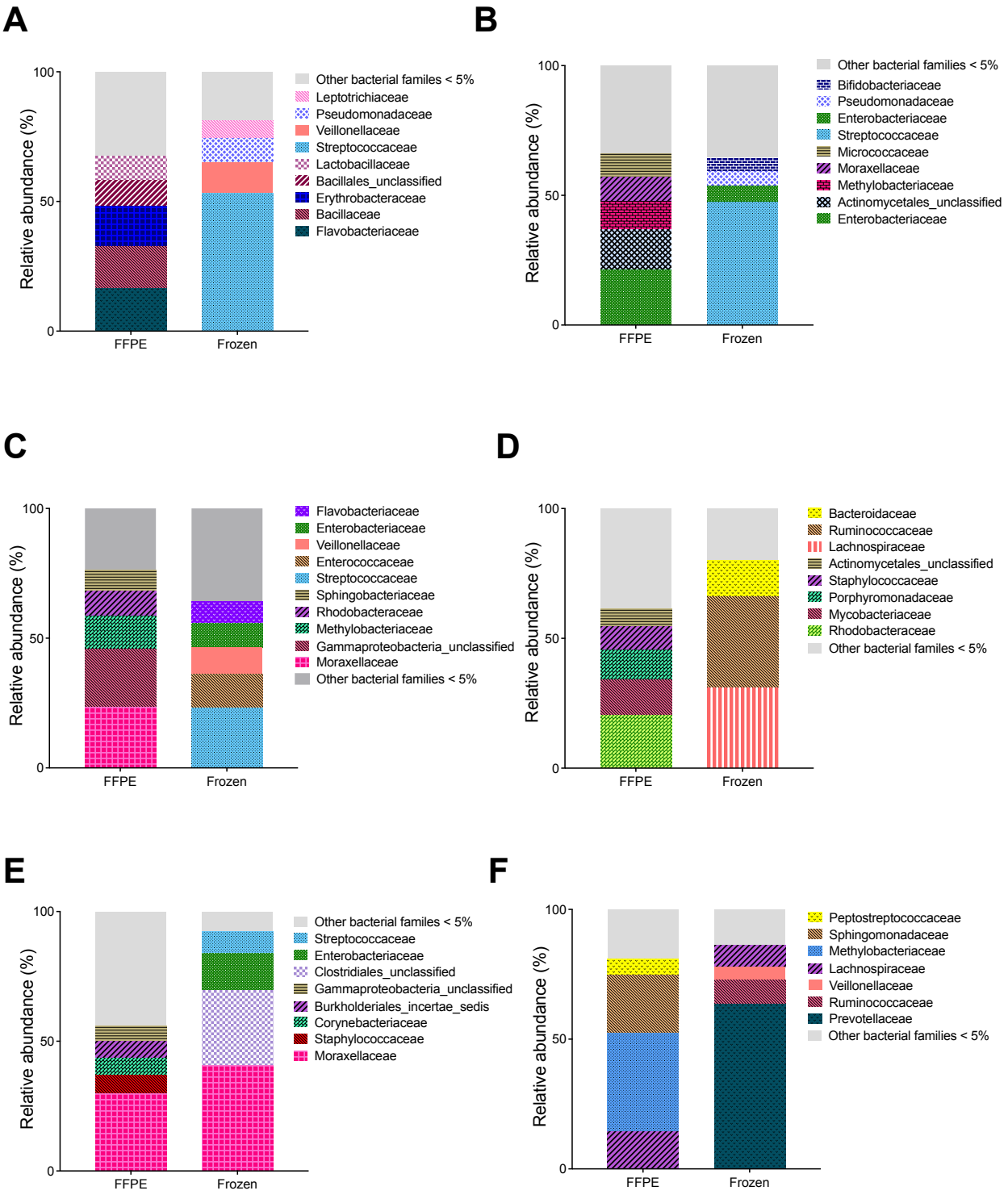

FIGURE 4.12: Relative abundance plots for matched FFPE and fresh-frozen material obtained from SUDIC patients whose death was (A and B) unexplained, (C and D) explained but non-infectious, (E and F) and due to a lung infection. The coloured bars represent different bacterial families. The x-axis describes the tissue processing.

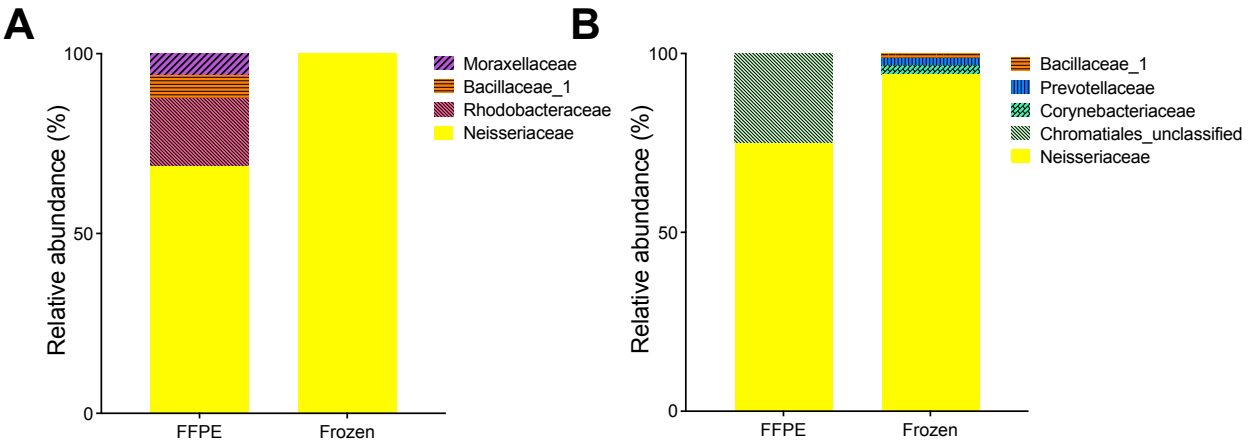

FIGURE 4.13: Relative abundance plots for matched FFPE and fresh-frozen material obtained from a patient whose final cause of death was septicaemia. Tissues obtained were (A) heart and (B) kidney and the x-axis describes the tissue processing

4.4 Discussion

In order to perform 16S rRNA gene sequencing on archived FFPE material, at least 292bp DNA fragments had to be obtained in order to amplify the V4 region of the 16S rRNA gene. In order to perform a deeper classification, the V3-V4 region which spans approximately 460 bps. For this reason the DNA extraction method from FFPE material was optimised.

Following acquisition of sequencing results, the effects of tissue processing had to be examined for result interpretation. This study therefore firstly focused on stages of DNA extraction, and secondly on the sequence processing and data interpretation.

4.4.1 Optimisation of DNA extraction

The highest quality, quantity, and fragment length of DNA was obtained using the traditional xylene method for deparaffinisation followed by extraction using the Qiagen QIAamp DNA FFPE Tissue kit (kit Q). Depending on the tissue type, between three and eight 10-micron sections gave the highest quality DNA by reducing the tissue to wax ratio. The DNA recovered from FFPE tissues was shown to have significantly lower quality and shorter fragment lengths than frozen tissue as expected. However, the combination of techniques optimised in this study still allowed the extraction of DNA from FFPE material that was suitable for amplification of the V3-V4 region of the 16S rRNA gene which spans approximately 460 bps.

DNA extracted from FFPE mouse placenta samples was shown to have fragment lengths >1500 bps demonstrated by the TapeStation, which if truly present should allow amplification of the desired region. However, the TapeStation measures lengths based on gel electrophoresis meaning this finding may be a result of DNA modifications. Cross-linking in the DNA resulting in secondary structure formation would cause the DNA to would migrate through the gel at a slower rate.

This slower migration would be misinterpreted as a higher molecular weight and therefore the fragment length may have been overestimated.

The PreCR repair mix treats DNA damage including abasic sites, thymidine dimers, oxidized guanine and pyrimidines, and deaminated cytosine, but cannot target fragmentation or protein-DNA cross-links. The use of this repair treatment on the FFPE tissues did not significantly increase the amount of amplifiable DNA. Together with the overestimation of fragment length using the TapeStation, these results suggest that cross-linkage is the major source of DNA modification preventing DNA amplification.

The investigation into input material demonstrated the importance of limiting the amount of paraffin obtained to ensure the full removal of paraffin in and around the tissue. The remainder of excess paraffin inhibits later stages of the DNA extraction. Although the use of eight 10-micron sections of mouse lung samples was found to provide the highest DNA concentration, the amount of tissue embedded into the paraffin block was relatively small compared to routinely collected clinical samples, such as the ones that will be used at later stages of this thesis. This will be considered when performing DNA extraction from FFPE SUDIC samples in subsequent chapters.

Amplification of varying lengths of the housekeeping β-actin gene allowed for the assessment of, not only attainable, but amplifiable lengths of DNA. The use of multiple amplicon lengths provided insight into the cut-off length of extracted DNA fragments. In a majority of clinical specimens all amplicon lengths studied were successfully amplified. In all of these samples both the V3-V4 and the V4 regions of the 16S rRNA gene could be amplified. Interestingly, in the two samples where the 385 bp β-actin product could not be amplified, the 16S rRNA gene was detected. This could be due to differing rates of degradation of genetic material between host and bacterial DNA. Once a tissue dies, autolysis commences which can result in DNA degradation and fragmentation [151, 152]. Bacteria present within, or on the surface of this tissue would continue to thrive, rendering their

DNA is more stable and protected. Alternatively, this may be due to contamination at sample collection or tissue processing. This would represent a false-positive result. From these results we can deduce that if the β-actin gene is amplifiable but the 16S rRNA gene is not, then the tissue is likely negative for bacteria. However, if the β-actin gene is not amplifiable, this does not mean the 16S gene is not obtainable. Therefore, although this targeted PCR can be used as a measure of quality host DNA it cannot be used to determine if a DNA extract is of sufficient quality to amplify the 16S rRNA gene.

It is well documented that the choice of 16S gene region affects bacterial identification and can often skew results [153–155]. The importance of primer selection for different communities and sample types requires careful consideration [156, 157]. Results from this study demonstrated that although the level of taxonomic classification may have a lower resolution in the V4 region compared to the V3-V4 region, the main bacterial families identified were the same using both regions. The V3-V4 region was better able to classify members of the community to Genus level than the V4 region, as has been shown previously [100, 158]. The inability of sequences of the V4 region to identify *Escherichia/Shigella* members is consistent with other studies that have shown that sequences belonging to the Enterobacteriaceae family are unable to be distinguished using this region alone [158]. Using the V4 region, although there may be difficulty in identifying bacteria to the Genus level, for the purpose of identifying whether a bacterium is colonising the tissue at a high relative abundance and therefore causing an infection this region should be sufficient in cases where the V3-V4 region cannot be sequenced.

Comparison of tissue processing on retrievable bacterial information from PM tissues shows that for low biomass material, the same bacteria are not identified from matched FFPE and frozen tissues for a majority of samples tested. These tissue samples were all collected from patients were the cause of death was not identified as infectious and therefore bacterial presence was not expected. The extremely low bacterial biomass in these tissues could have resulted in non-specific amplification during PCR or perhaps a higher rate of contamination at sample collection. In

the two cases where there was a known infection and therefore an overwhelming dominance, the causative bacteria was identified in both frozen and FFPE tissues by 16S gene sequencing. Although this sample size is too small to draw any solid conclusions, these observations are consistent with another study which found no significant difference between communities identified when comparing FFPE to frozen material in gastric samples with high bacterial biomass [150].

Higher rates of contamination were observed in the FFPE DNA extraction than the frozen DNA extraction, particularly with the detection of Burkholderiaceae in all negative controls which were introduced at the stage of DNA extraction. The identification of bacteria in these samples is likely part of the 'kitome' of the FFPE DNA extraction kit, a phenomena that has been discussed in the literature [159–161]. Such 'kitome' contaminants have been show to have particular impact on low bacterial biomass samples such as the ones used in the present study [161]. This finding highlights the importance of negative extraction controls to help account for this potential contamination and non-specific amplification.

This investigation has demonstrated that using the optimised methods both the fresh-frozen material and the FFPE material can be used for the intended purpose. Although the results differ in some cases, the method allows detection of overwhelming bacterial colonisation with a single bacterial species potentially representing an infection.

Bacterial presence in Sudden Unexpected Death in Infancy

5.1 Introduction

As discussed in Section 1.5, there is mounting evidence to suggest that infection is the cause of death for a subset of SUDIC that remain undiagnosed due to limitations in current diagnostic techniques. The development of an optimised protocol to perform 16S rRNA gene sequencing on both frozen and FFPE PM material from cases of SUDIC in Chapter 4 presents the opportunity to investigate the bacterial presence within these tissue to open up the possibility for retrospective diagnosis.

At autopsy, all major organs are sampled and undergo a range of clinical tests in order to identify the CoD. Extensive sampling includes lung, heart, thymus, pancreas, liver, spleen, lymph nodes, adrenal glands, muscle, and brain [162]. These samples are all formalin-fixed and paraffin embedded for histological analysis. As FFPE material is stable at room temperature, where permission is granted, these samples enter long-term storage. For investigation of lipid storage diseases, tissue samples are frozen and stained using Oil Red O dye. At GOSH, this is performed on 'composite' samples which include heart, liver, spleen, kidney, and muscle samples. Where permission for storage is granted by families, these samples are stored long-term at -80°C. The GOSH microbiology department performed primary and subculture on lung and spleen samples in all cases, with additional investigation of cerebrospinal fluid (CSF), faecal material, and lesions where requested by the pathologist. After testing, these samples are discarded.

Therefore for the purpose of the investigation, frozen heart, kidney, liver, muscle, and spleen samples were available. In cases where FFPE material was obtained, heart, kidney, liver, spleen, and lung samples were also obtained. The location of these samples within the body can be found in Figure 5.1.

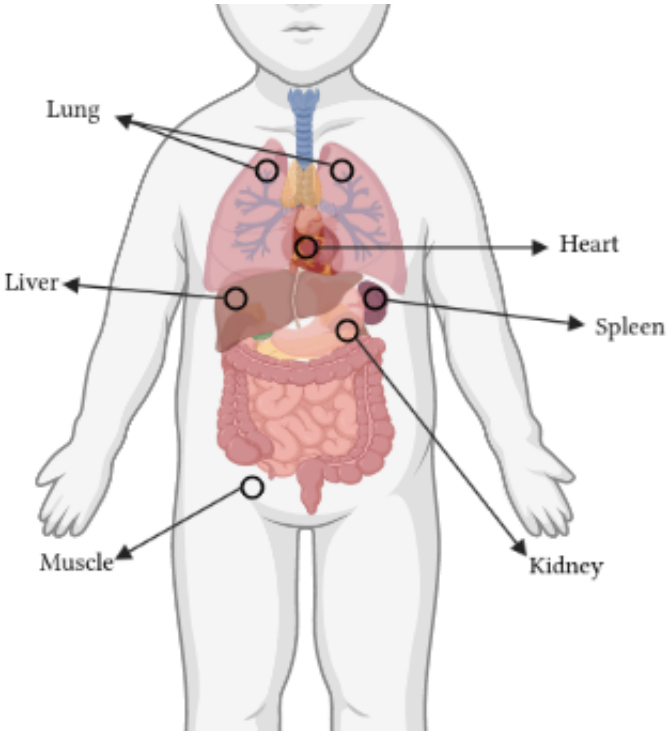

FIGURE 5.1: Location of tissue samples obtained during PM examination of SUDIC. Samples obtained from organs shown on this diagram were archived for SUDIC cases and these samples were available for this study. The muscle sample is collected from the psoas muscle.

5.2 Chapter Aims

The aims for this chapter are summarised below in Figure 5.2.

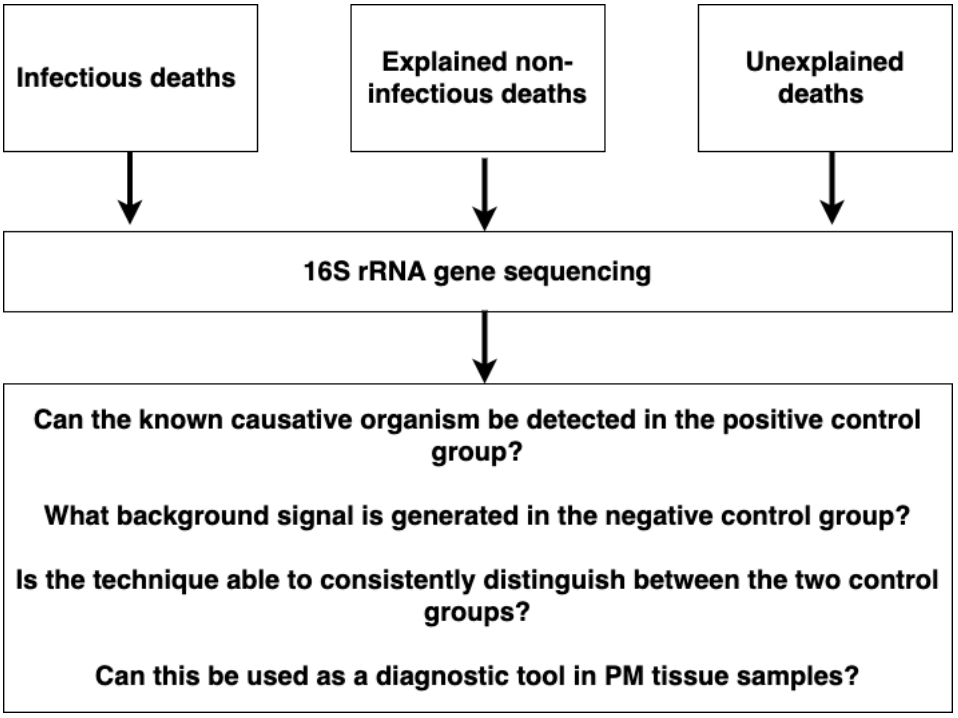

FIGURE 5.2: Flowchart of the aims for Chapter 5

5.3 Results

Forty-three cases of SUDIC that had undergone extensive autopsy at GOSH between 2014 and 2018 were used for this study. All cases held consent for scientific research. Frozen archival heart, kidney, liver, muscle, and spleen samples were obtained from each case where available with a total of 187 frozen tissue samples being collected. Archival FFPE material was collected from 10 cases with a total of 38 FFPE PM samples. Each case was categorised into one of the following groups depending on the final CoD as stated on the PM report:

1. Explained non-infectious- the CoD was determined but was not infectious
2. Infectious- the CoD was infectious
3. Unexplained- the CoD could not be identified following investigation

Information on samples collected from each case in the explained non-infectious and infectious groups can be found in Table 5.1 and those in the unexplained group can be found in Table 5.2.

TABLE 5.1: The final cause of death each for explained non-infectious cases (eC) and infectious cases (iC) of SUDIC included in this study along with the frozen and/or FFPE tissue samples obtained for analysis. Tissue samples are as follows H = heart, K = kidney, Li = liver, Lu = lung, S = spleen, T = tumour, M = muscle

	Case	Final CoD	Causative organism	Frozen tissues	FFPE tissues
Explained, non-infectious					
8	eC1	Allergic inflammation due to asthma		H, K, Li, M, S	H, Li, S, Lu
9	eC2	Severe head injury		H, K, Li, M, S	H, K, S, Lu
11	eC3	Allergic inflammation due to asthma		H, K, Li, M, S	H, K, Li, S, Lu
16	eC4	Post-cardiac arrest syndrome		H, K, Li, M, S	
21	eC5	Congenital heart disease		H, K, Li, M, S	
24	eC6	Myopathy		H, K	
38	eC7	Myocarditis		H, K, Li, M, S	
43	eC8	Metabolic disorder		H, K	
45	eC9	Cardiomyopathy		H, K, Li, M, S	
53	eC10	Metastatic tumour		Lu, S, *T1, T2	
56	eC11	Allergic inflammation due to asthma		H, K, Li, M, S	
57	eC12	Aortic stenosis		H, K, Li, M, S	
58	eC13	Hypoxic injury to brain		S	
59	eC14	Severe head injury		H, K, Li, M, S	
4	eC15	Rheumatic disease			H, Li, S, Lu
Infectious					
12	iC1	Septicaemia	Group A Streptococcus	S	H, K, Lu
13	iC2	Septicaemia	*Neisseria meningitidis*	H, K, Li, M, S	H, K, Lu
37	iC3	Septicaemia	*Neisseria meningitidis*	H, K, Li, M, S	
50	iC4	Septicaemia	Unidentified pathogen	H, K, Li, M, S	
55	iC5	Septicaemia	*Klebsiella pneumoniae*	H, K, Li, M, S	
5	iC6	Bronchopneumonia	Unidentified pathogen	H, K, Li, M	
6	iC7	Bronchopneumonia	Viral origin	H, K, Li, M, S	
14	iC8	Bronchopneumonia	Unidentified pathogen	H, K, Li, M	H, Li, Lu

TABLE 5.2: The final cause of death for each unexplained case (uC) of SUDIC included in this study along with the frozen and/or FFPE tissue samples obtained for analysis. Tissue samples are as follows H = heart, K = kidney, Li = liver, Lu = lung, S = spleen, M = muscle

	Case	**Final CoD**	**Frozen tissues**	**FFPE tissues**
1	uC1	Unexplained	H, K, Li, M, S	H, Li, S, Lu
2	uC2	Unexplained	K, Li, M, S	
3	uC3	Unexplained	H, K, Li, S	
7	uC4	Unexplained	H, K, Li, M	
15	uC5	Unexplained	H, K, Li, M, S	H, Li, S, Lu
17	uC6	Unexplained	H, K, Li, M, S	
19	uC7	Unexplained	H, K, Li, M, S	
20	uC8	Unexplained	H, K, Li, M,	H, Li, S, Lu
18	uC9	Unexplained	H, K, Li, M, S	
22	uC10	Unexplained	H, K, Li, M, S	
23	uC11	Unexplained	H, K, Li, M, S	
35	uC12	Unexplained	H, K, Li, S	
36	uC13	Unexplained	H, K, Li, M, S	
40	uC14	Unexplained	H, K, Li, M	
41	uC15	Unexplained	H, K, Li, M, S	
42	uC16	Unexplained	H, K, Li, M, S	
47	uC17	Unexplained	H, K, Li, M, S	
48	uC18	Unexplained	H, K, Li, M, S	
49	uC19	Unexplained	H, K, Li, M, S	
51	uC20	Unexplained	H, K, Li, M, S	

Of the 187 frozen PM tissue samples, 66 (35%) were positive for the bacterial 16S rRNA gene and were subsequently sequenced. A breakdown of total positive frozen samples within each CoD group can be found in Table 5.3. In both the infant and childhood cohorts there was a significantly higher percentage of positive samples in the infectious cohort than in the explained non-infectious cohort (68.4% vs. 11.8%, P = 0.0008 ; 73.3% vs. 31%, P =0.0063 respectively). A greater percentage of positive samples were also observed in the unexplained cohort compared to the explained non-infectious cohort in both the infant and childhood cases (34.7% vs. 11.8%, P = 0.0814 ; 72.7% vs. 31%, P = 0.0031) but these differences only reached significance in the childhood cohort.

Positive frozen PM tissue samples generated an average of 14,149 sequencing reads (range 25 - 190,027). Sequencing reads amongst positive samples were compared between CoD groups to identify whether a higher number of sequencing reads were associated with a particular group. Tissue samples with a higher number of sequencing reads were associated with infectious and unexplained groups. The average reads numbers for explained non-infectious, infectious, and unexplained deaths were 1,777, 12,026, and 16,695 respectively. No obvious patterns were observed in the number of sequencing reads with the PMI, but there were not enough samples to perform statistical analysis.

In order to identify whether 16S rRNA gene sequencing method can successfully distinguish between infected and non-infected tissues.

5.3.1 Sequencing PM tissues from explained non-infectious SUDIC

This study examined a total of 15 cases of explained non-infectious SUDIC (total of 58 frozen PM tissue samples and 17 FFPE samples). The cause of death included asthma, severe head injury, and post-cardiac arrest syndrome. Of the 58 frozen tissue samples collected, 11 (26%) were positive for the 16S rRNA gene. These 11 tissue samples were from seven cases of explained non-infectious deaths. The

Table 5.3: Total number of SUDIC samples that were positive for the 16S rRNA gene. Group comparison was performed using a Fishers extract test. A P-value <0.05 was considered to be significant

Group	Total samples	Positive	Negative	Group comparison
Infant				
(A) Explained	17	2 (11.8%)	15 (88.2%)	A vs. B: $P = 0.0008$ (***)
(B) Infectious	19	13 (68.4%)	6 (31.6%)	A vs. C: $P = 0.0814$ (ns)
(C) Unexplained	72	25 (34.7%)	47 (65.3%)	B vs. C: $P = 0.0101$ (*)
Child				
(A) Explained	42	13 (31%)	29 (69%)	A vs. B: $P = 0.0063$ (**)
(B) Infectious	15	11 (73.3%)	4 (26.7%)	A vs. C: $P = 0.0031$ (**)
(C) Unexplained	22	16 (72.7%)	6 (27.3%)	B vs. C: $P = >0.9999$ (ns)

average number of sequencing reads in the positive samples was 1,177 (range 44 - 8,045). Of the 17 FFPE samples collected, 6 (35%) were positive for the 16S rRNA gene. The average number of sequencing reads for the FFPE samples was 1,197 (range 46 - 3,458).

Each positive sample demonstrated mixed bacterial communities, rather than a single dominating OTU (frozen samples seen in Figure 5.3 A-D, and FFPE samples shown in Figure 5.4). Bacteria identified within these positive tissues were bacteria commonly found in the GI tract and included *Bacteroides*, *Escherichia/Shigella*, *Roseburia*, *Bifidobacterium*, *Prevotella*, *Clostridium*, and *Ruminococcus*. Although there were no trends observed in the PM tissues collected from different patients, there were similarities observed in bacteria identified in PM tissues collected from the same patient. The number of OTUs identified in each tissue sample average 11 (range 4-18).

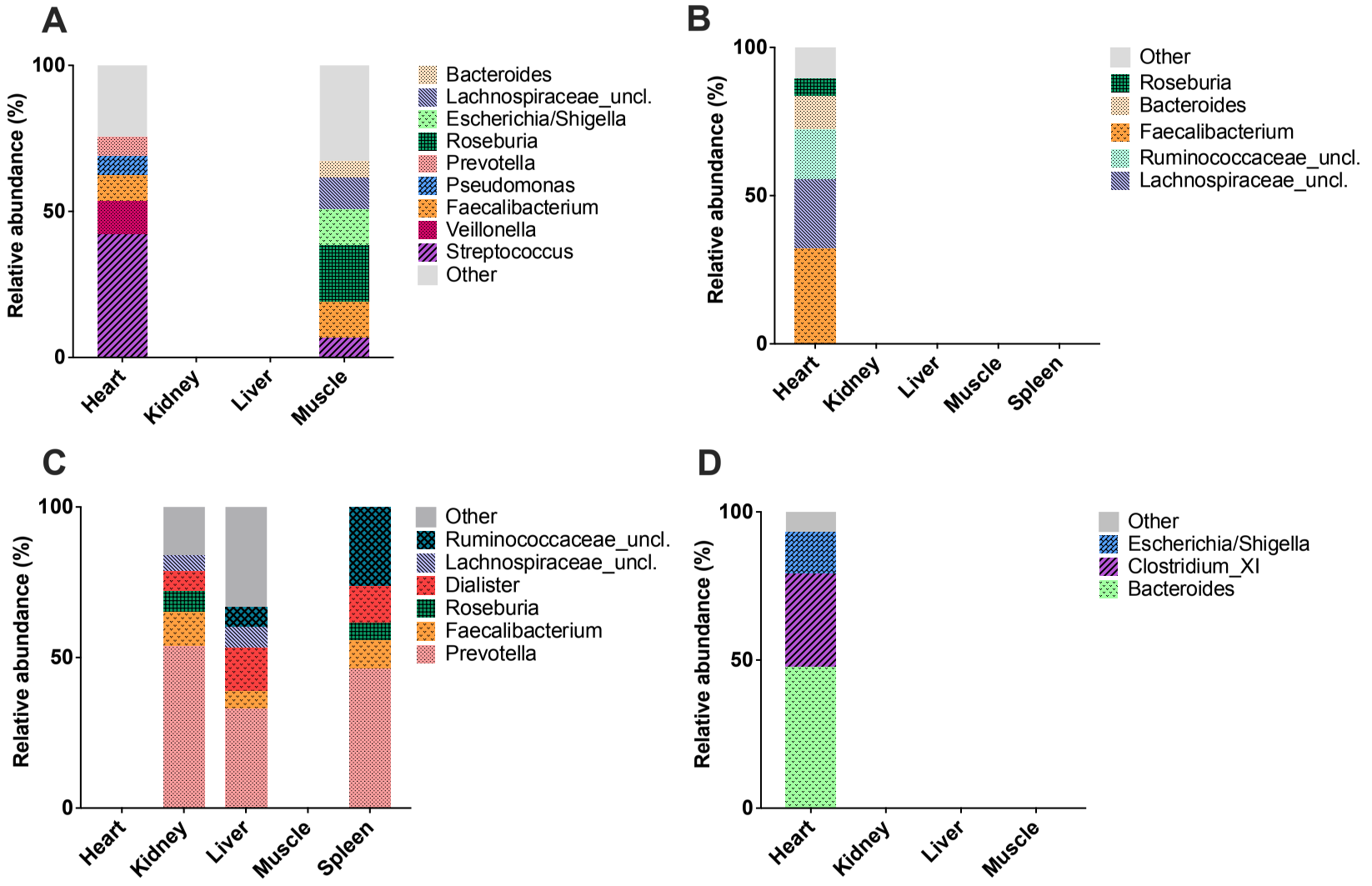

FIGURE 5.3: Relative abundance plots showing the bacteria identified in frozen PM tissues from explained non-infectious tissues using the 16S rRNA gene sequencing technique. Each plot represents a different case (A = eC1, B = eC2, C = eC3, D = eC4). The x-axis describes the PM tissue sample collected and the absence of a bar represents frozen PM samples that were obtained for analysis but negative for the 16S rRNA gene.

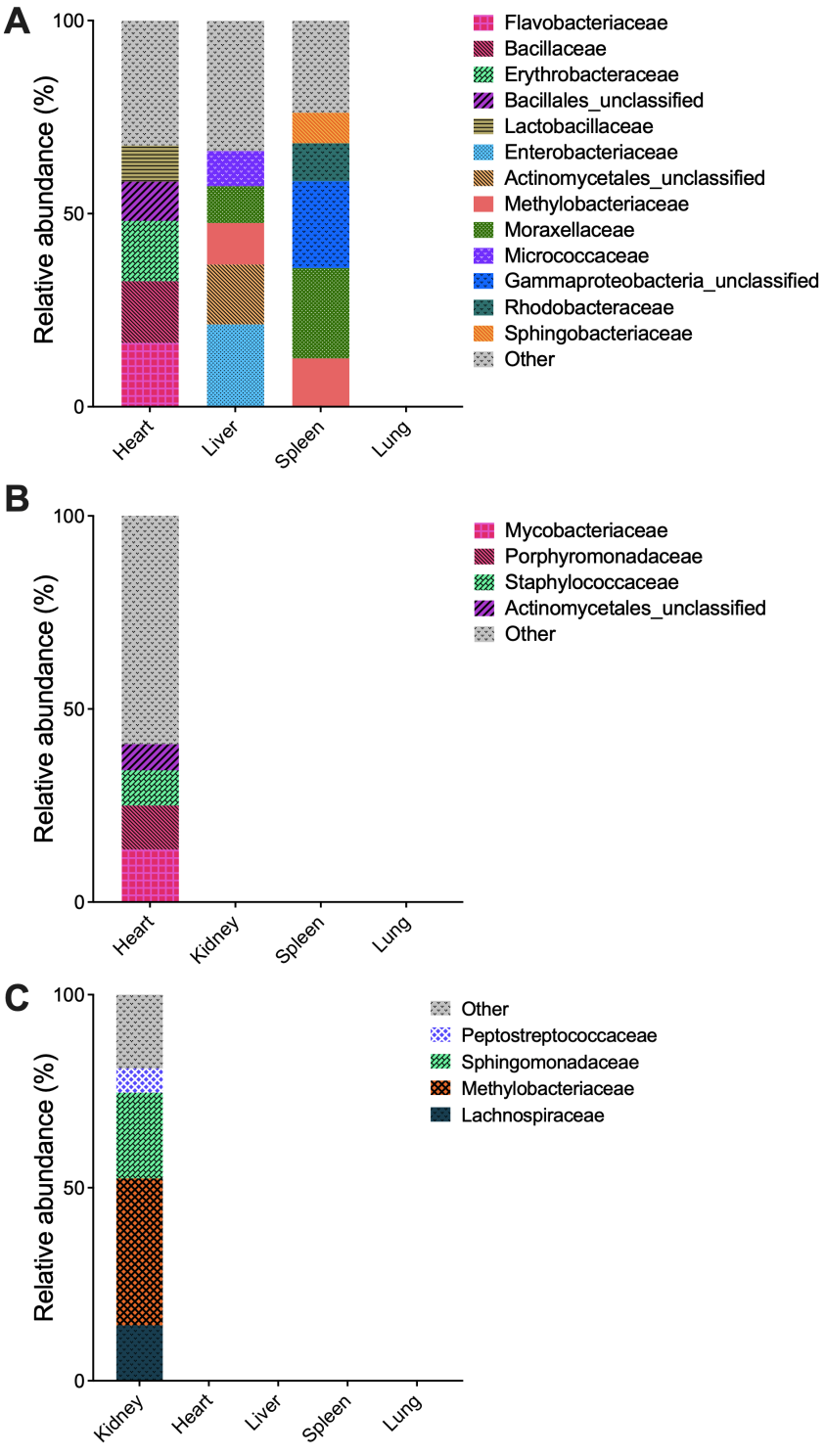

FIGURE 5.4: Relative abundance plots showing the bacteria identified in PM tissues from explained non-infectious tissues using the 16S rRNA gene sequencing technique. Each plot represents a different case (A = eC1, B = eC2, C = eC3). The x-axis describes the FFPE PM tissue sample collected and the absence of a bar represents samples that were obtained but negative for the 16S rRNA gene.

5.3.2 Sequencing PM tissues from infectious SUDIC

There were eight cases of infectious SUDIC examined as part of this study. These eight cases included five cases of septicaemia and three respiratory infections. As respiratory infections are often localised and frozen lung tissue is not routinely archived, these infections were examined separately.

A total of 21 frozen PM tissues were obtained from cases of systemic infection. Of these 21 tissues, 13 (62%) were considered positive for the 16S rRNA gene. The number of sequencing reads from these positive frozen tissues was in the range of 154 to 109,637 reads with a mean of 19,076 (±36,302). In all five cases of septicaemia a dominant bacterial member was identified, present at ≥80% in at least one frozen tissue sample from each PM case (total of 12/14 (86%) positive frozen tissue samples). A total of 9 FFPE tissues were obtained from two cases of systemic infection, of which 6 (67%) were positive for the 16S rRNA gene.

From SUDIC cases with a final CoD of respiratory infection 13 frozen PM tissues were obtained and 10 were considered positive and subsequently sequenced. Sequencing reads generated from these samples were in the range of 48 to 19,876 reads with a mean of 4,552 reads (±7,670). Of these three cases only one had a tissue with ≥80% relative abundance. It is important to note that these frozen PM tissue samples included heart, kidney, liver, spleen, and muscle samples. Frozen PM lung samples were not available. In one case PM FFPE heart, liver, and lung samples were also obtained. These samples generated 1918, 523, and 2363 reads respectively.

As each case displayed a unique bacterial profile these cases will be addressed individually below.

5.3.2.1 Systemic infections

Infectious case 1- Septicaemia caused by Group A *Streptococcus* (GAS)

iC1 was a childhood case of septicaemia caused by GAS. In the week leading up to death the child had experienced fever, cough, runny nose, and episodes of vomiting. PM histological examination showed bronchopneumonia associated with bilateral pleural effusions, bronchitis, and tracheitis. Generalised reactive lymphoid hyperplasia and moderate lymphoplasmacytic infiltrate were observed in the portal tracts of the liver, which indicate systemic illness. Microbiological investigation showed β-haemolytic GAS in the blood, CSF, lung, and spleen.

For this investigation, archival frozen spleen tissue was available. 16S rRNA gene sequencing revealed *Streptococcus* in the frozen spleen tissue at a relative abundance of 84.5% (Figure 5.5). Salmonella was also identified in this tissue at a relative abundance of 15.5%. FFPE heart, kidney, and lung samples were also obtained for analysis but all were negative for the 16S rRNA gene.

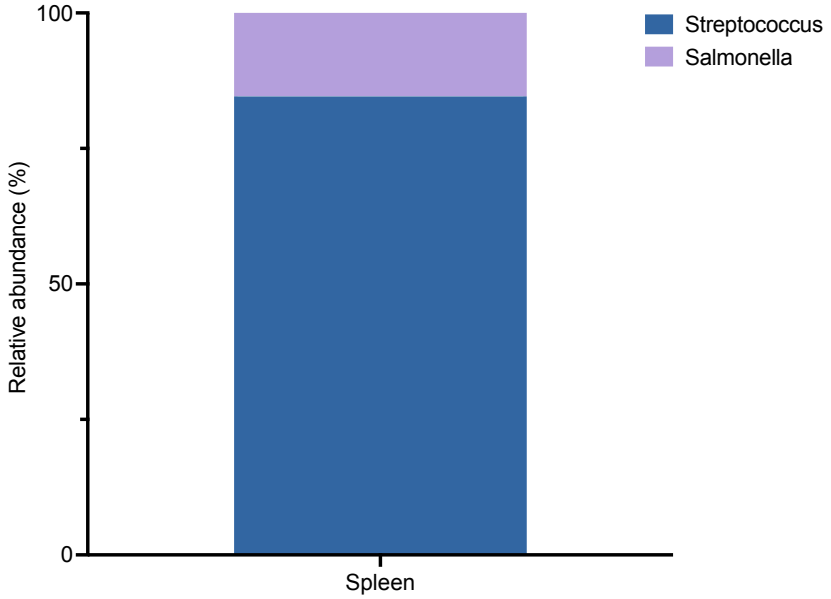

FIGURE 5.5: Relative abundance plot showing bacteria identified in frozen PM tissues collected from iC1

Infectious case 2- Septicaemia caused by *Neisseria meningitidis*

iC2 is a childhood case of septicaemia. In the weeks leading up to death the child had visited their general practitioner (GP) on two occasions with fever. The child was taken to the emergency department prior to death due to breathing difficulties and the development of a non-blanching skin rash. Broad-spectrum antibiotics were administered on arrival, but the child deteriorated rapidly. At PM investigation, histological findings included a generalised purpuric-haemorrhagic skin rash, multiple visceral haemorrhages and signs of acute tissue hypoxia. PM microbiological culture showed no bacterial growth in the blood, lung, spleen, and CSF. In a blood sample collected at the emergency department *Neisseria meningitidis* was detected by PCR.

This study examined frozen heart, kidney, liver, muscle, and spleen tissue samples and FFPE heart, kidney, and lung samples. Using the 16S rRNA gene sequencing technique, *Neisseria* was identified in the frozen heart, kidney, and liver samples present at 98.7% to 100% relative abundance (Figure 5.6). In the muscle, the dominating bacteria was *Enterococcus* (relative abundance 77.5%) followed by *Escherichia/Shigella* (relative abundance 16.3%) and *Neisseria* (relative abundance 1.4%). In the FFPE material, Neisseriaceae was present at 68.8%, 75%, and 87.5% in the heart, kidney, and lung samples respectively (Figure 5.7.

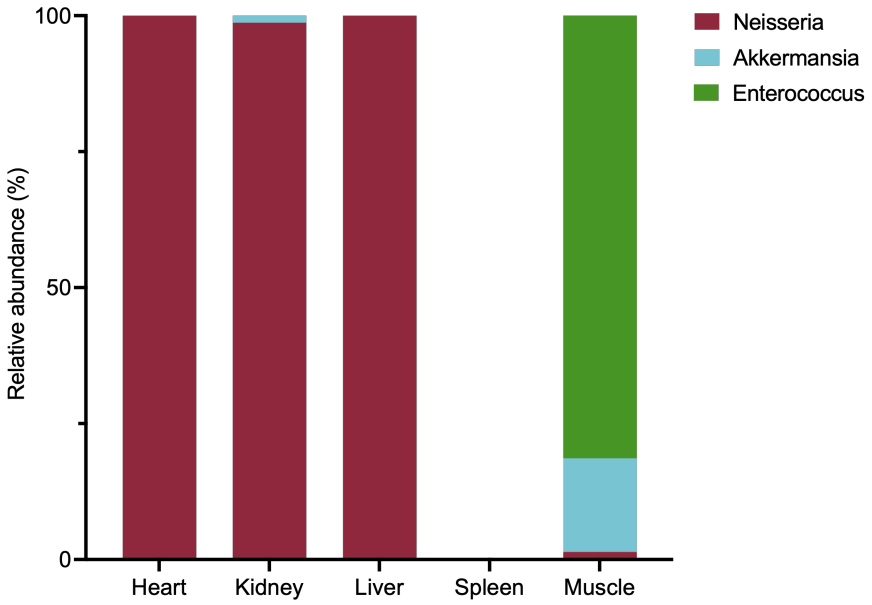

FIGURE 5.6: Relative abundance plot showing bacteria identified in frozen PM tissues collected from iC2

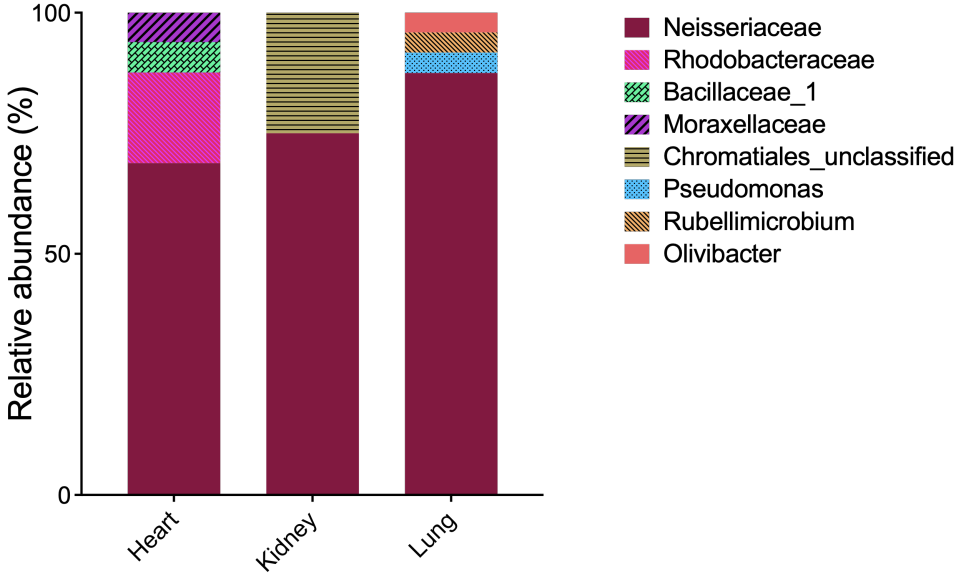

FIGURE 5.7: Relative abundance plot showing bacteria identified in FFPE PM tissues collected from iC2

Infectious case 3- Septicaemia caused by *Neisseria meningitidis*

iC3 is a case of septicaemia in an infant caused by *Neisseria meningitidis*. This infant had a short history of runny nose, cough, diarrhoea and fever, and had visited the GP in the two weeks leading up to death. PM histology found no evidence of meningitis but extensive bilateral adrenal haemorrhages were observed. The PM microbiology showed *Neisseria meningitidis* in the blood, CSF, and lung. URT commensals were detected in the spleen.

Frozen heart, kidney, liver, spleen and muscle samples were obtained from this case. The muscle tissue was negative for the 16S rRNA gene. *Neisseria* was identified in heart, liver, and spleen specimen at relative abundances of 99.6%, 96.7%, and 97.6% respectively (Figure 5.8). In the liver and spleen tissues, *Enterococcus* was also identified at 3.1% and 2.4% respectively. The kidney sample was dominated by *Enterococcus* with a relative abundance of 93% followed by *Neisseria* which was present at 7%.

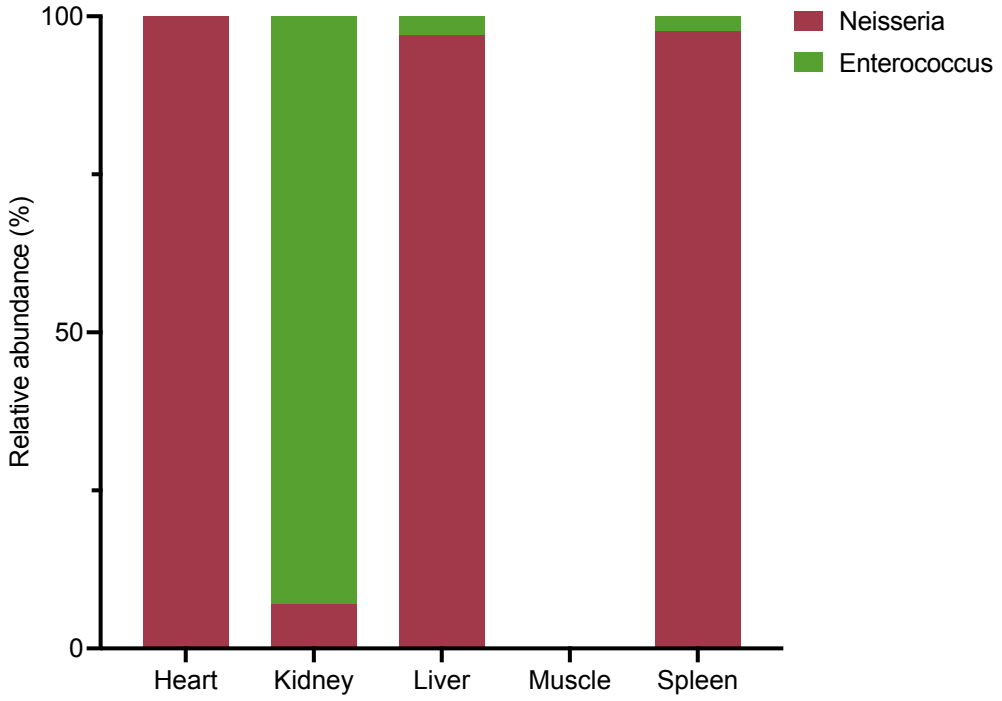

FIGURE 5.8: Relative abundance plot showing bacteria identified in frozen PM tissues collected from iC3

Infectious case 4- Septicaemia caused by an unidentified organism

iC4 was an infant case of septicaemia caused by an unidentified organism. This infant had been experiencing fever, cough, and vomiting. On arrival at the emergency department antibiotics and antivirals were administered. Histological findings included haemorrhagic adrenal glands, generalised vascular congestion, small visceral thrombi and hepatic microabscesses which are consistent with overwhelming bacterial sepsis. The histological changes in this infant were consistent with septicaemia typically caused by *Streptococcus pneumoniae*, GAS, and *Haemophilus influenzae*. However, PM microbiology results from investigation of the blood, lung, and CSF were negative potentially due to the administration of antibiotics in the hours leading up to death. The causative organism could not be identified.

Frozen heart, kidney, liver, muscle, and spleen samples were obtained from this patient. The heart tissue was positive for the 16S rRNA gene and *Streptococcus* was identified at a relative abundance of 100% (Figure 5.9).

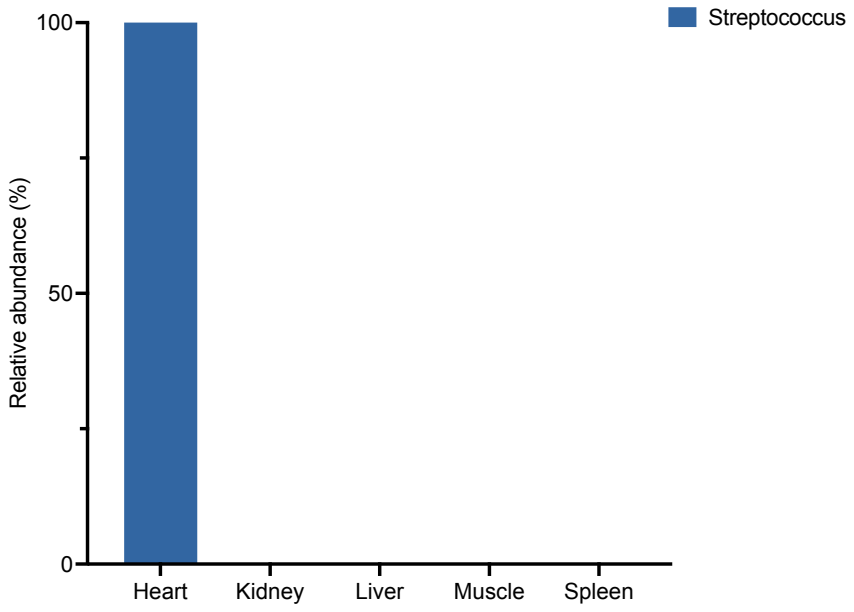

FIGURE 5.9: Relative abundance plot showing bacteria identified in frozen PM tissues collected from iC4

Infectious case 5- Septicaemia caused by *Klebsiella pneumoniae*

iC5 is an infant case of septicaemia caused by *Klebsiella pneumoniae*. A few days before death the infant visited the GP and was prescribed antibiotics for a throat infection. The following day they had fever and episodes of vomiting. Histology in this case showed petechiae of the heart, bilateral serosanguinous effusions, pleural petechiae in the lung, foci of haemorrhages in the tonsils and moderate acute pneumonia. Microbiological investigation revealed *Klebsiella pneumoniae* in the blood, right and left tonsils, lung, and spleen samples. No bacterial growth was observed in the CSF.

Frozen heart, kidney, liver, muscle, and spleen samples were obtained for this study. The frozen heart sample was negative. The kidney, liver, and spleen samples revealed Enterobacteriaceae members at relative abundances of 100%. The frozen muscle sample showed Enterobacteriaceae members at 81% followed by *Escherichia/Shigella* at 16% (Figure 5.10).

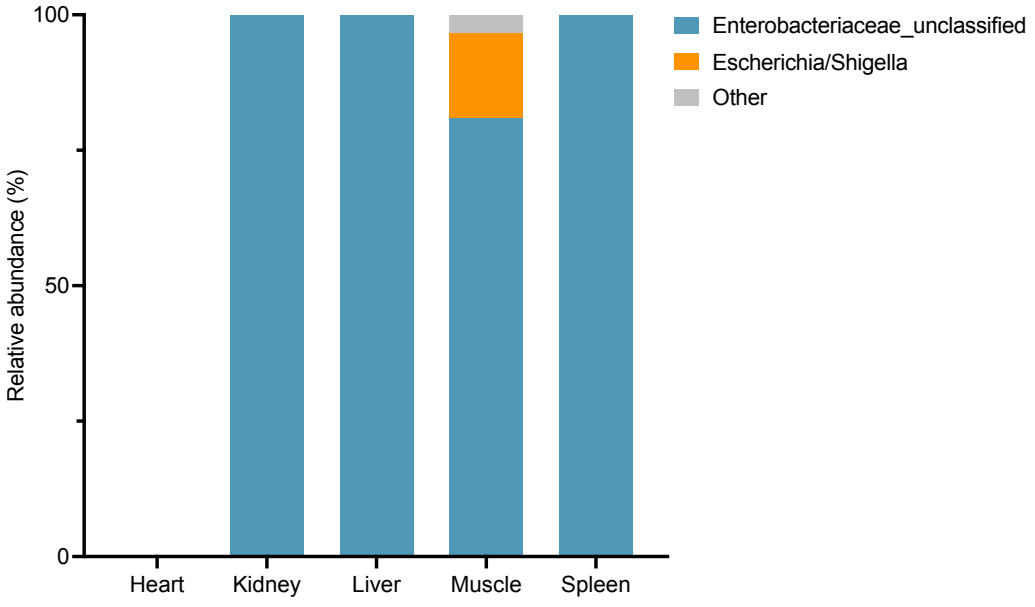

FIGURE 5.10: Relative abundance plot showing bacteria identified in frozen PM tissues collected from iC5

5.3.2.2 Respiratory infections

16S rRNA sequencing results from PM tissues collected from respiratory infection cases were less clear when examining frozen tissues alone as PM lung tissue samples are not routinely frozen and archived.

Infectious case 6- Pneumonia caused by an unidentified organism

iC6 is a childhood case of pneumonia. This child was experiencing a mild URT in the weeks leading up to death, with fever and a cough. The child was found lifeless in his cot in the prone sleeping position. Histological examination showed evidence of pneumonia and increased peribronchial lymphoid tissue. It was suspected that there was an initial viral infection superimposed by a bacterial infection. The causative organism in this case could not be identified by classical bacterial culture.

Frozen PM heart, kidney, liver, and muscle were obtained for the sequencing. The frozen heart and muscle tissues were positive for the 16S gene. In both of these tissues *Enterococcus* was dominating, representing 99.8% and 85.3% in the heart and muscle samples respectively (Figure 5.11). In the muscle tissue *Salmonella* and Enterobacteriaceae bacterium were also present at low relative abundances.

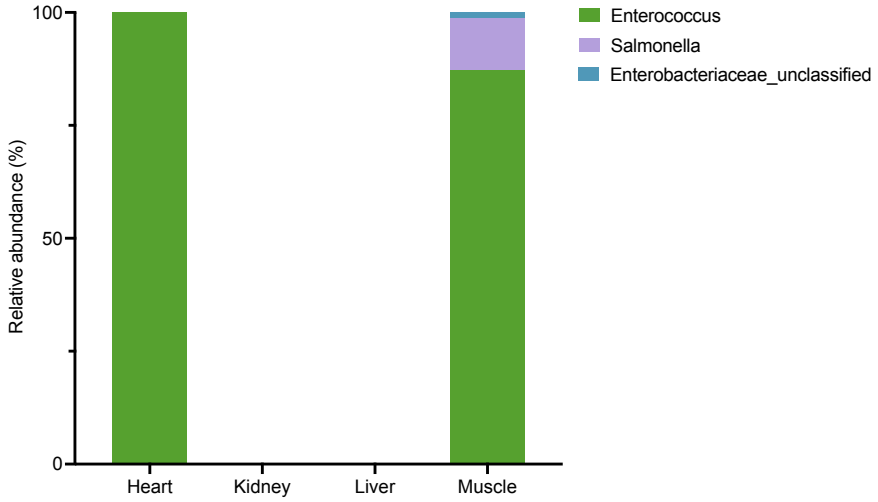

FIGURE 5.11: Relative abundance plot showing bacteria identified in frozen PM tissues collected from iC6

Infectious case 7- Bronchiolitis caused by Influenza A Virus

iC7 was a childhood case of bronchiolitis of viral origin. The child had been generally unwell in the weeks prior to death, with painful breathing and diarrhoea. At the time of death they were taking oral antibiotics. At PM investigation vascular congestion and oedema was observed in the heart and lungs. In the tracheal there was lymphoplasmacytic inflammation and bacterial colonies observed. Microbiological investigation showed no bacterial growth in the blood, spleen, or CSF. URT commensals were identified in the lung. Molecular virological investigation showed Influenza A virus present in the lung tissue.

For this study frozen heart, kidney, liver, spleen, and muscle samples were collected. The heart, spleen, and kidney samples were positive but all samples had less than 154 sequencing reads (Figure 5.12). This case FFPE tissue was not available and therefore lung tissue was not investigated.

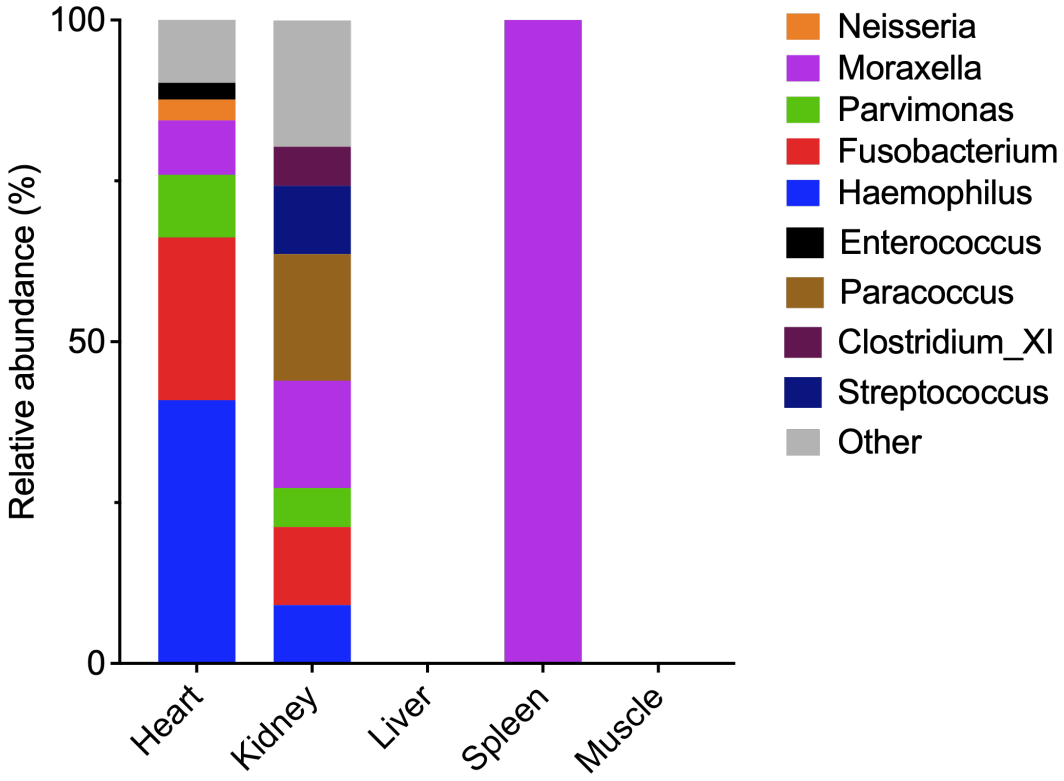

FIGURE 5.12: Relative abundance plot showing bacteria identified in frozen PM tissues collected from iC7

Infectious case 8- Pneumonia caused by an unidentified organism

iC8 was an infant case of pneumonia where the causative organism could not be identified at PM. The infant had been unsettled in the hours leading up to death and had been taking oral antibiotics for another condition. At PM, histologically there was inflammation of the trachea and small airways with multiple areas of infection observed in the lung. PM microbiology was performed on the blood, CSF, lung, and spleen. Viridians streptococci was identified in the blood, URT commensals in the lung and no growth was observed in the CSF or spleen.

Frozen heart, kidney, liver, and muscle where obtained for this study. Sequencing of the 16S rRNA gene in these tissues showed mixed communities of bacteria in the heart, kidney, liver, and muscle. These bacteria were mainly enteric and included Lachnospiraceae, Enterobacteriaceae, Bacteroides, and Veillonellaceae.

For this case, archived FFPE PM heart, liver, and lung material were also obtained. The heart and liver tissues showed mixed communities of enteric bacteria, but in the FFPE PM lung sample *Streptococcus* was identified at a relative abundance of 93% (Figure 5.13).

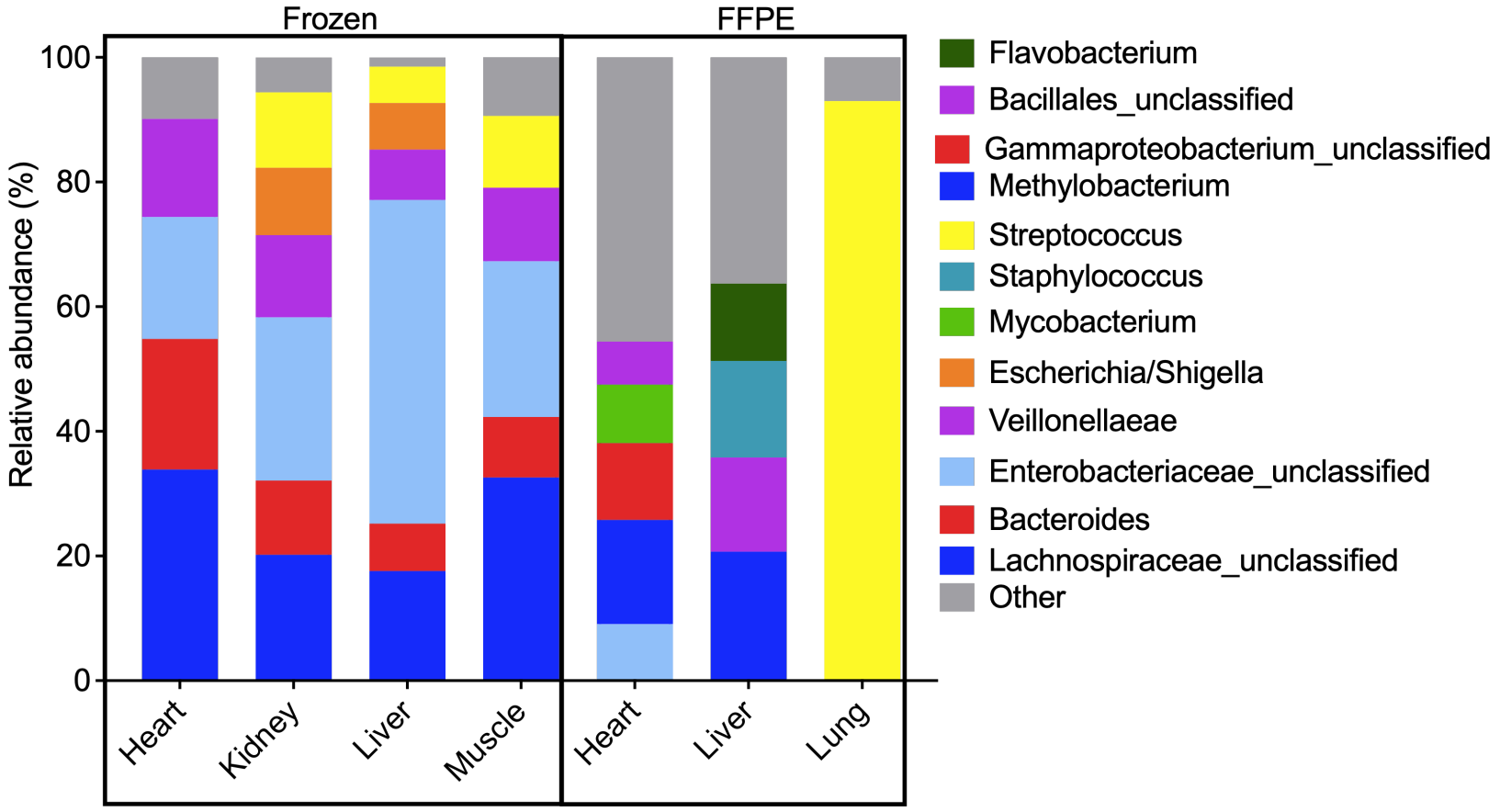

FIGURE 5.13: Relative abundance plots of iC8 with a final diagnosis of Pneumonia caused by an unidentified organism

Case	Final CoD	Causative organism by bacterial culture	Dominating bacterium identified by 16S rRNA gene sequencing
iC1	Septicaemia	Group A Streptococcus	*Streptococcus* (84.6%)
iC2	Septicaemia	Unidentified organism	*Neisseria* (98-100%)
iC3	Septicaemia	*Neisseria meningitidis*	*Neisseria* (97-100%)
iC4	Septicaemia	Unidentified organism	*Streptococcus* (100%)
iC5	Septicaemia	*Klebsiella pneumoniae*	Enterobacteriaceae_unclassified (81-100%)
iC6	Pneumonia	Unidentified organism	*Enterococcus* (87.2-100%)
iC7	Bronchiolitis	Viral origin	n/a
iC8	Pneumonia	Unidentified organism	*Streptococcus* (100%)

5.3.3 Defining a threshold using control groups

As there was a clear difference in the number of sequencing reads generated and the OTU patterns observed in the frozen PM material explained, non-infectious cause group and the infectious cause group, a threshold was established in order to differentiate between low-level bacterial colonisation that may occur as part of the PM process and a true infection-causing pathogen. As the sequencing results obtained from the FFPE material were not as consistent as those obtained from frozen material, in cases where both frozen and FFPE material were obtained, the FFPE material was not used for subsequent analysis.

The positive mock community controls of known bacterial colonisation generated an average of 33,572 sequencing reads and the negative extraction controls had an average of 67 reads. These findings were used to as basis to determine a cut-off for the number of sequencing reads to remove any generated data representing low-level bacterial detection that is not indicative of infection. Sequencing reads numbers $\leq$1% of the positive bacterial community were considered not representative of infection. Therefore frozen PM samples must have generated at least 335 sequencing reads to be considered potentially infected.

The second criteria that was set was based on observed bacterial dominance. In all SUDIC cases with an explained non-infectious cause mixed communities were observed. In SUDIC cases with an infectious cause, a dominant bacteria was present in a majority of samples positive for the 16S rRNA gene. For this reason for a PM sample to be considered infected, the dominant bacterial member must be present at $\geq$80% of the overall community.

The results of the implementation of these criteria in the explained non-infectious and infectious control groups are shown in Table 5.4 and Table 5.5 respectively. In the explained, non-infectious group, 8/15 (53.3%) positive samples had of a minimum of 350 sequencing reads. However, there were no PM tissue samples in this group that passed the second criteria of greater than 80% dominance. In the explained infectious group there were 18/24 (75%) PM samples that had $\geq$350

TABLE 5.4: Application of defined criteria to tissues from cases of explained non-infectious SUDIC

		Criteria 1 ≥350 sequencing reads		Criteria 2 ≥80% dominance	
Case	Tissue	Sequencing reads	Pass/Fail	Dominance (%)	Pass/Fail
eC13	S	26	Fail		
eC7	Li	40	Fail		
eC5	Li	44	Fail		
eC14	Li	151	Fail		
eC5	S	179	Fail		
eC12	S	265	Fail		
eC1	M	341	Fail		
eC4	H	419	Pass		
eC3	K	479	Pass	53.7	Fail
eC1	H	526	Pass	42.2	Fail
eC12	K	739	Pass	73.7	Fail
eC3	Li	781	Pass	33	Fail
eC5	K	844	Pass	43.1	Fail
eC3	S	1588	Pass	46.2	Fail
eC2	H	8045	Pass	32.2	Fail

sequencing reads, and 14 of these 18 (78%) had a dominant bacteria present at ≥80% relative abundance.

Table 5.5: Application of defined criteria to tissues from cases of SUDIC with an infectious cause of death

		Criteria 1 ≥350 sequencing reads		Criteria 2 ≥80% dominance	
Case	Tissue	Sequencing reads	Pass/Fail	Dominance (%)	Pass/Fail
iC6	K	48	Fail		
iC7	K	66	Fail		
iC7	S	97	Fail		
iC5	M	135	Fail		
iC7	H	154	Fail		
iC2	K	154	Fail		
iC2	Li	355	Pass	100	Pass
iC1	S	415	Pass	84.5	Pass
iC8	Li	642	Pass	51.9	Fail
iC8	H	817	Pass	33.9	Fail
iC8	K	841	Pass	26.2	Fail
iC5	Li	993	Pass	100	Pass
iC5	K	1129	Pass	100	Pass
iC5	S	1351	Pass	100	Pass
iC4	H	3098	Pass	100	Pass
iC8	M	3285	Pass	32.6	Fail
iC2	H	3897	Pass	100	Pass
iC2	M	5346	Pass	81.4	Pass
iC3	S	13496	Pass	97.6	Pass
iC6	H	19694	Pass	100	Pass
iC6	M	19876	Pass	87.2	Pass
iC3	Li	23457	Pass	96.9	Pass
iC3	H	103605	Pass	100	Pass
iC3	K	109637	Pass	93	Pass

5.3.4 Application of the defined threshold to cases of unexplained SUDIC

The unexplained cases used in this investigation included a total of 20 cases of unexplained SUDIC (15 infant cases and 5 childhood cases). The unexplained cases had a mean number of sequencing reads of 16,695 (range 26 - 190,052). Of the infant unexplained cases 36% were positive and 66% were negative. In the childhood cases 73% were positive and 27% were negative. The bacterial profiles of these PM tissues generally fell into one of two groups; those that resembled the explained, non-infectious cause cohort with mixed communities, and those

that resembled the infectious cohort with a single dominating bacteria. Of the 15 infants, six cases did not have any positive tissue samples. All five childhood cases had at least one positive tissue sample.

The criteria defined in Section 5.3.3 were applied to the PM tissues from the unexplained cohort to identify any tissues that may have been infected at the time of death. The results of application of the threshold to this cohort are shown in Table 5.6. This study obtained 94 frozen PM tissue samples and of these 41 were considered positive. A total of 15/41 (37%) frozen PM tissue samples had $\geq$350 sequencing reads and of these, 10 (67%) demonstrated a dominant bacterium present at $\geq$80% relative abundance. These 10 samples were collected from two infant cases and two childhood cases of unexplained death. Of the five FFPE samples that were investigated in this cohort, 0 (0%) passed the defined threshold. Each case that passed the defined threshold will be discussed.

Table 5.6: Application of defined criteria to tissues from cases of unexplained SUDIC

		Criteria 1 ≥350 sequencing reads		Criteria 2 ≥80% dominance	
Case	Tissue	Sequencing reads	Pass/Fail	Dominance (%)	Pass/Fail
uC16	H	26	Fail	-	-
uC18	S	26	Fail	-	-
uC19	K	26	Fail	-	-
uC10	S	29	Fail	-	-
uC20	H	33	Fail	-	-
uC17	Li	35	Fail	-	-
uC13	K	36	Fail	-	-
uC2	Li	44	Fail	-	-
uC17	K	46	Fail	-	-
uC10	K	62	Fail	-	-
uC8	K	77	Fail	-	-
uC2	K	79	Fail	-	-
uC19	H	82	Fail	-	-
uC3	S	85	Fail	-	-
uC18	H	103	Fail	-	-
uC19	M	105	Fail	-	-
uC11	S	113	Fail	-	-
uC1	K	118	Fail	-	-
uC11	Li	124	Fail	-	-
uC13	H	126	Fail	-	-
uC10	Li	128	Fail	-	-
uC3	H	168	Fail	-	-
uC11	K	186	Fail	-	-
uC9	K	295	Fail	-	-
uC10	M	333	Fail	-	-
uC1	S	346	Fail	-	-
uC10	H	377	Pass	75.1	Fail
uC11	H	409	Pass	50.9	Fail
uC18	M	1679	Pass	22.6	Fail
uC12	S	1911	Pass	100	Pass
uC1	M	23215	Pass	100	Pass
uC1	L	26497	Pass	100	Pass
uC4	H	32589	Pass	100	Pass
uC4	K	33140	Pass	100	Pass
uC4	L	39618	Pass	100	Pass
uC4	M	40659	Pass	100	Pass
uC12	H	45885	Pass	100	Pass
uC14	L	93238	Pass	84.1	Pass
uC14	K	94853	Pass	73.5	Fail
uC14	H	124637	Pass	94.1	Pass
uC14	M	190052	Pass	68.9	Fail

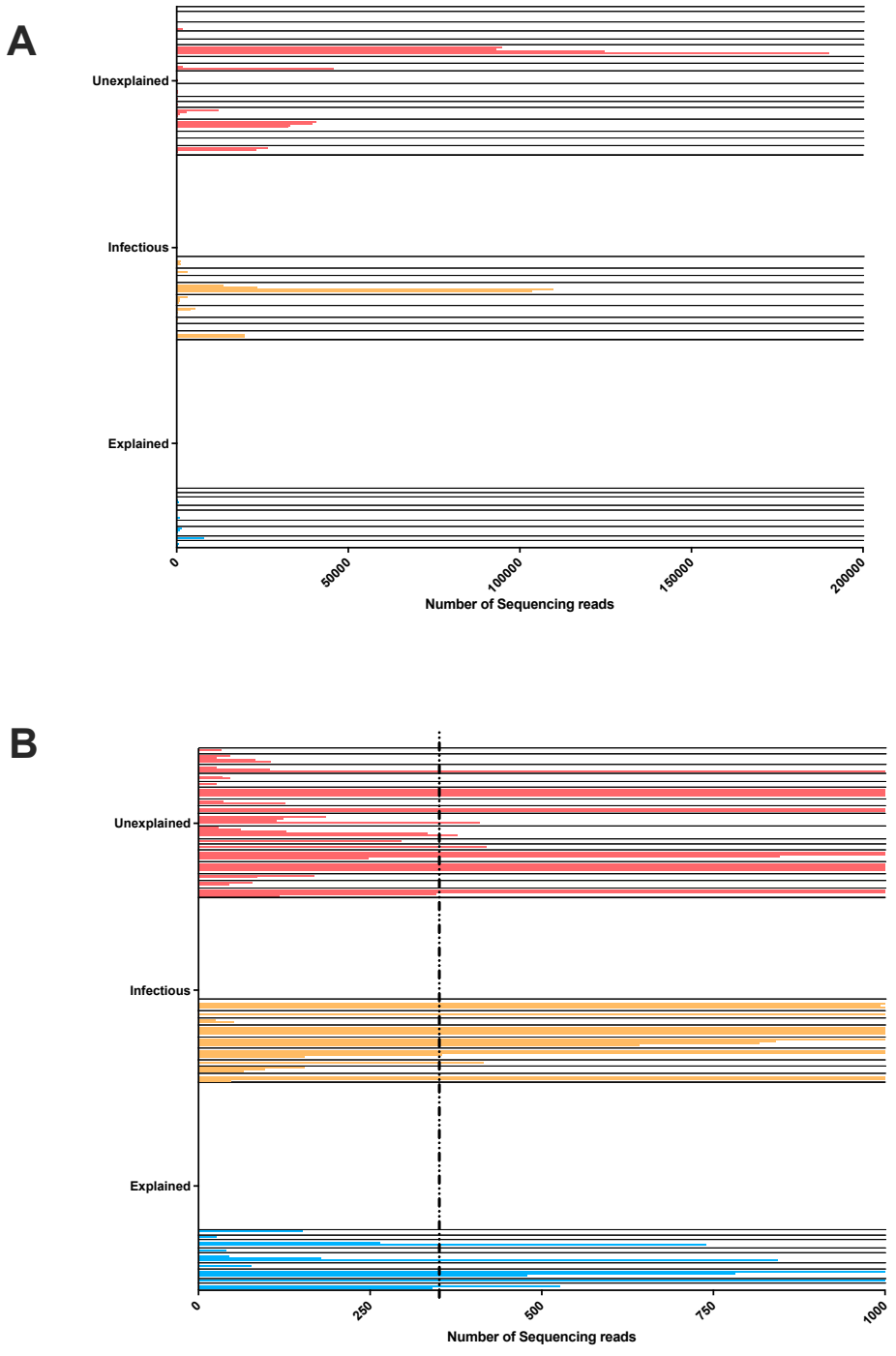

FIGURE 5.14: Number of sequencing reads in the frozen PM samples collected from each CoD group. (A) Total number of sequencing reads and (B) The number of sequencing reads with threshold applied.

Unexplained Case 12- SIDS

uC12 was an infant who had a history of bronchiolitis for which they had previously been hospitalised. In the weeks leading up to death the infant had been experiencing fever and a chesty cough. The child was found unresponsive after a period of sleep. PM histology was normal and there was no evidence of inflammation in any organ. PM bacterial culture of the blood, spleen, and CSF were negative. URT commensals were isolated from the lung. Molecular virology was performed on the PM lung tissue which was negative for the panel of respiratory viruses tested for at GOSH.

From this case frozen heart, kidney, liver, and spleen samples were obtained. The kidney and liver samples were negative, but the heart and spleen were both dominated by Enterococcus at 100% and 99.3% respectively.

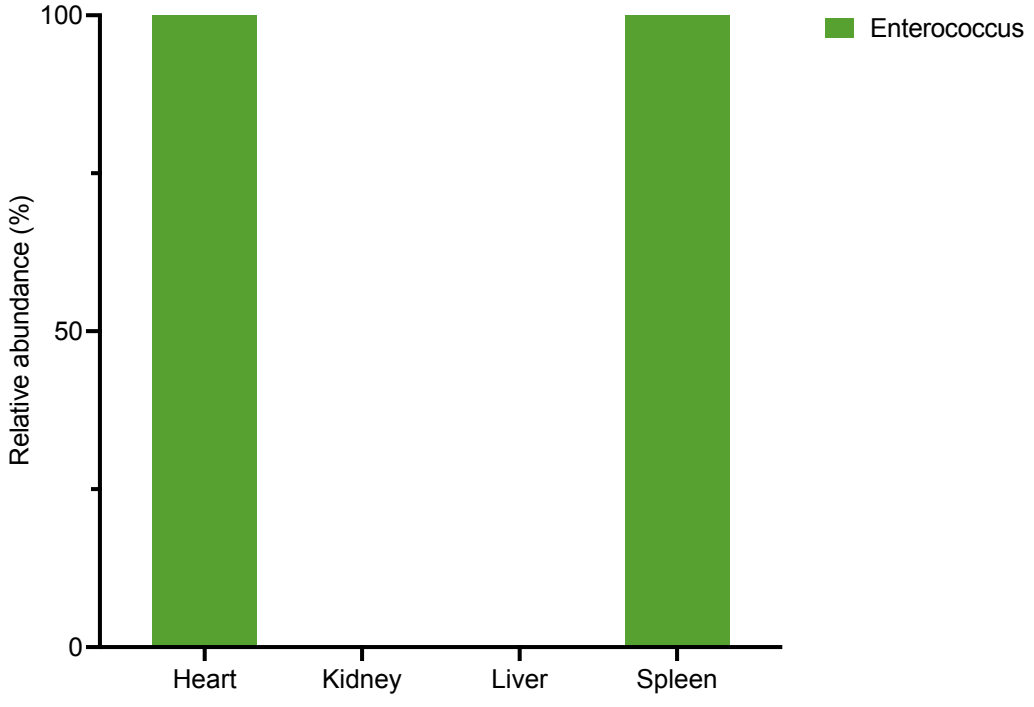

FIGURE 5.15: Relative abundance plot showing bacteria identified in frozen PM tissues collected from uC12

Unexplained Case 14- SIDS, Cardiac myopathy of unknown aetiology

In case uC14, the death was of unexplained aetiology but where cardiac myopathy was observed at PM investigation. This infant had no appetite, was unsettled, and a change in breathing observed prior to the last sleep. Histological examination of the lungs found vascular congestion, but no evidence of pneumonia or bronchiolitis. PM microbiology findings showed no growth in the blood, lung, spleen, or CSF. Virology was performed on the lung and heart, and both tissues were negative for the targeted panel.

In this case heart, kidney, liver and muscle samples were obtained. All samples were sequenced but only the heart and liver passed the defined threshold. These tissues were dominated by Enterobacteriaceae, present in the heart at 94.1% and the liver at 84.1%). In both samples *Enterococcus* was the second most abundant bacteria. *Enterococcus* relative abundance was higher in tissues from the abdominal cavity that the thoracic cavity. Interestingly, although the kidney and muscle did not pass the defined threshold they showed the same bacterial identification with Enterobacteriaceae dominating at 73.5% and 68.9% in the kidney and muscle respectively, followed by *Enterococcus*.

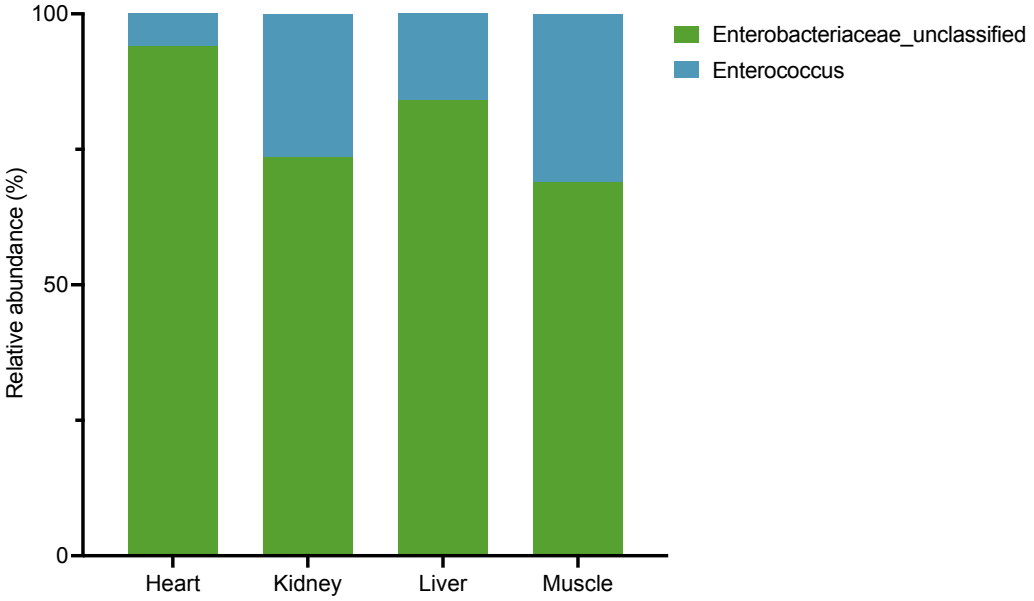

FIGURE 5.16: Relative abundance plot showing bacteria identified in frozen PM tissues collected from uC14

Unexplained case 1- SUDC

uC1 of unexplained death of a child under five years-of-age. The child was found unresponsive after a period of sleep. No other clinical history was available.

At PM examination there was focal alveolar oedema and mild intra-alveolar haemorrhages in the lungs. There was an overgrowth of bacteria in the airways but this was presumed to be aspirated during resuscitation efforts. PM microbiology showed growth of Clostridia in the blood and spleen samples, *Staphylococcus aureus* in the lung sample and no growth in the CSF sample.

For this case heart, kidney, liver, muscle, and spleen samples were obtained for 16S rRNA gene sequencing. Relative abundance plots can be found in Figure 5.17. The heart and kidney were negative for the 16S gene but the liver and muscle showed 100% relative abundance of *Escherichia/Shigella*.

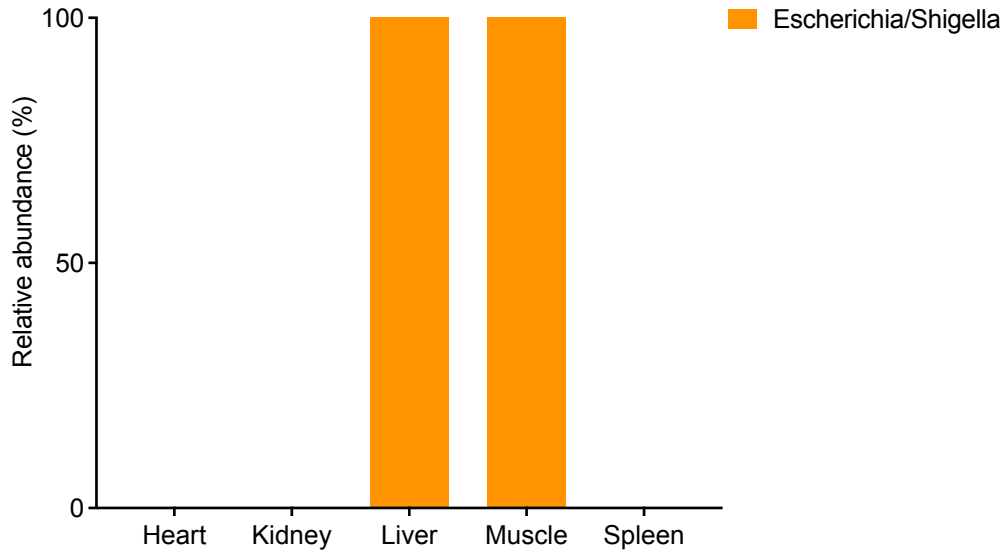

FIGURE 5.17: Relative abundance plot showing bacteria identified in frozen PM tissues collected from uC3

Unexplained case 4- SUDC

uC4 was an unexplained death of a child under five years of age who had a history of a impetigo for which they were being treated with antibiotics. Prior to death the child had been unwell with fever and irritability. At PM investigation, histologically there was bacterial overgrowth and congestion in the lungs, which were considered to be products of decomposition. PM microbiology found *Actinomyces lingnae* in the blood and *Clostridium sordelli* in the lung and swabs collected from both the left and right ear. No growth was observed in the CSF or spleen samples.

For this study, frozen heart, kidney, liver, and muscle samples were collected and the 16S rRNA gene sequenced. All tissues were positive and dominated by *Clostridium* as shown in Figure 5.18. *Clostridium* was present at 100% relative abundance in all samples.

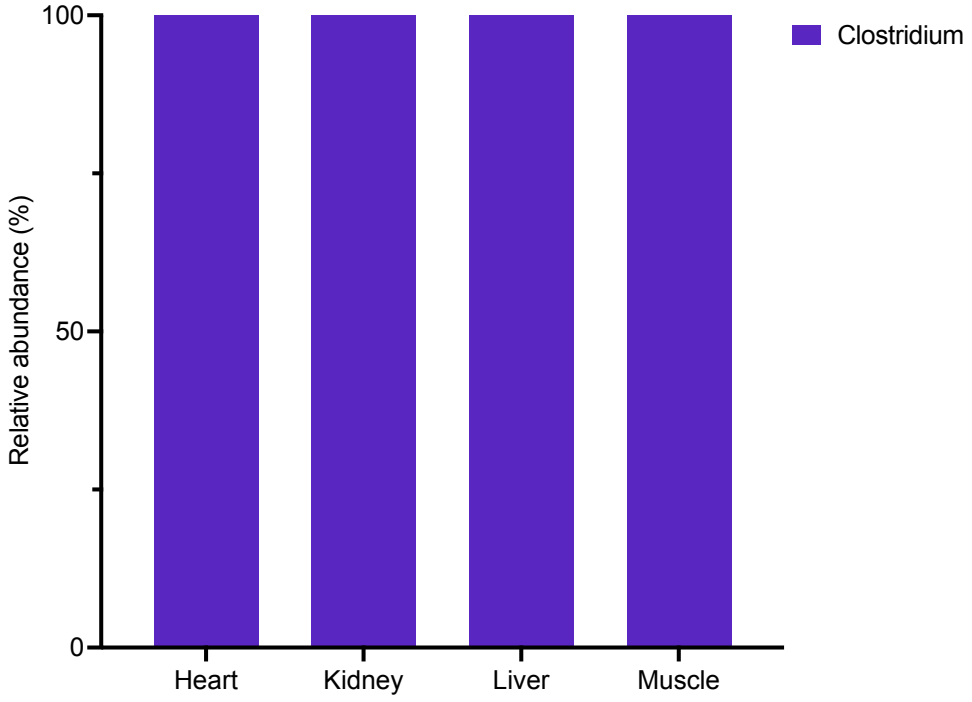

FIGURE 5.18: Relative abundance plot showing bacteria identified in frozen PM tissues collected from uC4

5.4 Discussion

This study aimed to assess the use of 16S rRNA gene sequencing in identifying infection in cases of SUDIC, and to investigate whether this technique provides an advantage over the current routine standard procedure of microbiological culture. The results from this investigation suggest that the 16S rRNA gene sequencing method can distinguish between infectious and non-infectious cases of SUDIC and can successfully identify a causative organism in infectious cases. In cases where infection was diagnosed only by histological examination and where the causative organism could not be identified using bacterial culture, the 16S rRNA gene sequencing method was able to identify the likely infection-causing bacterium. These results show that this method could therefore enhance the PM investigation into cases of SUDIC.

The demographics of this study population reflects the proportions of SUDIC in England and Wales. Unexplained death represents approximately 40% of SUDIC and represented 46% of the samples used in this study. The smallest cohort in this study were the infectious deaths which represented 18% of the study cohort. Explained deaths with a non-infectious cause represented 42% of the study cohort. Cases with an explained cause of death had lower rates of consent for scientific research than those cases in which the cause of death remained unknown.

The 16S rRNA gene sequencing technique was able to clearly distinguish between cases of SUDIC with infectious and non-infectious causes. A clear difference observed in both the number of sequencing reads generated and the OTU dominance pattern. In all explained deaths with a known non-infectious cause, a large proportion of tissue samples were negative. The samples that were successfully amplified and sequenced generated a low number of sequencing reads and showed no bacterial dominance. These mixed communities of bacteria suggest that there is very little, if any, bacterial presence in these tissues. Similar results were observed in a study examining microbial communities of internal organs from human cadavers using 16S rRNA gene sequencing. The cause of death for these cases was non-infectious.

On average each sample had 5,404 sequencing reads and mixed communities including Bacteroidales, Burkholderiales, Caulobacterales, and Lactobacillales were observed [163].

The absence of significant bacterial DNA may result in non-specific primer binding and off-target amplification [164]. The finding of similarity of bacteria identified within PM tissues obtained from the same case is consistent with findings in the piglet model in Chapter 3. According to these findings it is likely that these bacteria arise from the individual rather than the external environment during sample acquisition and handling. Although PM bacterial translocation causes difficulty in the interpretation of bacterial culture results obtained at PM, the 16S rRNA gene sequencing method appears to be able to differentiate between true infection-causing organisms and those that are a product of the PM process. This would therefore assist in the interpretation of results should the method be used in routine PM investigation.

In infected tissues from cases of confirmed systemic infection, sequencing read numbers were significantly higher than those from explained non-infectious cases and tissues generally demonstrated a single dominant bacterium. In cases where a causative organism was identified during routine PM investigation by bacterial culture, the same organism was identified in this study. An interesting case was iC2 where blood samples were collected prior to death when the infant arrived at the emergency department. Investigation of this blood sample using molecular methods identified the presence of *Neisseria meningitidis*. Bacterial culture on tissues collected during autopsy could not identify the causative organism, likely due to the administration antibiotics prior to death. In this study *Neisseria* was identified as the dominating bacteria in PM heart, kidney, and liver samples.

In cases of respiratory infection, a majority of frozen PM tissues, which did not include lung samples, did not pass the defined threshold. The bacterial communities identified within these PM tissues resembled those identified in explained non-infectious deaths suggesting that the respiratory infections were localised and therefore the tissues included in this investigation were found to be uninfected at

the time of death. In the single respiratory infection case where FFPE lung was obtained, an infectious pattern was observed where the read numbers surpassed the threshold and *Streptococcus* was observed at a relative abundance >80%. At PM investigation, the bacterial culture of Viridians streptococci from the blood and URT commensals from the lung were not considered to be relevant to the cause of death. Unfortunately, sequencing of the V3-V4 region of the 16S gene cannot provide reliable species level classification and therefore the relevance of this finding at PM cannot be determined. *Streptococcus pneumoniae* is the most common cause of pneumonia in the paediatric populations. Other *Streptococcus* species such as those in the Viridians streptococci group are common commensals and infections are rare. These bacteria are therefore often dismissed as a cause of pneumonia due to its relatively low pathogenicity [165]. However, they are of clinical relevance, particularly the paediatric population where the immune system is not fully developed. In the period between waning maternal antibodies acquired during gestation and the maturation of the infantile immune system, infants experience a short period where the are at an increased risk of infection. During this period cases of community-acquired pneumonia, bacteraemia, and infective endocarditis caused by Viridians streptococci have been reported [166–169]. Results from cases such as this suggest a hybrid approach may be preferred, where 16S rRNA gene sequencing is performed as an initial screen. In tissues were a dominant pathogen is identified, species-specific PCRs could be requested for species identification. This speciation may supplement the 16S rRNA gene sequencing data and allow the pathologist to make a more informed decision than with the 16S data alone.

In this study cohort there was only a single case of SUDIC caused by a respiratory infection case where FFPE lung sample was available and therefore conclusions on the reliability of results cannot be drawn. This study has shown that identification of the likely infectious-causing bacterial pathogen is possible in FFPE material obtained from SUDIC cases caused by systemic infections. However, the results obtained from FFPE material were often less clear than those obtained

from frozen material. In the FFPE tissues, the dominant bacteria had a lower relative abundance and more low abundance OTUs were identified potentially due to the tissue processing protocol. Whilst the likely infectious-causing pathogen was identified in both FFPE and frozen material, the clearer results from the frozen tissues make it the more desired sample type. Therefore, for future investigations frozen material is recommended.

These results highlight the importance of obtaining the correct tissue type for the diagnosis of different infectious illnesses. In cases of respiratory infection, samples such as lung tissue should be investigated for diagnosis. As lung tissue is not routinely archived as frozen material, this study was unable to identify potential cases of respiratory infection making it entirely possible for undiagnosed respiratory infections to remain in the unexplained cohort. These results support a change in routine procedure to include lung tissue in frozen archival material. The importance of tissue type was also demonstrated in systemic infections where the bacterial pathogen was not always identified in all tissue types. In PM muscle and kidney samples, the likely bacterial pathogen was not always identified. In two cases of systemic infection these samples displayed dominating enteric bacteria rather than the known infection-causing bacterium, potentially due to their close proximity to the GI tract. It is interesting to note that in the two cases of septicaemia caused by *Neisseria meningitidis*, enteric bacteria were the dominating bacteria in these tissues. The enteric bacteria was also identified in the liver and spleen which also have close proximity to the GI tract but these were present at low relative abundances. Interestingly, although not dominated by the causative pathogen, these samples still passed the defined threshold which was not observed in any case of explained non-infectious death. This could be representative of two scenarios; (i) a change in gut microbiome as a result of infectious disease, (ii) a different translocation process occurring in individuals with infection. It is well established that diseases associated with a host immune response causes dysbiosis in the gut. These colonising bacteria may be more pathogenic and therefore have more invasive properties allowing translocation to occur. Therefore, although this bacterium was not the causative pathogen, it could be used a marker of infectious

disease. A study performed by Pechal *et al.* investigated microbial diversity and taxa present in six external anatomic locations in forensic cases from homicide, suicide, and chronic illnesses such as heart disease. They found that using diversity statistics and indicator taxa, individuals with ante-mortem health conditions could be successfully distinguished from those whose death was a result of violent circumstances [170].

When comparing results generated from 16S sequencing to those obtained from bacterial culture, 16S sequencing provides a more objective view of the bacteria present within a tissue sample and can identify more bacteria within a given samples as has been shown in other published studies [91, 92, 95, 171]. The 16S sequencing method provides random universal amplification of all 16S genes present within a sample, avoiding the interference of bacterial growth curves and potential false over-representation of low abundance bacteria due to their increased culturability. In cases where patients had received antibiotic treatment prior to death, the bacterial culture at PM was negative. However, using the 16S sequencing technique the bacteria was identified in all cases. This is advantageous compared to bacterial culture which requires viable bacteria, as occasionally when infants arrive at the emergency department they are administered with IV antibiotics preventing bacterial growth on culture, as was demonstrated in this study. These findings support other published work that has shown the 16S method is more useful in identifying causative organisms in blood from septic infants and synovial fluid from children with suspected musculoskeletal infections [91–93], and assisting in clinical management decisions following antibiotic administration [172]. Other studies have shown similar cases where bacterial culture has been negative but positive when using 16S PCR [173].

This study has shown that the traditional culture method is sometimes incapable of identifying bacteria responsible for fatal infection in PM cases where histological evidence is unclear. Unclear histological changes could be due to rapid deterioration of the individual with no time for immune response, or the unidentified

immunocompromised status of a vulnerable infant or child. Using current techniques this would leave the infection undetected and therefore the death would present as SIDS. Using the differences observed in the control groups to establish a threshold for background sequencing, the 16S rRNA method was able to identify a dominant bacterium in all infectious cases and reveal a subset of currently unexplained cases that show overwhelming bacterial presence raising the possibility of infection as the cause of death.

This study identified four cases of unexplained SUDIC where the 16S rRNA gene sequencing results were suggestive of infection. In the three cases where clinical history was available, all infants exhibited symptoms of infectious disease such as a cough, fever, irritability, and loss of appetite prior to death which are also common observations in SIDS. All infants also died during a period of sleep another characteristic of SIDS deaths. In two cases where bacterial culture of the blood, spleen, and CSF was negative, *Enterococcus* (uC1) and Enterobacteriaceae (uC2) members dominated positive tissue samples. *Enterococcus* is known to cause infections in immunocompromised newborns and typically is associated with nosocomial bacteraemia although community acquired Enterococcal infection has been reported [174]. It has been reported to be a cause of SUDI [175]. A study investigating *Enterococcus faecium* and *Enterococcus faecalis* in blood of newborns found all blood cultures to be negative but positive using PCR methods. Fernandez *et al.* also reported 95 episodes of *Enterococcus faecalis* bacteraemia, all of which were associated with hospitalisation and catheter use [86]. Enterobacteriaceae is a large family of Gram-negative bacteria that consists of many human pathogens such as *Klebsiella*, *Escherichia*, *Enterobacter* and *Salmonella*. The technique was unable to classify this down to genus level, but *Escherichia*, *Klebsiella*, and *Enterobacter* have all been known to play a role in SIDS [86].

In uC4, *Clostridium* was identified at >99% in all frozen tissues sampled. Interestingly, at PM investigation *Clostridium sordelli* was cultured from the lung, the spleen, and swabs collected from both auditory canals. However, the identification of these organisms at PM was not considered relevant to the cause of

death and rather a product of the PM process. Given that *Clostridium* species are strict anaerobes, they have been found to undergo extremely rapid proliferation following death in studies of the thanatomicrobiome [128, 176–178], meaning the PM finding of *Clostridium* in the absence of other evidence is often interpreted as contamination. However, in this study *Clostridium* was not found to be of dominant presence in any deaths of known non-infectious cause. However, the identification of *Clostridium sordelli* in these tissues at PM investigation together with its dominance in all tissue samples as identified in this study suggest that it may in fact be relevant to the cause of death.

Interestingly, previous studies performing comprehensive autopsy reviews found that *Staphylococcus aureus* and *Escherichia coli* were cultured more often in PM tissues from SIDS cases than those from non-infectious deaths [58, 59]. Another study investigating *Staphylococcus aureus* prevalence in faecal material from SIDS infants found higher rate of detection using molecular methods in SIDS than in healthy live controls [65]. Similar results were observed in the present study where *Staphylococcus aureus* and CoNS were identified in a higher percentage of unexplained deaths than explained non-infectious deaths using bacterial culture methods as part of routine PM investigation. However the higher detection rate was not significant. Using the 16S sequencing method which is less biased, *Staphylococcus aureus* and CoNS were not identified in any PM tissues from the study cohort. The 16S method was performed on frozen heart, kidney, liver, muscle, and spleen, where as the bacterial culture was performed on blood, lung, spleen, and CSF so direct comparison could only be performed on spleen samples. *Staphylococcus* is well recognised skin coloniser and therefore high rates of contamination. It is also readily cultivable and so likely to outgrow other bacteria. Coagulase-negative staphylococci (CoNS) has a contamination rate between 62-82% in blood culture [179, 180].

To conclude, the results of this chapter have demonstrated that 16S rRNA gene sequencing can successfully identify infected tissues in PM tissues from SUDIC cases and could enhance the PM investigation of infection as a cause of SUDIC.

Discussion

Despite years of clinical research and medical advances, the cause of many SUDIC cases remains unknown. There is considerable evidence to suggest that a number of currently unexplained deaths may be caused by infection that remains unidentified due to current PM investigation methods [59]. Classical bacterial culture is the gold standard technique, which is associated with both methodical limitations due to the need for specific bacteria targeting and difficulty in determining the relevance of findings due to curation of only qualitative results [181]. This thesis aimed to investigate the potential use of 16S rRNA gene sequencing in PM tissues to enhance the diagnosis of bacterial infection in SUDIC. To achieve this, a 16S rRNA gene sequencing method was optimised for use on both frozen and FFPE PM tissues as it provides a less bias, quantitative alternative to current methods. The effects of collecting tissue samples from an individual PM were also assessed to gain a deeper understanding of how this may affect PM microbiology results generated in this study.

6.1 The clinical relevance of PM microbiology results

During routine PM microbiological investigations, the interpretation of bacteriology results is difficult due to the multiple sources of potential contamination. One such origin of contamination is PM bacterial translocation. It is believed that PM bacterial translocation occurs after death and therefore identification of bacteria may reflect bacterial colonisation and overgrowth occurring, rather than infection, as the cause of death [86, 119, 181]. Forensic studies have shown that under ambient conditions PM translocation may occur as soon as five minutes

PM [116, 117]. Therefore, if bacteria are identified at autopsy, it is often difficult to determine their clinical relevance. Pragmatic approaches have been suggested, with a monomicrobial growth more likely to be considered significant, especially if identified at multiple sites [79]. However, this is based on limited evidence and does not account for the possible differential effects of the cadaver being stored in refrigerated conditions.

In Chapter 3, PM bacterial translocation was investigated in two animal models free of known infection using 16S rRNA gene sequencing. A majority of PM tissues collected over the 14-day study period from the mouse model were negative for the bacterial 16S rRNA gene. In the piglet model, the 16S rRNA gene was detected in a higher percentage of PM tissue samples. In these samples, communities of mixed bacterial flora were identified and generated bacterial sequencing reads were low suggesting low levels of bacterial presence. These results suggest that PM bacterial translocation is minimal in these two animal models.

The finding of minimal PM bacterial translocation was echoed in the investigation of PM tissues collected from cases of SUDIC performed in Chapter 5. In cases of SUDIC with an explained non-infectious cause, a large proportion of PM samples were negative for the 16S rRNA gene. In those where the bacterial 16S gene was identified, a low number of sequencing reads and no bacterial dominance was observed. The finding of mixed bacterial flora in SUDIC with a non-infectious cause is in agreement with the results of the animal model study performed in Chapter 3, as well as other studies which have stated that mixed communities of physiological flora is characteristic of PM bacterial translocation [115, 182, 183]. It is also believed that an increasing post-mortem interval is associated with PM microbial changes which may affect PM bacterial translocation [163] and an increased difficulty in interpreting PM microbiology results. However, in the present study the PMI had no impact on the interpretation of sequencing results. This study has provided a greater insight to the PM process and the ways in which these conditions impact bacterial 16S rRNA gene sequencing results. With this information, performing 16S rRNA gene sequencing on PM tissues would allow a

more informed interpretation of PM results and therefore an improved diagnosis of infection in SUDIC.

Sample contamination that may occur during tissue collection at autopsy and subsequent processing is another concern when interpreting PM microbiology results [86]. In the study of PM bacterial translocation in animal models in Chapter 3, sampling was performed in a sterile manner with careful consideration to limit sample contamination. In Chapter 5, the PM tissues from SUDIC cases were collected at autopsy and therefore collection in a sterile manner was not guaranteed. However, even in this 'real-world' setting, notable sample contamination was not observed and therefore did not impact the interpretation of sequencing results.

6.2 16S rRNA gene sequencing and its potential role in clinic

The 16S rRNA gene sequencing technique is a more universal and less biased method than current routine bacterial culture and the results of this study suggest it may be advantageous for the identification of ante-mortem infection in cases of SUDIC. Bacterial culture is very sensitive and requires viable bacterium to obtain a positive culture. Therefore, if a highly cultivable contaminant is present at low levels it is still likely to to be cultured under laboratory conditions [184]. Using 16S rRNA gene sequencing the community of bacteria can be assessed with information on bacteria present and their abundance in relation to one another. Another advantage of this molecular method is its ability to detect the presence of bacterial genes following administration of antibiotics [172]. As bacterial culture requires viable bacteria for identification, if a patient has been administered antibiotics prior to death, as is often the case with SUDIC, then a negative culture will be obtained [185]. In cases of SUDIC, this bacteriological evidence may be the difference between confirming infection as the cause of death or classifying the death as SIDS.

In order to provide evidence for the use of the 16S rRNA gene sequencing method in SUDIC, this method was optimised for using in FFPE material. FFPE is a form of tissue preservation that allows tissue staining for histological examination. A large proportion of archived material is stored as FFPE blocks and given this was a retrospective study it was important to optimise the DNA extraction method for application of the 16S rRNA gene sequencing technique. In Chapter 4, DNA extraction from FFPE material was successfully optimised to obtain DNA on which the 16S rRNA gene sequencing technique could be performed. In cases of SUDIC with an infectious cause of death, the causative organism was identified. However, the results obtained from this material were not as clear as those from frozen PM tissues. Therefore, frozen PM material should be used where available.

When analysing results from frozen PM material, it was important to define simple criteria which could be used to filter out low-level bacterial presence that was not representative of a true infection. The basis of the criteria is shown in Figure 6.1. The ability of these simple criteria to successfully identify cases of infectious SUDIC demonstrate that this method could be used to diagnose infections PM. Setting criteria that must be met for a tissue to be considered infectious facilitates the ease of interpretation by pathologists. The results require less interpretation than routine bacterial culture and could allow a move towards a more standardised approach across institutions.

The 16S rRNA gene sequencing technique successfully detected the likely bacterial pathogen in all cases of infection where the causative bacterium was identified at routine PM investigation, as well as in cases where the infection-causing bacterium was not identified using routine bacterial culture. The technique had 100% specificity and 100% sensitivity when performed on cases of SUDIC caused by systemic infection in this study. However, for real-world use, prospective evaluation would be required to include a wide range of clinical settings and causes of death, including several different infectious causes, in order to determine true test performance for diagnostic use.

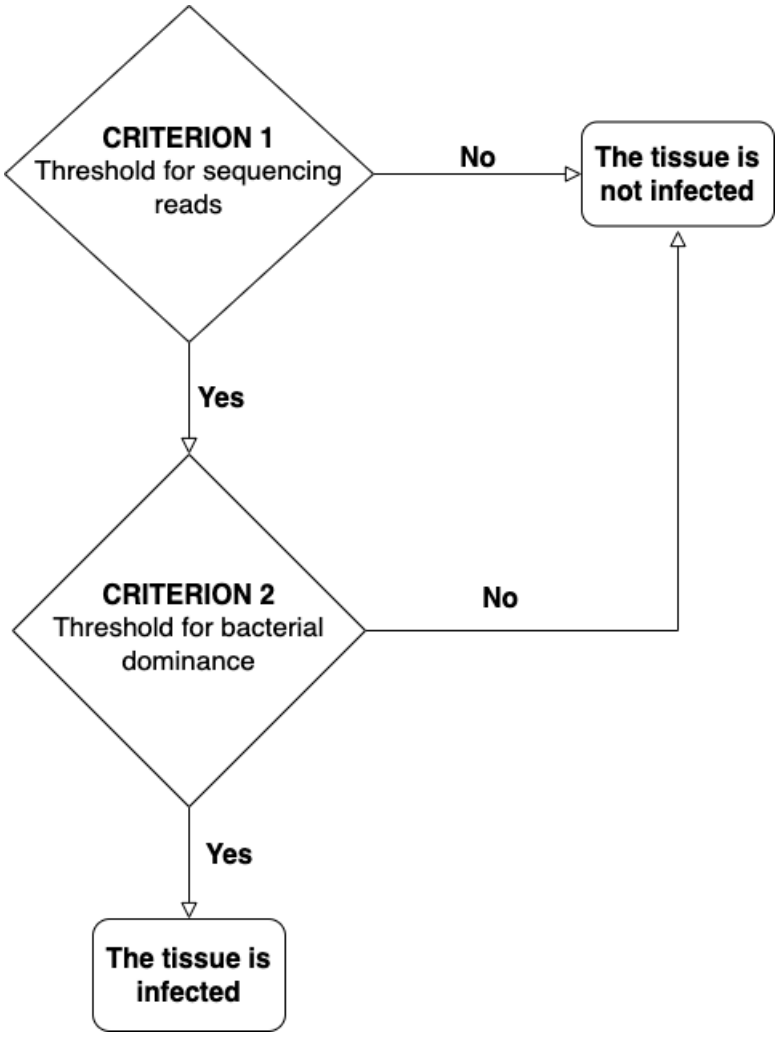

FIGURE 6.1: Criteria established to identify infected PM tissues. These criteria can be modified based on sequencing output. Setting thresholds enables simple sorting of cases to determine whether the tissue is infected

This study has been unable to provide evidence for the use of this technique in identifying potential respiratory infections resulting in SUDIC. This is likely due to the unavailability of frozen PM lung tissue samples. Provided these samples are available, it is expected that the application of this technique will enable identification of the bacterial pathogen in cases of respiratory infection.

6.3 Infection as a cause of SIDS

The development of a more consistent and reliable technique to identify and diagnose infection in the PM setting is important for a more accurate picture of infection that cause SUDIC, vulnerable infants, and better research into SIDS. Many researchers have hypothesised that infection is responsible for more SUDIC cases than currently diagnosed [36, 53, 136, 186], which is supported by results of this study. Application of the 16S rRNA gene sequencing technique to currently unexplained cases of SUDIC highlighted several cases of SUDIC that may have been caused by infection, all of which also displayed clinical symptoms suggestive of infection prior to death. These additional PM findings may provide pathologists with the evidence they require to confirm infection as the CoD.

Figure 6.2 shows the proportion of deaths that were explained non-infectious, infectious, and unexplained at (A) routine PM investigation and (B) during this study. The use of this sequencing technique on PM tissues could reduce the proportion of unexplained deaths if these findings are confirmed in larger prospective studies. By identifying a cause of death and reducing the rates of unexplained cases, the aetiology of SUDIC can be better understood. It is also important to reduce the burden of unexplained death on families by offering a diagnosis, reducing potential feelings of guilt and self-blame.

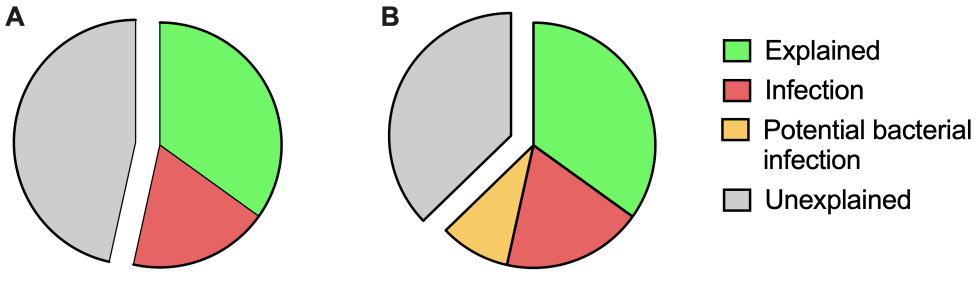

FIGURE 6.2: Pie charts demonstrating the final CoD diagnosis (A) using current routine methods and (B) using the 16S rRNA gene sequencing technique. The use of this technique reveals potential bacterial infections in cases of SUDIC which reduces the number of SUDIC deaths that remain unexplained.

6.4 Study limitations

Although this study has identified bacterial infections, it is important to note that this assay is targeted to bacterial pathogens and is unable to identify infections of viral or fungal origin. Viral infections are known to contribute to SUDIC and therefore should not be overlooked during routine PM investigation. Nevertheless, molecular methods are also applicable to detection of viruses, for example the use of targeted qPCR techniques which are currently used in routine investigation of SUDIC. However, similar to targeting bacteria using qPCR, these assays require targeting of specific viruses and therefore risk missing potential viral pathogens.

Whilst this study has provided evidence that supports the use of this technique in diagnosis of bacterial infection, this technique is not a stand-alone diagnostic technique. Results obtained from this type of investigation require careful consideration along with results obtained from other PM diagnostic tests, both those that are currently used and perhaps techniques such as cytokine profiling which are not yet clinically available.

6.5 Conclusions

To conclude, this study has firstly provided a greater depth of knowledge regarding the PM microbiological process under clinical conditions. This information could assist in the interpretation of microbiology results obtained from PM samples. In this study a universal, unbiased assay has been developed and the definition of simple criteria has shown that this technique could be used in routine practice to identify ante-mortem infection. This has the potential to develop the understanding of the role of infection in SUDIC.

This study has the potential to contribute to the understanding of the role of infection in SUDIC and there are several other avenues that can be explored. These will be discussed in Chapter 7.

Future work

In order to develop a better understanding of the practical use of the 16S rRNA gene sequencing technique in routine diagnostic practice, further investigation on more samples is required. Therefore, the next step in this investigation would be to perform a prospective study (outlined in Figure 7.1 and explained below).

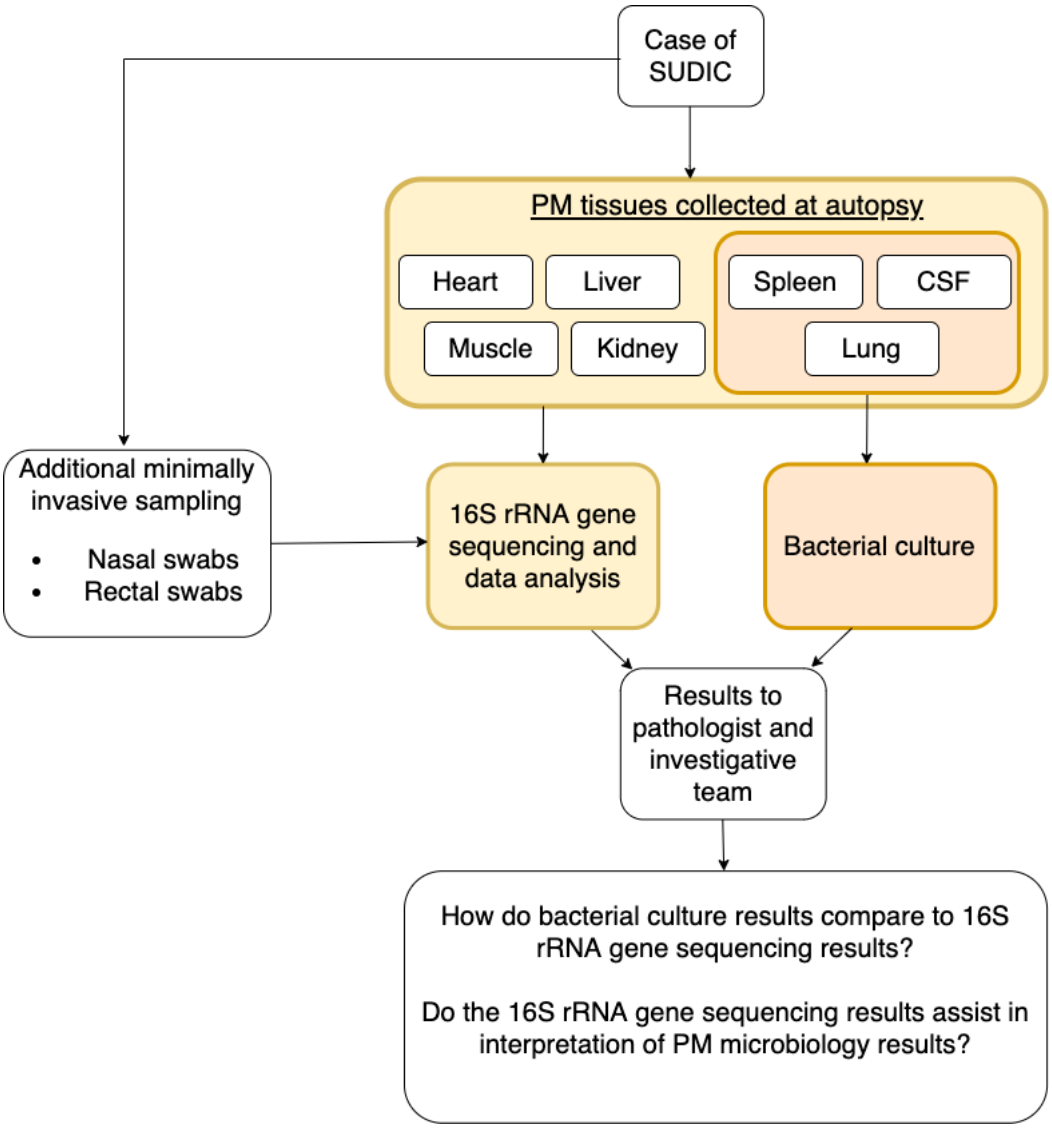

FIGURE 7.1: Outline of prospective study

The prospective study would consist of performing the 16S rRNA gene sequencing technique at the time of PM investigation alongside the current routine bacteriology procedures. In this study, as only a small amount of tissue is required to perform 16S rRNA gene sequencing, PM tissues that are collected during routine autopsy could be utilised and therefore no additional tissue sampling at autopsy would be required. This information would then be presented, together with all other evidence from the investigation, to the investigating professionals in order to evaluate whether this information assists with the interpretation of PM microbiology results. These results would also allow a direct comparison between results obtained from routine bacterial culture and 16S rRNA gene sequencing on PM spleen, lung, and CSF samples. In addition to this comparison, species-specific PCRs could be performed in cases where a dominant bacteria is identified in PM tissues. As the 16S rRNA gene sequencing method described in this thesis cannot identify bacteria to species level, a hybrid approach may prove beneficial and assist further in determining the relevance of results.

A prospective study would also allow the investigation of the use of the 16S rRNA gene sequencing technique on lung tissue to identify potential cases of respiratory infection. As lung tissue is routinely collected for microbiological investigation, the sequencing technique could be applied on these samples since they were missing from the current study. This would be performed as described above. As URT infections are common in this demographic, it would be important to investigate whether these transient infections are prevalent in deaths with a confirmed, non-infectious cause of death to ensure that they are not misinterpreted as being causative of death. Should 16S rRNA gene sequencing results suggestive of infection be identified in cases with confirmed non-infectious, confirmed infectious, and unexplained causes of death a nasal swab may assist in identifying whether this bacterial infection is present in the upper or lower respiratory tract which may help pathologists determine where the infection lies and therefore how relevant it is to the cause of death.

Another avenue that could be explored is the composition of the gut microbiome,

which could be achieved using rectal swabs. These provide a minimally invasive collection technique which may provide indicators of ante-mortem illness. Dysbiosis of the gut microbiota has been associated with many disease states and diversity measures of the gut microbiota have been used in a forensic study to successfully identify the manner of death [163]. An increase in the members of the Proteobacteria phylum, to which many human pathogens belong, has also been associated with infectious illness. Therefore, investigating the composition of the gut microbiota at autopsy could act as a biomarker for disease in cases of SUDIC.

Should these further investigations provide evidence to support the use of this technique in routine PM diagnostics for the investigation of SUDIC, there will be a need for an automated and robust bioinformatic pipeline. Ideally this data processing and visualisation would allow pathologists to easily access, view, and interpret results. As these would always be processed and presented using the same pipeline, this would reduce interpretation bias.

A limitation of the current study was that the sequencing of short regions of the 16S rRNA gene meant that bacteria could only be classified to genus level. This may be critical to the interpretation of whether the bacteria identified caused fatal infection. With the ongoing technological advancements in third generation sequencing techniques, new sequencing technologies are emerging that are able to sequence longer fragments of DNA potentially enabling the whole 16S rRNA gene to be sequenced. Technology such as the MinION by Oxford Nanopore Technologies is an accessible, relatively low-cost sequencing device that can sequence DNA strands greater than 100 kilobases in length, compared to the MiSeq which sequences short paired-end reads spanning just 250bps. Using this technology to sequence the entire 16S rRNA would allow classification of identified bacteria to species level which may assist in determining the likelihood of the identified bacteria causing a fatal infection. For these emerging technologies, automated bioinformatic pipelines have also been created which could further assist in implementing this method into routine clinical diagnostics.

Finally, rather than targeted amplicon sequencing of the 16S rRNA gene which can only detect bacteria, shotgun metagenomic sequencing could be performed to investigate all components within the PM tissue. Shotgun metagenomic sequencing allows sequencing of all genomic DNA in a sample, which means that bacteria, fungi, viruses and other microorganisms can be investigated simultaneously. This would allow assessment of not only the role of bacteria, but also other microorganisms.

To conclude, this thesis has shown that 16S rRNA gene sequencing may assist in the PM investigation of SUDIC. However, further validation of the method is required before this could be introduced into routine PM diagnostics.

Bibliography

[1] J. Bruce Beckwith. Defining the sudden infant death syndrome. *Archives of Pediatrics and Adolescent Medicine*, 2003. ISSN 10724710. doi: 10.1001/archpedi.157.3.286.

[2] C. E. A. Leach, P. S. Blair, P. J. Fleming, I. J Smith, M. W. Platt, P. J. Berry, B. FRCP, J. Golding, and the CESDI SUDI Research Group. Epidemiology of SIDS and Explained Sudden Infant Deaths. *Pediatrics*, 1999. ISSN 0031-4005. doi: 10.1542/peds.104.4.e43.

[3] Martin Ward Platt, Peter S. Blair, Peter J. Fleming, Iain J. Smith, Tim J. Cole, Charlotte E.A. Leach, P. J. Berry, and Jean Golding. A clinical comparison of SIDS and explained sudden infant deaths: How healthy and how normal? *Archives of Disease in Childhood*, 2000. ISSN 00039888. doi: 10.1136/adc.82.2.98.

[4] Office for National Statistics. Unexpected deaths in infancy, England and Wales (dataset). Technical report, Office of National Statistics, 2020. URL `https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/unexplaineddeathsininfancyenglandandwalesunexpecteddeaths`.

[5] John Bruce Beckwith. Discussion of terminology and defi nition of the sudden infant death syndrome. In *Sudden infant death syndrome: proceedings of the second international conference on the causes of sudden death in infants*, 1970.

[6] Office for National Statistics. Unexpected deaths in infancy, England and Wales (Dataset). Technical report, 2017.

[7] Torleiv Ole Rognum, Åshlid Vege, Arne Stray-Pedersen, and Lillian Bøylestad. A Scandanavian Perspective. In J.R. Duncan and R.W. Byard,

editors, *SIDS Sudden Infant and Early Childhood Death: The Past, the Present and the Future.*, chapter 20. University of Adelaide Press, 2018.

[8] Centers for Disease Control and Prevention. Trends in Sudden Unexpected Infant Death by Cause, 1990–2019, 2021. URL `https://www.cdc.gov/sids/data.htm`.

[9] Jacqueline Müller-Nordhorn, Alice Schneider, Ulrike Grittner, Konrad Neumann, Thomas Keil, Stefan N. Willich, and Sylvia Binting. International time trends in sudden unexpected infant death, 1969-2012. *BMC Pediatrics*, 2020. ISSN 14712431. doi: 10.1186/s12887-020-02271-x.

[10] Do baby boxes really save lives? *BBC*, 2017. URL `https://www.bbc.co.uk/news/magazine-39366596`.

[11] Helena Lee. Why Finnish Babies Sleep in Boxes. *BBC News*, pages 1–6, 2013. URL `http://www.bbc.co.uk/news/magazine-22751415`.

[12] Lucia Binding. Experts raise concerns over safety of baby boxes, 2018.

[13] Barry J. Taylor, Joanna Garstang, Adele Engelberts, Toshimasa Obonai, Aurore Cote, Jane Freemantle, Mechtild Vennemann, Matt Healey, Peter Sidebotham, Edwin A. Mitchell, and Rachel Y. Moon. International comparison of sudden unexpected death in infancy rates using a newly proposed set of cause-of-death codes. *Archives of Disease in Childhood*, 2015. ISSN 14682044. doi: 10.1136/archdischild-2015-308239.

[14] Peter Sidebotham, Francine Bates, Catherine Ellis, and Lucy Lyus. Preventive Strategies for Sudden Infant Death Syndrome. In Jhodie R. Duncan and Roger W. Byard, editors, *SIDS Sudden Infant and Early Childhood Death: The Past, the Present and the Future*, chapter 12. University of Adelaide Press, Adelaide, 2018.

[15] Office for National Statistics. Infant and perinatal mortality 2000: health areas. Technical report, Office for National Statistics, 2000.

[16] Office for National Statistics. Main tables: Unexplained deaths in infancy, England and Wales 2018 (dataset). Technical report, Office for National Statistics, 2020.

[17] Hazel Brooke, Angus Gibson, David Tappin, and Helen Brown. Case-control study of sudden infant death syndrome in Scotland, 1992-5. *British Medical Journal*, 1997. ISSN 09598146. doi: 10.1136/bmj.315.7111.812a.

[18] P. J. Fleming, R. Gilbert, Y. Azaz, P. J. Berry, P. T. Rudd, A. Stewart, and E. Hall. Interaction between bedding and sleeping position in the sudden infant death syndrome: A population based case-control study. *British Medical Journal*, 1990. ISSN 09598146. doi: 10.1136/bmj.301.6743.85.

[19] A. N. Stanton. Overheating and Cot Death. *The Lancet*, 324(8413):1199–1201, 1984. ISSN 01406736. doi: 10.1016/S0140-6736(84)92753-3.

[20] E. A.S. Nelson, B. J. Taylor, and I. L. Weatherall. Sleeping Position and Infant Bedding May Predispose To Hyperthermia and the Sudden Infant Death Syndrome, 1989. ISSN 01406736.

[21] P. J. Berry. Pathological findings in SIDS. In *Journal of Clinical Pathology*, 1992.

[22] The Lullaby Trust. The safest room temperature for babies. URL `https://www.lullabytrust.org.uk/safer-sleep-advice/baby-room-temperature/`.

[23] Patricia M. Dietz, Lucinda J. England, Carrie K. Shapiro-Mendoza, Van T. Tong, Sherry L. Farr, and William M. Callaghan. Infant Morbidity and Mortality Attributable to Prenatal Smoking in the U.S. *American Journal of Preventive Medicine*, 2010. ISSN 07493797. doi: 10.1016/j.amepre.2010.03.009.

[24] Tatiana M. Anderson, Juan M. Lavista Ferres, Shirley You Ren, Rachel Y. Moon, Richard D. Goldstein, Jan Marino Ramirez, and Edwin A. Mitchell.

Maternal smoking before and during pregnancy and the risk of sudden unexpected infant death. *Pediatrics*, 2019. ISSN 10984275. doi: 10.1542/peds.2018-3325.

[25] R Scragg, E A Mitchell, B J Taylor, A W Stewart, R P Ford, J M Thompson, E M Allen, and D M Becroft. Bed sharing, smoking, and alcohol in the sudden infant death syndrome. New Zealand Cot Death Study Group. *BMJ*, 1993. ISSN 0959-8138. doi: 10.1136/bmj.307.6915.1312.

[26] J. Schellscheidt, N. Øyen, and G. Jorch. Interactions between maternal smoking and other prenatal risk factors for sudden infant death syndrome (SIDS). *Acta Paediatrica, International Journal of Paediatrics*, 1997. ISSN 08035253. doi: 10.1111/j.1651-2227.1997.tb08612.x.

[27] Kui Zhang and Xianmin Wang. Maternal smoking and increased risk of sudden infant death syndrome: A meta-analysis, 2013. ISSN 13446223.

[28] Tushar Shah, Kevin Sullivan, and John Carter. Sudden infant death syndrome and reported maternal smoking during pregnancy. *American Journal of Public Health*, 2006. ISSN 00900036. doi: 10.2105/AJPH.2005.073213.

[29] M. M. Vennemann, T. Bajanowski, B. Brinkmann, G. Jorch, K. Yücesan, C. Sauerland, and E. A. Mitchell. Does breastfeeding reduce the risk of sudden infant death syndrome? *Pediatrics*, 2009. ISSN 00314005. doi: 10.1542/peds.2008-2145.

[30] Fern R. Hauck, John M.D. Thompson, Kawai O. Tanabe, Rachel Y. Moon, and Mechtild M. Vennemann. Breastfeeding and reduced risk of sudden infant death syndrome: A meta-analysis, 2011. ISSN 00314005.

[31] John M.D. Thompson, Kawai Tanabe, Rachel Y. Moon, Edwin A. Mitchell, Cliona McGarvey, David Tappin, Peter S. Blair, and Fern R. Hauck. Duration of breastfeeding and risk of SIDS: An individual participant data meta-analysis, 2017. ISSN 10984275.

[32] Carl E. Hunt. The Cardiorespiratory Control Hypothesis for Sudden Infant Death Syndrome. *Clinics in Perinatology*, 19(4):757–771, 1992.

[33] E. A. Phillipson and C. E. Sullivan. Arousal: The forgotten response to respiratory stimuli, 1978. ISSN 00030805.

[34] Carmel Harrington, Turkka Kirjavainen, Arthur Teng, and Colin E. Sullivan. Altered autonomic function and reduced arousability in apparent life-threatening event infants with obstructive sleep apnea. *American Journal of Respiratory and Critical Care Medicine*, 2002. ISSN 1073449X. doi: 10.1164/ajrccm.165.8.2102059.

[35] James S. Kemp and Bradley T. Thach. Sudden Death in Infants Sleeping on Polystyrene-Filled Cushions. *New England Journal of Medicine*, 1991. ISSN 0028-4793. doi: 10.1056/nejm199106273242605.

[36] Jane Blood-Siegfried. The role of infection and inflammation in sudden infant death syndrome, 2009. ISSN 08923973.

[37] James J. Filiano and Hannah C. Kinney. A perspective on neuropathologic findings in victims of the sudden infant death syndrome: The triple-risk model. *Neonatology*, 1994. ISSN 16617819. doi: 10.1159/000244052.

[38] David S. Newburg and W. Allan Walker. Protection of the neonate by the innate immune system of developing gut and of human milk, 2007. ISSN 00313998.

[39] Barbara M. Ostfeld, Ofira Schwartz-Soicher, Nancy E. Reichman, Julien O. Teitler, and Thomas Hegyi. Prematurity and sudden unexpected infant deaths in the United States. *Pediatrics*, 2017. ISSN 10984275. doi: 10.1542/peds.2016-3334.

[40] Charles A. Janeway, Paul Travers, Mark Walport, and Mark Shlomchick. *Immunobiology: The Immune System in Health and Disease*. Garland Science, New York, 6 edition, 2005.

[41] Laura E. Carr, Misty D. Virmani, Fernanda Rosa, Daniel Munblit, Katelin S. Matazel, Ahmed A. Elolimy, and Laxmi Yeruva. Role of Human Milk Bioactives on Infants' Gut and Immune Health, 2021. ISSN 16643224.

[42] Neil Molony, C. Caroline Blackwell, and Anthony Busuttil. The effect of prone posture on nasal temperature in children in relation to induction of staphylococcal toxins implicated in sudden infant death syndrome. *FEMS Immunology and Medical Microbiology*, 1999. ISSN 09288244. doi: 10.1016/S0928-8244(99)00078-4.

[43] Linda M. Harrison, James A. Morris, David R. Telford, Susan M. Brown, and Keith Jones. The nasopharyngeal bacterial flora in infancy: Effects of age, gender, season, viral upper respiratory tract infection and sleeping position. *FEMS Immunology and Medical Microbiology*, 1999. ISSN 09288244. doi: 10.1016/S0928-8244(99)00068-1.

[44] Linda M. Harrison, James A. Morris, Lisa A. Bishop, Robert M. Lauder, Christine A.M. Taylor, and David R. Telford. Detection of specific antibodies in cord blood, infant and maternal saliva and breast milk to staphylococcal toxins implicated in sudden infant death syndrome (SIDS). *FEMS Immunology and Medical Microbiology*, 2004. ISSN 09288244. doi: 10.1016/j.femsim.2004.06.010.

[45] C. Caroline Blackwell, Doris A.C. MacKenzie, Valerie S. James, Robert A. Elton, Abdulaziz A. Zorgani, Donald M. Weir, and Anthony Busuttil. Toxigenic bacteria and sudden infant death syndrome (SIDS): Nasopharyngeal flora during the first year of life. *FEMS Immunology and Medical Microbiology*, 1999. ISSN 09288244. doi: 10.1016/S0928-8244(99)00071-1.

[46] James A. Morris. The common bacterial toxins hypothesis of sudden infant death syndrome. *FEMS Immunology and Medical Microbiology*, 1999. ISSN 09288244. doi: 10.1016/S0928-8244(99)00067-X.

[47] J. A. Morris, D. Haran, and A. Smith. Hypothesis: Common bacterial toxins are a possible cause of the sudden infant death syndrome. *Medical Hypotheses*, 1987. ISSN 03069877. doi: 10.1016/0306-9877(87)90145-9.

[48] Paul Nathan Goldwater. SIDS, prone sleep position and infection: An overlooked epidemiological link in current SIDS research? Key evidence for the "Infection Hypothesis". *Medical Hypotheses*, 2020. ISSN 15322777. doi: 10.1016/j.mehy.2020.110114.

[49] Paul Nathan Goldwater. Sudden infant death syndrome, infection, prone sleep position, and vagal neuroimmunology. *Frontiers in Pediatrics*, 2017. ISSN 22962360. doi: 10.3389/fped.2017.00223.

[50] David Tappin, Hazel Brooke, Russell Ecob, and Angus Gibson. Used infant mattresses and sudden infant death syndrome in Scotland: Case-control study. *British Medical Journal*, 2002. ISSN 09598146. doi: 10.1136/bmj.325.7371.1007.

[51] Richard E. Sherburn and Richard O. Jenkins. Cot mattresses as reservoirs of potentially harmful bacteria and the sudden infant death syndrome. *FEMS Immunology and Medical Microbiology*, 2004. ISSN 09288244. doi: 10.1016/j.femsim.2004.06.011.

[52] Richard O. Jenkins and Richard E. Sherburn. Growth and survival of bacteria implicated in sudden infant death syndrome on cot mattress materials. *Journal of Applied Microbiology*, 2005. ISSN 13645072. doi: 10.1111/j.1365-2672.2005.02620.x.

[53] Caroline Blackwell, Sophia Moscovis, Sharron Hall, Christine Burns, and Rodney J. Scott. Exploring the risk factors for sudden infant deaths and their role in inflammatory responses to infection, 2015. ISSN 16643224.

[54] R.W. Byard and L. Moore. Can thymic pefechlae be used to separate slds infants from controls? *Pathology*, 1993. ISSN 00313025. doi: 10.1016/s0031-3025(16)35738-5.

[55] Roger W Byard. The Autopsy and Pathology of Sudden Infant Death Syndrome. In *SIDS Sudden Infant and Early Childhood Death: The Past, the Present and the Future.*, chapter 24. University of Adelaide Press, 2018.

[56] Paul Nathan Goldwater. Intrathoracic petechial hemorrhages in sudden infant death syndrome and other infant deaths: Time for re-examination? *Pediatric and Developmental Pathology*, 2008. ISSN 10935266. doi: 10.2350/08-01-0404.1.

[57] Y. Aoki. Histopathological findings of the lung and trachea in sudden infant death syndrome: Review of 105 cases at Dade County Medical Examiner Department. *Japanese Journal of Legal Medicine*, 1994. ISSN 00471887.

[58] Paul Nathan Goldwater. Sterile site infection at autopsy in sudden unexpected deaths in infancy. *Archives of Disease in Childhood*, 2009. ISSN 00039888. doi: 10.1136/adc.2007.135939.

[59] MA Weber, NJ Klein, JC Hartley, PE Lock, M. Malone, and NJ Sebire. Infection and sudden unexpected death in infancy: a systematic retrospective case review. *The Lancet*, 2008. ISSN 01406736. doi: 10.1016/S0140-6736(08)60798-9.

[60] R. Gilbert, P. Rudd, P. Jeremy Berry, P. J. Fleming, E. Hall, D. G. White, V. O.C. Oreffo, P. James, and J. A. Evans. Combined effect of infection and heavy wrapping on the risk of sudden unexpected infant death. *Archives of Disease in Childhood*, 1992. ISSN 00039888. doi: 10.1136/adc.67.2.171.

[61] J. E. Malam, G. F. Carrick, D. R. Telford, and J. A. Morris. Staphylococcal toxins and sudden infant death syndrome. *Journal of Clinical Pathology*, 1992. ISSN 00219746. doi: 10.1136/jcp.45.8.716.

[62] D. R. Telford, J. A. Morris, P. Hughes, A. R. Conway, S. Lee, A. J. Barson, and D. B. Drucker. The nasopharyngeal bacterial flora in the sudden infant death syndrome. *Journal of Infection*, 1989. ISSN 01634453. doi: 10.1016/S0163-4453(89)91094-3.

[63] Abdulaziz Zorgani, Stephen D. Essery, Osama Al Madani, Alastair J. Bentley, Valerie S. James, Doris A.C. MacKenzie, Jean W. Keeling, Caroline Rambaud, John Hilton, C. Caroline Blackwell, Donald M. Weir, and Anthony Busuttil. Detection of pyrogenic toxins of Staphylococcus aureus in sudden infant death syndrome. *FEMS Immunology and Medical Microbiology*, 1999. ISSN 09288244. doi: 10.1016/S0928-8244(99)00077-2.

[64] Melanie J. Newbould, Jacqui Malam, Joanna M. McIllmurray, J. A. Morris, D. R. Telford, and A. J. Barson. Immunohistological localisation of staphylococcal toxic shock syndrome toxin (TSST-1) antigen in sudden infant death syndrome. *Journal of Clinical Pathology*, 1989. ISSN 00219746. doi: 10.1136/jcp.42.9.935.

[65] Amanda R. Highet and Paul N. Goldwater. Staphylococcal enterotoxin genes are common in Staphylococcus aureus intestinal flora in Sudden infant death syndrome (SIDS) and live comparison infants. *FEMS Immunology and Medical Microbiology*, 2009. ISSN 09288244. doi: 10.1111/j.1574-695X.2009.00592.x.

[66] Karin Helweg-Larsen, J. B. Lundemose, N. Øyen, R. Skjærven, B. Alm, G. Wennergren, T. Markestad, and L. M. Irgens. Interactions of infectious symptoms and modifiable risk factors in sudden infant death syndrome. The nordic epidemiological SIDS study. *Acta Paediatrica, International Journal of Paediatrics*, 1999. ISSN 08035253. doi: 10.1080/08035259950169521.

[67] C. Caroline Blackwell, Doris A.C. MacKenzie, Valerie S. James, Robert A. Elton, Abdulaziz A. Zorgani, Donald M. Weir, and Anthony Busuttil. Toxigenic bacteria and sudden infant death syndrome (SIDS): Nasopharyngeal flora during the first year of life. *FEMS Immunology and Medical Microbiology*, 1999. ISSN 09288244. doi: 10.1016/S0928-8244(99)00071-1.

[68] A. S. Douglas, T. M. Allan, and P. J. Helms. Seasonality and the sudden infant death syndrome during 1987-9 and 1991-3 in Australia and Britain.

British Medical Journal, 1996. ISSN 09598146. doi: 10.1136/bmj.312.7043.1381a.

[69] Carrie Shapiro-Mendoza, Sharyn Parks, Alexa Erck Lambert, Lena Camperlengo, Carri Cottengim, and Christine Olson. The Epidemiology of Sudden Infant Death Syndrome and Sudden Unexpected Infant Deaths: Diagnostic Shift and other Temporal Changes. In *SIDS Sudden infant and early childhood death: The past, the present and the future*. University of Adelaide Press, 2018. ISBN 9781925261677. doi: 10.20851/sids-13.

[70] E. C. Uren, A. L. Williams, I. Jack, and J. W. Rees. Association of respiratory virus infections with sudden infant death syndrome. *Medical Journal of Australia*, 1980. ISSN 0025729X. doi: 10.5694/j.1326-5377.1980.tb134996.x.

[71] Gulhan Yagmur, Nihan Ziyade, Neval Elgormus, Taner Das, M. Feyzi Sahin, Muzaffer Yildirim, Ayse Ozgun, Arzu Akcay, Ferah Karayel, and Sermet Koc. Postmortem diagnosis of cytomegalovirus and accompanying other infection agents by real-time PCR in cases of sudden unexpected death in infancy (SUDI). *Journal of Forensic and Legal Medicine*, 2016. ISSN 18787487. doi: 10.1016/j.jflm.2015.11.008.

[72] S. C.M. de Crom, J. W.A. Rossen, R. A. de Moor, E. J.M. Veldkamp, A. M. van Furth, and C. C. Obihara. Prospective assessment of clinical symptoms associated with enterovirus and parechovirus genotypes in a multicenter study in Dutch children. *Journal of Clinical Virology*, 2016. ISSN 18735967. doi: 10.1016/j.jcv.2016.01.014.

[73] Enagnon Kazali Alidjinou, Marine Dorsi Di Meglio, Antoine Biron, Marion Jeannoel, Isabelle Schuffenecker, and Ann Claire Gourinat. Enterovirus and parechovirus coinfection in a sudden unexpected infant death. *Pediatrics*, 2020. ISSN 10984275. doi: 10.1542/PEDS.2019-3686.

[74] Krow Ampofo, Jeffrey Bender, Xiaoming Sheng, Kent Korgenski, Judy Daly, Andrew T. Pavia, and Carrie L. Byington. Seasonal invasive pneumococcal

disease in children: Role of preceding respiratory viral infection. *Pediatrics*, 2008. ISSN 00314005. doi: 10.1542/peds.2007-3192.

[75] A. G.S.C. Jansen, E. A.M. Sanders, A. Van der Ende, A. M. Van Loon, A. W. Hoes, and E. Hak. Invasive pneumococcal and meningococcal disease: Association with influenza virus and respiratory syncytial virus activity? *Epidemiology and Infection*, 2008. ISSN 09502688. doi: 10.1017/S0950268807000271.

[76] Jeanne Marie Hament, Piet C. Aerts, Andre Fleer, Hans Van Dijk, Theo Harmsen, Jan L.L. Kimpen, and Tom F.W. Wolfs. Direct binding of respiratory syncytial virus to pneumococci: A phenomenon that enhances both pneumococcal adherence to human epithelial cells and pneumococcal invasiveness in a murine model. *Pediatric Research*, 2005. ISSN 00313998. doi: 10.1203/01.pdr.0000188699.55279.1b.

[77] Bart E. Van Ewijk, Tom F.W. Wolfs, Piet C. Aerts, Kok P.M. Van Kessel, Andre Fleer, Jan L.L. Kimpen, and Cornelis K. Van Der Ent. RSV mediates pseudomonas aeruginosa binding to cystic fibrosis and normal epithelial cells. *Pediatric Research*, 2007. ISSN 00313998. doi: 10.1203/pdr.0b013e3180332d1c.

[78] Andreas Handel, Ira M. Longini, and Rustom Antia. Intervention strategies for an influenza pandemic taking into account secondary bacterial infections. *Epidemics*, 2009. ISSN 17554365. doi: 10.1016/j.epidem.2009.09.001.

[79] Victoria Bryant and Neil Sebire. Natural Diseases Causing Sudden Death in Infancy and Early Childhood. In *SIDS Sudden infant and early childhood death: The past, the present and the future*. University of Adelaide Press, 2018. ISBN 9781925261677. doi: 10.20851/sids-25.

[80] Daniel D. Rhoads, Stephen B. Cox, Eric J. Rees, Yan Sun, and Randall D. Wolcott. Clinical identification of bacteria in human chronic wound infections: Culturing vs. 16S ribosomal DNA sequencing. *BMC Infectious Diseases*, 2012. ISSN 14712334. doi: 10.1186/1471-2334-12-321.

[81] Xiaoling Yu, Wenqian Jiang, Yang Shi, Hanhui Ye, and Jun Lin. Applications of sequencing technology in clinical microbial infection, 2019. ISSN 15821838.

[82] Andreas Hiergeist, Joachim Gläsner, Udo Reischl, and André Gessner. Analyses of intestinal microbiota: Culture versus sequencing. *ILAR Journal*, 2015. ISSN 19306180. doi: 10.1093/ilar/ilv017.

[83] Jörg Overmann, Birte Abt, and Johannes Sikorski. Present and Future of Culturing Bacteria. *Annual Review of Microbiology*, 2017. ISSN 0066-4227. doi: 10.1146/annurev-micro-090816-093449.

[84] Sien Ombelet, Barbara Barbé, Dissou Affolabi, Jean-Baptiste Ronat, Palpouguini Lompo, Octavie Lunguya, Jan Jacobs, and Liselotte Hardy. Best Practices of Blood Cultures in Low- and Middle-Income Countries. *Frontiers in Medicine*, 2019. ISSN 2296-858X. doi: 10.3389/fmed.2019.00131.

[85] Jeremy W. Pryce, Sebastian E.A. Roberts, Martin A. Weber, Nigel J. Klein, Michael T. Ashworth, and Neil J. Sebire. Microbiological findings in sudden unexpected death in infancy: Comparison of immediate postmortem sampling in casualty departments and at autopsy. *Journal of Clinical Pathology*, 2011. ISSN 00219746. doi: 10.1136/jcp.2011.089698.

[86] A. Fernández-Rodríguez, J. L. Burton, L. Andreoletti, J. Alberola, P. Fornes, I. Merino, M. J. Martínez, P. Castillo, B. Sampaio-Maia, I. M. Caldas, V. Saegeman, and M. C. Cohen. Post-mortem microbiology in sudden death: sampling protocols proposed in different clinical settings, 2019. ISSN 14690691.

[87] Stefan Riedel. The value of postmortem microbiology cultures, 2014. ISSN 1098660X.

[88] Daniel McDonald, Morgan N. Price, Julia Goodrich, Eric P. Nawrocki, Todd Z. Desantis, Alexander Probst, Gary L. Andersen, Rob Knight, and

Philip Hugenholtz. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME Journal*, 2012. ISSN 17517362. doi: 10.1038/ismej.2011.139.

[89] Pelin Yilmaz, Laura Wegener Parfrey, Pablo Yarza, Jan Gerken, Elmar Pruesse, Christian Quast, Timmy Schweer, Jörg Peplies, Wolfgang Ludwig, and Frank Oliver Glöckner. The SILVA and "all-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Research*, 2014. ISSN 03051048. doi: 10.1093/nar/gkt1209.

[90] Qiong Wang, George M. Garrity, James M. Tiedje, and James R. Cole. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, 2007. ISSN 00992240. doi: 10.1128/AEM.00062-07.

[91] S. Hashavya, I. Gross, A. Michael-Gayego, N. Simanovsky, and R. Lamdan. The efficacy of 16S ribosomal DNA sequencing in the diagnosis of bacteria from blood, bone and synovial fluid samples of children with musculoskeletal infections. *Journal of Children's Orthopaedics*, 2018. ISSN 18632548. doi: 10.1302/1863-2548.12.170049.

[92] Silvana K. Rampini, Guido V. Bloemberg, Peter M. Keller, Andrea C. Büchler, Günter Dollenmaier, Roberto F. Speck, and Erik C. Böttger. Broad-range 16S rRNA gene polymerase chain reaction for diagnosis of culture-negative bacterial infections. *Clinical Infectious Diseases*, 2011. ISSN 10584838. doi: 10.1093/cid/cir692.

[93] Shiqiang Shang, Guoxian Chen, Yidong Wu, Lizhong Du, and Zhengyan Zhao. Rapid diagnosis of bacterial sepsis with PCR amplification and microarray hybridization in 16S rRNA gene. *Pediatric Research*, 2005. ISSN 00313998. doi: 10.1203/01.PDR.0000169580.64191.8B.

[94] Clarissa Oeser, Marcus Pond, Philip Butcher, Alison Bedford Russell, Philipp Henneke, Ken Laing, Timothy Planche, Paul T. Heath, and Kathryn

Harris. PCR for the detection of pathogens in neonatal early onset sepsis. *PLoS ONE*, 2020. ISSN 19326203. doi: 10.1371/journal.pone.0226817.

[95] Dinesh Aggarwal, Tanmay Kanitkar, Michael Narouz, Berge S. Azadian, Luke S.P. Moore, and Nabeela Mughal. Clinical utility and cost-effectiveness of bacterial 16S rRNA and targeted PCR based diagnostic testing in a UK microbiology laboratory network. *Scientific Reports*, 2020. ISSN 20452322. doi: 10.1038/s41598-020-64739-1.

[96] Travis C. Glenn. Field guide to next-generation DNA sequencers. *Molecular Ecology Resources*, 2011. ISSN 1755098X. doi: 10.1111/j.1755-0998.2011.03024.x.

[97] Ruud H. Deurenberg, Erik Bathoorn, Monika A. Chlebowicz, Natacha Couto, Mithila Ferdous, Silvia García-Cobos, Anna M.D. Kooistra-Smid, Erwin C. Raangs, Sigrid Rosema, Alida C.M. Veloo, Kai Zhou, Alexander W. Friedrich, and John W.A. Rossen. Application of next generation sequencing in clinical microbiology and infection prevention. *Journal of Biotechnology*, 2017. ISSN 18734863. doi: 10.1016/j.jbiotec.2016.12.022.

[98] Duncan MacCannell. Next Generation Sequencing in Clinical and Public Health Microbiology. *Clinical Microbiology Newsletter*, 2016. ISSN 18734391. doi: 10.1016/j.clinmicnews.2016.10.001.

[99] Yair Motro and Jacob Moran-Gilad. Next-generation sequencing applications in clinical bacteriology, 2017. ISSN 22147535.

[100] Soumitesh Chakravorty, Danica Helb, Michele Burday, Nancy Connell, and David Alland. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of Microbiological Methods*, 2007. ISSN 01677012. doi: 10.1016/j.mimet.2007.02.005.

[101] Stephen J. Salipante, Dhruba J. Sengupta, Christopher Rosenthal, Gina Costa, Jessica Spangler, Elizabeth H. Sims, Michael A. Jacobs, Samuel I.

Miller, Daniel R. Hoogestraat, Brad T. Cookson, Connor McCoy, Frederick A. Matsen, Jay Shendure, Clarence C. Lee, Timothy T. Harkins, and Noah G. Hoffman. Rapid 16S rRNA Next-Generation Sequencing of Polymicrobial Clinical Samples for Diagnosis of Complex Bacterial Infections. *PLoS ONE*, 2013. ISSN 19326203. doi: 10.1371/journal.pone.0065226.

[102] Saskia Decuypere, Conor J. Meehan, Sandra Van Puyvelde, Tessa De Block, Jessica Maltha, Lompo Palpouguini, Marc Tahita, Halidou Tinto, Jan Jacobs, and Stijn Deborggraeve. Diagnosis of Bacterial Bloodstream Infections: A 16S Metagenomics Approach. *PLoS Neglected Tropical Diseases*, 2016. ISSN 19352735. doi: 10.1371/journal.pntd.0004470.

[103] Li Ping Xia, Long Yan Bian, Min Xu, Ying Liu, Ai Ling Tang, and Wen Qin Ye. 16s rRNA gene sequencing is a non-culture method of defining the specific bacterial etiology of ventilator-associated pneumonia. *International Journal of Clinical and Experimental Medicine*, 2015. ISSN 19405901.

[104] Jill E. Clarridge. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases, 2004. ISSN 08938512.

[105] Nahla Heikal, Roberto H. Nussenzveig, and Archana M. Agarwal. Deparaffinization with mineral oil: A simple procedure for extraction of high-quality DNA from archival formalin-fixed paraffin-embedded samples. *Applied Immunohistochemistry and Molecular Morphology*, 2014. ISSN 15334058. doi: 10.1097/PAI.0b013e3182a77bfe.

[106] Mitsuo Sakamoto, Yi Huang, Mayuko Ohnishi, Makoto Umeda, Isao Ishikawa, and Yoshimi Benno. Changes in oral microbial profiles after periodontal treatment as determined by molecular analysis of 16S rRNA genes. *Journal of Medical Microbiology*, 2004. ISSN 00222615. doi: 10.1099/jmm.0.45576-0.

[107] James J. Kozich, Sarah L. Westcott, Nielson T. Baxter, Sarah K. Highlander, and Patrick D. Schloss. Development of a dual-index sequencing strategy and

curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Applied and Environmental Microbiology*, 2013. ISSN 00992240. doi: 10.1128/AEM.01043-13.

[108] Tanja Magoč and Steven L. Salzberg. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, 2011. ISSN 13674803. doi: 10.1093/bioinformatics/btr507.

[109] Torbjørn Rognes, Tomáš Flouri, Ben Nichols, Christopher Quince, and Frédéric Mahé. VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, 2016. ISSN 21678359. doi: 10.7717/peerj.2584.

[110] Patrick D. Schloss, Sarah L. Westcott, Thomas Ryabin, Justine R. Hall, Martin Hartmann, Emily B. Hollister, Ryan A. Lesniewski, Brian B. Oakley, Donovan H. Parks, Courtney J. Robinson, Jason W. Sahl, Blaz Stres, Gerhard G. Thallinger, David J. Van Horn, and Carolyn F. Weber. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, 2009. ISSN 00992240. doi: 10.1128/AEM.01541-09.

[111] James R. Cole, Qiong Wang, Jordan A. Fish, Benli Chai, Donna M. McGarrell, Yanni Sun, C. Titus Brown, Andrea Porras-Alfaro, Cheryl R. Kuske, and James M. Tiedje. Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Research*, 2014. ISSN 03051048. doi: 10.1093/nar/gkt1244.

[112] Paul J. McMurdie and Susan Holmes. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE*, 8(4):e61217, 2013. URL `http://dx.plos.org/10.1371/journal.pone.0061217`.

[113] Lily Gates, Nigel J. Klein, Neil J. Sebire, and Dagmar G. Alber. Characterising Post-mortem Bacterial Translocation Under Clinical Conditions Using 16S rRNA Gene Sequencing in Two Animal Models. *Frontiers in Microbiology*, 2021. ISSN 1664302X. doi: 10.3389/fmicb.2021.649312.

[114] Vadim Mesli, Christel Neut, and Valery Hedouin. Postmortem bacterial translocation. In David O. Carter, Jeffery K. Tomberlin, M. Eric Benbow, and Jessica L Metcalf, editors, *Forensic Microbiology*. John Wiley & Sons Ltd., 2017.

[115] J. A. Morris, L. M. Harrison, and S. M. Partridge. Postmortem bacteriology: A re-evaluation. *Journal of Clinical Pathology*, 59(1):1–9, 2006. ISSN 00219746. doi: 10.1136/jcp.2005.028183.

[116] Markus M. Heimesaat, Silvia Boelke, André Fischer, Lea Maxie Haag, Christoph Loddenkemper, Anja A. Kühl, Ulf B. Göbel, and Stefan Bereswill. Comprehensive postmortem analyses of intestinal microbiota changes and bacterial translocation in human flora associated mice. *PLoS ONE*, 2012. ISSN 19326203. doi: 10.1371/journal.pone.0040758.

[117] Z. M. Burcham, J. A. Hood, J. L. Pechal, K. L. Krausz, J. L. Bose, C. J. Schmidt, M. E. Benbow, and H. R. Jordan. Fluorescently labeled bacteria provide insight on post-mortem microbial transmigration. *Forensic Science International*, 2016. ISSN 18726283. doi: 10.1016/j.forsciint.2016.03.019.

[118] Cristian Palmiere, Coraline Egger, Guy Prod'Hom, and Gilbert Greub. Bacterial Translocation and Sample Contamination in Postmortem Microbiological Analyses. *Journal of Forensic Sciences*, 2016. ISSN 15564029. doi: 10.1111/1556-4029.12991.

[119] Martin A. Weber, John C. Hartley, Ivan Brooke, Paul E. Lock, Nigel J. Klein, Marian Malone, and Neil J. Sebire. Post-mortem interval and bacteriological culture yield in sudden unexpected death in infancy (SUDI). *Forensic Science International*, 2010. ISSN 03790738. doi: 10.1016/j.forsciint.2010.02.002.

[120] R Core Team. R: A Language and Environment for Statistical Computing, 2020. URL `https://www.r-project.org/`.

[121] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. URL `https://ggplot2.tidyverse.org`.

[122] Harvey Motulsky. GraphPad Prism5.0 Statistics Guide. *Clases*, 2007. ISSN 0003-1305.

[123] Sari Tuomisto, Pekka J. Karhunen, and Tanja Pessi. Time-dependent post mortem changes in the composition of intestinal bacteria using real-time quantitative PCR. *Gut Pathogens*, 2013. ISSN 17574749. doi: 10.1186/1757-4749-5-35.

[124] Kelsey E. Lawrence, Khiem C. Lam, Andrey Morgun, Natalia Shulzhenko, and Christiane V. Löhr. Effect of temperature and time on the thanatomicrobiome of the cecum, ileum, kidney, and lung of domestic rabbits. *Journal of Veterinary Diagnostic Investigation*, 2019. ISSN 19434936. doi: 10.1177/1040638719828412.

[125] Nihal Hasan and Hongyi Yang. Factors affecting the composition of the gut microbiota, and its modulation, 2019. ISSN 21678359.

[126] Embriette R. Hyde, Daniel P. Haarmann, Aaron M. Lynne, Sibyl R. Bucheli, and Joseph F. Petrosino. The living dead: bacterial community structure of a cadaver at the onset and end of the bloat stage of decomposition. *PloS one*, 2013. ISSN 19326203. doi: 10.1371/journal.pone.0077733.

[127] Jessica L Metcalf, Laura Wegener Parfrey, Antonio Gonzalez, Christian L Lauber, Dan Knights, Gail Ackermann, Gregory C Humphrey, Matthew J Gebert, Will Van Treuren, Donna Berg-Lyons, Kyle Keepers, Yan Guo, James Bullard, Noah Fierer, David O Carter, and Rob Knight. A microbial clock provides an accurate estimate of the postmortem interval in a mouse model system. *eLife*, 2013. ISSN 2050-084X. doi: 10.7554/elife.01104.

[128] Jennifer L. Pechal, Tawni L. Crippen, M. Eric Benbow, Aaron M. Tarone, Scot Dowd, and Jeffery K. Tomberlin. The potential use of bacterial community succession in forensics as described by high throughput metagenomic sequencing. *International Journal of Legal Medicine*, 2014. ISSN 09379827. doi: 10.1007/s00414-013-0872-1.

[129] Carolina Probst, Jörn Gethmann, Jens Amendt, Lena Lutz, Jens Peter Teifke, and Franz J. Conraths. Estimating the postmortem interval of wild boar carcasses. *Veterinary Sciences*, 2020. ISSN 23067381. doi: 10.3390/vetsci7010006.

[130] António Raposo1, Esteban Pérez, Catarina Tinoco de Faria, María Antonia Ferrús, and Conrado Carrascosa. Food Spoilage by Pseudomonas spp.- An Overview. In *Food Borne Pathogens and Antibiotic Resistance*. John Wiley & Sons, Inc., 2016.

[131] Juan Carlos Hurtado, Llorenç Quintó, Paola Castillo, Carla Carrilho, Fabiola Fernandes, Dercio Jordao, Lucilia Lovane, Mireia Navarro, Isaac Casas, Rosa Bene, Tacilta Nhampossa, Paula Santos Ritchie, Sónia Bandeira, Calvino Sambo, Valeria Chicamba, Sibone Mocumbi, Zara Jaze, Flora Mabota, Mamudo R. Ismail, Cesaltina Lorenzoni, Assucena Guisseve, Natalia Rakislova, Lorena Marimon, Natalia Castrejon, Ariadna Sanz, Anelsio Cossa, Inacio Mandomando, Khátia Munguambe, Maria Maixenchs, Carmen Muñoz-Almagro, Eusebio Macete, Pedro Alonso, Jordi Vila, Quique Bassat, Clara Menéndez, Miguel J. Martínez, and Jaume Ordi. Postmortem Interval and Diagnostic Performance of the Autopsy Methods. *Scientific Reports*, 2018. ISSN 20452322. doi: 10.1038/s41598-018-34436-1.

[132] I. V.K. Lobmaier, Å Vege, P. Gaustad, and T. O. Rognum. Bacteriological investigation-significance of time lapse after death. *European Journal of Clinical Microbiology and Infectious Diseases*, 2009. ISSN 09349723. doi: 10.1007/s10096-009-0762-0.

[133] Beth Lowe, Terence Marsh, Natasha Isaacs-Cosgrove, Roy Kirkwood, Matti Kiupel, and Martha Mulks. Defining the "core microbiome" of the microbial communities in the tonsils of healthy pigs. *BMC Microbiology*, 2012. ISSN 14712180. doi: 10.1186/1471-2180-12-20.

[134] Matthew K. McIntyre, Trent J. Peacock, Kevin S. Akers, and David M. Burmeister. Initial characterization of the pig skin bacteriome and its effect

on in vitro models of wound healing. *PLoS ONE*, 2016. ISSN 19326203. doi: 10.1371/journal.pone.0166176.

[135] Mikael Lenz Strube, Julie Elvekjær Hansen, Sophia Rasmussen, and Karl Pedersen. A detailed investigation of the porcine skin and nose microbiome using universal and Staphylococcus specific primers. *Scientific Reports*, 2018. ISSN 20452322. doi: 10.1038/s41598-018-30689-y.

[136] A. R. Highet. An infectious aetiology of sudden infant death syndrome, 2008. ISSN 13645072.

[137] John B.A. Okello, Jaymi Zurek, Alison M. Devault, Melanie Kuch, Andrew L. Okwi, Nelson K. Sewankambo, Gabriel S. Bimenya, Debi Poinar, and Hendrik N. Poinar. Comparison of methods in the recovery of nucleic acids from archival formalin-fixed paraffin-embedded autopsy tissues. *Analytical Biochemistry*, 2010. ISSN 00032697. doi: 10.1016/j.ab.2010.01.014.

[138] Vincent J. Gnanapragasam. Unlocking the molecular archive: The emerging use of formalin-fixed paraffin-embedded tissue for biomarker research in urological cancer, 2010. ISSN 14644096.

[139] Sarah M. Hykin, Ke Bi, and Jimmy A. McGuire. Fixing formalin: A method to recover genomic-scale DNA sequence data from formalin-fixed museum specimens using high-throughput sequencing. *PLoS ONE*, 2015. ISSN 19326203. doi: 10.1371/journal.pone.0141579.

[140] Cornelis J.J. Huijsmans, Jan Damen, Johannes C. Van Der Linden, Paul H.M. Savelkoul, and Mirjam H.A. Hermans. Comparative analysis of four methods to extract DNA from paraffin-embedded tissues: Effect on downstream molecular applications. *BMC Research Notes*, 2010. ISSN 17560500. doi: 10.1186/1756-0500-3-239.

[141] Eric Bonnet, Marie Laure Moutet, Céline Baulard, Delphine Bacq-Daian, Florian Sandron, Lilia Mesrob, Bertrand Fin, Marc Delépine, Marie Ange Palomares, Claire Jubin, Hélène Blanché, Vincent Meyer, Anne Boland,

Robert Olaso, and Jean François Deleuze. Performance comparison of three DNA extraction kits on human whole-exome data from formalin-fixed paraffin-embedded normal and tumor samples. *PLoS ONE*, 2018. ISSN 19326203. doi: 10.1371/journal.pone.0195471.

[142] Jianghai Lin, Stephen H. Kennedy, Therese Svarovsky, Jeffrey Rogers, Joseph W. Kemnitz, Anlong Xu, and Krina T. Zondervan. High-quality genomic DNA extraction from formalin-fixed and paraffin-embedded samples deparaffinized using mineral oil. *Analytical Biochemistry*, 2009. ISSN 00032697. doi: 10.1016/j.ab.2009.08.016.

[143] Burcu Sengüven, Emre Baris, Tulin Oygur, and Mehmet Berktas. Comparison of methods for the extraction of DNA from formalin-fixed, paraffin-embedded archival tissues. *International Journal of Medical Sciences*, 2014. ISSN 14491907. doi: 10.7150/ijms.8842.

[144] Samantha J. Mcdonough, Aditya Bhagwate, Zhifu Sun, Chen Wang, Michael Zschunke, Joshua A. Gorman, Karla J. Kopp, and Julie M. Cunninghamid. Use of FFPE-derived DNA in next generation sequencing: DNA extraction methods. *PLoS ONE*, 2019. ISSN 19326203. doi: 10.1371/journal.pone.0211400.

[145] Pauline Robbe, Niko Popitsch, Samantha J.L. Knight, Pavlos Antoniou, Jennifer Becq, Miao He, Alexander Kanapin, Anastasia Samsonova, Dimitrios V. Vavoulis, Mark T. Ross, Zoya Kingsbury, Maite Cabes, Sara D.C. Ramos, Suzanne Page, Helene Dreau, Kate Ridout, Louise J. Jones, Alice Tuff-Lacey, Shirley Henderson, Joanne Mason, Francesca M. Buffa, Clare Verrill, David Maldonado-Perez, Ioannis Roxanis, Elena Collantes, Lisa Browning, Sunanda Dhar, Stephen Damato, Susan Davies, Mark Caulfield, David R. Bentley, Jenny C. Taylor, Clare Turnbull, and Anna Schuh. Clinical whole-genome sequencing from routine formalin-fixed, paraffin-embedded specimens: pilot study for the 100,000 Genomes Project. *Genetics in Medicine*, 2018. ISSN 15300366. doi: 10.1038/gim.2017.241.

[146] Yuzhu Zhi, Daisuke Sasai, Yoichiro Okubo, Minoru Shinozaki, Haruo Nakayama, Somay Yamagata Murayama, Megumi Wakayama, Tadashi Ide, Zean Zhang, and Kazutoshi Shibuya. Comparison between the effectiveness of polymerase chain reaction and In situ hybridization in detecting the presence of pathogenic fungi by using the preserved DNA in formalin-fixed and paraffin-embedded tissues. *Japanese Journal of Infectious Diseases*, 2013. ISSN 13446304. doi: 10.7883/yoken.66.173.

[147] Noriko Watanabe, Shin-ichiro Ohno, Moe Sakuma, Mayo Kuriwaki, Mai Miura, and Masahiko Kuroda. A case report on death from acute bacterial cholangitis accompanied by von Meyenburg complexes. *Medicine*, 2021. ISSN 0025-7974. doi: 10.1097/md.0000000000025526.

[148] Zsolt Jobbagy, Claire B. Fabian, Vincent A. Memoli, and Joseph D. Schwartzman. A novel Streptococcus organism identified in a case of fulminant endocarditis using 16S rDNA sequencing. *Journal of Molecular Diagnostics*, 2004. ISSN 15251578. doi: 10.1016/S1525-1578(10)60503-X.

[149] Christopher J. Stewart, Roxana Fatemizadeh, Pamela Parsons, Christopher A. Lamb, Deborah A. Shady, Joseph F. Petrosino, and Amy B. Hair. Using formalin fixed paraffin embedded tissue to characterize the preterm gut microbiota in necrotising enterocolitis and spontaneous isolated perforation using marginal and diseased tissue. *BMC Microbiology*, 2019. ISSN 14712180. doi: 10.1186/s12866-019-1426-6.

[150] Ines Pinto-Ribeiro, Rui M. Ferreira, Joana Pereira-Marques, Vanessa Pinto, Guilherme Macedo, Fátima Carneiro, and Ceu Figueiredo. Evaluation of the use of formalin-fixed and paraffin- embedded archive gastric tissues for microbiota characterization using next-generation sequencing. *International Journal of Molecular Sciences*, 2020. ISSN 14220067. doi: 10.3390/ijms21031096.

[151] Pamela Tozzo, Salvatore Scrivano, Matteo Sanavio, and Luciana Caenazzo. The role of DNA degradation in the estimation of post-mortem interval: A

systematic review of the current literature. *International Journal of Molecular Sciences*, 2020. ISSN 14220067. doi: 10.3390/ijms21103540.

[152] Walter Bär, Adelgunde Kratzer, Marco Mächler, and Werner Schmid. Post-mortem stability of DNA. *Forensic Science International*, 1988. ISSN 03790738. doi: 10.1016/0379-0738(88)90118-1.

[153] Fiona Fouhy, Adam G. Clooney, Catherine Stanton, Marcus J. Claesson, and Paul D. Cotter. 16S rRNA gene sequencing of mock microbial populations-impact of DNA extraction method, primer choice and sequencing platform. *BMC Microbiology*, 2016. ISSN 14712180. doi: 10.1186/s12866-016-0738-z.

[154] Jethro S. Johnson, Daniel J. Spakowicz, Bo Young Hong, Lauren M. Petersen, Patrick Demkowicz, Lei Chen, Shana R. Leopold, Blake M. Hanson, Hanako O. Agresta, Mark Gerstein, Erica Sodergren, and George M. Weinstock. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, 2019. ISSN 20411723. doi: 10.1038/s41467-019-13036-1.

[155] David A.W. Soergel, Neelendu Dey, Rob Knight, and Steven E. Brenner. Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. *ISME Journal*, 2012. ISSN 17517362. doi: 10.1038/ismej.2011.208.

[156] Anna Klindworth, Elmar Pruesse, Timmy Schweer, Jörg Peplies, Christian Quast, Matthias Horn, and Frank Oliver Glöckner. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*, 2013. ISSN 03051048. doi: 10.1093/nar/gks808.

[157] Alan W. Walker, Jennifer C. Martin, Paul Scott, Julian Parkhill, Harry J. Flint, and Karen P. Scott. 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice. *Microbiome*, 2015. ISSN 20492618. doi: 10.1186/s40168-015-0087-4.

[158] Juan Jovel, Jordan Patterson, Weiwei Wang, Naomi Hotte, Sandra O'Keefe, Troy Mitchel, Troy Perry, Dina Kao, Andrew L. Mason, Karen L. Madsen, and Gane K.S. Wong. Characterization of the gut microbiome using 16S or shotgun metagenomics. *Frontiers in Microbiology*, 2016. ISSN 1664302X. doi: 10.3389/fmicb.2016.00459.

[159] Susannah J. Salter, Michael J. Cox, Elena M. Turek, Szymon T. Calus, William O. Cookson, Miriam F. Moffatt, Paul Turner, Julian Parkhill, Nicholas J. Loman, and Alan W. Walker. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, 2014. ISSN 17417007. doi: 10.1186/s12915-014-0087-z.

[160] Eliana P. Velásquez-Mejía, Jacobo de la Cuesta-Zuluaga, and Juan S. Escobar. Impact of DNA extraction, sample dilution, and reagent contamination on 16S rRNA gene sequencing of human feces. *Applied Microbiology and Biotechnology*, 2018. ISSN 14320614. doi: 10.1007/s00253-017-8583-z.

[161] L. F. Stinson, J. A. Keelan, and M. S. Payne. Identification and removal of contaminating microbial DNA from PCR reagents: impact on low-biomass microbiome analyses. *Letters in Applied Microbiology*, 2019. ISSN 1472765X. doi: 10.1111/lam.13091.

[162] Baroness Helena Kennedy. Sudden unexpected death in infancy and childhood: Multi-agency guidelines for care and investigation. *The Royal College of Pathologists*, 1(November):1–107, 2016. URL `www.rcpath.org`.

[163] Holly Lutz, Alexandria Vangelatos, Neil Gottel, Antonio Osculati, Silvia Visona, Sheree J. Finley, Jack A. Gilbert, and Gulnaz T. Javan. Effects of Extended Postmortem Interval on Microbial Communities in Organs of the Human Cadaver. *Frontiers in Microbiology*, 2020. ISSN 1664302X. doi: 10.3389/fmicb.2020.569630.

[164] Sidney P. Walker, Maurice Barrett, Glenn Hogan, Yensi Flores Bueso, Marcus J. Claesson, and Mark Tangney. Non-specific amplification of human

DNA is a major challenge for 16S rRNA gene sequence analysis. *Scientific Reports*, 2020. ISSN 20452322. doi: 10.1038/s41598-020-73403-7.

[165] Ian C. Michelow, Kurt Olsen, Juanita Lozano, Nancy K. Rollins, Lynn B. Duffy, Thedi Ziegler, Jaana Kauppila, Maija Leinonen, and George H. McCracken. Epidemiology and Clinical Characteristics of Community-Acquired Pneumonia in Hospitalized Children. *Pediatrics*, 2004. ISSN 00314005. doi: 10.1542/peds.113.4.701.

[166] Fang Yih Liaw, Chih Chien Wang, Yaw Wen Chang, and Shyi Jou Chen. Community-acquired streptococcus viridans pneumonia in a healthy child. *Indian Pediatrics*, 2012. ISSN 00196061.

[167] Christopher D. Doern and Carey Ann D. Burnham. It's not easy being green: The viridans group streptococci, with a focus on pediatric clinical manifestations, 2010. ISSN 00951137.

[168] M. Freitas, A. Castelo, G. Petty, C. E. Gomes, and E. Carvalho. Viridans streptococci causing community acquired pneumonia. *Archives of Disease in Childhood*, 2006. ISSN 00039888. doi: 10.1136/adc.2006.094847.

[169] Sun Ha Choi, Seung Ick Cha, Keum Ju Choi, Jae Kwang Lim, Hyewon Seo, Seung Soo Yoo, Jaehee Lee, Shin Yup Lee, Chang Ho Kim, and Jae Yong Park. Clinical characteristics of community-acquired viridans streptococcal pneumonia. *Tuberculosis and Respiratory Diseases*, 2015. ISSN 17383536. doi: 10.4046/trd.2015.78.3.196.

[170] Jennifer L. Pechal, Carl J. Schmidt, Heather R. Jordan, and M. Eric Benbow. A large-scale survey of the postmortem human microbiome, and its potential to provide insight into the living health condition. *Scientific Reports*, 2018. ISSN 20452322. doi: 10.1038/s41598-018-23989-w.

[171] C. L. Liu, H. W. Ai, W. P. Wang, L. Chen, H. B. Hu, T. Ye, X. H. Zhu, F. Wang, Y. L. Liao, Y. Wang, G. Ou, L. Xu, M. Sun, C. Jian, Z. J. Chen, L. Li, B. Zhang, L. Tian, B. Wang, S. Yan, and Z. Y. Sun. Comparison of 16S

rRNA gene PCR and blood culture for diagnosis of neonatal sepsis. *Archives de Pediatrie*, 2014. ISSN 0929693X. doi: 10.1016/j.arcped.2013.11.015.

[172] Axel Ursenbach, Frédéric Schramm, François Séverac, Yves Hansmann, Nicolas Lefebvre, Yvon Ruch, and Xavier Argemi. Revised version (INFD-D-20-00242): impact of 16S rDNA sequencing on clinical treatment decisions: a single center retrospective study. *BMC Infectious Diseases*, 21(1): 1–7, dec 2021. ISSN 14712334. doi: 10.1186/S12879-021-05892-4/TABLES/5. URL `https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-021-05892-4`.

[173] Patrick C.Y. Woo, Dorothy M.W. Tam, Susanna K.P. Lau, Ami M.Y. Fung, and Kwok Yung Yuen. Enterococcus cecorum Empyema Thoracis Successfully Treated with Cefotaxime. *Journal of Clinical Microbiology*, 2004. ISSN 00951137. doi: 10.1128/JCM.42.2.919-922.2004.

[174] Jan Evans Patterson, Anne H. Sweeney, Michael Simms, Nina Carley, Richard Mangi, James Sabetta, and Robert W. Lyons. An analysis of 110 serious enterococcal infections: Epidemiology, antibiotic susceptibility, and outcome. *Medicine (United States)*, 1995. ISSN 15365964. doi: 10.1097/00005792-199507000-00003.

[175] Caroline Rambaud, Michèle Guibert, Elisabeth Briand, Liliane Grangeot-Keros, Aurore Coulomb-L'Herminé, and Michel Dehan. Microbiology in sudden infant death syndrome (SIDS) and other childhood deaths. *FEMS Immunology and Medical Microbiology*, 1999. ISSN 09288244. doi: 10.1016/S0928-8244(99)00072-3.

[176] Gulnaz T. Javan, Sheree J. Finley, Ismail Can, Jeremy E. Wilkinson, J. Delton Hanson, and Aaron M. Tarone. Human Thanatomicrobiome Succession and Time since Death. *Scientific Reports*, 2016. ISSN 20452322. doi: 10.1038/srep29598.

[177] Ismail Can, Gulnaz T. Javan, Alexander E. Pozhitkov, and Peter A. Noble. Distinctive thanatomicrobiome signatures found in the blood and internal

organs of humans. *Journal of Microbiological Methods*, 2014. ISSN 18728359. doi: 10.1016/j.mimet.2014.07.026.

[178] Torri B. Thomas, Sheree J. Finley, Jeremy E. Wilkinson, Daniel J. Wescott, Azriel Gorski, and Gulnaz T. Javan. Postmortem microbial communities in burial soil layers of skeletonized humans. *Journal of Forensic and Legal Medicine*, 2017. ISSN 18787487. doi: 10.1016/j.jflm.2017.05.009.

[179] Ron B. Schifman, Calvin L. Strand, Frederick A. Meier, and Peter J. Howanitz. Blood culture contamination: A College of American Pathologist Q-Probes study involving 640 institutions and 497 134 specimens from adult patients. *Archives of Pathology and Laboratory Medicine*, 1998. ISSN 00039985.

[180] Melvin P. Weinstein, Michael L. Towns, Seth M. Quartey, Stanley Mirrett, Larry G. Reimer, Giovanni Parmigiani, and L. Barth Relier. The clinical significance of positive blood cultures in the 1990s: A prospective comprehensive evaluation of the microbiology, epidemiology, and outcome of bacteremia and fungemia in adults. *Clinical Infectious Diseases*, 1997. ISSN 10584838. doi: 10.1093/clind/24.4.584.

[181] Jeremy W. Pryce, Martin A. Weber, John C. Hartley, Michael T. Ashworth, Marian Malone, and Neil J. Sebire. Difficulties in interpretation of post-mortem microbiology results in unexpected infant death: Evidence from a multidisciplinary survey. *Journal of Clinical Pathology*, 2011. ISSN 00219746. doi: 10.1136/jclinpath-2011-200056.

[182] S. Christoffersen. The importance of microbiological testing for establishing cause of death in 42 forensic autopsies. *Forensic Science International*, 250: 27, may 2015. ISSN 18726283. doi: 10.1016/J.FORSCIINT.2015.02.020. URL `/pmc/articles/PMC7130849//pmc/articles/PMC7130849/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7130849/`.

[183] Katarzyna Jermakow and Marta Rorat. Post-Mortem Microbiology: Retrospective Analysis of Infections Caused by Enterococcus Strains. *Pathogens*,

11(2), feb 2022. ISSN 20760817. doi: 10.3390/PATHOGENS11020204. URL /pmc/articles/PMC8880551//pmc/articles/PMC8880551/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC8880551/.

[184] Gary V. Doern, Karen C. Carroll, Daniel J. Diekema, Kevin W. Garey, Mark E. Rupp, Melvin P. Weinstein, and Daniel J. Sextong. Practical Guidance for Clinical Microbiology Laboratories: A Comprehensive Update on the Problem of Blood Culture Contamination and a Discussion of Methods for Addressing the Problem. *Clinical microbiology reviews*, 33(1), jan 2019. ISSN 1098-6618. doi: 10.1128/CMR.00009-19. URL https://pubmed.ncbi.nlm.nih.gov/31666280/.

[185] Tinzar Basein, Bradley J. Gardiner, Gabriela M. Andujar Vazquez, Andrew S. Joel Chandranesan, Arthur R. Rabson, Shira Doron, and David R. Snydman. Microbial identification using DNA target amplification and sequencing: Clinical utility and impact on patient management. *Open Forum Infectious Diseases*, 2018. ISSN 23288957. doi: 10.1093/ofid/ofy257.

[186] Paul Nathan Goldwater. Infection: The neglected paradigm in SIDS research, 2017. ISSN 14682044.

Appendix

Appendix 1

Case ID	Age	PMI	PM microbiology culture results
*uC1	1.5y	3d	Clostridia (B, S)
			Staphylococcus aureus (Lu)
			Negative (CSF)
uC2	1.5y	5d	Coagulase negative Staphylococci (S)
			Negative (B, Lu, CSF)
uC3	2y	7d	*Streptococcus viridans* (B)
			GAS, GBS (Lu, S)
*uC4	1.5y	6d	*Clostridium sordelli* (Lu, S, ear swabs)
			Actinomyces lignae (B)
			URT commensals (Lu)
			Coagulase negative Staphylococci (S)
uC5	1mo	5d	*Staphylococcus aureus* (B)
			Coagulase negative Staphylococci (B)
			Streptococcus lutetiensis (B, S)
			Negative (Lu, CSF)
uC6	5mo	4d	Negative (B, Lu, S, CSF)
uC7	7d	2d	Negative (B)
			Staphylococcus aureus (S)
uC8	8mo	11d	Negative (B, Lu)
			Coliforms (CSF, laryngeal swab)
uC9	6mo	5d	*Escherichia coli* (Lu)
			Entero faecalis (S)
uC10	3mo	3d	Negative (B, S)
			URT, S. aureus (Lu)
uC11	3y	4d	*Haemophilus parainfluenzae* (Lu)
			Staphylococcus hominis (B)
			Negative (Spleen, CSF)
*uC12	9mo	4d	URT commensals (Lu)

			Negative (S, B, CSF)
uC13	3mo	5d	*Staphylococcus aureus* (Lu)
			Negative (B, S, CSF)
*uC14	1mo	2d	Coliforms (Lu)
uC15	3mo	3d	Coagulase-negative staphylococci (Lu)
			Negative (B, CSF)
uC16	6mo	7d	*Staphylococcus aureus* (Lu)
			Negative (B, S, CSF)
uC17	10mo	1d	*Enterococcus* sp. (B)
			URT commensals (Lu)
			Negative (S, CSF)
			Haemophilus parainfluenzae (ear swab)
uC18	3mo	1d	*Enterococcus faecalis* (B)
			Staphylococcus aureus (Lu)
			Negative (S, CSF)
uC19	10mo	1d	Negative (B, S, CSF)
			URT commensals (Lu)
uC20	4mo	4d	*Staphylococcus aureus* (Lu)
			Klebsiella oxytoca (S)
			Enterobacter clocae (CSF)

Appendix 1

Appendix 2

Case	Age	PMI	Culture
eC1	1.5y	3d	Streptococcus salivarius (B, Lu)
eC2	3.5y	8d	Streptococcus oralis (B, Lu)
			Enterococcus (B)
			Streptococcus pyogenes (CSF)
			Negative (S)
eC3	13y	4d	Enterococcus faecalis (B)
			Negative (Lu)
eC4	7mo	1d	Negative (B, Lu, S, brain swab)
eC5	1y	2d	URT commensals (Lu)
			Negative (S, CSF)
eC6	2y	1d	URT commensals (Lu)
			Negative (B, S, CSF)
eC7	1mo	6d	URT commensals (Lu)
			Coliforms (S)
			Negative (CSF)
eC8	2mo	1d	Negative (B, Lu, S, CSF)
eC9	9mo	1d	Streptococcus (B)
			URT commensals (Lu)
			S.aureus (Lu)
			Negative (S, CSF)
eC10	3y	9d	not performed
eC11	7y	4d	URT commensals (Lu)
			Negative (B, S, CSF)
eC12	11y	4d	Negative (B, Lu, S, CSF)
eC13	4y	1d	not performed
eC14	3y	3d	URT commensals (Lu)
			Coliforms (S)
			Negative (CSF)

eC15	4y	13d	Negative (B, Lu, S, CSF)